

Zastosowanie przetwarzania sygnałów w fuzji danych strumieniowych

Krzysztof Brzostowski



Oficyna Wydawnicza Politechniki Wrocławskiej

Krzysztof Brzostowski

**Zastosowanie
przetwarzania sygnałów
w fuzji danych strumieniowych**



**Oficyna Wydawnicza Politechniki Wrocławskiej
Wrocław 2020**

Recenzenci

Piotr AUGUSTYNIAK

Andrzej DOBROWOLSKI

Opracowanie redakcyjne i korekta

Hanna JUREK

Projekt okładki

Janusz M. SZAFRAN

Wszelkie prawa zastrzeżone. Niniejsza książka, zarówno w całości, jak i we fragmentach, nie może być reprodukowana w sposób elektroniczny, fotograficzny i inny bez zgody wydawcy i właściciela praw autorskich.

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2020

OFICyna WYDAWNICZA POLITECHNIKI WROCLAWSKIEJ

Wybrzeże Wyspiańskiego 27, 50-370 Wrocław

<http://www.oficyna.pwr.edu.pl>

e-mail: oficwyd@pwr.edu.pl

zamawianie.książek@pwr.edu.pl

ISBN 978-83-7493-099-4

Druk i oprawa: beta-druk, www.betadruk.pl

Żonie Katarzynie, córce Tosi

Spis treści

| | |
|--|-----------|
| Przedmowa | 13 |
| Wykaz ważniejszych oznaczeń | 15 |
| I. Wprowadzenie | 19 |
| 1. Fuzja danych strumieniowych | 21 |
| 1.1. Wstęp | 21 |
| 1.2. Osobiste sieci czujnikowe | 23 |
| 1.3. Big data – dane masowe, zmienne i różnorodne | 27 |
| 1.3.1. Big data w sporcie i medycynie | 30 |
| 1.3.2. Big data – wyzwania | 31 |
| 1.4. Fuzja danych | 33 |
| 1.5. Fuzja danych strumieniowych | 35 |
| 1.6. Analiza ruchu człowieka | 35 |
| 1.6.1. Śledzenie aktywności fizycznej | 36 |
| 1.6.2. Rozpoznawanie aktywności fizycznej | 38 |
| 1.6.3. Kinematyczna analiza układu ruchu człowieka | 39 |
| 1.7. Przykłady wykorzystania fuzji danych w analizie ruchu człowieka | 41 |
| II. Fuzja danych strumieniowych | 47 |
| 2. Elementy fuzji danych | 49 |
| 2.1. Wstęp | 49 |
| 2.2. Poziom abstrakcji w systemach fuzji danych | 52 |
| 2.3. Typy konfiguracji źródeł danych | 59 |
| 2.4. Przegląd architektur dla systemów fuzji danych | 61 |
| 2.5. Architektura systemu fuzji danych | 63 |
| 2.6. Modele fuzji danych | 64 |
| 2.6.1. Model JDL | 64 |
| 2.6.2. Model Dasarathy | 66 |
| 2.6.3. Model Boyda | 67 |

| | |
|---|------------|
| 2.6.4. Model typu wodospad | 69 |
| 2.6.5. Model typu Omnibus | 70 |
| 3. Sygnały, pomiary, szumy | 73 |
| 3.1. Wstęp | 73 |
| 3.2. Sygnał | 74 |
| 3.3. Metody reprezentacji sygnałów | 77 |
| 3.3.1. Reprezentacja czasowa sygnałów | 77 |
| 3.3.2. Reprezentacja częstotliwościowa sygnałów | 80 |
| 3.3.3. Reprezentacja czasowo-częstotliwościowa sygnałów | 83 |
| 3.4. Pomiary | 87 |
| 3.5. Szumy | 90 |
| 3.5.1. Szum biały | 93 |
| 3.5.2. Szum kolorowy | 94 |
| 3.5.3. Szum impulsowy | 99 |
| 3.6. Miary jakości sygnałów pomiarowych | 99 |
| 4. Czujniki pomiarowe | 101 |
| 4.1. Wstęp | 101 |
| 4.2. Czujnik przyśpieszenia | 105 |
| 4.2.1. Zakłócenia w czujnikach przyśpieszenia | 107 |
| 4.2.2. Model pomiarowy czujnika przyśpieszenia | 108 |
| 4.2.3. Charakterystyka czujnika przyśpieszenia | 109 |
| 4.2.4. Przykładowe pomiary z wykorzystaniem czujnika przyśpieszenia | 109 |
| 4.3. Żyroskop | 111 |
| 4.3.1. Zakłócenia w czujnikach żyroskopowych | 113 |
| 4.3.2. Model pomiarowy czujnika żyroskopowego | 114 |
| 4.3.3. Charakterystyka | 115 |
| 4.3.4. Przykładowe pomiary z wykorzystaniem czujnika żyroskopowego | 115 |
| 4.4. Czujnik pola magnetycznego (magnetometr) | 116 |
| 4.4.1. Zakłócenia w czujnikach pola magnetycznego | 119 |
| 4.4.2. Model pomiarowy czujnika pola magnetycznego | 119 |
| 4.4.3. Charakterystyka | 119 |
| 4.4.4. Przykładowe pomiary z wykorzystaniem czujnika magnetometrycznego | 120 |
| 4.5. Metody analizy zakłóceń w czujnikach pomiarowych | 120 |
| 4.5.1. Reprezentacja szumu z wykorzystaniem wariancji Allana | 123 |

| | | |
|-----------|--|------------|
| 4.5.2. | Podsumowanie metod modelowania zakłóceń w czujnikach pomiarowych..... | 127 |
| 4.5.3. | Zastosowanie wariancji Allana oraz gęstości widmowej mocy do analizy pomiarów z czujników inercyjnych i pola magnetycznego | 129 |
| 5. | Przetwarzanie wstępne | 133 |
| 5.1. | Wstęp | 133 |
| 5.2. | Normalizacja i standaryzacja sygnałów pomiarowych | 134 |
| 5.3. | Usuwanie szumów o niskich częstotliwościach..... | 135 |
| 5.4. | Usuwanie szumów o wysokich częstotliwościach | 136 |
| 5.5. | Filtr Wienera..... | 138 |
| 5.6. | Metody przetwarzania wstępnego | 145 |
| 5.7. | Podstawy przetwarzania wstępnego z modelem | 147 |
| 5.7.1. | Modele liniowe | 149 |
| 5.7.2. | Modele odcinkowo-liniowe | 151 |
| 5.7.3. | Szereg Volterra | 152 |
| 5.7.4. | Modele blokowo-zorientowane..... | 153 |
| 5.7.5. | Modele NARMAX | 156 |
| 5.7.6. | Model TVARMAX | 159 |
| 5.7.7. | Modele z rzadką reprezentacją sygnału | 161 |
| 5.7.8. | Modele w dziedzinie czasowo-częstotliwościowej | 164 |
| 5.7.9. | Modele w przestrzeni stanów | 175 |
| 5.8. | Modelowanie szumów pomiarowych | 176 |
| 5.8.1. | Modelowanie z wykorzystaniem transformacji falkowej ... | 176 |
| 5.8.2. | Modelowanie z wykorzystaniem empirycznej dekompozycji mod (EMD)..... | 178 |
| 5.8.3. | Modelowanie z wykorzystaniem rzadkiej reprezentacji ... | 180 |
| 5.8.4. | Modelowanie z wykorzystaniem uczenia głębokiego..... | 182 |
| 5.9. | Usuwanie składowych o wysokich częstotliwościach | 184 |
| 6. | Poziom podstawowy: fuzja danych | 189 |
| 6.1. | Wstęp | 189 |
| 6.2. | Estymacja stanu..... | 190 |
| 6.2.1. | Model w przestrzeni stanu | 192 |
| 6.2.2. | Zadanie estymacji stanu | 194 |
| 6.2.3. | Filtr Bayesa | 197 |
| 6.3. | Filtracja optymalna – stan ciągle | 197 |
| 6.3.1. | Filtr Kalmana | 199 |
| 6.3.2. | Rozszerzony filtr Kalmana | 204 |

| | | |
|-----------|---|------------|
| 6.3.3. | Bezśladowy filtr Kalmana | 207 |
| 6.3.4. | Filtr Cząsteczkowy | 214 |
| 6.4. | Filtracja optymalna – stan dyskretny | 217 |
| 6.4.1. | Ukryte Modele Markowa | 218 |
| 6.5. | Fuzja danych z wykorzystaniem filtra Kalmana..... | 220 |
| 7. | Poziom średni: fuzja właściwości | 225 |
| 7.1. | Wstęp | 225 |
| 7.2. | Ekstrakcja cech..... | 226 |
| 7.3. | Segmentacja | 228 |
| 7.3.1. | Metody segmentacji | 233 |
| 7.4. | Generowanie cech..... | 237 |
| 7.4.1. | Metody generowania cech w dziedzinie czasu | 238 |
| 7.4.2. | Metody generowania cech w dziedzinie częstotliwości | 245 |
| 7.4.3. | Metody generowania cech w dziedzinie czasowo-częstotliwościowej | 261 |
| 7.5. | Metody generowania cech nieliniowych | 264 |
| 7.6. | Redukcja wymiaru..... | 280 |
| 7.6.1. | Sformułowanie problemu | 281 |
| 7.6.2. | Selekcja cech | 282 |
| 7.6.3. | Redukcja cech | 285 |
| 7.7. | Klasyfikacja | 300 |
| 7.7.1. | Okno czasowe | 301 |
| 7.7.2. | Metody przetwarzania danych | 302 |
| 7.7.3. | Zbiory danych | 303 |
| 7.7.4. | Metody uczenia | 305 |
| 7.8. | Algorytmy klasyfikacji | 307 |
| 7.8.1. | Algorytm k -Najbliższych Sąsiadów..... | 308 |
| 7.8.2. | Regresja logistyczna | 311 |
| 7.8.3. | Maszyna wektorów nośnych | 311 |
| 7.8.4. | Głębokie sieci neuronowe | 312 |
| 7.9. | Metody oceny jakości algorytmów fuzji właściwości..... | 313 |
| 7.10. | Rozpoznawanie aktywności ruchowej człowieka – fuzja właściwości | 316 |
| 7.10.1. | Algorytm fuzji właściwości | 317 |
| 8. | Poziom wysoki: fuzja decyzji | 325 |
| 8.1. | Wstęp | 325 |
| 8.2. | Podstawy fuzji decyzji | 325 |
| 8.3. | Metody fuzji decyzji | 330 |
| 8.3.1. | Fuzja decyzji na podstawie wielkości dyskretnych | 331 |

| | |
|---|------------|
| 8.3.2. Fuzja decyzji na podstawie wielkości ciągłych | 336 |
| III. Modelowanie układu ruchu człowieka | 341 |
| 9. Ruch ciała sztywnego | 343 |
| 9.1. Wstęp | 343 |
| 9.2. Ciało sztywne w przestrzeni trójwymiarowej | 343 |
| 9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3 | 346 |
| 9.3.1. Reprezentacja oś–kąt | 346 |
| 9.3.2. Kąty Eulera | 347 |
| 9.3.3. Kwaternion | 352 |
| 9.3.4. Współrzędne wykładnicze | 355 |
| 9.4. Ruch ciała sztywnego w przestrzeni trójwymiarowej | 357 |
| 9.4.1. Przekształcenia jednorodne | 358 |
| 9.4.2. Reprezentacja wykładnicza w ruchu ciała sztywnego | 359 |
| 9.5. Prędkość ruchu ciała sztywnego | 362 |
| 10. Kinematyka systemu ciał sztywnych | 365 |
| 10.1. Wstęp | 365 |
| 10.2. Łańcuch kinematyczny dla systemu ciał sztywnych | 365 |
| 10.3. Reprezentacja łańcucha kinematycznego dla systemu ciał sztywnych | 367 |
| 10.3.1. Reprezentacja Denavita–Hartenberga | 367 |
| 10.3.2. Reprezentacja wykładnicza | 368 |
| 10.4. Metody reprezentacji łańcucha kinematycznego: podsumowanie.. | 371 |
| 10.5. Zadanie proste kinematyki | 371 |
| 11. Kinematyka układu ruchu człowieka | 373 |
| 11.1. Wstęp | 373 |
| 11.2. Mechanika układu ruchu człowieka | 373 |
| 11.3. Kinematyczna analiza układu ruchu człowieka | 374 |
| 11.4. Zadanie proste kinematyki w analizie układu ruchu człowieka ... | 377 |
| 11.4.1. Model kinematyki prostej kończyny górnej | 379 |
| 11.4.2. Model kinematyki prostej kończyny dolnej | 383 |
| 11.5. Antropometria | 386 |
| 12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej | 389 |
| 12.1. Wstęp | 389 |
| 12.2. Estymacja orientacji ciała sztywnego | 390 |
| 12.2.1. Zadanie estymacji orientacji ciała sztywnego | 390 |

| | |
|---|------------|
| 12.3. Algorytmy estymacji orientacji ciała sztywnego..... | 391 |
| 12.4. Wykorzystanie filtrów Kalmana do szacowania orientacji obiektu w przestrzeni | 403 |
| 12.4.1. Rozwiązanie z wykorzystaniem filtra Kalmana | 406 |
| 12.4.2. Rozwiązanie z wykorzystaniem rozszerzonego filtra Kalmana | 408 |
| 13. Ilościowa analiza układu ruchu człowieka | 411 |
| 13.1. Wstęp | 411 |
| 13.2. Główne zadania w ilościowej analizie układu ruchu człowieka ... | 411 |
| 13.2.1. Kalibracja | 412 |
| 13.2.2. Dryft | 415 |
| 13.2.3. Zakłócenia magnetyczne | 416 |
| 13.2.4. Kinematyka i ograniczenia biomechaniczne..... | 417 |
| 13.2.5. Ocena wyników ilościowej analizy układu ruchu | 417 |
| 13.3. Śledzenie ruchu w przestrzeni stawu łokciowego | 418 |
| 13.3.1. Estymacja położenia segmentów bez uwzględniania ogra- niczeń anatomicznych | 421 |
| 13.3.2. Estymacja położenia segmentów z uwzględnieniem ogra- niczeń anatomicznych | 424 |
| 13.3.3. Estymacja położenia segmentów bazująca na modelu ki- nematyki | 425 |
| 13.4. Estymacja położenia kończyny górnej i dolnej w przestrzeni | 427 |
| 13.4.1. Kończyna górna | 427 |
| 13.4.2. Kończyna dolna | 429 |
| 13.5. Przykład ilościowej analizy układu ruchu człowieka | 430 |
| 13.5.1. Dane pomiarowe | 430 |
| 13.5.2. Algorytm estymacji parametrów łańcucha kinematyczne- go kończyny górnej | 431 |
| IV. Zastosowania | 433 |
| 14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym | 435 |
| 14.1. Wstęp | 435 |
| 14.2. Algorytm estymacji prędkości chodu | 437 |
| 14.2.1. Estymacja orientacji | 437 |
| 14.2.2. Estymacja prędkości | 438 |
| 14.2.3. Usuwanie dryftu | 438 |
| 14.3. Dane pomiarowe | 443 |

| | |
|---|------------|
| 14.4. Badania eksperymentalne | 444 |
| 14.4.1. Kryteria oceny | 444 |
| 14.4.2. Porównanie z metodą ZUPT | 444 |
| 14.4.3. Porównanie z systemem <i>motion capture</i> | 444 |
| 14.5. Uwagi końcowe | 446 |
| 15. Odszumianie bazujące na reprezentacji rzadkiej | 449 |
| 15.1. Wstęp | 449 |
| 15.2. Algorytm estymacji prędkości chodu oraz pozycji | 450 |
| 15.3. Estymacja dryftu | 451 |
| 15.3.1. Dobór parametrów algorytmu TARA | 453 |
| 15.4. Analiza algorytmu | 454 |
| 15.5. Badania eksperymentalne | 456 |
| 15.6. Wyniki | 457 |
| 15.7. Uwagi końcowe | 460 |
| 16. Odszumianie bazujące na reprezentacji czasowo-częstotliwościowej | 463 |
| 16.1. Wstęp | 463 |
| 16.2. Metodyka | 464 |
| 16.2.1. Algorytm EMD | 464 |
| 16.2.2. Usuwanie zakłóceń z wykorzystaniem metody EMD | 464 |
| 16.3. Algorytm usuwania zakłóceń z wykorzystaniem metody EMD ... | 465 |
| 16.3.1. Sformułowanie problemu | 465 |
| 16.3.2. Algorytm TVD | 465 |
| 16.3.3. Algorytm FLSA | 466 |
| 16.4. Badania eksperymentalne | 467 |
| Bibliografia | 471 |
| Skorowidz | 501 |
| Streszczenie w języku angielskim | 507 |

Przedmowa

Fuzja danych to łączenie danych pochodzących z wielu różnych źródeł w celu uzyskania kompletnego opisu obserwowanego zjawiska. Jest to podejście, które w ostatnich latach zyskało duże znaczenie, a jego przydatność została potwierdzona w wielu obszarach, takich jak: bezprzewodowe sieci czujnikowe, inteligentne systemy transportowe, biometria, systemy telemedycyny, systemy monitorowania środowiska, robotyka, monitorowanie aktywności człowieka i wiele innych. Przez lata wypracowano wiele technik, architektur oraz algorytmów, które umożliwiają konstrukcję systemów fuzji danych ułatwiających osiągnięcie założonych celów w możliwie najlepszy sposób.

Pierwsza część książki dotyczy zagadnień podstawowych związanych z fuzją danych. Omówiono w niej podstawowe architektury systemów fuzji danych oraz algorytmy. Dokonano systematycznego podziału algorytmów fuzji danych, po przyjęciu klasycznego podziału, tj. na fuzję na poziomie podstawowym (fuzja danych), na poziomie średnim (fuzja właściwości) oraz na poziomie wysokim (fuzja decyzji).

W pierwszej części omówiono również zagadnienia związane z metodami reprezentacji sygnałów. Rozpatrzono nie tylko klasyczne metody reprezentacji sygnałów, tj. reprezentację czasową oraz częstotliwościową, ale również metodę stosunkowo nową, tj. reprezentację czasowo-częstotliwościową. Reprezentacja czasowo-częstotliwościowa została omówiona nie tylko na przykładach metod typowych dla tego przypadku, tj. krótkoczasowej transformacji Fouriera czy transformacji falkowej, ale również nowego podejścia, jakim jest transformacja Hilberta–Huanga.

W tej części książki omówiono również zagadnienia związane z przetwarzaniem wstępnym danych pomiarowych, a w szczególności metody usuwania niepożądanych składowych z pomiarów. Po krótkim wprowadzeniu do tematyki tej części omówiono klasyczne narzędzia do rozwiązania tego zadania, a następnie skupiono się na nowoczesnych metodach bazujących na modelowaniu matematycznym. W tej części zaprezentowano również podejścia do modelowania szumów, przy czym skupiono się na metodach nowoczesnych opartych na transformacji falkowej oraz transformacji Hilberta–Huanga, jak również metodach opartych na rzadkiej reprezentacji sygnałów oraz uczeniu głębokim.

Druga część książki podzielona została na dwie części. W pierwszej z nich omówiono podstawowe zagadnienia związane z kinematyczną analizą układu ruchu człowieka. W części tej przedstawiono podstawowe zagadnienia z zakresu robotyki,

Przedmowa

które wykorzystano do konstrukcji m.in. odpowiednich modeli matematycznych. Z kolei w drugiej części omówiono metody ilościowej analizy układu ruchu człowieka oraz sposoby wykorzystania metod fuzji danych w tym obszarze. W ilościowej analizie układu ruchu człowieka można wykorzystać m.in. nowoczesne czujniki pomiarowe, które umożliwiają pomiary przyspieszeń oraz prędkości kątowych. W książce przedstawiono, w jaki sposób dane pochodzące z tych czujników, w połączeniu z odpowiednio zaprojektowanymi algorytmami, mogą zostać wykorzystane do analizy układu ruchu człowieka.

W pracy zaprezentowano również oryginalne wyniki prac autora w zakresie fuzji danych i przetwarzania sygnałów. Jest to tematem trzeciej części książki. Omówiono w niej algorytmy i metody usuwania szumów nisko- i wysokoczęstotliwościowych na potrzeby ilościowej analizy układu ruchu człowieka.

Pragnę podziękować osobom, które przyczyniły się do powstania tej książki. Przede wszystkim dziękuję Profesorowi Jerzemu Świątkowi za cenne uwagi przekazane podczas pracy nad rękopisem. Chciałbym również podziękować studentom – Panom mgr. inż. Łukaszowi Leśniczowskiemu oraz mgr. inż. Piotrowi Szwachowi za owocną współpracę, której wyniki zostały wykorzystane w książce do zilustrowania poruszanych w niej zagadnień.

Krzysztof Brzostowski

Wykaz ważniejszych oznaczeń

- $s(k)$ – sygnał
- \mathbb{R} – zbiór liczb rzeczywistych
- S_{FT} – transformata Fouriera sygnału $s(k)$
- f – częstotliwość
- t – czas
- $A(f)$ – widmo amplitudowe sygnału $s(k)$
- $\rho(f)$ – widmo fazowe sygnału $s(k)$
- w – okno czasowe
- τ – przesunięcie na osi czasu
- Ψ – funkcja falkowa
- $h(k)$ – funkcja opisująca tor pomiarowy
- $f(k)$ – funkcja opisująca proces
- $n(k)$ – szum procesowy
- $v(k)$ – szum pomiarowy (zakłócenie)
- $y(k)$ – pomiar sygnału wyjściowego
 - \mathbf{a} – wektor przyśpieszenia
 - \mathbf{y}_a – pomiar przyśpieszenia
 - ω – wektor prędkości kątowych
 - \mathbf{y}_ω – pomiar prędkości kątowej
 - \mathbf{v} – wektor prędkości liniowej
 - \mathbf{y}_v – pomiar prędkości liniowej
 - m – natężenie pola magnetycznego
 - \hbar – inklinacja (nachylenie) magnetyczna
- \mathbf{y}_m – pomiar natężenia pola magnetycznego
 - b – współczynnik tarcia elementu tłumiącego
 - m – masa ciała
 - k – współczynnik sprężystości elementu sprężystego
 - F – siła wymuszająca
 - \mathbf{x} – wektor stanu
 - x – oś x układu współrzędnych w przestrzeni
 - y – oś y układu współrzędnych w przestrzeni
 - z – oś z układu współrzędnych w przestrzeni

Wykaz ważniejszych oznaczeń

- G – transmitancja operatorowa
- $\{G\}$ – układ odniesienia
- $\{B\}$ – układ lokalny
- S_{yy} – gęstość widmowa mocy sygnału pomiarowego y
- R_{yy} – funkcja autokorelacji sygnału pomiarowego y
- σ_{Av}^2 – wariancja Allana
- K – liczba próbek pomiarowych
- Q_{qn} – współczynnik dla błędu kwantyzacji
- Q_{arw} – współczynnik dla błędu ARW
- Q_{bi} – współczynnik dla błędu BI
- C_{bi} – funkcja kosinus całkowity
- Q_{rrw} – współczynnik dla błędu RRW
- Q_{rr} – współczynnik dla błędu RR
- \mathcal{F} – ogólna postać algorytmu estymacji sygnału z zakłóconych pomiarów
- Q – kryterium jakości
- M – liczba parametrów modelu
- e – błąd
- k_{imp} – odpowiedź impulsowa
- \mathcal{P} – liczba punktów pracy
- \mathcal{D}_u – wymiar wektora wejściowego
- \mathcal{D}_x – wymiar wektora stanu
- \mathcal{D}_y – wymiar wektora wyjściowego
- Ξ – macierz rzadka
- \mathbf{A} – macierz systemowa
- \mathbf{B} – macierz wejść
- \mathbf{C} – macierz wyjść
- \mathbf{D} – macierz przenoszenia
- \mathcal{N} – rozkład normalny
- P – funkcja prawdopodobieństwa
- μ – wartość oczekiwana
- \mathcal{W}_{pf} – funkcja ważności (metoda *Monte Carlo*)
- Υ – algorytm ekstrakcji cech
- \mathcal{D}_z – wymiar wektora cech po ekstrakcji
- Υ_{rw} – algorytm redukcji wymiaru
- Υ_{rc} – algorytm redukcji cech
- \mathbf{W}_{pca} – macierz w metodzie PCA
- \mathbf{W}_{ica} – macierz separująca w metodzie ICA
- \mathbf{A}_{mix} – macierz mieszająca w metodzie ICA
- \mathbf{W}_{lda} – macierz rzutowania (ang. *projection matrix*) w metodzie LDA
- \mathbf{p} – wektor położenia punktu w przestrzeni

- R** – macierz obrotu
- SO(3)** – specjalna grupa obrotów w \mathbb{R}^3
- so(3)** – algebra Liego specjalnej grupy obrotów w \mathbb{R}^3
- SE(3)** – specjalna grupa euklidesowa w \mathbb{R}^3
- se(3)** – algebra Liego specjalnej grupy euklidesowej w \mathbb{R}^3
- θ – kąt obrotu
- α – kąt obrotu wokół osi z (kąty Eulera ZYZ)
- β – kąt obrotu wokół osi y (kąty Eulera ZYZ)
- γ – kąt obrotu wokół osi z (kąty Eulera ZYZ)
- φ – kąt obrotu wokół osi z (kąty Eulera RPY)
- ϑ – kąt obrotu wokół osi y (kąty Eulera RPY)
- ψ – kąt obrotu wokół osi z (kąty Eulera RPY)
- q** – kwaternion jednostkowy
- $\check{\omega}$ – wektor osi rotacji
- T** – macierz transformacji jednorodnej (przekształcenie we współrzędnych jednorodnych pomiędzy dwoma układami)
- p** – współrzędne jednorodne (wektorowa reprezentacja punktu w \mathbb{R}^4)
- ξ – wektor przestrzennych prędkości kątowych
- $\hat{\xi}$ – macierz przestrzennych prędkości kątowych
- R̂** – estymata macierzy obrotu
- q̂** – estymata kwaternionu jednostkowego
- \hat{v} – estymator wektora prędkości liniowej
- $\hat{\theta}$ – estymata kąta obrotu

Część I

Wprowadzenie

1. Fuzja danych strumieniowych

Zacznij od robienia tego, co konieczne; potem zrób to, co możliwe; nagle odkryjesz, że dokonałeś niemożliwego.

św. Franciszek z Asyżu

1.1. Wstęp

Rzeczywisty rozwój współczesnych technologii produkcji układów elektronicznych ma wpływ na postęp w różnych obszarach, takich jak technika pomiarowa, metody bezprzewodowej transmisji danych czy też algorytmy przetwarzania danych. Głównymi przyczynami popularności tych technologii są m.in. relatywnie niska cena, niewielkie rozmiary oraz niski pobór energii urządzeń wykonanych w tych technologiach. Istotnym walorem wyróżniającym te urządzenia wykonane na bazie współczesnych technologii wytwarzania jest również to, że umożliwiają one bezprzewodową transmisję danych.

Urządzenia pomiarowe bazujące na tych technologiach znalazły zastosowanie m.in. w urządzeniach technicznych (np. układy regulacji temperatury), medycznych (np. elektrokardiografia), dla sportowców (np. pulsometry), codziennego użytku (np. telefony komórkowe), jak również w rozrywce (np. okulary wirtualnej rzeczywistości). Różnorodność urządzeń pomiarowych, z której wynika wielorakość danych pomiarowych, nakłada dodatkowe wymagania na metody ich przetwarzania. W przypadku takich danych istotną kwestią jest m.in. konieczność łącznego przetwarzania danych pochodzących z różnych czujników. Dodatkową trudnością podczas ich przetwarzania może być również to, że dane takie często zmieniają się, a zbiory, w których są gromadzone, mają duże rozmiary. Za przykład mogą posłużyć zbiory danych, w których gromadzone są dane z urządzeń do przechwytywania obrazu ruchomego.

W pracy tej skupimy się na dwóch obszarach zastosowań nowoczesnych technologii wytwarzania urządzeń pomiarowych oraz metod przetwarzania danych pomiarowych, tj. na sporcie i na medycynie. Obszar zastosowań nowoczesnych

1. Fuzja danych strumieniowych

technologii oraz metod w medycynie rozciąga się od medycznych baz danych, w których przechowywane są dane pacjenta oraz historia jego choroby, przez metody transmisji danych medycznych, aż do metod ich przetwarzania oraz algorytmów wspomagających podejmowanie decyzji. Z jednej strony wymienione technologie wpływają na rozwój medycyny, z drugiej natomiast wymuszają poprawę jakości systemów ochrony zdrowia. Pożądanym skutkiem tak w pierwszym, jak i w drugim przypadku jest wpływ na polepszenie stanu zdrowia ludzi [132, 323, 427].

Wzbogacanie nowo projektowanych systemów ochrony zdrowia o nowoczesne technologie usprawnia zarówno profilaktykę zdrowotną, jak i leczenie osób chorych. Umiejętne ich wykorzystanie umożliwi również poszerzenie zakresu ich zastosowania, na przykład o promocję zdrowego trybu życia (profilaktyka wczesna).

Chcąc wspomagać leczenie osób chorych lub wykrywać choroby na wczesnym etapie rozwoju, a także zachęcać ludzi do zdrowego trybu życia, potrzebne są odpowiednie dane oraz metody ich przetwarzania. Dane te mogą być pewnymi wielkościami mającymi odniesienie np. do fizjologii lub biochemii człowieka. Ułatwiają one rozpoznanie stanu zdrowia pacjenta, ocenę postępu w jego leczeniu lub zmierzenie stopnia aktywności fizycznej monitorowanej osoby.

Wielkości te pozyskiwane są głównie z wykorzystaniem specjalistycznego sprzętu medycznego, z którego pacjent może skorzystać w szpitalu. Rozwój technologii pomiarowych, z którym związana jest miniaturyzacja urządzeń pomiarowych oraz spadek ich cen, sprawia, że są one coraz częściej wykorzystywane w leczeniu osób chorych. Niskie ceny zwiększają dostępność tych urządzeń, natomiast ich niewielkie rozmiary zapewniają wygodę w ich użytkowaniu. Przykładem jest urządzenie wykorzystywane w terapii osób z problemami kardiologicznymi, tj. elektrokardiograf.

W przypadku wykonywania pomiarów z dala od szpitala może okazać się, że konieczna będzie transmisja tych danych. Do transmisji mogą zostać wykorzystane np. współczesne technologie transmisji bezprzewodowej. Zaletą takiego rozwiązania jest to, że transfer danych może być wykonywany w tle, bez wiedzy użytkownika. Rozwiązania takie przydatne są na przykład podczas monitorowania osób chorych, które kontynuują leczenie poza szpitalem. Technologie bezprzewodowej transmisji danych stosowane są z powodzeniem m.in. do monitorowania osób starszych, które mieszkają samotnie.

Gromadzone dane medyczne muszą być nie tylko bezpiecznie przechowywane w medycznych bazach danych, ale również przetwarzane w celu pozyskania informacji istotnych z perspektywy prowadzonej terapii. Dzięki temu możliwa jest ocena postępów w leczeniu pacjenta czy też wczesna reakcja na objawy chorobowe. Projektując algorytmy przetwarzania danych medycznych, należy brać pod uwagę nie tylko to, że są to **dane wrażliwe** (tj. zawierające dane o pacjentach i ich

chorobach), ale również to, że zgromadzonych danych może być dużo oraz mogą one charakteryzować się dużą różnorodnością.

Zastosowanie nowoczesnych technologii informatycznych oraz teleinformatycznych w systemach ochrony zdrowia nie tylko wpływa na usprawnienie procesu pozyskiwania danych oraz ich przetwarzania, ale również na zmianę postawy zarówno pacjentów (podczas leczenia), jak i osób zdrowych (profilaktyka). Jest to krok w stronę **medycyny zaangażowanej** [163]. Medycyna zaangażowana charakteryzuje się tym, że pacjent nie jest biernym odbiorcą zaleceń lekarskich, ale staje się aktywnym użytkownikiem systemu ochrony zdrowia.

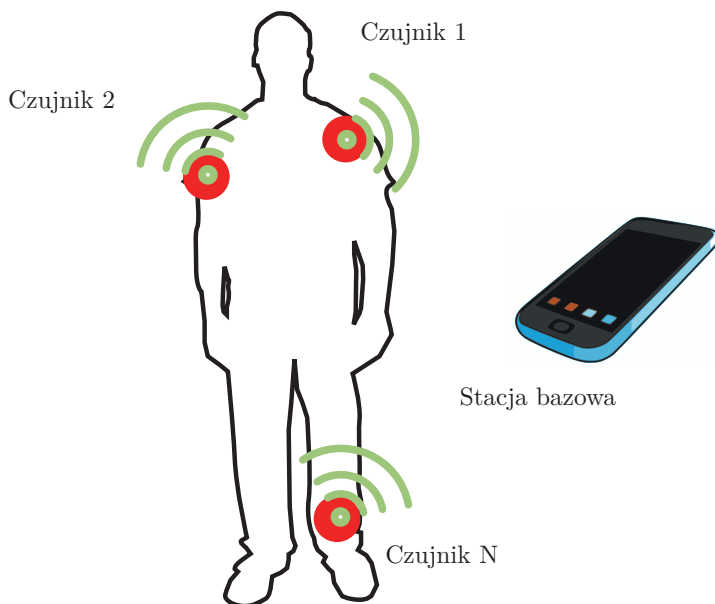
W dalszej części tego rozdziału, stanowiącego wprowadzenie w tematykę współczesnych metod przetwarzania danych na potrzeby medycyny i sportu, omówione zostaną podstawowe zagadnienia związane z pozyskiwaniem oraz transmisją danych pomiarowych. Omówiona zostanie natura tych danych oraz opisane zostaną metody przetwarzania tych danych. Rozdział zakończono przeglądem przykładowych aplikacji z zakresu medycyny i sportu, które są potencjalnymi obszarami zastosowań zaproponowanego w niniejszym opracowaniu podejścia.

1.2. Osobiste sieci czujnikowe

Jedną z technologii, która ma znaczący wpływ na rozwój i upowszechnienie urządzeń elektronicznych są **mikrouządzenia elektromechaniczne** (ang. *Micro-electromechanical Systems*, MEMS) [296]. Współczesna technologia ich wytwarzania umożliwia konstrukcję oraz produkcję półprzewodnikowych przyrządów elektronicznych o zróżnicowanym przeznaczeniu, tj. przetworników sygnałów, czujników (sensorów), elementów wykonawczych (aktuatorów), modułów do bezprzewodowej transmisji danych czy też elementów do przetwarzania danych. Elementy te mogą być łączone w różnych konfiguracjach, zależnych od wymagań projektowych. Tak połączone elementy nazywane są **mikrosystemami**. Nieduży pobór mocy oraz niewielkie rozmiary umożliwiają łączenie ze sobą wielu mikrosystemów, które komunikują się pomiędzy sobą np. przez moduły do bezprzewodowej transmisji danych.

Połączone ze sobą mikrosystemy mogą być podstawą do budowy bezprzewodowych sieci czujnikowych WSN (ang. *Wireless Sensor Networks*). Sieci WSN składają się z wielu elementów, które rozlokowane są na pewnym obszarze i komunikują się pomiędzy sobą z wykorzystaniem technologii bezprzewodowych. Początkowo sieci typu WSN były rozwijane na potrzeby militarne [150]. Małe koszty produkcji oraz miniaturyzacja urządzeń, które związane są z doskonaleniem technologii wytwarzania, sprawiły, że bezprzewodowe urządzenia aktualnie są wykorzystywane masowo również w zastosowaniach cywilnych.

1. Fuzja danych strumieniowych

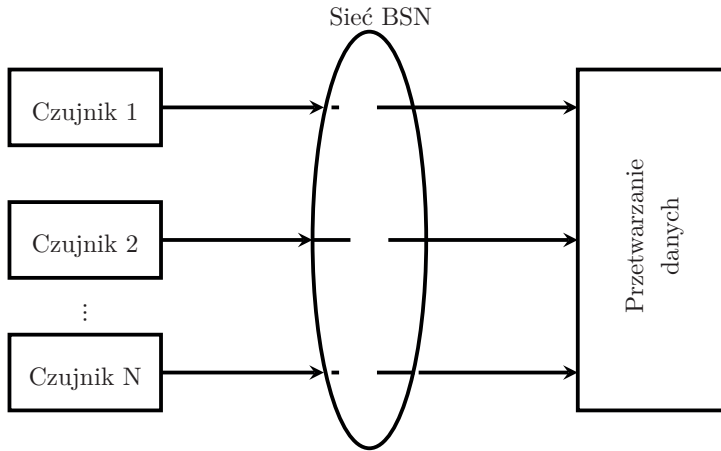


Rys. 1.1. Przykład osobistej sieci czujnikowej (BSN)

Postęp w obszarze technologii WSN umożliwił rozwój m.in. osobistych sieci czujnikowych BSN (ang. *Body Sensor Networks*) (rys. 1.1). Sieci BSN to sieci, w których elementy wchodzące w ich skład, wykorzystywane są do akwizycji danych z czujników służących do monitorowania parametrów fizjologicznych (np. elektrokardiografia, elektroencefalografia czy elektromiografia), kinematycznych człowieka (przyśpieszenie lub prędkość kątowna) lub kinetycznych (związane z generowanymi siłami i momentami sił).

Wykorzystanie sieci BSN do monitorowania człowieka ma różnorodne zastosowania. Sieci takie mogą być stosowane np. do śledzenia aktywności ruchowej człowieka (z wykorzystaniem czujników do pomiaru wielkości fizjologicznych oraz kinematycznych), rozpoznawania wykonywanych przez człowieka czynności (również bazują na czujnikach do pomiaru wielkości fizjologicznych oraz kinematycznych), analizą biomechaniczną (w tym przypadku wykorzystywane są głównie czujniki kinematyczne), czy też analizą generowanych podczas ruchu sił i momentów sił (tutaj zastosowanie mają czujniki kinetyczne).

Przykład sieci BSN przedstawiono na rys. 1.1. Głównym elementem tej sieci jest stacja bazowa. W sieciach BSN stacją bazową może być telefon komórkowy lub palmtop. Integralną częścią sieci są również czujniki, których rodzaj oraz liczba



Rys. 1.2. Koncepcja osobistej sieci czujnikowej (BSN)

zależą od konkretnego zastosowania sieci BSN. Na rysunku 1.2 przedstawiono z kolei schemat sieci BSN dla przykładu z rys. 1.1.

Zaletą rozwiązań wykorzystujących sieci BSN jest to, że akwizycja danych jest nieuciążliwa dla użytkownika, dyskretna i nie ma negatywnego wpływu na komfort codziennego jej użytkowania. Przyczyniają się do tego niewielkie rozmiary czujników pomiarowych, ich waga oraz niskie zużycie energii elektrycznej. Wymienione cechy połączonych w sieci BSN bezprzewodowych czujników sprawiają, że są one często wykorzystywane w medycynie i sporcie [456].

Zastosowanie BSN w sporcie

Zastosowanie osobistych sieci czujnikowych w sporcie związane jest głównie z treningiem sportowym. Warto podkreślić, że powszechność urządzeń mogących wchodzić w skład sieci BSN sprawia, że rozwiązania takie wykorzystywane są w treningu zarówno sportowców profesjonalnych, jak i osób, które sport uprawiają rekreacyjnie.

Rozwiązania bazujące na sieciach BSN mogą być wykorzystywane na różnych etapach treningu sportowego. Trening sportowy to zorganizowany proces, w którego trakcie zawodnik uczy się zarówno techniki (trening techniczny), jak i taktyki danej dyscypliny sportowej (trening taktyczny). Celem treningu jest również rozwinięcie funkcji ustrojowych (trening wytrzymałościowy), które umożliwią uzyskiwanie możliwie najlepszych wyników w danej dyscyplinie [125, 376]. W celu przygotowania mentalnego zawodnika lub grupy zawodników stosuje się również trening psychologiczny [125].

1. Fuzja danych strumieniowych

Z treningiem sportowym związane są takie zadania, jak planowanie, analiza i monitorowanie. Do prawidłowej realizacji każdego z tych zadań potrzebne są dane dotyczące konkretnego sportowca. Źródłem danych mogą być czujniki pomiarowe (np. pracujące w sieciach BSN), jak również człowiek (np. trener oraz inne osoby ze sztabu szkoleniowego).

Sieci BSN mogą być wykorzystywane zarówno w treningu technicznym, jak i wytrzymałościowym. W pracach [21] oraz [135] przedstawiono rozwiązania bazujące na sieciach BSN do wspomagania treningu technicznego golfisty. Innymi istotnymi zagadnieniami są śledzenie aktywności sportowca w trakcie treningu, rozpoznawanie typów wykonywanych ćwiczeń wraz ze zliczaniem liczby ich powtórzeń. W pracach [7, 8] oraz [436] przedstawiono przykładowe rozwiązania, w których wykorzystano sieci BSN w treningu sportowym.

Cechą charakterystyczną rozwiązań bazujących na sieciach BSN jest to, że w trakcie ich działania mogą być gromadzone duże ilości danych. By możliwe było wykorzystanie zebranych pomiarów na potrzeby treningu sportowego, konieczne jest ich wcześniejsze przetworzenie. Przetwarzanie zgromadzonych danych może odbywać się na różnych poziomach. W zależności od wybranego poziomu na danych pomiarowych mogą być wykonywane takie operacje, jak usuwanie niepożądanych składników z pomiarów, normalizacja pomiarów, ich segmentacja czy też ekstrakcja cech z pomiarów.

Gromadzone z wykorzystaniem sieci BSN dane mogą tworzyć nie tylko zbiory o dużych wolumenach. Dane te charakteryzują się również tym, że mogą być to zbiory różnorodne. Często występują w nich również dane nadmiarowe, które mogą szybko się dezaktualizować. Natura tych danych sprawia, że istotnym zagadnieniem staje się opracowanie nowych algorytmów, które będą mogły sprostać zadaniu przetwarzania danych o takiej charakterystyce. Zagadnienie to jest szczególnie istotne w zadaniach, w których konieczne jest przetwarzanie danych w czasie rzeczywistym.

Zastosowanie BSN w medycynie

Osobiste sieci czujnikowe wykorzystywane są również z powodzeniem w medycynie. Często w tym kontekście pojawia się termin **telemedycyna**. Telemedycyna jest formą świadczenia usług medycznych i opieki zdrowotnej, która łączy elementy nowoczesnych technologii teleinformatycznych (np. bezprzewodowa transmisja danych) i informatycznych (np. algorytmy przetwarzania danych) z medycyną [308]. Częstym wymaganiami nakładanymi na rozwiązania telemedyczne jest konieczność pracy w reżimie czasu rzeczywistego. Oznacza to, że na transmisję danych oraz na algorytmy przetwarzania tych danych mogą być nałożone specyficzne dla tego obszaru wymagania.

1.3. *Big data* – dane masowe, zmienne i różnorodne

Telemedycyna może być wykorzystywana m.in. do diagnozowania pacjentów na odległość, zdalnego śledzenia stanu zdrowia pacjenta oraz wspomagania wykonywania wybranych zabiegów lekarskich. Telemedycyna wykorzystywana jest również w profilaktyce i edukacji medycznej oraz do promowania zdrowego trybu życia.

W wielu pracach wskazuje się na korzyści wynikające z wykorzystania sieci BSN w ramach świadczenia usług telemedycznych. W efekcie zastosowania sieci BSN w tym obszarze opracowano wiele systemów do monitorowania pacjentów przewlekle chorych [76, 140, 288, 414], do wspomagania rehabilitacji [70, 149, 188] oraz do śledzenia osób starszych [24, 229, 297].

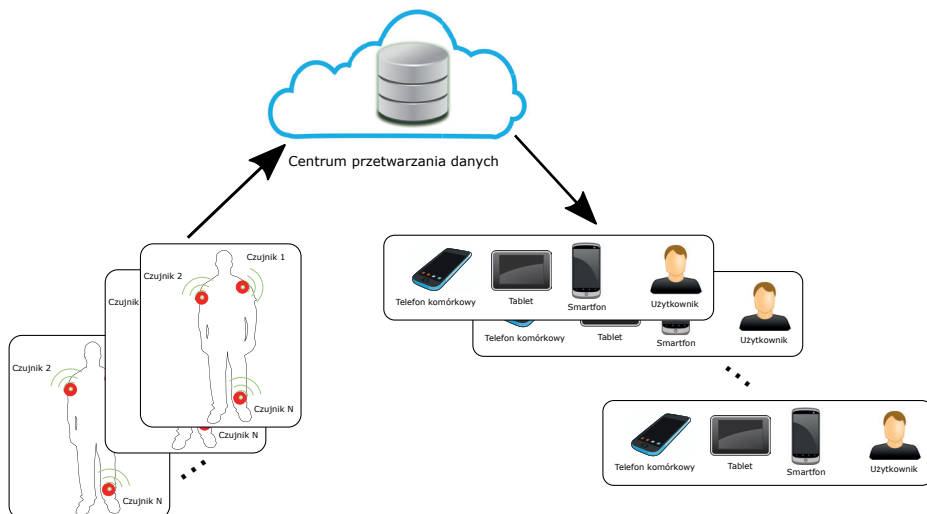
Również i w tym przypadku zastosowanie sieci BSN może sprawić, że zbiory danych będą miały duże rozmiary. Z kolei różnorodność dostępnych urządzeń pomiarowych, które mogą wchodzić w skład sieci BSN, prowadzi do tego, że gromadzone dane będą charakteryzowały się dużą różnorodnością. W wielu przypadkach mogą pojawić się w nich również dane nadmiarowe. Systemy telemedyczne powinny umożliwiać transmisję takich danych, ich przetwarzanie oraz późniejsze przechowywanie.

1.3. **Big data** – dane masowe, zmienne i różnorodne

Rozwój metod i technik pozyskiwania danych z wykorzystaniem czujników pomiarowych wykorzystujących np. sieci BSN, jak również dane gromadzone przez wyszukiwarki internetowe czy też serwisy społecznościowe sprawiają, że liczba generowanych każdego dnia danych jest ogromna. Na przykład w pracy [279] wskazano, że w roku 2012 każdego dnia wytworzono 10^{18} bajtów danych. Przewiduje się, że liczba ta będzie podwajała się co 40 miesięcy. Eksplozja danych związana jest z upowszechnieniem systemów baz danych, sieci bezprzewodowych oraz Internetu. Systemy baz danych stosowane są m.in. w bankowości, handlu, medycynie i innych dziedzinach działalności człowieka. Omawiane już sieci bezprzewodowe, w tym bezprzewodowe sieci czujnikowe BSN, mogą generować duże zbiory danych (rys. 1.3). Również użytkownicy sieci Internet każdego dnia wytwarzają ogromne ilości danych. Dane te generowane są zarówno przez usługę poczty elektronicznej, jak i przez popularne w ostatnich latach różnorodne serwisy społecznościowe.

Dane wytworzone w wyniku działalności człowieka, zarówno w przestrzeni publicznej (np. gospodarka), jak i prywatnej, zawierają wiedzę przydatną w przedsiębiorstwach m.in. do wspomagania podejmowania decyzji biznesowych. Dane te mogą być również wykorzystywane w celu poprawy świadczonych usług lub oferowanych produktów. W wielu przypadkach wiedzę z pozyskanych danych można uzyskać jedynie przez odpowiednie ich przetworzenie.

1. Fuzja danych strumieniowych

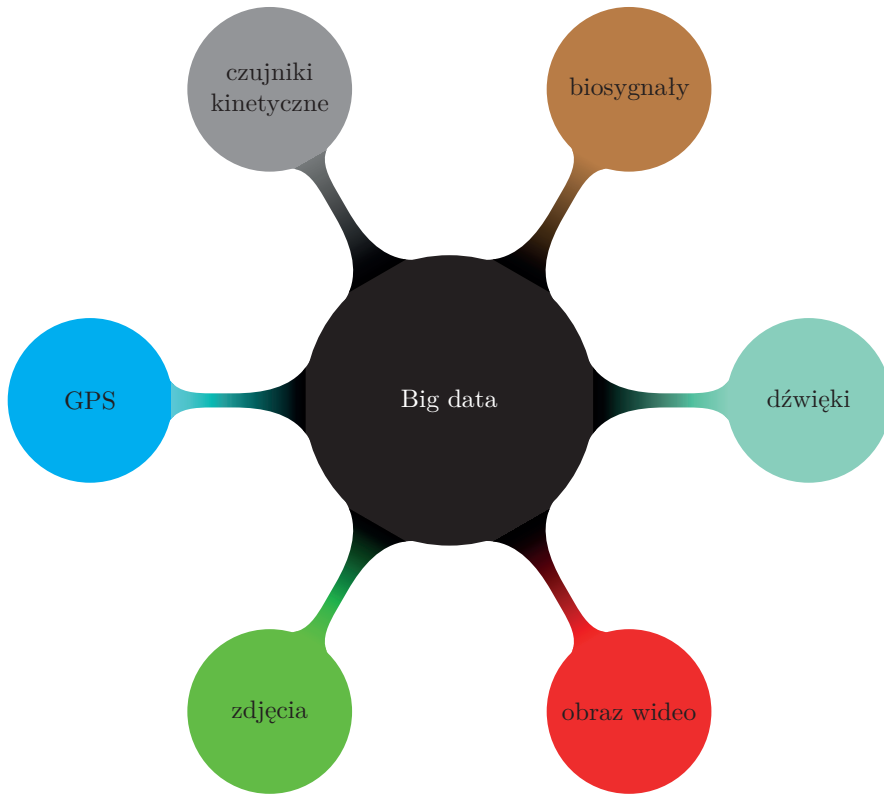


Rys. 1.3. Big data i osobiste sieci czujnikowe

Różnorodność zbiorów danych o dużych wolumenach wynika z tego, że w ich skład mogą wchodzić m.in. zdjęcia, wiadomości przesyłane w serwisach społecznościowych, dane pozyskane z czujników pomiarowych, kamer wideo czy pochodzące z systemów nawigacji satelitarnej takich, jak GPS (ang. *Global Positioning System*, system nawigacji satelitarnej), czy dane zgromadzone w bazach danych sieci handlowych (rys. 1.4). Obserwowany dzisiaj gwałtowny rozwój technologii informatycznych wskazuje na to, że w przyszłości będą pojawiały się kolejne źródła generujące dane nowego typu. Warto w tym miejscu również podkreślić, że duże zbiory danych, dla których projektowane są algorytmy ich przetwarzania, w przyszłości mogą różnić się od aktualnie wykorzystywanych zbiorów nie tylko ilościowo, ale również jakościowo.

Kolejną cechą charakteryzującą gromadzone dzisiaj zbiory danych jest to, że część z nich szybko staje się nieaktualna. Przetwarzanie danych o dużych wolumenach, z których część szybko traci na ważności, staje się krytyczne w przypadku, gdy obliczenia muszą być wykonane w czasie rzeczywistym. Przykładem, gdzie istotne jest przetwarzanie dużych wolumenów danych, które szybko stają się nieaktualne, są metody rozwijane na potrzeby analizy danych z giełdy papierów wartościowych. W innych obszarach, takich jak medycyna i sport, dane pomiarowe dotyczące na przykład stanu zdrowia pacjenta lub poziomu wytrenowania sportowca mogą w wielu przypadkach szybko się dezaktualizować.

1.3. Big data – dane masowe, zmienne i różnorodne



Rys. 1.4. Big data – dane masowe, zmienne oraz różnorodne

Oprócz dwóch podanych cech charakteryzujących wspólnie pozyskiwane dane jest również kolejna, która dotyczy ich nadmiarowości. Nadmiarowość danych wiąże się z m.in. z częstotliwością pozyskiwania danych, która może być wyższa od tej, jaka jest wymagana w rzeczywistości. Jeżeli obniżenie częstotliwości nie jest możliwe, np. z powodów technicznych, rozwiązaniem jest selekcja pozyskiwanych danych. Nadmiarowość danych może nakładać dodatkowe wymagania na algorytmy przetwarzania takich danych. W tym przypadku istotną kwestią staje się zwiększanie liczby wymaganych obliczeń. Przykładowymi źródłami danych, które mogą być nadmiarowe, są m.in. zbiory z obrazami wideo. W celu ograniczenia napływających danych stosuje się np. metody wykrywające istotne zmiany w analizowanym obrazie. Przykładem zastosowania metod selekcji danych oraz wykrywania w nich zmian są systemy monitorowania otoczenia, np. na lotniskach lub w miastach, przez analizę obrazu z wielu kamer wideo.

1. Fuzja danych strumieniowych

Obszarami, które już teraz wymagają przetwarzania danych o podanej powyżej charakterystyce, są m.in.:

- wojsko,
- bezpieczeństwo, np. monitorowanie miast lub lotnisk,
- finanse, bankowość i giełda,
- transport, np. towarowy,
- monitorowanie środowiska,
- sport, zdrowie i służba zdrowia.

Prekursorami w obszarze przetwarzania dużych wolumenów danych charakteryzujących się zmiennością i różnorodnością są firmy takie, jak Google[®], Amazon[™] czy Facebook[®]. Potrzebę przetwarzania tego typu danych zauważają coraz częściej również instytucje finansowe, takie jak banki czy firmy telekomunikacyjne (np. operatorzy telefonii komórkowej). W wymienionych firmach generatorem danych są głównie klienci tych podmiotów gospodarczych. Przetwarzanie i późniejsza analiza danych generowanych przez klientów może ułatwić m.in. przewidywanie zachowań klienta, co może wpłynąć na poprawę jakości oferowanych usług.

1.3.1. Big data w sporcie i medycynie

Problem dużych zbiorów danych, które m.in. ze względu na brak ujednoczonych standardów zapisu i przechowywania danych charakteryzują się różnorodnością, jest również widoczny w obszarze związanym ze sportem i medycyną. Za przykład mogą posłużyć dane pochodzące z ultrasonografu (tzw. USG), obrazowania metodą rezonansu magnetycznego (ang. *Magnetic Resonance Imaging*, MRI), tomografii pozytonowo-emisyjnej (ang. *Positron Emission Tomography*, PET), kardiomonitorów itp. Takimi urządzeniami są również np. długoczasowe rejestratory EKG (holtery EKG). Dane pochodzące z tych urządzeń, ze względu na specyfikę obszaru ich zastosowania, mogą być źródłem zbiorów o dużych wolumenach, jednakże nie są to zbiory szybko dezaktualizujące się. Natomiast w zaawansowanych konstrukcjach robotów chirurgicznych, takich jak ZEUS [386] lub *da Vinci* [424], wymaga się, by opóźnienia związane z przetwarzaniem danych (np. z kamer) były niewielkie. Kamery te przechwytyują obraz z pola operacyjnego, który przekazują po ich przetworzeniu do specjalnych okularów lub binokularów używanych przez chirurgów. Tego typu zadania bardzo często wymagają przetwarzania dużych wolumenów danych, które w tym przypadku szybko zmieniają się.

Również w sporcie coraz częściej różnorodne dane wykorzystywane są m.in. do wspomagania treningu sportowego. Zarówno w laboratoriach akademickich, jak i klubach sportowych stosuje się systemy typu *motion capture* (np. BTS) do analizy ruchu sportowca, platformy do analizy chodu i postawy ciała (np.

1.3. Big data – dane masowe, zmienne i różnorodne

platformy Kistler) czy do pomiarów sił (np. *BiodeX*). W treningu stosuje się również kamery wideo (np. system *SportVU* firmy STATS).

Alternatywą dla typowych rozwiązań bazujących na przetwarzaniu obrazu wideo są czujniki pomiarowe typu IMU (ang. *Inertial Measurement Unit*) oraz MARG (ang. *Magnetic Angular Rate and Gravity*). Czujniki pomiarowe typu IMU charakteryzują się tym, że do ich budowy wykorzystuje się trójosiowe czujniki przyspieszenia oraz trójosiowe czujniki żyroskopowe. Urządzenia pomiarowe MARG oprócz akcelerometrów i żyroskopów wyposażone są dodatkowo w czujnik natężenia pola magnetycznego (magnetometr). Czujniki te zyskują dużą popularność i są chętnie stosowane w zagadnieniach związanych m.in. z analizą ruchu na potrzeby treningu sportowego (niska cena, wygodne w użytkowaniu, możliwość stosowania poza laboratorium). W typowych aplikacjach wykorzystujących czujniki IMU oraz MARG dane pomiarowe charakteryzują się dużą zmiennością. Jest to związane z dużą częstotliwością próbkowania, która może wynosić, w aktualnie dostępnych na rynku urządzeniach, nawet 1 kHz [367].

Dane pozyskane z czujników typu IMU lub MARG mogą być nie tylko stosowane jako alternatywa do innych rozwiązań, ale również jako uzupełnienie innych danych, np. pochodzących z kamer wideo. Wówczas tak pozyskane dane będą charakteryzować się nie tylko zmiennością, ale również różnorodnością. Generowane przez systemy pomiarowe składające się zarówno z czujników typu IMU lub MARG, jak i kamer wideo, zbiory danych będą również charakteryzowały się dużą objętością.

1.3.2. Big data – wyzwania

Duże nadmiarowe zbiory danych, które często zmieniają się, należy przetwarzać, wykorzystując metody uwzględniające specyfikę tych zbiorów. Omówimy teraz istotne zagadnienia, które związane są z projektowaniem algorytmów przetwarzania zbiorów danych o takiej charakterystyce [204].

Jedną z ważniejszych kwestii związanych z przetwarzaniem dużych zbiorów zawierających dane pomiarowe jest problem niepożądanych składników w tych danych. Zagadnienie to dotyczy zarówno danych pochodzących z systemu GPS, czujników pomiarowych (np. IMU, MARG etc.) czy specjalistycznych kamer wideo oraz różnorodnych urządzeń medycznych, takich jak wspomniane wcześniej MRI lub PET. Projektując algorytmy przetwarzania danych pomiarowych, należy uwzględnić specyfikę niepożądanych składników, jak również charakterystykę samych danych (np. obraz statyczny lub ruchomy, szereg czasowy).

Kolejne zagadnienie, które należy uwzględnić podczas przetwarzania danych, dotyczy ich niespójności. Problem ten jest szczególnie istotny, gdy wyniki prze-

1. Fuzja danych strumieniowych

tworzenia tych danych będą wykorzystane np. do wspomagania podejmowania decyzji.

Ponieważ gromadzone dane mogą pochodzić z różnych źródeł, kolejnym zagadnieniem do rozpatrzenia przez projektanta jest problem lokalizacji źródeł danych. W ogólności źródła te mogą być skupione, oznacza to wówczas, że mamy do czynienia ze zwartym i scentralizowanym systemem gromadzenia i przetwarzania danych. Z drugiej strony źródła danych mogą być rozproszone, np. w sieci komputerowej (np. sieci BSN). Taki system nazywamy systemem rozproszonym. Oczywiście, metody i techniki przetwarzania pozyskanych danych pomiarowych należy dostosować do projektowanego systemu. W systemach scentralizowanych wyniki pomiarów przesyłane są do jednostki centralnej, gdzie następnie są przetwarzane. W systemach zdecentralizowanych zazwyczaj nie przesyła się wszystkich wyników pomiarów pozyskanych z czujników pomiarowych, a jedynie wyniki ich przetwarzania. Takie podejście może być wymuszone ograniczonymi możliwościami transmisji danych, redukcją czasu ich przesyłu lub koniecznością racjonalnego wykorzystania energii w urządzeniach pomiarowych.

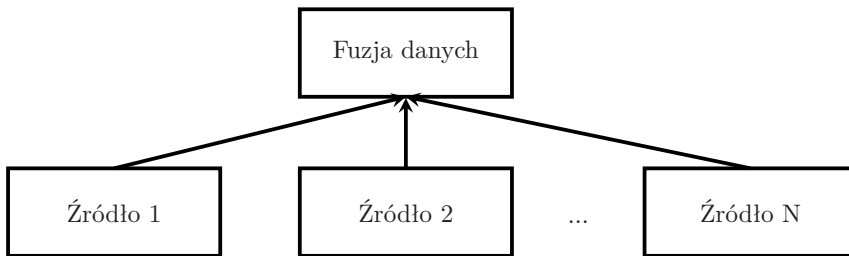
Innym równie ważnym zagadnieniem, szczególnie w rozwiązaniach, gdzie istotne jest pozyskanie wyników w czasie rzeczywistym, jest uwzględnienie częstotliwości, z jaką pozyskiwane są dane pomiarowe (ang. *multirate systems*). Jest to kwestia szczególnie istotna w przypadku systemów, w których dane pochodzą z różnych źródeł, ponieważ każde ze źródeł może pozyskiwać dane z inną częstotliwością. Ponadto, jeżeli dane przesyłane są do jednostki centralnej, kolejnym problemem mogąącym komplikować akwizycję i późniejsze ich przetwarzanie są opóźnienia związane z transmisją danych.

Innym zagadnieniem związanym z gromadzeniem danych z różnych źródeł jest to, że dane te mogą być ze sobą skorelowane. Przetwarzanie danych skorelowanych wiąże się z dodatkowymi obliczeniami, które wpływają niekorzystnie na wydajność systemów. Dlatego istotne jest wykrywanie korelacji i uwzględnianie ich na etapie transmisji lub wstępnego przetwarzania danych.

Istotnym problemem, który należy również rozważyć, projektując algorytmy dla zbiorów typu *big data*, jest wiarygodność źródła danych [278, 447]. Wiarygodność źródła danych oceniana jest np. na podstawie jakości dostarczanych danych oraz na podstawie jakości kanału, którym są one dostarczane. Istotne stają się wówczas kwestie związane z konstrukcją metod szacowania wiarygodności źródeł. Warto zaznaczyć, że metody te mogą być wykorzystane z jednej strony do badania jakości źródeł danych, z drugiej strony natomiast mogą służyć poprawie jakości tych danych, które dostarczane są przez różne źródła. Kwestie te są bardzo istotne w ujęciu jakości uzyskiwanych wyników, ponieważ wykorzystanie niepewnych danych z niezwyfikowanego źródła może prowadzić do błędnych obliczeń, które wykorzystują dane tego typu.

1.4. Fuzja danych

Opisanymi zagadnieniami związanymi z organizacją przetwarzania danych masowych, zmiennych, różnorodnych oraz nadmiarowych zajmuje się m.in. **fuzja danych** (ang. *data fusion*) (rys. 1.5).



Rys. 1.5. Idea fuzji danych

Definicja 1.4.1 (Fuzja danych). *Fuzja danych obejmuje teorie, metody, techniki i narzędzia wykorzystywane do łączenia danych pochodzących z różnych źródeł w celu uzyskania opisu lub zrozumienia obiektu, które wpływają na poprawę jakości wspomaganie podejmowania decyzji lub sterowania [337].*

Cechą charakterystyczną fuzji danych jest to, że uzyskany opis obiektu jest pełniejszy od opisu uzyskanego na podstawie niezależnego przetwarzania i niezależnej analizy danych pochodzących z różnych źródeł. Źródłem danych w fuzji danych mogą być [337]:

- czujniki pomiarowe,
- bazy danych,
- bazy wiedzy,
- wiedza człowieka.

Warto zaznaczyć, że fuzja danych jest również możliwa w przypadku danych pochodzących z jednego czujnika pomiarowego. W tym przypadku do fuzji wykorzystuje się dane z czujnika pochodzące z różnych momentów [289].

Ważnym pojęciem w fuzji danych jest **synergia**. Zgodnie z [114] synergia to „współdziałanie różnych czynników, skuteczniejsze niż suma ich oddzielnych działań”. Metody fuzji danych umożliwiają osiągnięcie efektu synergii z wykorzystaniem danych pochodzących z wielu źródeł. Dzięki temu można pozyskać nowe dane, których otrzymanie nie byłoby możliwe, gdyby dane z wielu źródeł przetwarzano innymi metodami.

Ponieważ terminologia z zakresu fuzji danych nie jest w pełni uzgodniona, niektórzy autorzy używają terminu fuzja danych zamiennie z terminem **fuzja in-**

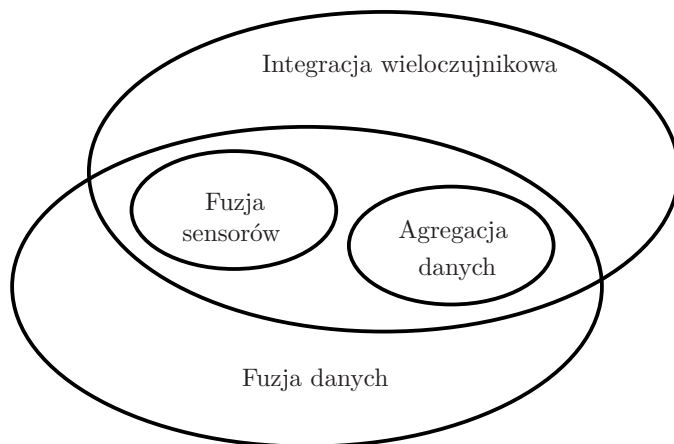
1. Fuzja danych strumieniowych

formacji [59, 337]. Można jednak zauważyć tendencję, że termin fuzja danych jest częściej wykorzystywany w przypadku fuzji danych, gdzie źródłem danych jest czujnik pomiarowy, natomiast termin fuzja informacji wykorzystuje się w przypadku danych, które zostały już wstępnie przetworzone.

Wielu autorów używa terminów pokrewnych do fuzji danych, takich jak **fuzja sensoryczna** lub **fuzja wielosensoryczna** (polisensoryczna), **agregacja danych** oraz **integracja wielosensoryczna**.

Fuzja sensoryczna lub fuzja wielosensoryczna tym różni się od integracji wielosensorycznej, że w procesie fuzji wydobywa się z danych dodatkowe informacje. Natomiast w przypadku integracji wielosensorycznej odbywa się jedynie proces łączenia danych w jeden spójny opis [254]. To co odróżnia fuzję wielosensoryczną o integracji wielosensorycznej to charakterystyczny dla fuzji efekt synergii.

Z kolei agregacja danych odnosi się do zadania gromadzenia danych z różnych źródeł oraz usunięcia danych nadmiarowych. Celem jest przygotowanie zbiorów danych o niewielkich rozmiarach zawierających istotne dane [1]. Wykonanie tych operacji na zebranych danych może utrudnić lub uniemożliwić wystąpienie zjawiska synergii, ponieważ w procesie usuwania nadmiarowych danych może dojść do usunięcia tych istotnych. Na rysunku 1.6 przedstawiono zależność między fuzją danych, fuzją sensoryczną, integracją sensoryczną oraz agregacją danych.



Rys. 1.6. Zależność pomiędzy fuzją danych a fuzją sensorów, agregacją danych i integracją wieloczujnikową

W pracy rozpatrywane będą głównie zagadnienia związane z fuzją danych oraz zagadnieniami z nią związanymi, tj. fuzją sensoryczną i fuzją wielosensoryczną.

Zagadnienia te można traktować jako proces kilkuetapowy. Etapy te możemy podzielić następująco:

- pozyskanie danych z wielu źródeł,
- generacja i ekstrakcja cech z pozyskanych danych,
- konkluzje z wyekstrahowanych cech,
- decyzje na podstawie konkluzji.

1.5. Fuzja danych strumieniowych

Tematami wiodącymi w pracy są zagadnienia związane z przetwarzaniem danych o specyficznej naturze, tj. **strumień danych**.

Definicja 1.5.1 (Strumień danych). *Strumień danych jest uporządkowaną parą (s, t) , gdzie s jest skalarem lub wektorem, natomiast $t \in T$, gdzie T jest zbiorem indeksów [17, 301].*

Fuzja danych strumieniowych jest szczególnym przypadkiem fuzji danych (patrz definicja 1.4.1), gdzie źródłem danych są strumienie danych (patrz definicja 1.5.1). Ponieważ strumienie danych są danymi pojawiającymi się sekwencyjnie w (potencjalnie) zmiennych chwilach czasowych, ich fuzja może okazać się zadaniem trudnym [131]. Warto nadmienić, że w wielu praktycznych zadaniach strumienie danych mogą mieć charakterystykę taką, jak omówione już dane typu *big data*, tj. mogą być to zbiory masowe, zmienne, różnorodne oraz nadmiarowe. Przetwarzanie takich zbiorów nakłada na metody i algorytmy służące do ich przetwarzania dodatkowe wymagania. Na przykład algorytmy przetwarzania takich zbiorów powinny umożliwiać wyszukiwanie w dużych zbiorach danych elementów, które są istotne i powinny być przetworzone w pierwszej kolejności. Jest to również kluczowa kwestia w przypadku, gdy konieczna jest szybka reakcja na zmiany zachodzące w dużych zbiorach danych. W tej pracy omówione zostaną podejścia, które umożliwiają spełnienie wymienionych wymagań nakładanych na algorytmy przetwarzania danych strumieniowych.

1.6. Analiza ruchu człowieka

Omawiane wcześniej zagadnienia dotyczące telemedycyny związane były m.in. z monitorowaniem osób chorych lub starszych. Wskazaliśmy również, że dane pozyskane z wykorzystaniem sieci BSN mogą być stosowane m.in. w monitorowaniu aktywności fizycznej człowieka. Jednym z typów danych, z których korzysta się w tych zadaniach, są dane związane z motoryką człowieka. Do podstawowych danych tego typu zaliczamy dane zebrane z czujników przyspieszenia i żyroskopów (czujniki IMU), które podczas pomiarów przymocowane są do odpowiednich

1. Fuzja danych strumieniowych

segmentów ciała osoby monitorowanej. W ogólności analiza ruchu człowieka dotyczy [5]:

- śledzenia aktywności fizycznej,
- rozpoznawanie aktywności fizycznej,
- kinematycznej analizy układu ruchu człowieka.

Wymienione zadania w ostatnich latach rozpatrywano głównie w ramach badań nad metodami przetwarzania obrazów [5, 429]. Jednak postęp w dziedzinie bezprzewodowej transmisji danych oraz rozwój technologii wytwarzania mikroelektromechanicznych czujników pomiarowych sprawił, że przedstawiono wiele metod wykorzystujących nowe rozwiązania technologiczne w tych obszarach [63, 92, 346]. W literaturze omówiono również wiele rozwiązań stosujących zarówno metody widzenia komputerowego, jak i bazujące na danych z inercyjnych czujników pomiarowych [239, 332, 385, 444] czy też sygnałach EMG [13, 74].

Celem metod do analizy ruchu jest: a) dostarczenie narzędzi do budowy zależności w czasie i w przestrzeni dla obiektów znajdujących się w ruchu oraz b) narzędzi do interpretacji tych zależności. Budowane zależności mogą dotyczyć pojedynczego punktu, grupy punktów należących do jednego obiektu lub zależności pomiędzy grupami punktów.

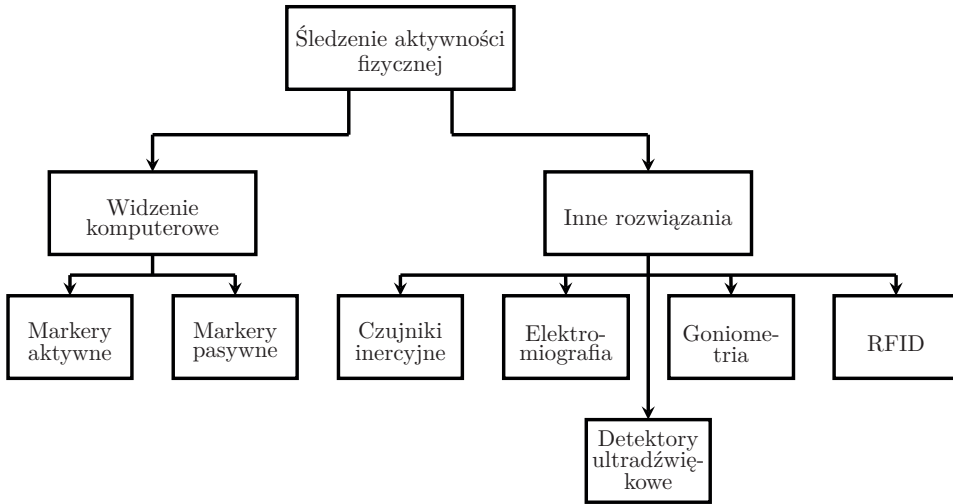
1.6.1. Śledzenie aktywności fizycznej

Zadaniem algorytmów do śledzenia aktywności fizycznej człowieka jest m.in. budowa zależności czasowo-przestrzennych dla ludzkiej sylwetki. Na rysunku 1.7 przedstawiono systematykę systemów do śledzenia aktywności fizycznej człowieka.

Śledzenie aktywności fizycznej można dzielić według różnych kryteriów. Wśród najistotniejszych wyróżniamy śledzenie wybranych części ciała człowieka (np. ręka, noga, całe ciało). Śledzenie aktywności może być również rozpatrywane w przestrzeni dwu- lub trójwymiarowej, może być wykonywane w pomieszczeniach zamkniętych (np. laboratorium) lub na otwartych przestrzeniach (np. boisko). Może dotyczyć jednej lub wielu osób.

Monitorowanie aktywności fizycznej może być wykorzystane w medycynie, np. w celu wspomagania rehabilitacji [57, 98, 479], w sporcie do analizy i wspomagania treningu technicznego [167, 303, 382], treningu osobistego [220, 439] i w rozrywce [97].

Głównym problemem, który należy rozwiązać w systemach do śledzenia aktywności fizycznej jest przetwarzanie zakłóconych danych pomiarowych. W przypadku czujników inercyjnych typu IMU pomiary obciążone są tzw. **dryftem**. Dryft jest wolnozmiennym niepożądanym składnikiem sygnału. Warto w tym miejscu zaznaczyć, że usuwanie tego typu składników, w przypadku rzeczywistych pomiarów, jest zadaniem trudnym. Metody usuwające niepożądane składniki pomiarów



Rys. 1.7. Klasyfikacja metod śledzenia aktywności fizycznej człowieka

możemy podzielić na takie, które niwelują ich wpływ, wykorzystując dane z dodatkowych źródeł, np. czujników GPS, kamer wideo. Inna grupa metod wykorzystuje specyfikę aktywności fizycznej człowieka. Jednym z przykładów w tej kategorii jest algorytm ZUPT (ang. *Zero UPdaTe*). W metodzie tej niepożądane wolnozmiennie składowe eliminuje się, używając do tego wiedzę o specyfice chodu człowieka. Wymienione metody zostaną szerzej opisane w dalszej części pracy.

Pomimo omówionych wad, systemy bazujące na czujnikach inercyjnych są popularnym i wygodnym narzędziem do akwizycji i przetwarzania danych w obszarze analizy aktywności ruchowej człowieka. Zaletą rozwiązań jest to, że niwelują wady systemów wizyjnych, m.in. eliminując zjawisko **okluzji**, czyli efektu zakrywania obiektów znajdujących się dalej od kamery (obserwatora) przez obiekty znajdujące się bliżej kamery. Rozwiązania bazujące na czujnikach IMU nie są, tak jak jest to w przypadku systemów wizyjnych, wrażliwe na warunki świetlne. Czas przetwarzania danych w systemach opartych na czujnikach IMU jest również znacznie krótszy niż w przypadku systemów wizyjnych. Niezaprzeczalną zaletą tych rozwiązań jest również to, że systemy bazujące na czujnikach inercyjnych umożliwiają śledzenie obiektów nie tylko w pomieszczeniach zamkniętych, ale również w przestrzeniach otwartych.

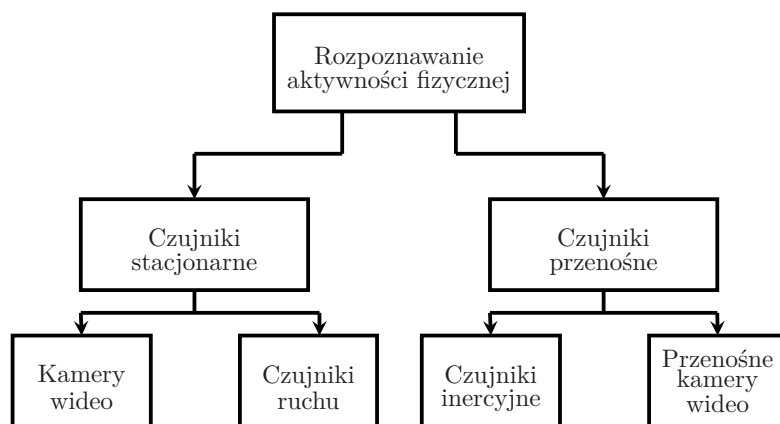
1.6.2. Rozpoznawanie aktywności fizycznej

Problem rozpoznawania aktywności fizycznej człowieka od wielu lat jest rozpatrywany w wielu obszarach badawczych takich, jak wojskowość, medycyna, sport i bezpieczeństwo. Na przykład osoby chore na cukrzycę powinny dbać o swoją formę fizyczną przez systematyczne wykonywanie ćwiczeń. Informacje o długości spacerów, przebytych dystansach podczas marszów, biegów lub przejażdżek rowerowych są istotne zarówno dla pacjenta, jak i dla lekarza.

W problematyce rozpoznawania aktywności fizycznej rozróżniamy dwa główne podejścia, w których wykorzystywane są czujniki pomiarowe:

- stacjonarne,
- przenośne.

Czujniki stacjonarne to czujniki na stałe przymocowane w pewnych ustalonych miejscach, np. kamery wideo lub czujniki ruchu zamontowane w pomieszczeniach lub na zewnątrz budynków. W tym przypadku, by możliwe było rozpoznanie aktywności, osoba musi znaleźć się w zasięgu działania czujnika. Czujniki przenośne zamontowane są na użytkowniku, więc rozpoznawanie aktywności może być wykonane w dowolnym momencie. Na rysunku 1.8 zaproponowano systematykę systemów do rozpoznawania aktywności fizycznej.



Rys. 1.8. Systematyka metod rozpoznawania aktywności fizycznej człowieka

Przykładem wykorzystania czujników stacjonarnych do rozpoznawania aktywności fizycznej są inteligentne domy [406]. Dostępne czujniki pomiarowe oraz metody przetwarzania danych umożliwiają rozpoznawanie różnorodnych form aktywności fizycznej, np. jedzenia, mycia naczyń lub brania prysznica. Niestety, z dużą skutecznością tych rozwiązań często związany jest wysoki koszt montażu wymaga-

nych urządzeń oraz koszty związane z późniejszym utrzymaniem infrastruktury pomiarowej.

Alternatywą do przedstawionego rozwiązania są systemy bazujące na czujnikach przenośnych, których rozwój związany jest z postępowaniem technologii wytwarzania mikrosystemów takich, jak mikrouządzenia elektromechaniczne. Na rynku dostępne są czujniki, które umożliwiają pomiary wybranych wielkości charakterystycznych dla użytkownika znajdującego się w ruchu (np. system GPS lub czujnik przyspieszenia), wielkości fizjologicznych (np. pomiar EKG lub pulsu) lub wielkości w otoczeniu człowieka (np. pomiar temperatury lub wilgotności).

Przykładem wykorzystania technologii bezprzewodowych są systemy wspomagające leczenie osób przewlekle chorych lub w podeszłym wieku. Jednym z takich rozwiązań są systemy detekcji i sygnalizacji upadku monitorowanego użytkownika [36].

Z konstrukcją systemów rozpoznawania aktywności fizycznej człowieka związanych jest wiele zagadnień badawczych i konstrukcyjnych. Wśród najważniejszych wyróżnić możemy wybór wielkości pomiarowych, które będą mierzone oraz cech, które zostaną z nich wyodrębnione (wyekstrahowane), wybór metod przetwarzania wstępnego oraz wnioskowania, jak również konstrukcja zbiorów uczących i testujących. Istotnym zagadnieniem związanym ze zbiorami uczącymi jest ich właściwe przygotowanie, tj. takie, by możliwe było wykorzystanie ich w odniesieniu do nowych użytkowników. Wymienione zadania są trudne i wymagają szczególnej uwagi ze względu na ich wpływ na jakość systemów rozpoznawania aktywności fizycznej człowieka. Oddzielnymi zagadnieniami są kwestie dotyczące urządzeń pomiarowych, które powinny być energooszczędne, charakteryzować się niewielkimi rozmiarami oraz nie przeszkadzać w codziennych aktywnościach użytkownika.

1.6.3. Kinematyczna analiza układu ruchu człowieka

Celem analizy układu ruchowego człowieka jest ocena jego potencjału, a w szczególności jego funkcji lokomocyjnych [46]. Ocena ta może być przydatna zarówno w przypadku osób chorych (np. analiza na potrzeby rehabilitacji), jak i osób zdrowych, a w szczególności sportowców (np. analiza na potrzeby treningu technicznego). W tych przypadkach ocenie podlegają głównie układy kostno-stawowy oraz mięśniowy.

Chcąc dokonać oceny potencjału ruchowego człowieka, należy przeprowadzić analizę zebranych danych pomiarowych. W przypadku układu kostno-stawowego wyróżniamy następujące typy danych pomiarowych:

- czasowo-przestrzenne,
- kinematyczne,
- kinetyczne.

1. Fuzja danych strumieniowych

Do grupy danych czasowo-przestrzennych zaliczamy prędkość (np. chodu), długość kroków i częstotliwość ich stawiania. Do grupy danych kinematycznych zaliczamy położenia kątowe segmentów ciała, prędkości oraz przyspieszenia poszczególnych segmentów ciała. Do oceny kinetyki ruchu człowieka potrzebne są pomiary sił oraz momentów sił wytwarzanych podczas ruchu człowieka.

Do analizy układu mięśniowego wykorzystywane są pomiary elektromiograficzne. Elektromiografia jest metodą pomiaru potencjału czynnościowego mięśnia. Dane pomiarowe pochodzące z elektromiografii wykorzystywane są np. do oceny reakcji mięśni podczas wykonywanych ruchów [46]. Analiza danych tego typu może ułatwić ocenę pracy układu mięśniowego badanej osoby.

Pomiary wymienionych wielkości, dotyczących zarówno układu kostno-szkieletowego, jak i mięśniowego przy aktualnym stanie rozwoju techniki są możliwe z wykorzystaniem odpowiednich urządzeń pomiarowych. Wykorzystanie części z tych narzędzi pomiarowych nie nastęrcza wielu problemów, inne natomiast są kłopotliwe w obsłudze. Do pierwszej kategorii zaliczyć można pomiary elektromiograficzne, momentów sił oraz długość i częstotliwość kroków. Do drugiej z nich zaliczamy pomiary położenia kątowych segmentów ciała, prędkości i przyspieszeń w tych segmentach, jak i prędkość poruszania się (np. chodu).

Pomimo wyraźnego podziału metod pomiarowych, mają one jedną wspólną negatywną właściwość, tj. trudność przeprowadzenia pomiarów w warunkach nielaboratoryjnych. Alternatywą są rozwiązania bazujące na urządzeniach wykonanych z wykorzystaniem nowych technologii. Niewielkie wymiary oraz możliwość bezprzewodowego transferu danych sprawiły, że możliwe jest wykonywanie badań poza specjalistycznymi laboratoriami.

Typowe narzędzia do kinematycznej analizy układu ruchu człowieka bazują na systemach wykorzystujących wiele kamer oraz zestaw znaczników, rozmieszczonych na ciele człowieka. Systemy takie, nazywane systemami *motion capture* (ang. *Motion Capture System*, pol. systemy przechwytywania ruchu) są uznanymi rozwiązaniami wykorzystywanymi do analizy układu ruchu człowieka. Niestety, umożliwiają badania jedynie w warunkach laboratoryjnych. Przy czym są to również drogie rozwiązania, stąd wykorzystywane tylko w specjalistycznych laboratoriach. Jednakże systemy te znalazły zastosowanie m.in. w badaniach nad układem ruchu człowieka na potrzeby analizy patologii chodu [412] czy też w sporcie [86, 330].

Alternatywą dla systemów bazujących na kamerach są systemy wykorzystujące czujniki inercyjne (tj. IMU) lub inercyjne i magnetyczne (tj. MARG). Wśród istniejących rozwiązań możemy wyróżnić takie, które skupiają się na kończynie dolnej [227, 248, 389], górnej [47, 265, 481] oraz całej sylwetce człowieka [152, 470]. Na rysunku 1.9 przedstawiono systematykę systemów wykorzystywanych do kinematycznej analizy układu ruchu człowieka.

1.7. Przykłady wykorzystania fuzji danych w analizie ruchu człowieka



Rys. 1.9. Klasyfikacja metod kinematycznej analizy układu ruchu człowieka

Jednym z głównych ograniczeń systemów bazujących na czujnikach inercyjnych jest występowanie niepożądanych składników w danych pomiarowych. Przykładem są już wymienione składniki wolnozmiennne związane z pomiarami z czujników typu IMU lub MARG, czyli dryft. Problem eliminacji dryftu z oszacowań orientacji lub pozycji poszczególnych segmentów ludzkiego ciała nie jest zadaniem trywialnym. W ostatnich latach opracowano wiele metod, które umożliwiają kompensowanie negatywnego wpływu dryftu w systemach analizy układu ruchu człowieka. Niestety brak jest ogólnej metody eliminacji tego niekorzystnego zjawiska. Podejścia, które do tej pory zaproponowano, działają tylko w szczególnych przypadkach, dla wybranych typów aktywności ruchowej człowieka. Jednym z takich przykładów jest wymieniona już metoda ZUPT, która eliminuje dryft jedynie w przypadku chodu człowieka i tylko w szczególnych przypadkach [314, 401].

1.7. Przykłady wykorzystania fuzji danych w analizie ruchu człowieka

Przykład 1.1 (Analiza chodu człowieka). *Typowymi metodami pomiaru wielkości niezbędnych w analizie chodu człowieka są platformy dynamograficzne, systemy do analizy obrazu wideo oraz systemy do analizy ruchu. Platformy dynamograficzne dokonują pomiaru składowych sił reakcji podłoża. Natomiast idea systemów do analizy ruchu opiera się na wykorzystaniu markerów odblaskowych, które umieszczone są w charakterystycznych punktach na ciele człowieka. W systemach do analizy obrazu wideo dokonuje się śledzenia odpowiednich punktów anatomicznych na ciele człowieka, klatka po klatce. Ruch markerów jest rejestrowany przez kamery wideo, a następnie przetwarzany. Wykorzystanie odblaskowych markerów znacznie upraszcza późniejszą analizę obrazu wideo. Jednocześnie jest*

1. Fuzja danych strumieniowych

jednym ze źródeł błędów pomiarowych, związanych z nieprawidłowym umocowaniem markerów lub ich przesunięciami podczas wykonywania badań.

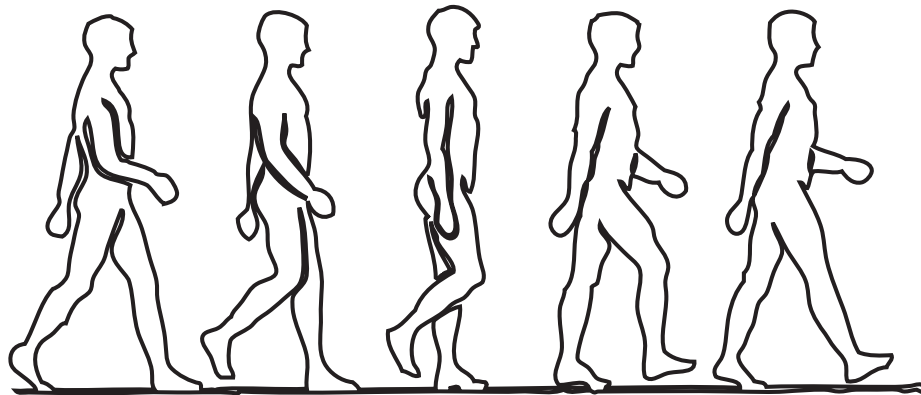
W ostatnich latach rozwój mikrosystemów sprawił, że tanie urządzenia o niewielkim poborze mocy oraz niewielkich rozmiarach stały się dostępne dla szerokiego grona odbiorców. Znalazły one również zastosowanie w analizie chodu człowieka. Czujniki, które wykorzystywane są najczęściej to wymienione już akcelerometry, żyroskopy i magnetometry [395].

Chód człowieka

Chód człowieka jest serią cyklicznych, naprzemiennych ruchów kończyn dolnych (rys. 1.10). Wyniki analizy danych charakterystycznych dla chodu, np. prędkość chodu oraz długość kroku, są wykorzystywane m.in. w diagnostyce na potrzeby rehabilitacji pacjentów. Opracowano również metody wykrywające nieprawidłowości w chodzie na podstawie zebranych danych pomiarowych.

Oprócz klasycznych metod pomiarowych wykorzystywanych na potrzeby analizy chodu człowieka w ostatnich latach zaproponowano wiele rozwiązań bazujących na systemach opartych na technologii MEMS [160, 395].

Czujniki typu akcelerometr, żyroskop i magnetometr mogą być wykorzystywane zarówno do pozyskiwania danych czasowo-przestrzennych, jak i kinematycznych.



Rys. 1.10. Kolejne fazy chodu człowieka

Przykład 1.2 (Ocena poziomu wytrenowania sportowca). Jednym z celów treningu sportowego jest zwiększanie wydolności fizycznej organizmu (por. rozdz. 1.2). Osiągnięcie tego, jak i innych celów treningowych warunkowane jest prawidłowo zaplanowanym treningiem sportowym. Odpowiednio zaplanowany trening to,

1.7. Przykłady wykorzystania fuzji danych w analizie ruchu człowieka

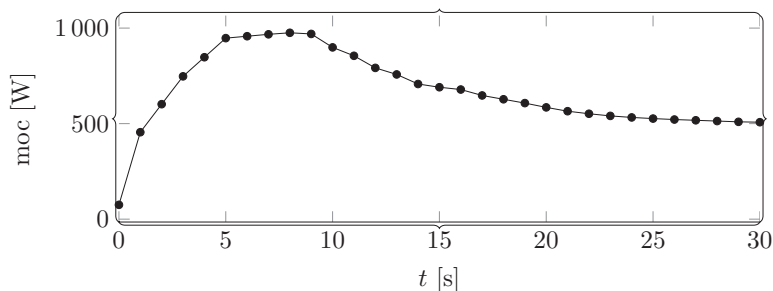
z jednej strony właściwa procedura treningowa, dobrana do dyscypliny sportowej, natomiast z drugiej strony, to właściwy dobór obciążeń treningowych, które muszą uwzględniać aktualny stan wytrenowania sportowca.

Do prawidłowego prowadzenia treningu potrzebne są dane ułatwiające scharakteryzowanie aktualnego stanu wytrenowania sportowca. W przypadku treningu ukierunkowanego na zwiększenie wydolności fizycznej organizmu istotne są tzw. wskaźniki wydolności.

W przypadku wydolności organizmu rozróżniamy wydolność aerobową (tlenową) i anerobową (beztlenową). Popularnym testem umożliwiającym ocenę wydolności organizmu na poziomie anerobowym jest test *wingate* [373]. Na podstawie tego testu możliwe jest właściwe dobranie obciążeń treningowych. Przykładową krzywą wyznaczoną dla testu *wingate* przedstawiono na rys. 1.11.

Istotą testu *wingate* jest wykonanie, w określonym czasie (ok. 30 s), maksymalnego wysiłku. Zazwyczaj wykorzystuje się w tym celu cykloergonometry. Jednak jak wskazano, m.in. w pracy [373], możliwe jest przeprowadzenie testu *wingate*, podczas którego dokonuje się pomiaru prędkości poruszającego się sportowca.

W ogólności pomiar prędkości nie jest prosty. Zazwyczaj wykorzystuje się w tym celu pomiary pośrednie, które można pozyskać m.in. z takich czujników jak akcelerometry, żyroskopy i magnetometry. Opracowano wiele metod estymacji prędkości poruszającego się obiektu. Metody te różnią się od siebie typem wykorzystanych danych pomiarowych oraz charakterem analizowanej aktywności (chód, bieg, pływanie itp.).



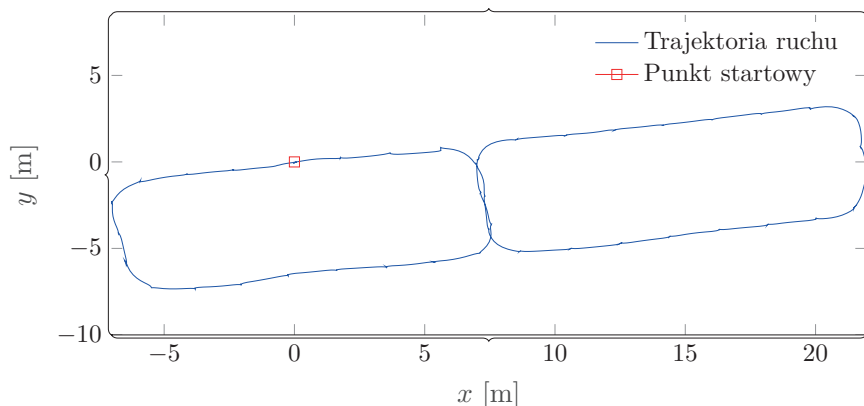
Rys. 1.11. Przykładowy wynik z testu *wingate*

Przykład 1.3 (Nawigacja osób niewidomych). *Przykładem wykorzystania współczesnych technologii transmisji i przetwarzania danych w nowych obszarach są systemy do nawigacji osób niewidomych [26, 27, 383]. Systemy takie, w przypadku nawigacji osób niewidomych w terenie otwartym, wykorzystują czujniki GPS. Niestety, stosowanie tych rozwiązań na terenach miejskich lub w budynkach jest ograniczone. Związane jest to z osłabieniem sygnału w takich miejscach z powodu*

1. Fuzja danych strumieniowych

m.in. odbić sygnału GPS, które są charakterystyczne dla terenów zabudowanych. Istotnym problemem zawężającym wykorzystanie omawianego podejścia w tych terenach jest również tłumienie sygnału GPS. Rozwiązaniem są systemy, które wspomagają istniejące rozwiązania oparte na czujnikach GPS, np. wykorzystujące sieci bezprzewodowe. Trwają również prace nad systemami, które w nawigacji nie wykorzystują sygnałów GPS.

W celu poprawnego nawigowania osób niewidzących należy jednocześnie rozwiązać następujące zadania, tj. ustalanie aktualnej lokalizacji osoby oraz budowa mapy otoczenia, w którym dana osoba aktualnie się znajduje. Potrzeba budowy mapy w otoczeniu osoby niewidzącej oraz ustalanie jej pozycji na tej mapie jest konieczne do nawigacji, która umożliwi bezpieczne przemieszczanie się z jednego miejsca w drugie. Systemy takie nazywane są systemami typu SLAM (ang. Simultaneous Localization and Mapping).



Rys. 1.12. Przykład estymacji trajektorii ruchu z wykorzystaniem czujników inercyjnych

W celu rozwiązania zadania lokalizacji mogą zostać wykorzystane czujniki takie, jak akcelerometr, żyroskop i magnetometr [192]. Umieszczenie tych czujników na ciele osoby niewidomej ułatwia wykonanie pomiarów, które umożliwią wyznaczenie jej aktualnej pozycji w obszarach zabudowanych. Zadanie ustalenia lokalizacji z wykorzystaniem czujników IMU lub MARG znane jest w literaturze jako zadanie typu PDR (ang. Pedestrian Dead Reckoning). Wyznaczanie pozycji użytkownika na podstawie danych z czujników IMU lub MARG nie jest zadaniem prostym, ponieważ wymagane jest w tym przypadku dwukrotne numeryczne całkowanie odpowiednio przetworzonych sygnałów z tych urządzeń pomiarowych. Jednym z przykładów rozwiązania problemu niedokładności w oszacowaniu pozycji użytkownika na podstawie

1.7. Przykłady wykorzystania fuzji danych w analizie ruchu człowieka

danych z czujników IMU lub MARG jest wykorzystanie dodatkowych czujników (np. kamery wideo).

Przykład oszacowania trajektorii ruchu dla poruszającego się człowieka przedstawiono na rys. 1.12.

Przykład 1.4 (Egzoszkielety medyczne). Współczesne metody leczenia wykorzystujące proste roboty rehabilitacyjne umożliwiają wielokrotne powtarzanie ćwiczeń przez pacjenta. Rozwój technologii wytwarzania urządzeń pomiarowych i wykonawczych oraz algorytmów przetwarzania danych stwarza warunki do budowy interakcyjnych robotów rehabilitacyjnych. Interakcja w takich rozwiązaniach może zachodzić zarówno w obszarze fizycznym, jak i poznawczym (np. interfejsy dotykowe) [287]. Przykładem interaktywnego robota rehabilitacyjnego jest egzoszkielec (ang. exoskeleton).

Egzoszkielec jest konstrukcją mechaniczno-elektryczną, która mocowana jest do ciała użytkownika. Elementy konstrukcji przytwierdzone są do wybranych jego części w celu wspomagania jego ruchów. Do wspomagania tych ruchów wykorzystywane są specjalistyczne elementy konstrukcji egzoszkieletu nazywane efektorami.

Do budowy egzoszkieletu niezbędna jest wiedza z takich obszarów jak mechanika, inżynieria biomedyczna, robotyka, elektronika czy informatyka. W takich konstrukcjach czujniki typu IMU lub MARG można wykorzystać do ustalenia wzajemnego położenia poszczególnych segmentów egzoszkieletu, prędkości kątowych itp. Wielkości te niezbędne są np. podczas sterowania urządzeniem rehabilitacyjnym.

Przykład 1.5 (Monitorowanie treningu technicznego sportowca). W sporcie techniką określa się sposób wykonania zadania ruchowego związanego z daną dyscypliną sportową. Technikę sportową wypracowuje się podczas przygotowania technicznego [376].

Sportowiec podczas treningu technicznego powinien dysponować metodami do oceny wykonywanych ruchów, jak i do śledzenia własnych postępów. W tych zadaniach sportowcowi powinien pomagać trener. Zazwyczaj uwagi przekazywane są sportowcowi w postaci oceny słownej (tj. informacja jakościowa) podczas lub bezpośrednio po wykonaniu ćwiczenia. Uwagi przekazywane są przez trenera, który obserwuje trening sportowca. Rozwój technik pomiarowych sprawia, że oprócz oceny słownej zawodnik może otrzymać dodatkowe informacje (tj. informacje jakościowe) dotyczące treningu w postaci zapisu wideo lub innych pomiarów (np. pomiary EKG, pulsu oraz wielkości czasowo-przestrzenne, kinematyczne lub kinetyczne). Cechą wspólną tych technik jest to, że muszą one w sposób przejrzysty przekazywać sportowcowi najważniejsze informacje dotyczące jego techniki.

Również i w tym przypadku przydatne może okazać się zastosowanie czujników typu IMU lub MARG.

Część II

Fuzja danych strumieniowych

2. Elementy fuzji danych

Całość to więcej niż suma jej składników.

Arystoteles

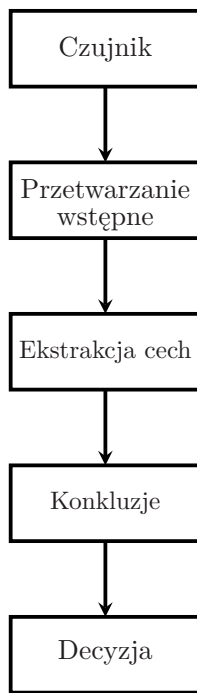
2.1. Wstęp

Głównym składnikiem bezprzewodowych sieci czujnikowych (np. sieci BSN) są urządzenia pomiarowe. Natomiast głównym zadaniem bezprzewodowych sieci typu BSN jest transfer danych z czujników pomiarowych do jednostek przetwarzających przesyłane dane pomiarowe. W przypadku danych pochodzących z wielu czujników pomiarowych istotnym zagadnieniem staje się organizacja procesu przetwarzania pomiarów. Ogólny schemat tego procesu został przedstawiony na rys. 2.1. W tym opracowaniu skupimy się na danych pomiarowych z czujników pracujących w ramach sieci typu BSN. Warto w tym miejscu podkreślić, że przedstawiony schemat przetwarzania danych nie ogranicza się jedynie do opisanego przypadku. Może być on również wykorzystany do innych typów danych. W ogólności w skład procesu przetwarzania danych wchodzi następujące elementy: **pomiar**, **przetwarzanie wstępne**, **ekstrakcja cech**, **konkluzje** oraz **decyzja**.

W przypadku bezprzewodowych sieci czujnikowych należy wziąć pod uwagę wiele urządzeń pomiarowych oraz ustalić metodę przetwarzania pozyskanych danych z wielu źródeł, tj. należy wskazać, w jaki sposób czujniki będą ze sobą połączone oraz jaka będzie kolejność przetwarzania danych. Przykładowe połączenie czterech czujników pomiarowych przedstawiono na rys. 2.2. Schemat ten wskazuje, że dane z czujników mogą być przetwarzane na różnych poziomach. To, na którym poziomie lub poziomach dane z czujników będą przetwarzane, zależy od typu zadania, dla którego opracowywane jest rozwiązanie.

Przykład z rysunku 2.2 pokazuje, że dane z czujników mogą być przetwarzane niezależnie albo łącznie. Przypadek przetwarzania niezależnego jest typowym przypadkiem, który jest najczęściej stosowany. Zastosowanie przetwarzania łącznego wiąże się z ustaleniem odpowiedniej kolejności obliczeń oraz właściwym doбором metod przetwarzania danych. W przypadku przetwarzania łącznego organizacja

2. Elementy fuzji danych

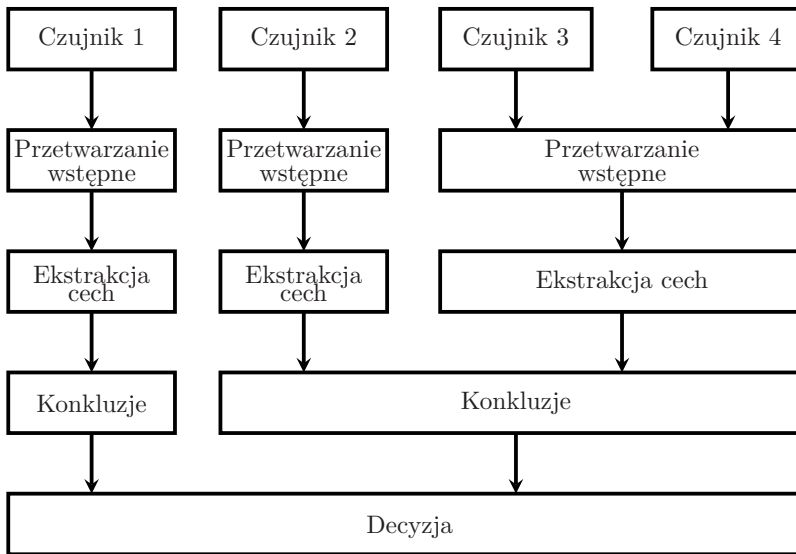


Rys. 2.1. Schemat przetwarzania danych w przypadku jednego czujnika

obliczeń może być dwuwariantowa, tzn. obliczenia mogą być rozpoczęte na najniższym poziomie (przypadek dla pierwszego i drugiego czujnika na rys. 2.2) lub mogą rozpoczynać się na wyższych poziomach (porównaj czujniki numer trzy i cztery na rys. 2.2).

Osobiste sieci czujnikowe złożone z wielu czujników charakteryzują się pewnymi ograniczeniami, które związane są z czujnikami wchodzącymi w skład tej sieci. Wśród najważniejszych ograniczeń wyróżniamy [299]:

- uszkodzenie czujnika (ang. *sensor deprivation*); awaria czujnika ogranicza możliwości pomiarowe BSN,
- ograniczony zasięg czujnika w przestrzeni (ang. *limited spatial coverage*); urządzenia pomiarowe mają ograniczony zasięg, np. akcelerometr umieszczony na jednej z kończyn dolnych nie może mierzyć przyspieszeń na kończynie górnej,
- niedokładność (ang. *imprecision*); każdy czujnik ma ograniczoną dokładność pomiarową,
- niepewność (ang. *uncertainty*); może być spowodowana: a) ograniczonymi możliwościami pomiarowymi czujnika, np. czujnik typu IMU nie może zmierzyć



Rys. 2.2. Schemat przetwarzania danych w przypadku wykorzystania wielu czujników

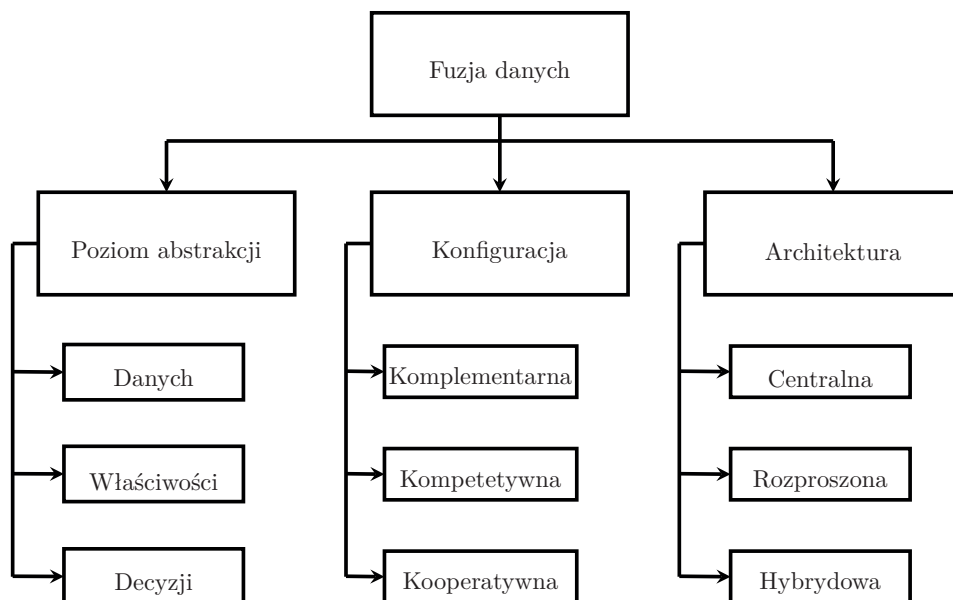
natężenia pola magnetycznego lub b) część z mierzonych wielkości staje się niedostępna, np. z powodu okluzji.

Wpływ wymienionych czynników na wyniki przetwarzania danych można minimalizować, wykorzystując np. metody fuzji danych. Tak jak wspomnieliśmy w rozdziale 1, fuzją danych nazywamy proces łączenia danych z różnych czujników. Rozróżniamy trzy typy fuzji danych. Pierwszy typ fuzji danych dotyczy poziomu **abstrakcji**, na którym realizowana jest fuzja danych. Drugi typ związany jest ze sposobem **konfiguracji** czujników pomiarowych. Trzeci typ dotyczy **architektury** systemu fuzji danych. Klasyfikacja rodzajów fuzji danych została przedstawiona na rys. 2.3.

Wśród najważniejszych przesłanek, które uzasadniają stosowanie fuzji danych, wyróżniamy m.in.

- poprawę jakości danych pomiarowych; pomiar tej samej wielkości wykonany przez zestaw czujników o odmiennej zasadzie działania (tj. systemy wielomodalne) usuwa niepożądane składowe z pomiarów,
- zwiększenie odporności oraz niezawodności; redundancja zapewniana przez pomiar z wielu czujników umożliwia działanie systemu w przypadku awarii jednego z nich,
- poprawę rozdzielczości, tj. możliwość redukcji błędów związanych z kwantyzacją,
- redukcję czasu pomiarów; grupa różnych czujników umożliwia wykonywanie pomiarów z różnych perspektyw,

2. Elementy fuzji danych



Rys. 2.3. Klasyfikacja rodzajów fuzji danych

- redukcję niejednoznaczności oraz niepewności; jest to przykład wykorzystania danych zebranych z różnych źródeł,
- zwiększenie pewności; w wielu przypadkach jeden czujnik nie jest w stanie zapewnić wystarczająco wysokiego poziomu pewności zebranych danych,
- łączenie niezależnych cech i wiedzy apriorycznej; możliwość budowy kompletnego obrazu systemu, uwzględniającego różne perspektywy uchwycone przez różne czujniki pomiarowe wchodzące w jego skład.

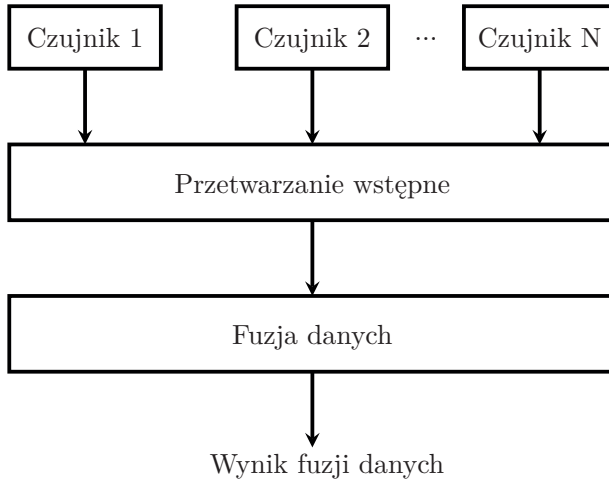
Wykorzystanie metod fuzji danych w procesie ich przetwarzania wpływa, jak wykazano to w pracy [144], na złożoność projektowanego systemu. Typowe systemy, w których nie zastosowano metod znanych z fuzji danych, charakteryzują się większą złożonością. Jest to związane z tym, że muszą one zawierać dodatkowe moduły do przetwarzania danych pomiarowych. Warto również podkreślić, że przetwarzanie danych z wykorzystaniem metod fuzji danych upraszcza proces uwzględniania dodatkowej wiedzy w procesie przetwarzania danych.

2.2. Poziom abstrakcji w systemach fuzji danych

Fuzję danych można rozpatrywać z różnych punktów widzenia (rys. 2.3). Na przykład może być ona rozpatrywana ze względu na poziom abstrakcji. Rozróżnie-

nie takie zostało zaproponowane i omówione m.in. w [59, 150, 253, 289]. W tym przypadku wyróżniamy następujące typy fuzji:

- a) danych (fuzja bezpośrednia) (patrz rys. 2.4),
- b) cech (fuzja właściwości) (patrz rys. 2.5),
- c) decyzji (patrz rys. 2.6).

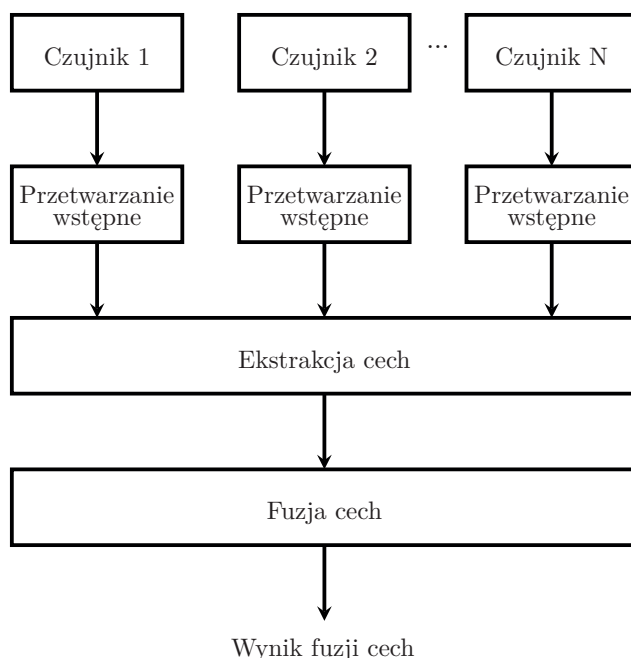


Rys. 2.4. Fuzja danych (poziom podstawowy) [151]

Fuzja danych

Przedmiotem fuzji danych na pierwszym stopniu są dane pochodzące z różnych źródeł tego samego typu (rys. 2.4). Na przykład możliwa jest fuzja danych z czujników przyspieszenia i żyroskopu. Natomiast na tym poziomie nie jest możliwa fuzja danych w przypadku, gdy dane pochodzą np. z czujników przyspieszenia i EKG. Zarówno w pierwszym, jak i w drugim przypadku źródła danych są różne, natomiast typy danych są takie same (dla pierwszego przypadku) i różne dla drugiego przypadku. Dane pochodzące z różnych źródeł, ale będące tego samego typu, nazywamy **jednomodalnymi danymi pomiarowymi** (ang. *commensurate data*). Dane jednomodalne są tego samego typu, jeśli pochodzą z różnych czujników, ale związane są z pomiarami tego samego zjawiska. Przykładem są pomiary pochodzące z czujników przyspieszenia oraz żyroskopu, które wykorzystano w pomiarach poruszającego się obiektu. Z drugiej strony, dane z czujników przyspieszenia oraz EKG nie są danymi jednomodalnymi, ponieważ dotyczą pomiarów dwóch różnych wielkości. Problem ustalenia typu danych jest jednym

2. Elementy fuzji danych



Rys. 2.5. Fuzja właściwości (poziom średni) [151]

z istotniejszych elementów mających wpływ zarówno na organizację pomiarów, ich akwizycję, jak również późniejsze przetwarzanie.

Teraz omówione zostaną kluczowe elementy, które należy brać pod uwagę, projektując systemy fuzji danych na poziomie podstawowym.

Mając ustalony typ źródeł danych, kolejnym zadaniem projektowym na poziomie podstawowym fuzji danych jest ustalenie **liczby źródeł danych**. Dane mogą pochodzić z różnych kanałów tego samego czujnika (np. w przypadku trójosiowego żyroskopu) lub z różnych węzłów zawierających czujniki tego samego typu. Możliwa jest również sytuacja, gdy dane pochodzą zarówno z różnych węzłów, jak i z różnych kanałów. Przykładem może być pomiar z czujnika przyspieszenia oraz kilku czujników EMG.

Kolejne zadanie projektowe dotyczy doboru **częstotliwości próbkowania**. Wielkość ta jest kluczowa dla systemów fuzji danych przetwarzających dane na poziomie podstawowym. Podczas doboru częstotliwości próbkowania istotne jest, m.in. ustalenie czy dane pomiarowe z czujników mają być próbkowane z jednakową czy z różną częstotliwością. Jest to ważna kwestia, ponieważ konieczność **synchronizacji pomiarów** próbkowanych z różną częstotliwością może wpłynąć

zarówno na jakość fuzji danych, jak i na złożoność projektowanego systemu oraz samych algorytmów.

Projektując algorytm fuzji dla danych, które pochodzą z różnych źródeł, istotne może okazać się również ustalenie **okresu pomiaru**. Dane z czujników mogą być przesyłane regularnie (tj. w ustalonych chwilach czasu) lub nieregularnie (tj. w dowolnych chwilach czasu). Wybór metody przesyłania danych z czujników może wiązać się z koniecznością **buforowania danych**. Fuzja danych może być wykonywana na tym poziomie dla pojedynczej próbki (w przypadku, gdy pomiary są synchronizowane) lub dla zbiorów zebranych pomiarów (w przypadku braku synchronizacji).

Po zebraniu danych pomiarowych, kolejny etap (por. rys. 2.4) związany jest z ich wstępnym przetworzeniem (ang. *pre-processing*). Główne zadania, które wykonywane są na tym etapie to **normalizacja danych** pomiarowych oraz **usuwanie z pomiarów składowych o niskich lub wysokich częstotliwościach** (tj. wolno- i szybkozmiennych). W niektórych przypadkach konieczne może być usunięcie składowych o konkretnych częstotliwościach.

Ostatnim etapem fuzji danych na poziomie podstawowym jest fuzja zebranych i wstępnie przetworzonych danych pomiarowych. Klasycznym algorytmem fuzji danych na pierwszym poziomie jest **filtr Kalmana**. Przykładem praktycznego wykorzystania filtra Kalmana jest zadanie estymacji orientacji obiektu w przestrzeni na podstawie danych pomiarowych pochodzących z czujników przyspieszenia i żyroskopu. Wadą filtra Kalmana jest ograniczony zakres jego zastosowań. Ograniczenia te narzucane są głównie przez dane pomiarowe. Oznacza to, że poprawne wyniki fuzji danych na tym poziomie można uzyskać jedynie w przypadku, gdy dane pomiarowe spełniają pewne warunki dotyczące zawartości niepożądanych składników. Zagadnienie to zostanie szerzej omówione w kolejnych rozdziałach książki.

Alternatywą dla filtra Kalmana są rozszerzony filtr Kalmana (ang. *Extended Kalman Filter*) oraz bezśladowy filtr Kalmana (ang. *Unscented Kalman Filter*). Na tym poziomie fuzji danych można również wykorzystać filtr cząsteczkowy (ang. *Particle Filter*) [253, 337]. Zastosowanie na podstawowym poziomie fuzji danych filtra Kalmana lub jednego z jego rozszerzeń ma jeszcze taką zaletę, że umożliwia estymację wielkości, których bezpośredni pomiar z wykorzystaniem dedykowanego czujnika nie jest możliwy. Za przykład niech posłuży estymacja orientacji obiektu w przestrzeni na podstawie danych z czujników przyspieszenia oraz żyroskopowych.

Innym praktycznym zastosowaniem fuzji danych na poziomie podstawowym jest możliwość wykonania automatycznej lub półautomatycznej kalibracji czujników pomiarowych [123, 357, 402].

Fuzja właściwości

Kolejny z wyróżnionych poziomów fuzji danych to poziom **właściwości (cech)** (rys. 2.5). Główne zadania, które należy rozwiązać na tym poziomie to **generowanie cech, konstrukcja wektora cech** bazująca na **redukcji cech** i/lub **selekcji cech** oraz **klasyfikacja**. Wymienione zadania, które należy wykonać w ramach fuzji właściwości nazywamy **ekstrakcją cech**.

Fuzję właściwości należy rozważyć w przypadku projektowania systemów, w których dane pochodzą ze źródeł różnego typu. Dane tego typu nazywamy danymi wielomodalnymi (ang. *noncommensurate data*).

Omówimy teraz kolejne etapy konstrukcji algorytmów fuzji właściwości. Pierwsze z zadań dotyczy generowania cech na podstawie danych pozyskanych ze źródeł różnego typu (np. czujników pomiarowych). Efektem działania algorytmów generowania cech jest zestaw wielkości opisujących badane zjawisko. Wielkości te mogą zostać wyznaczone w dziedzinie:

- czasu,
- częstotliwości,
- czasowo-częstotliwościowej.

Przykładowe cechy, które mogą zostać wyznaczone w dziedzinie czasu, to wartość minimalna i maksymalna oraz liczba przejść przez zero itp. Do tej kategorii zaliczamy również parametry statystyczne takie, jak: wartość średnia, wariancja, kurtoza, skośność czy też entropia. Z kolei w dziedzinie częstotliwości typowe przykłady cech to energia i moc sygnału, współczynniki Fouriera i melowo-częstotliwościowe współczynniki cepstralne (ang. *Mel Frequency Cepstral Coefficient*, MFCC). Natomiast najczęściej wykorzystywanymi cechami wyznaczanymi w dziedzinie czasowo-częstotliwościowej są współczynniki falkowe. Wybór dziedziny, w której będą wyznaczone cechy (**przestrzeń cech**), jest jedną z decyzji projektowych, którą należy podjąć na etapie projektowania algorytmu fuzji właściwości. Innym zagadnieniem, które należy rozważyć na etapie projektowania algorytmów przetwarzania danych, jest ustalenie **częstotliwości wyznaczenia cech**. Częstotliwość ta ma kluczowe znaczenie, ponieważ przekłada się na jakość fuzji właściwości oraz na liczbę wymaganych obliczeń, która, jeśli będzie znacząca, może wpłynąć na wydajność projektowanego systemu.

Zadaniem pomocniczym związanym z generowaniem cech jest zadanie **segmentacji**. Segmentacja to podział zbioru danych na podzbiory o podobnych właściwościach (np. statystycznych). Również w przypadku segmentacji, projektując algorytm fuzji właściwości, należy podjąć kilka decyzji projektowych.

W przypadku metod segmentacji istotnymi parametrami są szerokość i typ okna czasowego. Typowe podejście do ekstrakcji cech realizowane jest periodycznie z wykorzystaniem okna czasowego o ustalonej lub zmiennej szerokości. Z doбором

szerokości okna czasowego związany jest kolejny parametr określający, w jaki sposób okna bieżące i poprzednie sąsiadują ze sobą. Sąsiednie okna mogą nie nakładać się na siebie, mogą się ze sobą jedynie stykać lub częściowo się na siebie nakładać. Jeśli nałożenie dwóch sąsiadujących okien jest częściowe, to należy również określić stopień nałożenia tych okien na siebie. Stopień, w jaki dwa okna nakładają się na siebie traktowany jest jako kolejny parametr konstrukcyjny algorytmów segmentacji. Zarówno rozmiar okna czasowego, jak i typ sąsiedztwa okien są ważnymi parametrami konstrukcyjnymi, które mogą w sposób istotny wpłynąć tak na jakość segmentacji, jak i na jakość fuzji właściwości.

W wyniku zastosowania algorytmów generujących cechy na podstawie danych pomiarowych możemy uzyskać bardzo liczny zbiór cech. Przydatność tych cech może być różna i w dużym stopniu zależy od konkretnego zadania. Stąd może się okazać, że zasadne będzie zmniejszenie liczby wyznaczonych cech przed ich wykorzystaniem na dalszych etapach fuzji danych. Dwa główne podejścia umożliwiające redukcję liczby wygenerowanych cech to **redukcja cech** oraz **selekcja cech**. Istota każdego z tych podejść jest inna, natomiast efektem ich zastosowania jest eliminacja wymiarowości przestrzeni cech oraz uformowanie **wektora cech**. Wektor ten powinien opisywać możliwie najlepiej badany obiekt z wykorzystaniem możliwie najmniejszej liczby cech. Wynika to z tego, że celem tych podejść jest eliminacja ze zbioru wygenerowanych cech tych z nich, które są redundantne. Stąd w skład wektora cech powinny wchodzić jedynie cechy najbardziej informacyjne ze zbioru wcześniej wygenerowanych cech. Opisane zadanie konstrukcji wektora zawierającego minimalną liczbę cech możliwie najlepiej opisujących analizowany obiekt jest ważne ze względu na minimalizację nakładu obliczeń na dalszych etapach fuzji danych.

Zadanie redukcji cech polega na poszukiwaniu niskowymiarowej reprezentacji dla zbioru punktów leżących na pewnej wysokowymiarowej przestrzeni cech wygenerowanej na podstawie zbioru cech wejściowych. W wyniku jej zastosowania zostaje ustalony nowy wektor cech o niższym wymiarze, który jednocześnie wystarczająco dobrze opisuje analizowany obiekt. Z kolei selekcja cech polega na redukcji oryginalnego wektora cech przez wybór podzbioru cech z oryginalnego zbioru cech. Należy podkreślić, że w przypadku zadania selekcji cech należy ustalić kryterium, według którego selekcja ta będzie wykonana.

Efektom redukcji cech jest zestaw cech, który może wymagać normalizacji. Normalizacja jest niezbędna w sytuacji, gdy np. rozkłady rozpatrywanych cech znacznie się od siebie różnią. Normalizacja umożliwia również ocenę wpływu każdej z cech, które mają podlegać fuzji na tym poziomie.

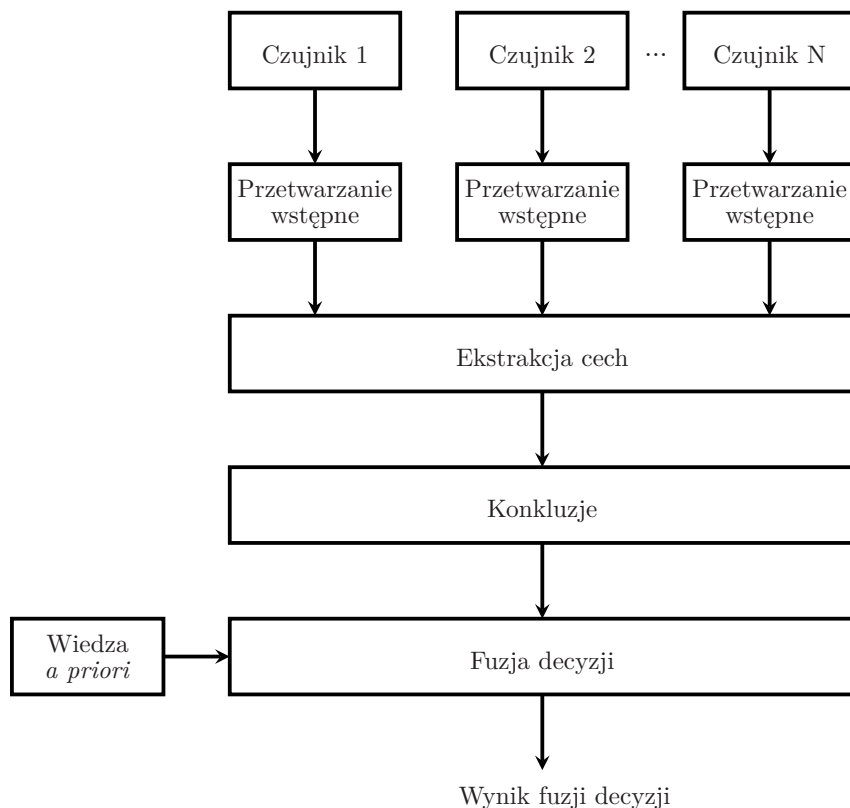
Ostatnie zadanie na etapie fuzji właściwości to **konkludowanie** na podstawie uformowanych wektorów cech. Główne podejście stosowane w tej fazie bazuje na algorytmach klasyfikacji. Algorytmy klasyfikacji wykorzystywane w zadaniu fuzji

2. Elementy fuzji danych

właściwości można podzielić na **parametryczne** i **nieparametryczne**. Metody parametryczne charakteryzują się tym, że wymagają m.in. wstępnego ustalenia typów rozkładów prawdopodobieństwa. Do grupy metod parametrycznych zaliczamy *gaussowskie* modele mieszane (ang. *Gaussian Mixture Models*) oraz metodę *k-średnich* (ang. *k-Means*). Z kolei do głównych metod nieparametrycznej fuzji właściwości zaliczamy metodę *k-NN* (ang. *k-nearest neighbors*), drzewa decyzyjne, Maszyny Wektorów Nośnych (ang. *Support Vector Machines*) oraz Sztuczne Sieci Neuronowe (ang. *Artificial Neural Networks*).

Fuzja decyzji

Ostatnim poziomem, który pozostał do omówienia to poziom **fuzji decyzji** (rys. 2.6). Głównym zadaniem na tym poziomie jest uzgodnienie decyzji na bazie danych oraz cech zgromadzonych na niższych poziomach jak również na podstawie



Rys. 2.6. Fuzja decyzji (poziom wysoki) [151]

konkluzji z poziomu drugiego. Algorytmy wykorzystywane na tym poziomie muszą charakteryzować się odpornością na zakłócenia, niekompletność oraz niespójność danych lub cech (właściwości), które zostały zebrane lub wyznaczone na niższych poziomach.

Najczęściej wykorzystywane algorytmy na poziomie trzecim bazują na metodach *bayesowskich*, metodzie Dempstera–Shafera lub logice rozmytej, metodach heurystycznych oraz klasycznym wnioskowaniu [253, 337].

Jednym z ważniejszych parametrów projektowych, dla algorytmów na trzecim poziomie, jest częstotliwość wykonania fuzji decyzji. Wybór częstotliwości fuzji danych może wiązać się z koniecznością przeprowadzania synchronizacji wyników klasyfikacji na niższym poziomie, tj. na poziomie drugim.

Kolejnym parametrem konstrukcyjnym jest wybór typu fuzji decyzji. Wyróżniamy trzy główne typy: lokalny, centralny i hybrydowy. W przypadku **fuzji lokalnej** zadanie klasyfikacji rozwiązywane jest oddzielnie w każdym czujniku lub węźle, a następnie wyniki klasyfikacji przesyłane są do węzła centralnego. Z kolei w **fuzji centralnej** wszystkie dane pomiarowe oraz wyznaczone cechy przesyłane są do węzła centralnego, gdzie dokonywana jest fuzja decyzji. W literaturze przedmiotu rozważa się również typ **fuzji hybrydowej**, to jest takiej, która jest kombinacją dwóch już omówionych typów fuzji decyzji. W tym przypadku w węźle centralnym fuzji decyzji dokonuje się na podstawie danych pomiarowych, cech, jak również konkluzji, które zebrano lub wyznaczono na niższych poziomach.

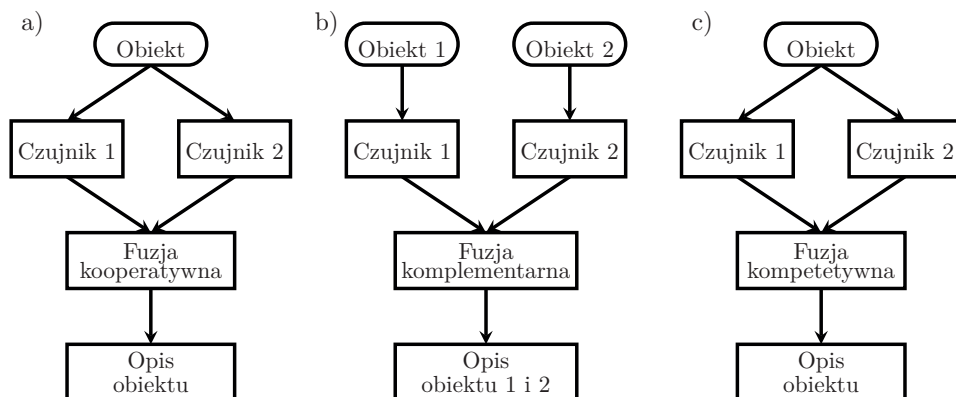
2.3. Typy konfiguracji źródeł danych

Proces fuzji danych można rozpatrywać również w ujęciu organizacji pomiarów oraz fizycznego rozmieszczenia czujników. Sposób porządkowania typów konfiguracji czujników został zaproponowany w pracy [116]. Zgodnie z tą propozycją wyróżniamy trzy konfiguracje połączeń czujników pomiarowych (rys. 2.7):

- kooperatywna (współpracy) (ang. *cooperative*),
- komplementarna (ang. *complementary*),
- kompetytywna (nadmiarowa) (ang. *redundant*).

Konfiguracja kooperatywna (współpracy) charakteryzuje się wykorzystaniem wielu różnych czujników dostarczających dane, które po ich przetworzeniu ułatwią wyznaczenie nowego opisu obiektu. Należy zaznaczyć, że opis ten byłby niemożliwy do uzyskania przy oddzielnym wykorzystaniu tych czujników. W ogólności realizacja fuzji danych w systemach z konfiguracją kooperatywną jest trudnym zadaniem. Trudności te spowodowane są m.in. tym, że niedokładność każdego z czujników wpływa na jakość rozwiązania dostarczanego przez system fuzji danych tak skonfigurowanych [48]. Przykładem konfiguracji współpracy jest

2. Elementy fuzji danych



Rys. 2.7. Klasyfikacja konfiguracji czujników pomiarowych: (a) fuzja kooperatywna, (b) fuzja komplementarna, (c) fuzja kompetetywna [59]

przetwarzanie danych z czujników typu IMU lub MARG. Na przykład dysponując danymi o przyspieszeniu obiektu, prędkościach kątowych oraz natężeniu pola magnetycznego, możliwe jest wyznaczenie orientacji obiektu w przestrzeni. Konfiguracja ta jest również wykorzystywana w systemach przechwytywania ruchu (tj. *motion capture*), gdzie kilka kamer wideo filmuje z różnych pozycji jeden plan.

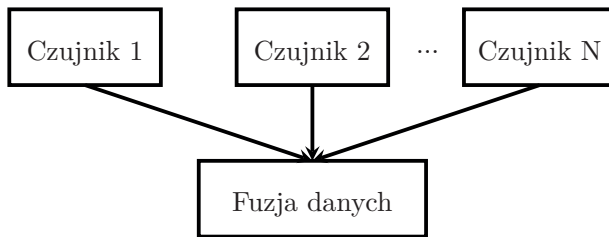
Konfiguracja komplementarna charakteryzuje się tym, że w tak skonfigurowanych systemach fuzji danych do obserwacji różnych obiektów wykorzystywane są odpowiednie czujniki. W ten sposób możliwe jest zdobycie danych, które dostarczają bogatszego opisu badanych obiektów. Przykładem takiej konfiguracji jest system, w skład którego wchodzi zestaw kamer wideo do monitorowania różnych pomieszczeń w budynku. Innym przykładem tej konfiguracji są algorytmy filtracji, w których wykorzystuje się dane z czujników, z których część mierzy szybkozmienne składowe sygnału, a część odpowiedzialna jest za pomiar składowych wolnozmiennych.

Konfiguracja kompetetywna (nadmiarowa) dotyczy pomiarów tego samego obiektu lub wybranej jego właściwości z wykorzystaniem różnych czujników pomiarowych tego samego typu. Celem doboru urządzeń pomiarowych oraz późniejszej ich konfiguracji jest zwiększenie odporności systemu na błędy pomiarowe. Skutkiem zastosowania konfiguracji nadmiarowej jest poprawa jakości fuzji danych. Przykładem fuzji nadmiarowej jest lokalizacja w systemach typu GPS, gdzie do określenia pozycji geograficznej obiektu wykorzystuje się dane przesyłane drogą radiową z kilku różnych satelitów.

2.4. Przegląd architektur dla systemów fuzji danych

Zadanie fuzji danych może być realizowane z wykorzystaniem różnych architektur. Architektury te powinny być dostosowane do problemu, a także możliwości pomiarowych i obliczeniowych projektowanego systemu. Architektura systemu fuzji danych określa konfigurację czujników pomiarowych oraz ich relacje z innymi elementami działającymi na tym samym poziomie, jak również na innych poziomach. Wyróżniamy następujące architektury fuzji danych:

- centralna (ang. *centralized*),
- rozproszona (ang. *distributed*),
- hybrydowa (ang. *hybrid*).

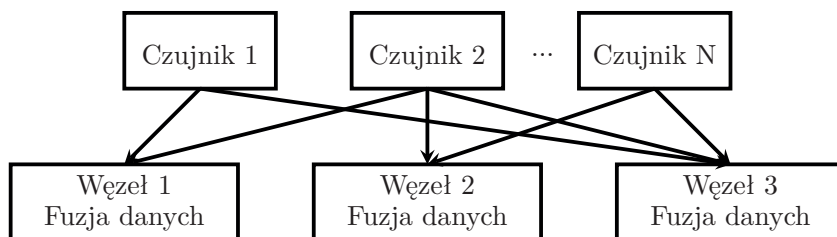


Rys. 2.8. Fuzja centralna [337]

Architektura centralna (rys. 2.8) charakteryzuje się tym, że wyróżniona jest w niej główna jednostka obliczeniowa, do której przesyłane są dane z urządzeń pomiarowych. W tym przypadku czujniki wykorzystywane są jedynie w celach pomiarowych. Natomiast przetwarzanie tych danych odbywa się w jednostce centralnej. Główne zadania jednostki centralnej to synchronizacja przesyłanych danych oraz grupowanie danych pomiarowych (asocjacja). Taka organizacja pomiarów, transmisji i przetwarzania danych jest organizacją optymalną z perspektywy jakości procesu fuzji danych. Wadą tego rozwiązania jest konieczność przesyłania danych pomiarowych ze wszystkich czujników pomiarowych do głównej jednostki obliczeniowej. Rozpatrując rozwiązanie procesu fuzji danych bazujące na architekturze centralnej, należy uwzględnić opóźnienia w transmisji danych. W wielu problemach z obszaru medycyny opóźnienia w transmisji danych są jednym z kluczowych zagadnień.

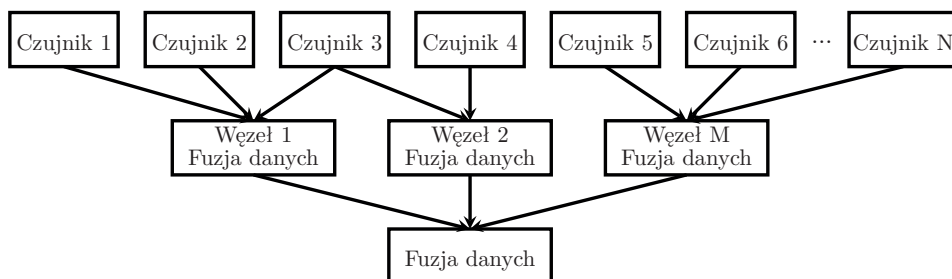
Architektura rozproszona (rys. 2.9) przeznaczona jest głównie dla systemów, w których wykorzystuje się czujniki pomiarowe zintegrowane z autonomicznymi podsystemami przetwarzania danych pomiarowych. Dane z układu pomiarowego czujnika przesyłane są do modułu wstępnego przetwarzania danych oraz dalej, do modułu fuzji danych. W tej architekturze fuzja danych wykonywana jest lokalnie

2. Elementy fuzji danych



Rys. 2.9. Fuzja rozproszona [337]

w podsystemach wykorzystujących dane z czujników pomiarowych. Po wykonaniu fuzji na poziomie lokalnym wyniki z podsystemu przesyłane są do centralnej jednostki obliczeniowej. Cechą charakterystyczną architektury rozproszonej jest redukcja wymaganego pasma na transmisję danych z czujnika do jednostki centralnej. Wadą tego rozwiązania jest to, że jednostka centralna przeprowadza fuzję danych na przetworzonych już danych. Przetwarzanie danych pomiarowych może je zubożyć, co jest głównym źródłem niedokładności procesu fuzji danych w tej architekturze i ma wpływ na końcową jakość wyniku fuzji danych.



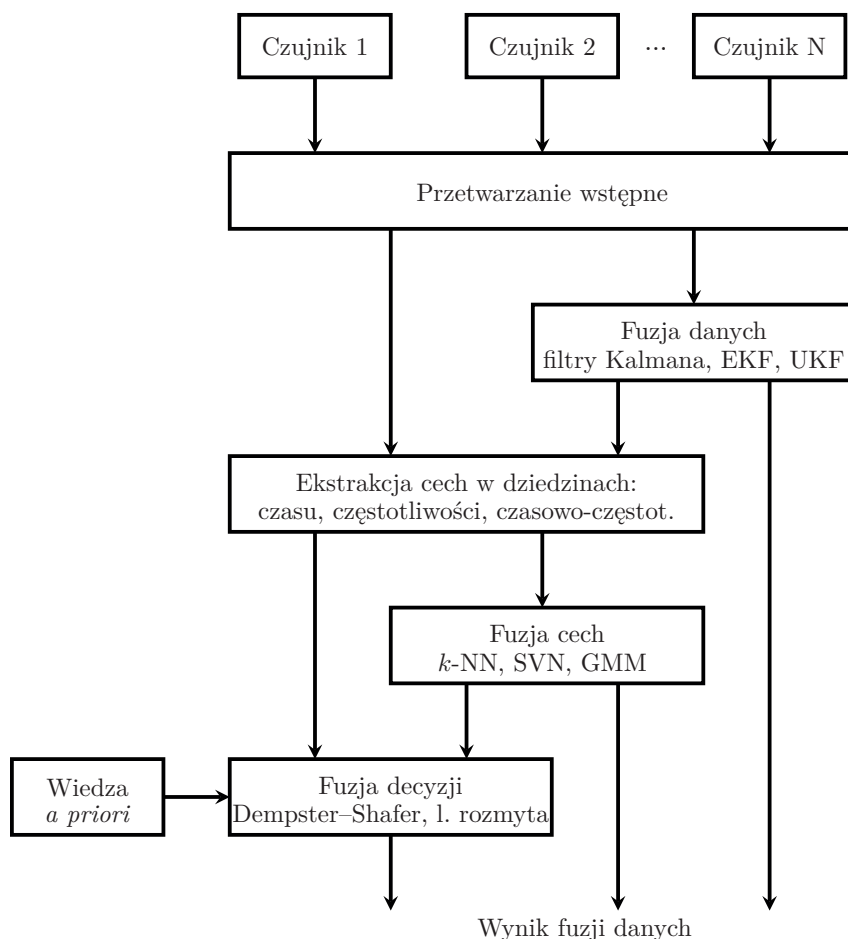
Rys. 2.10. Fuzja hybrydowa [337]

Architektura hybrydowa (rys. 2.10) jest kombinacją architektury centralnej i rozproszonej. Architektura hybrydowa wyróżnia się tym, że umożliwia realizację fuzji danych łącznie w architekturze centralnej i rozproszonej lub tylko w jednej z nich. To, który sposób zostanie wykorzystany, zależy od specyficznych wymagań nałożonych na projektowany system fuzji danych. Proste operacje, których wpływ na jakość końcowego efektu fuzji danych nie jest kluczowy, można realizować w architekturze rozproszonej. W przypadku danych kluczowych dla procesu fuzji danych lub takich, których przetworzenie wymaga znacznych mocy obliczeniowych, pomiary należy dostarczyć do jednostki centralnej w postaci takiej, jakiej zostały pozyskane z czujnika pomiarowego. Takie podejście zapewnia z jednej strony

minimalizację kosztów związanych z przesyłaniem danych, z drugiej natomiast obniża zapotrzebowanie na moc obliczeniową centralnej jednostki fuzji danych. Wadą architektury hybrydowej jest to, że wzrasta złożoność systemu wykonanego w tej architekturze, co może wpłynąć na wzrost zawodności tego rozwiązania.

2.5. Architektura systemu fuzji danych

Analiza poszczególnych poziomów fuzji danych oraz zadań związanych z każdym z nich umożliwiła zaproponowanie ogólnej architektury systemu fuzji danych.



Rys. 2.11. Architektura systemu fuzji danych

2. Elementy fuzji danych

W opracowanym systemie możemy wyróżnić trzy główne warstwy, które są powiązane z poziomami fuzji danych, tj. poziomem fuzji podstawowej, poziomem fuzji właściwości i poziomem fuzji decyzji. Architektura ta została przedstawiona na rys. 2.11.

Zaproponowana architektura systemu fuzji danych jest ogólna i podlega dostosowaniu do konkretnego zadania fuzji danych. Oczywiście nie wszystkie elementy przedstawione na rys. 2.11 muszą być wykorzystane w każdej realizacji systemu fuzji danych. To, które z nich zostaną wykorzystane, zależy od konkretnego zadania projektowego.

2.6. Modele fuzji danych

W ostatniej części tego rozdziału omówimy najważniejsze modele systemów fuzji danych, które są realizacjami zaproponowanej w podrozdziale 2.5 ogólnej architektury systemu fuzji danych. Elementy wchodzące w skład tej architektury omówiono w poprzednich podrozdziałach (tj. 2.1 oraz 2.2). Podkreślmy, że na przestrzeni lat opracowano wiele modeli fuzji danych o różnych cechach, które znalazły wiele zastosowań praktycznych. Modele fuzji danych można podzielić na następujące kategorie:

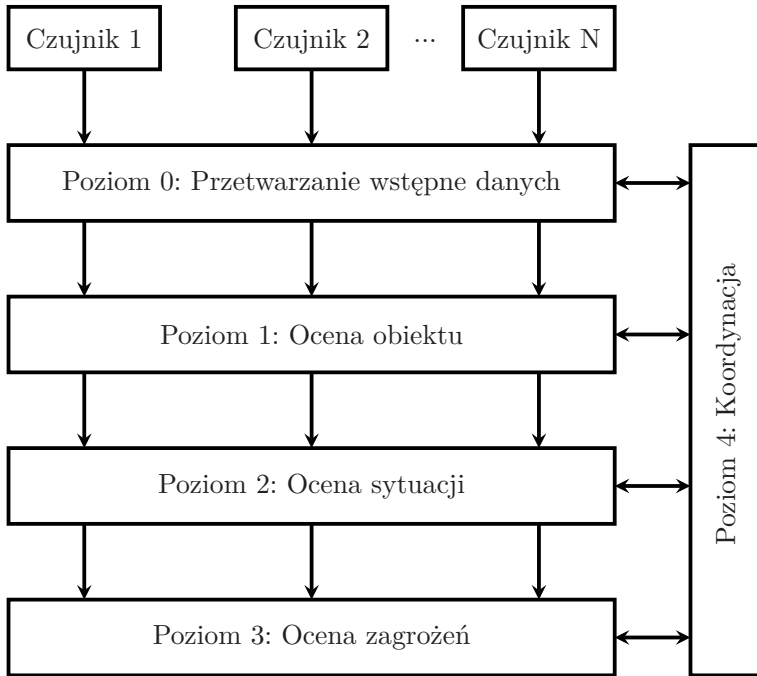
- modele bazujące na danych (ang. *data-based models*),
- modele bazujące na działaniach (ang. *activity-based models*).

W modelach bazujących na danych wyraźny jest podział ze względu na poziom abstrakcji. Do tej kategorii należą modele JDL oraz Dasarathy. Druga kategoria modeli charakteryzuje się tym, że wyraźnie wskazują, jakiego typu działania powinny być wykonane przez system fuzji danych pracujący w tym modelu. Przykładami takich modeli są modele Boyda, wodospad, zmodyfikowany wodospad oraz Omnibus. Omówimy i scharakteryzujemy teraz każdy z tych modeli.

2.6.1. Model JDL

Jednym z pierwszych modeli dla systemów fuzji danych jest model *Joint Directors' of Laboratories Model* (JDL). Model ten został zaproponowany przez Departament Obrony USA na przełomie lat 80. i 90. ubiegłego wieku [150, 151]. W zaproponowanym modelu wyróżniamy pięć poziomów (rys. 2.12).

Na poziomie 0 wykonywane są wszelkie zadania mające na celu przygotowanie danych, pochodzących z różnych źródeł, do ich dalszego przetwarzania na kolejnych poziomach (ang. *pre-processing of source data*). Do głównych zadań na tym poziomie zaliczamy zadania związane ze wstępnym przetwarzaniem danych, tj. usuwanie niepożądanych składowych o różnych charakterystykach, jak również normalizację oraz kompresję danych. Działania podejmowane na tym etapie mają



Rys. 2.12. Model JDL

głównie na celu takie przygotowanie danych, by możliwa była redukcja obliczeń koniecznych do wykonania na kolejnych etapach fuzji danych w tym modelu.

Kolejny poziom, tj. 1, dotyczy oceny obiektu (ang. *object refinement*). Do głównych zadań na tym poziomie w modelu JDL zaliczamy estymację stanu, grupowanie danych oraz ekstrakcję cech. Z kolei zadania na poziomie 2 (ocena sytuacji, ang. *situation refinement*) związane są z ustaleniem związków pomiędzy obiektami i obserwowanymi zdarzeniami. Poziom 3 to ocena wpływu aktualnego stanu obiektu, którą przeprowadzono na poziomie drugim. Dodatkowo na tym poziomie dokonuje się oceny potencjalnych zagrożeń i/lub awarii. Poziom ten nazywany jest poziomem oceny zagrożeń (ang. *threat refinement*). Na poziomie 4 monitoruje się procesy na poziomach niższych (od 0 do 3), zarządza się zasobami (np. czujnikami) oraz poprawia procesy na niższych poziomach. Poziom ten nazywany jest poziomem koordynacji (ang. *refinement of the process*).

W przypadku modelu JDL fuzja niskopoziomowa dotyczy poziomów 0 oraz 1, a fuzja wysokiego poziomu (decyzji) dotyczy poziomów od 2 do 4. W modelu tym również zwraca uwagę wykorzystywana terminologia, która nawiązuje do

2. Elementy fuzji danych

wojskowości. Oczywiście jest to związane z tym, że model JDL w swoim pierwotnym zastosowaniu był wykorzystywany w celach wojskowych.

2.6.2. Model Dasarathy

W przedstawionej klasyfikacji fuzji danych wskazaliśmy, że możliwy jest m.in. podział metod fuzji danych ze względu na poziom abstrakcji. Omówimy teraz model bazujący na tym podziale, tj. model Dasarathy [100]. Model ten różni się od wcześniej omówionego tym, że fuzja nie odbywa się przez wykonanie zadań na poszczególnych etapach lub poziomach, ale przez wykorzystanie odpowiednich funkcji. Funkcje w modelu Dasarathy zostały przedstawione na rys. 2.13.

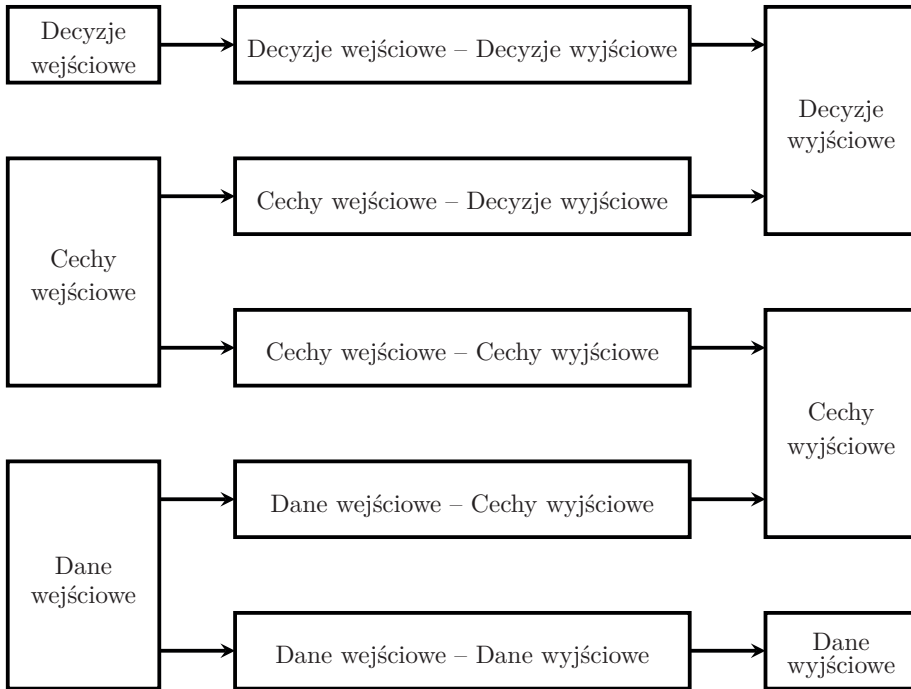
Pierwsza z funkcji modelu Dasarathy (patrz rys. 2.13) wymaga zarówno na wejściu, jak i na wyjściu pewnego zbioru decyzji. Celem wykonywanych operacji jest fuzja decyzji, tj. wyznaczenie nowej decyzji (lub nowych decyzji) na bazie innych (decyzje wejściowe). Na tym poziomie można również wykonać inną operację, tj. wzmocnić jedną z tych decyzji, które pojawiły się w zbiorze decyzji wejściowych. Operacje wykonywane przez tę funkcję w modelu *Dasarathy* jest tożsama z fuzją decyzji.

Funkcja mająca na wejściu zbiór cech, a na wyjściu decyzję (lub zbiór decyzji) jest drugą z omawianych funkcji w modelu Dasarathy. W tym przypadku funkcja ta bazuje głównie na algorytmach klasyfikacji, których wyjście traktowane jest jako decyzja wybierana z pewnego zbioru możliwych decyzji. Operacje wykonywane na tym poziomie odpowiadają konkludowaniu na poziomie drugim.

Funkcja, w której na wejściu wymagane są cechy, a na wyjściu otrzymujemy inny zbiór cech odpowiada operacji konstrukcji wektora cech na podstawie wygenerowanych cech. W ramach zadania konstrukcji wektora cech rozwiązuje się zadania redukcji oraz selekcji cech.

Kolejną funkcją w modelu Dasarathy jest funkcja, na której wejściu wymagane są nieprzetworzone dane pomiarowe, a na wyjściu otrzymujemy zestaw cech opisujących obiekt. W tym przypadku do generowania cech na podstawie nieprzetworzonych danych pomiarowych związanych z opisywanym obiektem wykorzystywane są odpowiednio zaprojektowane algorytmy do generowania cech.

Ostatnia z funkcji modelu Dasarathy wymaga na wejściu danych nieprzetworzonych. Z kolei na jej wyjściu dostępne są dane, które zostały już wstępnie przetworzone. Tak przygotowane dane umożliwiają znacznie bardziej precyzyjny i rzetelny opis obiektu, z którym związane są dostarczone na wejście funkcji nieprzetworzone dane pomiarowe. Na tym etapie wykorzystuje się głównie algorytmy znane z przetwarzania sygnałów, np. wykorzystywane do detekcji zmian oraz usuwania niepożądanych składowych z sygnałów itp.



Rys. 2.13. Model Dasarathy

Model Dasarathy, ze względu na wykorzystane w nim funkcje, często nazywany jest **funkcyjnym modelem Dasarathy**. W modelu tym wyraźnie też są zaznaczone typy danych wejściowych i wyjściowych, tj. dane pomiarowe, cechy i decyzje.

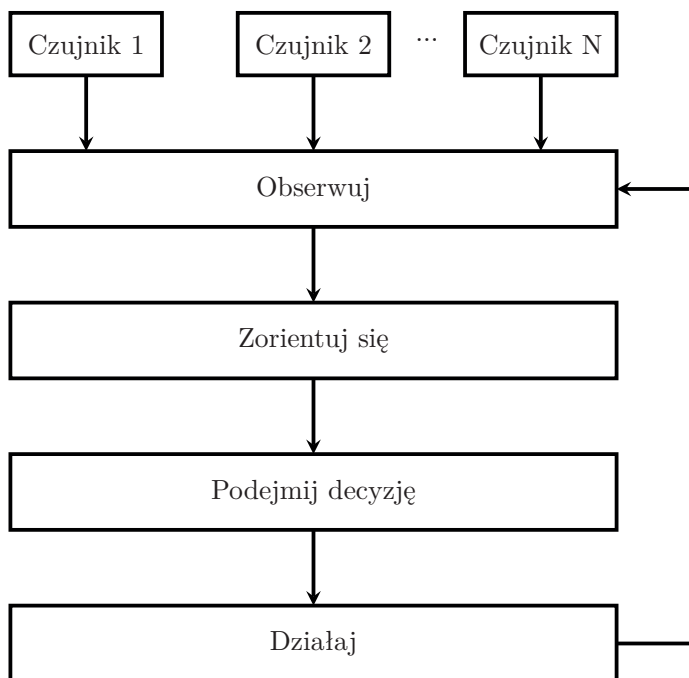
2.6.3. Model Boyda

Model **Boyda**, nazywany również pętlą Boyda, podobnie jak model JDL, opracowany został na potrzeby militarne. W przypadku tego modelu wyróżniamy następujące cztery etapy (rys. 2.14):

- obserwuj (ang. *observe*),
- zorientuj się (ang. *orient*),
- podejmij decyzję (ang. *decide*),
- działaj (ang. *act*).

Etap pierwszy – **obserwuj** – odpowiada poziomowi 1 w modelu JDL i związany jest z przygotowaniem danych pochodzących z różnych źródeł w celu ich dalszego przetwarzania na kolejnych etapach.

2. Elementy fuzji danych



Rys. 2.14. Model Boyda

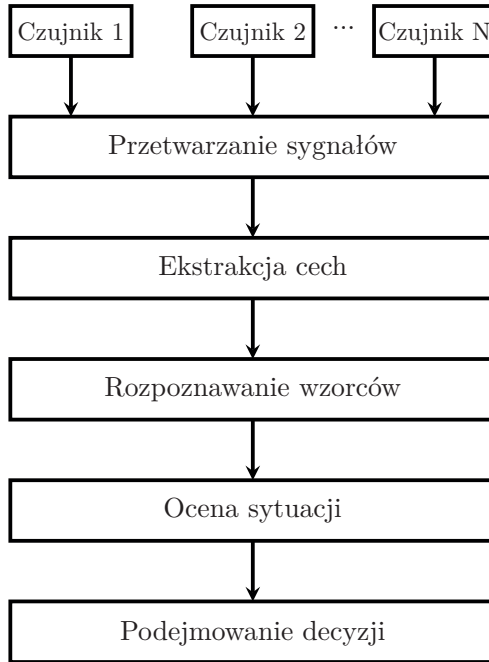
Etap drugi – **zorientuj się** – odpowiada poziomom od 2 do 4 w modelu JDL. Etap ten związany jest z przetwarzaniem danych zgromadzonych na wcześniejszym etapie. Na tym etapie dokonuje się fuzji przetworzonych danych oraz, jeśli jest to konieczne, wprowadza się wiedzę specjalistyczną. Celem tego etapu jest analiza sytuacji oraz ocena zagrożeń lub ewentualnych awarii.

Etap **podjęcia decyzji** to ocena wariantów decyzyjnych, ustalenie ograniczeń i kryteriów wyborów spośród dostępnych wariantów. Następnie dokonuje się wyboru optymalnej decyzji na podstawie ustalonego kryterium z uwzględnieniem ograniczeń. Etap podejmowania decyzji odpowiada etapowi piątemu w modelu JDL.

Ostatni etap, tj. (**działaj**), to etap realizacji podjętej/podjętych decyzji na etapie wcześniejszym oraz ocena jej/ich wpływu na otoczenie. Jest to również etap, po którego wykonaniu następuje powrót do etapu pierwszego.

Konstrukcja modelu Boyda wskazuje na to, że proces fuzji danych może być procesem iteracyjno-przyrostowym, tzn. kolejne etapy przechodzi się wielokrotnie, dokonując weryfikacji podjętych działań. Jeśli jest to konieczne, w każdej iteracji gromadzi się nowe dane lub wiedzę oraz uzupełnia zgromadzone już dane oraz wiedzę dotyczącą obiektu.

2.6.4. Model typu wodospad



Rys. 2.15. Model typu wodospad

Model typu **wodospad** (rys. 2.15) – w przeciwieństwie do modelu Boyda – nie ma pętli, natomiast podobnie jak model JDL charakteryzuje się budową hierarchiczną. Odmianą modelu typu wodospad jest model typu **zmodyfikowany wodospad** (rys. 2.16). W modelach tych wyróżniamy następujące poziomy:

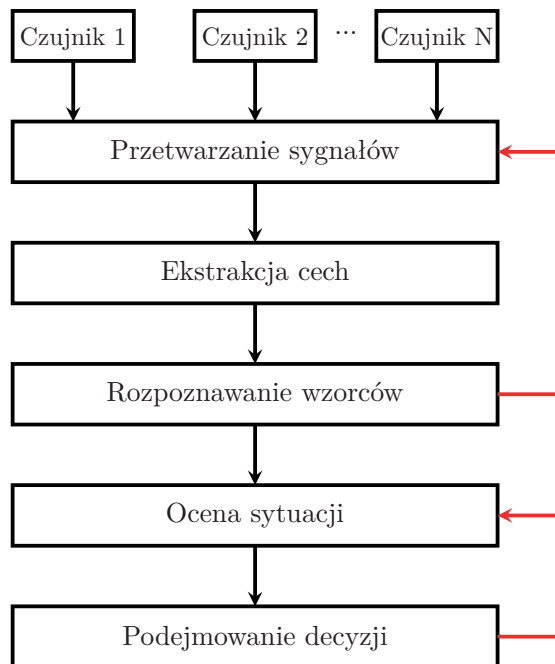
- pomiar (ang. *sensing*),
- przetwarzanie sygnałów (ang. *signal processing*),
- ekstrakcja cech (ang. *feature extraction*),
- rozpoznawanie wzorców (ang. *pattern processing and recognition*),
- ocena sytuacji (ang. *situation assesment*),
- podejmowanie decyzji (ang. *decision making*).

W zmodyfikowanym modelu typu wodospad pojawiają się dwie lokalne pętle sprzężenia zwrotnego (patrz rys. 2.16).

Warstwa pierwsza w modelu typu wodospad dotyczy pomiarów z wykorzystaniem czujników. Warstwy kolejne (2–5), odpowiadają zarówno warstwom 2–4 w modelu JDL, jak i etapowi drugiemu w modelu Boyda.

2. Elementy fuzji danych

Warstwa 5 (podejmowanie decyzji) odpowiada warstwie 5 w modelu JDL i etapowi podejmowania decyzji w modelu Boyda.



Rys. 2.16. Zmodyfikowany model typu wodospad

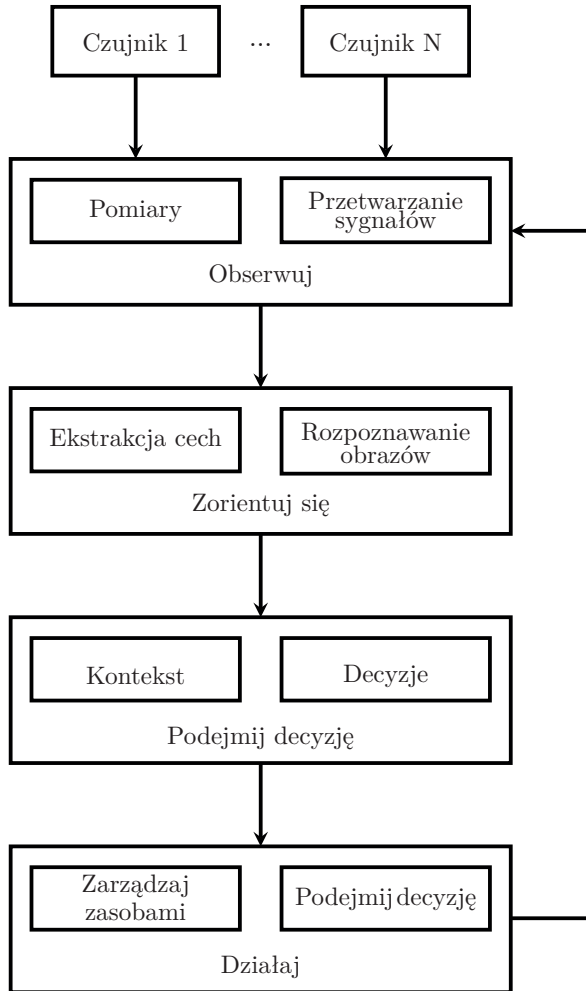
2.6.5. Model typu Omnibus

Model typu **Omnibus** (rys. 2.17) jest połączeniem wybranych elementów z modelu Boyda, Dasarathy i modelu typu wodospad. W modelu tym nie wykorzystuje się terminologii wojskowej, co nadaje mu charakter modelu uniwersalnego.

W konstrukcji modelu typu Omnibus przyjęto, że model:

- powinien umożliwiać ustalenie kolejności wykonywanych obliczeń,
- powinien wspierać iteracyjne wykonywanie operacji,
- powinien wskazywać wady i zalety różnych metod fuzji danych,
- powinien wykorzystywać typową terminologię,
- nie powinien być zorientowany na zastosowania wojskowe.

Model typu Omnibus jest ogólnym modelem, zawierającym najlepsze elementy wchodzące w skład omówionych wcześniej modeli [337]. Warto również podkreślić, że model ten jest w swojej budowie zbliżony do przedstawionego w podrozdziale 2.1



Rys. 2.17. Model typu Omnibus

schematu przetwarzania danych tak dla jednego czujnika (patrz rys. 2.1), jak i dla przypadku z wieloma czujnikami (rys. 2.2). W modelu Omnibus, podobnie jak to zostało przedstawione w zaproponowanych schematach, cykl obliczeniowy rozpoczyna się od przeprowadzenia pomiarów i ich wstępnego przetworzenia. W następnym kroku wykonywana jest ekstrakcja cech, po której, bazując na wyznaczonych cechach, podejmowana jest decyzja. Ostatni etap w modelu Omnibus to realizacja zaplanowanych działań zgodnych z decyzjami podjętymi na wcześniejszym etapie.

2. Elementy fuzji danych

Warto zauważyć, że model Omnibus w przeciwieństwie do modelu JDL wprowadza numerację kolejnych etapów fuzji danych, co wyraźnie wskazuje na cykliczność tego procesu. Innym elementem, o którym już wspomniano, jest terminologia. Jest ona ogólniejsza i nie nawiązuje do terminologii militarnej jak to było chociażby w przypadku modelu JDL. Jest to różnica mniej istotna, jednakże warta podkreślenia, ponieważ nie narzuca niepotrzebnych skojarzeń wojskowych.

Wadą modelu Omnibus jest to, że nie umożliwia on dekompozycji na moduły, które mogłyby być oddzielnie implementowane i testowane. Umożliwiałoby to z jednej strony uproszczenie implementacji modelu w projektowanym systemie fuzji danych, a z drugiej wykorzystanie tych modułów w innych implementacjach.

Przedstawione w tym rozdziale modele dla fuzji danych mają wiele wspólnych elementów. Nazewnictwo dla części z nich jest bardzo podobne. Twórcy niektórych z nich (np. model Omnibus) starali się odejść od oryginalnego wojskowego nazewnictwa. Dzięki temu modele stają się uniwersalne. Przedstawiony w tym rozdziale przegląd modeli fuzji danych ma na celu ułatwienie doboru spośród nich modelu fuzji danych dostosowanego do projektowanego systemu przetwarzania danych.

3. Sygnały, pomiary, szумы

Organizm przeciwstawia się chaosowi, dezintegracji, śmierci, podobnie jak sygnał przeciwstawia się szumowi.

Norbert Wiener

3.1. Wstęp

W rozdziale omówiono podstawowe zagadnienia dotyczące sygnałów, metod ich reprezentacji oraz technik ich pozyskiwania z otoczenia. Przedmiotem naszego zainteresowania będą m.in. systemy fuzji danych, w których sygnały pozyskiwane są z wykorzystaniem czujników pomiarowych (np. działających w ramach sieci typu BSN).

Pozyskiwanie sygnałów z otoczenia nazywamy **akwizycją danych**. Do głównych zadań akwizycji danych zaliczamy rejestrację pomiaru, jego próbkowanie i kwantyzację oraz transmisję. To jaki sygnał otrzymamy w wyniku przeprowadzonego procesu akwizycji zależy zarówno od typu obserwowanego zjawiska, jak i zastosowanych czujników pomiarowych. Na przykład jeśli przedmiotem naszego zainteresowania jest człowiek, to możemy skupić się albo na wykonywanych ruchach, albo pracy mięśni wykonywanej podczas tych ruchów. Do śledzenia ruchu człowieka możemy zastosować na przykład kamerę wideo, natomiast do monitorowania aktywności mięśni podczas ruchu wykorzystamy czujniki elektromiograficzne. Zauważmy, że rodzaj zjawiska i zastosowana metoda pomiarowa wpływa na to, jakiego typu sygnał będzie przetwarzany na dalszych etapach.

Niezależnie od metody pomiarowej, dane zgromadzone z wykorzystaniem czujników pomiarowych zawierają pewne niepożądane składniki. Ich źródła są różnorodne i mogą być związane m.in. z:

- niepewnością obiektu lub procesu pomiarowego,
- działaniem czujnika pomiarowego,
- błędami podczas transmisji,
- próbkowaniem,

3. Sygnały, pomiary, szумы

- kwantyzacją,
- czynnikami zewnętrznymi.

Niepożądane składniki sygnałów pomiarowych nazywamy szumami. **Szумы** dzielimy na **szумы wewnętrzne (procesowe)** oraz **szумы zewnętrzne**. Szумы zewnętrzne nazywa się również **zakłóceniami** [155]. Szумы wewnętrzne związane są z obserwowanym obiektem. Z kolei szумы zewnętrzne dotyczą zjawisk, które nie mają bezpośredniego związku z obserwowanym obiektem. Szумы te mogą być wywołane np. zmianami temperatury w otoczeniu obserwowanego obiektu. W efekcie ich działania w sygnale pomiarowym mogą pojawić się niepożądane składniki, które utrudniają ich późniejsze przetwarzanie.

Inną, równie istotną kwestią, która związana jest z sygnałami, są metody ich **reprezentacji**. W rozdziale omówione zostaną zarówno klasyczne metody reprezentacji sygnałów, takie jak **reprezentacja czasowa i częstotliwościowa**, jak też nowsze podejście, tj. **reprezentacja czasowo-częstotliwościowa**.

3.2. Sygnał

Sygnałem nazywamy pewną wielkość, której wartość może zostać zmierzona i która umożliwia przenoszenie informacji. Przykładem sygnału jest temperatura, ponieważ jej wartość jest zmienna, może więc być mierzona i jest zdolna do przenoszenia informacji. Wprowadźmy teraz definicję pojęcia **sygnału** [258, 417]: **Definicja 3.2.1 (Sygnał)**. *Sygnałem nazywamy zmianę pewnej mierzalnej wielkości fizycznej bądź stanu obiektu fizycznego, która jest nośnikiem informacji.*

Sygnały takie jak wymieniona temperatura, czyli takie, których dziedziną jest jednowymiarowa, będziemy oznaczać jako $s(t)$, gdzie $s \in \mathbb{R}^1$. Zaznaczmy, że w praktyce sygnały mogą być funkcjami więcej niż jednej zmiennej. Na przykład zdjęcie może być traktowane jako sygnał będący funkcją dwóch zmiennych, tj. współrzędnych odciętych i rzędnych. A jego dziedziną jest dwuwymiarowa, tj. $s \in \mathbb{R}^2$. Kolejnym przykładem jest nagranie wideo. W tym przypadku mamy nie jedno zdjęcie, a sekwencję zdjęć, które zmieniają się w czasie. Dziedzina takiego sygnału jest trójwymiarowa, a sam sygnał jest funkcją trzech zmiennych. Dwie z nich są zmiennymi przestrzennymi (tak jak dla zdjęcia), a trzecia jest zmienną w czasie.

Zmiany wielkości mogą zachodzić tak jak z przykładu powyżej, tj. w czasie. Wówczas takie sygnały nazywamy **sygnałami czasowymi** (ang. *temporal signal*). Sygnał może zmieniać się również w przestrzeni. Wówczas takie sygnały nazywamy **sygnałami przestrzennymi** (ang. *spatial signal*). Przykładem takiego sygnału jest zdjęcie. Zmiany wielkości fizycznych mogą zachodzić również jednocześnie w czasie i w przestrzeni. Takie sygnały nazywamy **sygnałami przestrzenno-czasowymi** (ang. *spatial-temporal signals*). Przykładem sygnału

przestrzenno-czasowego jest nagranie wideo, ponieważ składa się on z sekwencji zdjęć zmieniających się w czasie.

Sygnały mogą być klasyfikowane na wiele sposobów według różnych kryteriów. Na przykład sygnały można dzielić np. ze względu na zdolność do przewidywania przyszłych wartości sygnału. W tym przypadku wyróżniamy sygnały **zdeterminowane** i **losowe**. Sygnałem zdeterminowanym nazywamy sygnał, który może być opisany za pomocą ścisłych zależności matematycznych [33]. Natomiast sygnały, których nie można opisać za pomocą ścisłych zależności matematycznych, nazywamy sygnałami niezdeterminowanymi (losowymi) [33]. Sygnały te charakteryzują się tym, że nie jest możliwe ustalenie dokładnej wartości tego sygnału w pewnej chwili w przyszłości. Natura tych sygnałów jest losowa i mogą być one jedynie opisywane przez charakterystyki statystyczne.

W wielu przypadkach trudno ustalić, czy analizowany sygnał jest sygnałem zdeterminowanym czy losowym. W praktyce przyjmuje się, że sygnał jest zdeterminowany wówczas, gdy wielokrotnie przeprowadzone doświadczenie w identycznych warunkach daje jednakowe wyniki (po uwzględnieniu błędu pomiarowego). W przeciwnym przypadku rozpatrywany sygnał nazywamy sygnałem losowym [33].

Inne kryterium podziału sygnałów to podział ze względu na zależność charakterystyk statystycznych od czasu. Jeżeli charakterystyki te nie zależą od czasu, to takie sygnały nazywamy sygnałami **stacjonarnymi**. Jeżeli charakterystyki zależą od czasu, to taki sygnał nazywamy **niestacjonarnym** [33].

Definicja 3.2.2 (Sygnał stacjonarny). *Sygnał stacjonarny to sygnał, którego właściwości statystyczne lub struktura częstotliwościowa nie zmieniają się w czasie.*

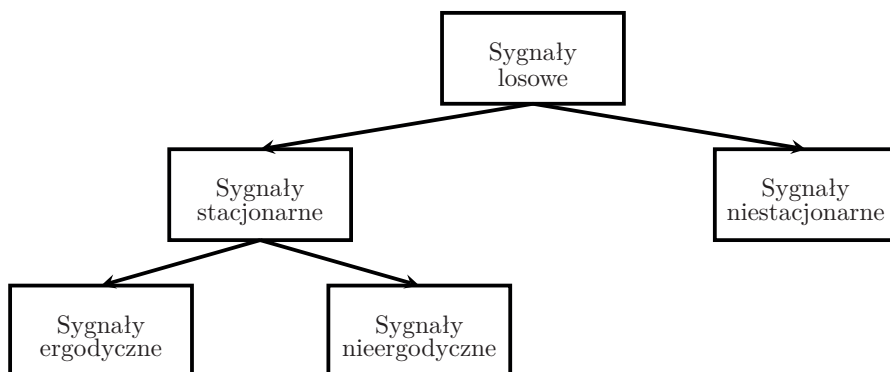
Sygnały stacjonarne dzielimy na sygnały **ergodyczne** i **nieergodyczne**. Sygnały ergodyczne charakteryzują się tym, że wybrana charakterystyka statystyczna sygnału otrzymana dla dowolnego zbioru realizacji jest równa tej charakterystyce, którą otrzymaliśmy dla jednej realizacji. Oznacza to, że jedna realizacja jest reprezentatywna dla całego zbioru realizacji. Sygnały nieergodyczne nie mają tej właściwości. Na rysunku 3.1 przedstawiono klasyfikację sygnałów losowych.

Sygnały zdeterminowane można natomiast podzielić np. ze względu na typ powtarzalności. W takim podziale wyróżniamy sygnały **okresowe** i **nieokresowe**. Z kolei sygnały okresowe dzielimy na sygnały **harmoniczne** i **poliharmoniczne**. Natomiast sygnały nieokresowe dzielimy na sygnały **prawie okresowe** i **przejsściowe**. Klasyfikację sygnałów zdeterminowanych przedstawiono na na rys. 3.2.

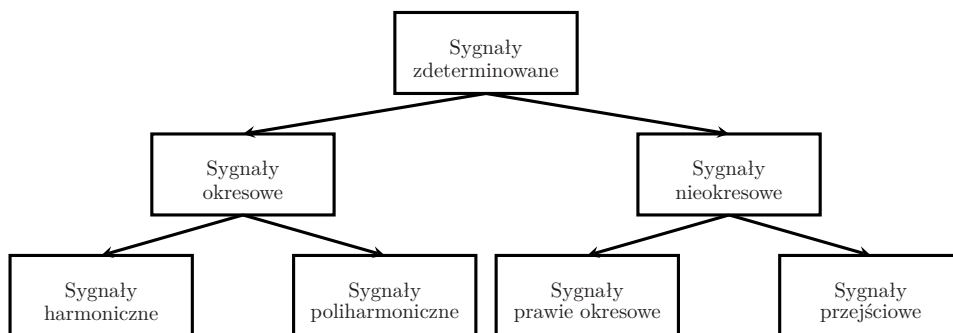
Sygnał harmoniczny jest sygnałem, którego zmienność można opisać funkcją sinusoidalną. Natomiast sygnały poliharmoniczne to sygnały, które złożone są z wielu sygnałów harmonicznych.

Sygnał prawie okresowy to sygnał, który powstaje przez zsumowanie co najmniej dwóch niezależnych sygnałów harmonicznych. Ponadto istotne jest w tym przypadku, by nie wszystkie ilorazy możliwych par częstotliwości sygnałów har-

3. Sygnały, pomiary, szумы



Rys. 3.1. Klasyfikacja sygnałów losowych [33]



Rys. 3.2. Klasyfikacja sygnałów zdeterminowanych [33]

monicznych były liczbami wymiernymi [33]. Sygnały przejściowe (nieustalone) to wszystkie sygnały nieokresowe, które nie są sygnałami prawie okresowymi.

Sygnały możemy również usystematyzować ze względu na dziedzinę określoności. Jeżeli sygnał określony jest w zbiorze ciągłym czasu, przestrzeni lub czasoprzestrzeni to taki sygnał nazywamy sygnałem **ciągłym**. Natomiast sygnały, które określone są w dyskretnym zbiorze czasu, przestrzeni lub czasoprzestrzeni nazywamy sygnałami **dyskretnymi**. W przypadku gdy dyskretna jest również dziedzina wartości sygnału dyskretnego, to taki sygnał nazywamy sygnałem **cyfrowym**.

Tak jak wspomniano, sygnały, których właściwości statystyczne lub struktura częstotliwościowa zmieniają się w czasie, nazywamy sygnałami **niestacjonarnymi**. Większość sygnałów fizycznych to sygnały niestacjonarne. Sygnały niestacjonarne otrzymuje się np. w przypadku występowania nagłych zmian warunków otoczenia

pracy systemu (sygnały nieustalone, przejściowe) lub gdy czas pracy systemu jest tak długi, że zmieniają się właściwości tego systemu.

W zasadzie nie istnieje metoda analizy sygnałów niestacjonarnych. Każdy z takich sygnałów należy rozpatrywać osobno. Podejmowane są również próby analizy sygnałów niestacjonarnych należących do pewnych klas sygnałów niestacjonarnych. Wśród klas sygnałów niestacjonarnych wyróżniamy sygnały o zmiennej w czasie wartości średniej lub sygnały o zmiennej w czasie strukturze częstotliwościowej.

Oddzielnym zagadnieniem jest budowa zależności pomiędzy niestacjonarnym sygnałem wejściowym a niestacjonarnym sygnałem wyjściowym systemu. Zagadnienie budowy takich zależności zostanie omówione w dalszej części książki.

3.3. Metody reprezentacji sygnałów

Ten sam sygnał można zapisać w różny sposób, czyli może mieć on różne reprezentacje [39, 387]. Wybór metody reprezentacji sygnału ma znaczenie zarówno praktyczne, jak i teoretyczne. W pierwszym przypadku może wpłynąć m.in. na wynik przetwarzania sygnału, w drugim może mieć znaczenie podczas formalnej analizy sygnału.

Reprezentacja sygnału jest rodzajem symbolicznego opisu sygnału, który to charakteryzuje się znacznym stopniem abstrakcji [387]. Cechą charakterystyczną reprezentacji sygnału jest to, że zawiera on pełną informację o sygnale. Reprezentacje sygnałów różnią się od siebie sposobem jej zapisu.

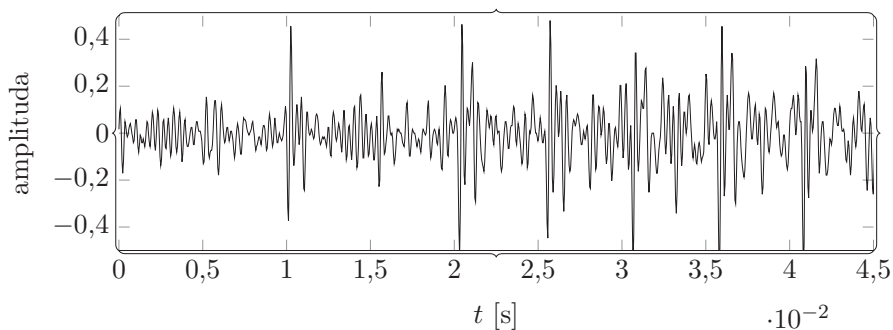
Podziału metod reprezentacji sygnałów można dokonać np. ze względu na dziedzinę, w której sygnał jest reprezentowany. W tym przypadku reprezentacje dzielimy na reprezentacje w dziedzinie **czasu** (t), **częstotliwości** (f) i **czasowo-częstotliwościowej** ($t-f$) oraz **reprezentację rzadką**. Taki podział został przyjęty w tej pracy i będzie wykorzystywany w dalszej jej części. Dodajmy jeszcze, że ostatni typ reprezentacji sygnału zostanie omówiony szczegółowo w dalszej części książki.

Przykłady reprezentacji sygnału $s \in \mathbb{R}^1$ w dziedzinach t , f , $t-f$ przedstawiono na rys. 3.3, 3.8 oraz 3.13. W dalszej części tego rozdziału omówiono poszczególne sposoby reprezentacji sygnałów oraz ich cechy charakterystyczne.

3.3.1. Reprezentacja czasowa sygnałów

Reprezentacja czasowa odpowiada typowemu sposobowi reprezentacji sygnału. Reprezentacja ta wskazuje początek i koniec sygnału, określa sposób zmiany sygnału w czasie, wyznacza jego chwilową amplitudę, jak również ocenia rozkład energii sygnału w czasie oraz ilustruje zmiany bieżącego kształtu sygnału. Przykład reprezentacji sygnału w dziedzinie czasu przedstawiono na rys. 3.3.

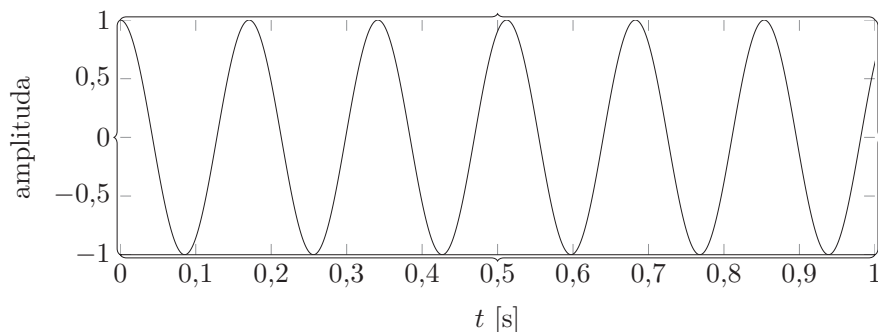
3. Sygnały, pomiary, szумы



Rys. 3.3. Przykład reprezentacji czasowej sygnału

Na rysunkach 3.4–3.7 przedstawiono kilka innych przykładów sygnałów oraz ich reprezentacje w dziedzinie czasu.

Pierwszy z sygnałów (rys. 3.4) jest sygnałem sinusoidalnym, do którego opisu wystarczy podanie wartości amplitudy, przesunięcia fazowego oraz pulsacji (ściśle związanej z częstotliwością).

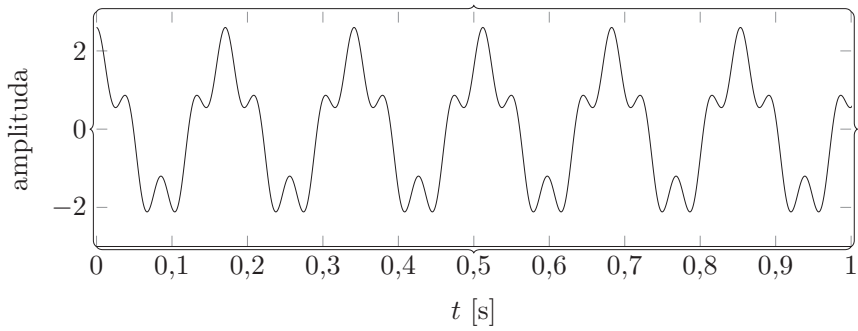


Rys. 3.4. Przykład sygnału sinusoidalnego

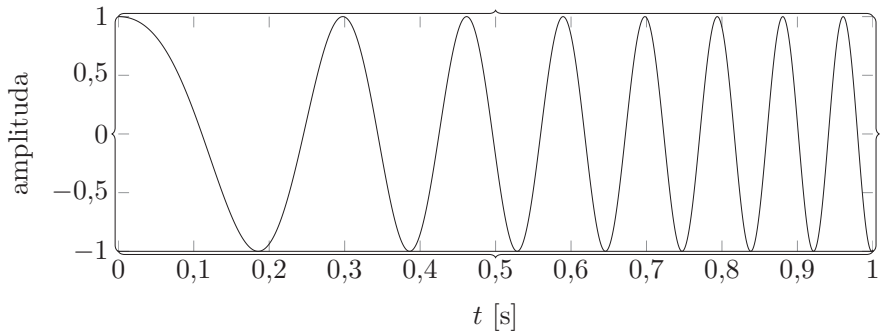
Na rysunku 3.5 przedstawiono sygnał deterministyczny będący złożeniem kilku sygnałów sinusoidalnych. W sygnale tym, oprócz amplitudy, przesunięcia fazowego oraz częstotliwości podstawowej, wyróżniamy również częstotliwości harmoniczne. Częstotliwości te są wielokrotnościami częstotliwości podstawowej.

Kolejnym przykładem jest sygnał o modulowanej częstotliwości (rys. 3.6). Jedną z wielkości, która określa ten sygnał jest częstotliwość chwilowa. W przypadku sygnału o modulowanej częstotliwości cechą charakterystyczną jest to, że jego częstotliwość chwilowa jest funkcją czasu. Przykładową metodą szacowania często-

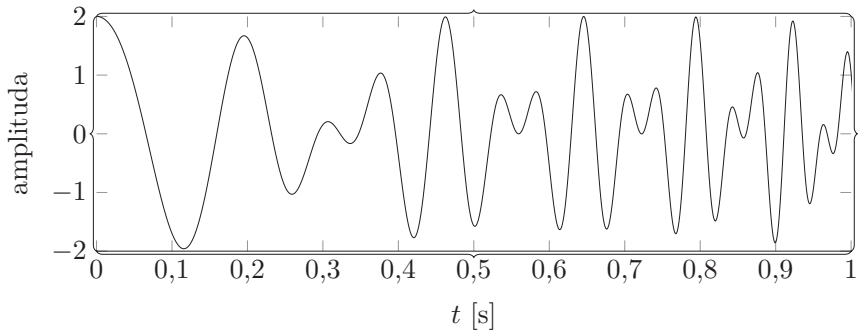
3.3. Metody reprezentacji sygnałów



Rys. 3.5. Przykład sygnału będącego złożeniem kilku sygnałów sinusoidalnych



Rys. 3.6. Przykład sygnału o modulowanej częstotliwości



Rys. 3.7. Przykład sygnału o modulowanej częstotliwości i amplitudzie

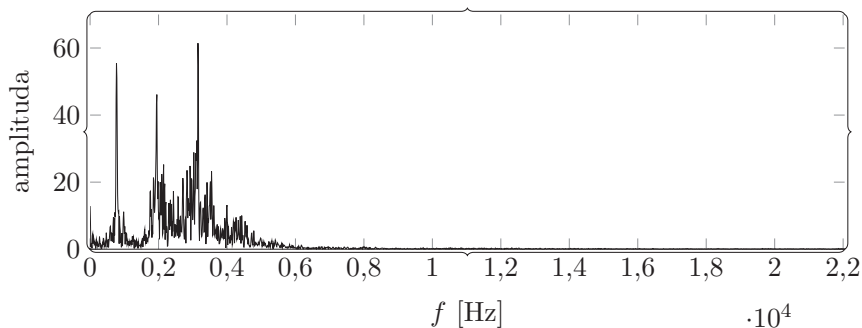
3. Sygnały, pomiary, szумы

tliwości chwilowej tego sygnału jest zliczanie liczby przejść przez zero (ang. *zero crossings rate*). Istnieją również inne, dokładniejsze, metody służące do estymacji chwilowej częstotliwości sygnału, które zostaną omówione w dalszej części tej pracy. Przedstawiony sygnał nazywamy sygnałem niestacjonarnym **jednokomponentowym**.

Na kolejnym z rysunków (rys. 3.7) przedstawiono sygnał, w którym zarówno amplituda, jak i częstotliwość są zmienne w funkcji czasu. Sygnał ten jest składową sygnałów przedstawionych na rys. 3.5, które to sygnały składowe pomnożono przez odpowiednie współczynniki. Każdy z komponentów opisany jest amplitudą i częstotliwością chwilową. Sygnał ten jest przykładem sygnału niestacjonarnego **wielokomponentowego**.

3.3.2. Reprezentacja częstotliwościowa sygnałów

Analizując sygnały omówione w poprzednim podrozdziale, zauważamy, że zmiany w sygnałach mogą zachodzić z różną prędkością. Intensywność tych zmian oraz to, w jaki sposób one zachodzą, stanowią jego ważną charakterystykę. Analizę tych zjawisk w sygnale ułatwia alternatywna do reprezentacji czasowej sygnału reprezentacja w dziedzinie częstotliwości. Reprezentacja częstotliwościowa jest ściśle powiązana z reprezentacją czasową i w wielu przypadkach stanowi efektywniejsze narzędzie badania sygnałów. Chcąc uzyskać reprezentację sygnału $s(t)$ w dziedzinie częstotliwości, należy dokonać jego konwersji z dziedziny czasu. Do tego przekształcenia wykorzystać można np. transformację Fouriera. Przekształcenie to charakteryzuje się tym, że ustala się relację pomiędzy dziedziną czasu a dziedziną częstotliwości, w których analizowany sygnał może być reprezentowany. Na rysunku 3.8 przedstawiono przykład takiej reprezentacji, tj. częstotliwościową reprezentację sygnału z rys. 3.3.



Rys. 3.8. Przykład reprezentacji sygnału w dziedzinie częstotliwości (moc widmowa)

3.3. Metody reprezentacji sygnałów

Przekształcenie Fouriera sygnału $s(t)$ zdefiniowane jest następującym przekształceniem całkowym:

$$S_{\text{FT}}(f) = \int_{-\infty}^{\infty} s(t)e^{-j2\pi ft} dt \quad (3.1)$$

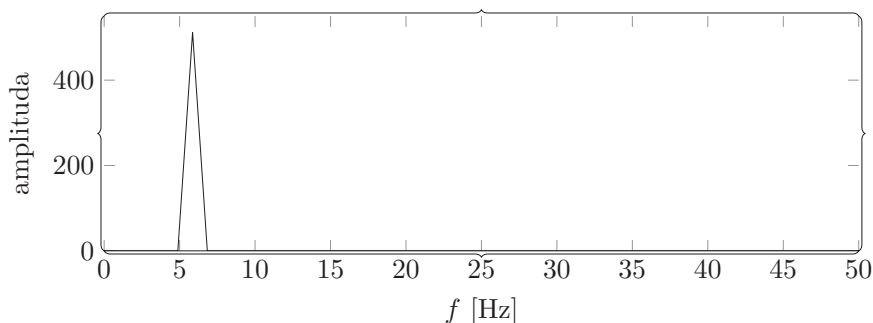
Przekształcenie (3.1) przyporządkowuje sygnałowi $s(t)$ funkcję zespoloną $S_{\text{FT}}(f)$. Reprezentacja $S_{\text{FT}}(f)$ sygnału $s(t)$ wskazuje częstotliwości komponentów składowych analizowanego sygnału oraz ich amplitudy. Ze względu na specyficzne właściwości reprezentacji $S_{\text{FT}}(f)$ sygnału, reprezentacja ta nazywana jest **gęstością widmową** lub krótko **widmem** [387].

Widmo stanowi podstawową metodę opisu sygnału w dziedzinie częstotliwości. Warto podkreślić, że pojęcie widma nie ma bezpośredniej interpretacji fizycznej [387]. Interpretację fizyczną mają natomiast widma **amplitudowe** i **fazowe**, które wyznaczamy z następujących zależności:

$$\begin{aligned} A(f) &= |S_{\text{FT}}(f)| \\ \rho(f) &= \arg(S_{\text{FT}}(f)) \end{aligned} \quad (3.2)$$

gdzie $A(f)$ jest widmem amplitudowym, a $\rho(f)$ jest widmem fazowym sygnału. Wielkości te opisują gęstość amplitudy i odpowiednio gęstość fazy składowych harmonicznym sygnału $s(t)$ w funkcji częstotliwości f . Na rysunkach 3.9–3.12 przedstawiono częstotliwościowe reprezentacje sygnałów z rysunków 3.4–3.7.

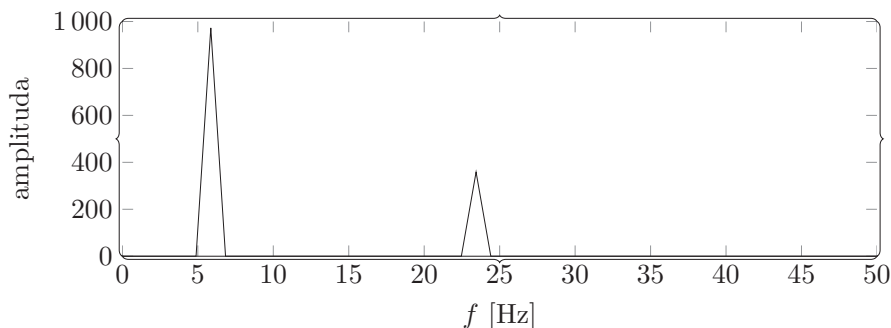
Widmo amplitudowe sygnału sinusoidalnego przedstawiono na rys. 3.9. Z przedstawionego wykresu wynika, że sygnał sinusoidalny w dziedzinie częstotliwości reprezentowany jest przez pojedynczy impuls o określonej amplitudzie dla określonej częstotliwości.



Rys. 3.9. Reprezentacja sygnału sinusoidalnego w dziedzinie częstotliwości

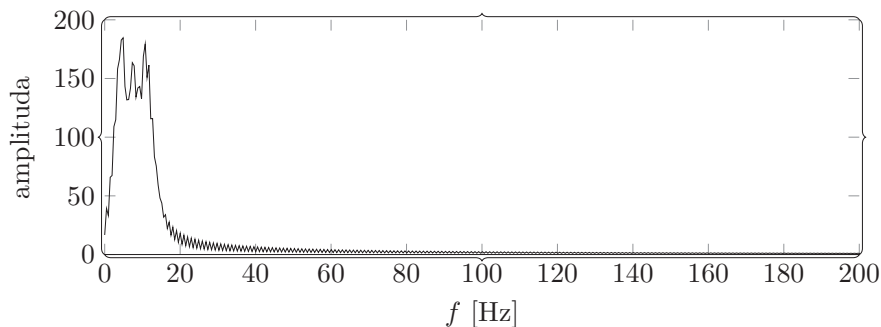
3. Sygnały, pomiary, szумы

Z kolei widmo amplitudowe sygnału będącego złożeniem kilku sygnałów sinusoidalnych przedstawiono na rys. 3.10. Jeżeli sygnał, tak jak w tym przypadku, składa się z kilku komponentów, to każdy z tych komponentów będzie reprezentowany przez pojedynczy impuls o określonej amplitudzie dla określonej częstotliwości. Analizując widmo z rys. 3.10, zauważamy, że wypadkowy sygnał jest złożeniem dwóch sygnałów sinusoidalnych o różnych częstotliwościach.



Rys. 3.10. Reprezentacja sygnału będącego złożeniem dwóch sygnałów sinusoidalnych w dziedzinie częstotliwości

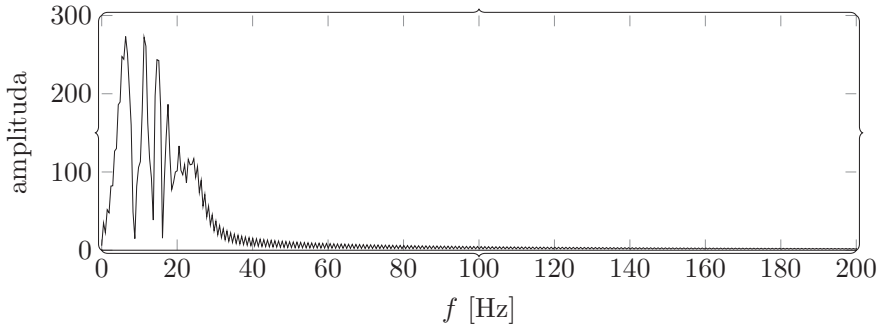
Kolejny sygnał, tj. o modulowanej częstotliwości, został przedstawiony w formie widma amplitudowego na rys. 3.11. W tym przypadku analiza częstotliwościowa wskazuje, że częstotliwość chwilowa sygnału zmienia się w sposób ciągły w pewnym określonym zakresie.



Rys. 3.11. Reprezentacja sygnału o modulowanej częstotliwości w dziedzinie częstotliwości

3.3. Metody reprezentacji sygnałów

Natomiast na rysunku 3.12 przedstawiono widmo amplitudowe sygnału o modulowanej częstotliwości i amplitudzie. Również i w tym przypadku nie jest możliwe rozdzielenie poszczególnych komponentów sygnału oraz precyzyjne wskazanie wartości częstotliwości. Zauważmy, że częstotliwość chwilowa sygnału zmienia się w sposób ciągły w pewnym zakresie częstotliwości. Dlatego dla sygnału możliwe jest jedynie wskazanie częstotliwości minimalnej i maksymalnej.



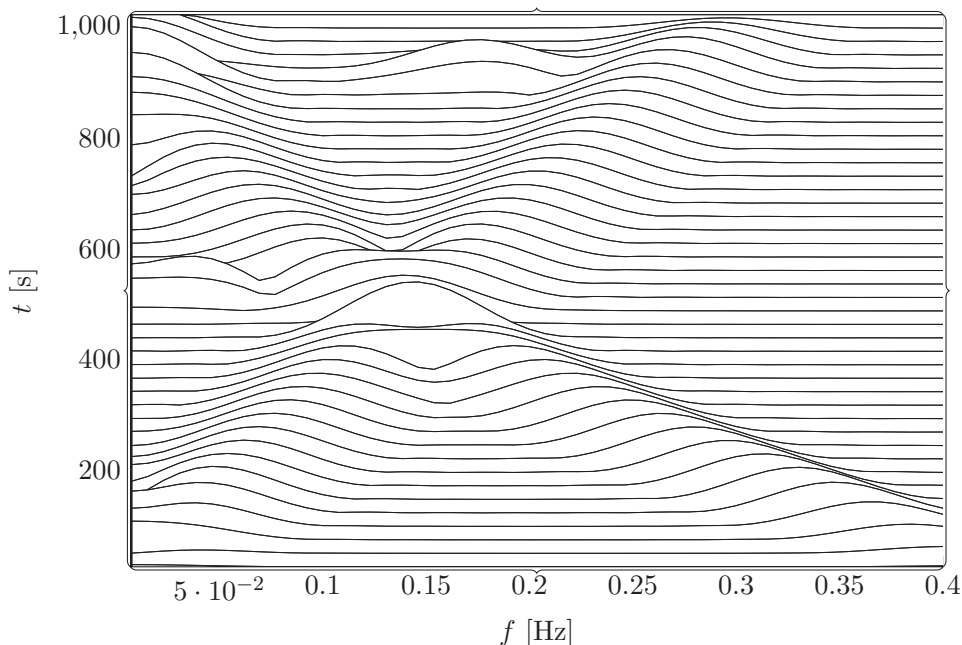
Rys. 3.12. Reprezentacja sygnału o modulowanej częstotliwości i amplitudzie w dziedzinie częstotliwości

Wadą reprezentacji sygnałów w dziedzinie częstotliwości jest brak możliwości obserwacji zmian w czasie składowych widma częstotliwości w analizowanym sygnale. Stąd reprezentacja częstotliwościowa sygnałów jest niewystarczająca dla sygnałów przedstawionych na rys. 3.6 oraz 3.7. Brak możliwości analizy zmian składowych częstotliwości jest szczególnie dotkliwy w analizie sygnałów niestacjonarnych. Rozwiązaniem jest zastosowanie reprezentacji umożliwiającej łączną analizę czasowo-częstotliwościową sygnałów.

3.3.3. Reprezentacja czasowo-częstotliwościowa sygnałów

Zarówno reprezentacja czasowa, jak i częstotliwościowa nie sprawdza się w przypadku sygnałów niestacjonarnych, ponieważ cechą charakterystyczną tych sygnałów jest to, że zawartość składowych w dziedzinie częstotliwości zmienia się np. w czasie. Rozwiązaniem tego problemu jest zastosowanie reprezentacji czasowo-częstotliwościowej. Przykład takiej reprezentacji przedstawiono na rys. 3.13. Reprezentacja sygnału w dziedzinie czasowo-częstotliwościowej możliwa jest z wykorzystaniem m.in. *krótkoczasowej transformacji Fouriera* (będącej modyfikacją transformacji Fouriera), *transformacji Gabora*, transformacji falkowej czy też *transformacji Hilberta-Huanga*.

3. Sygnały, pomiary, szумы



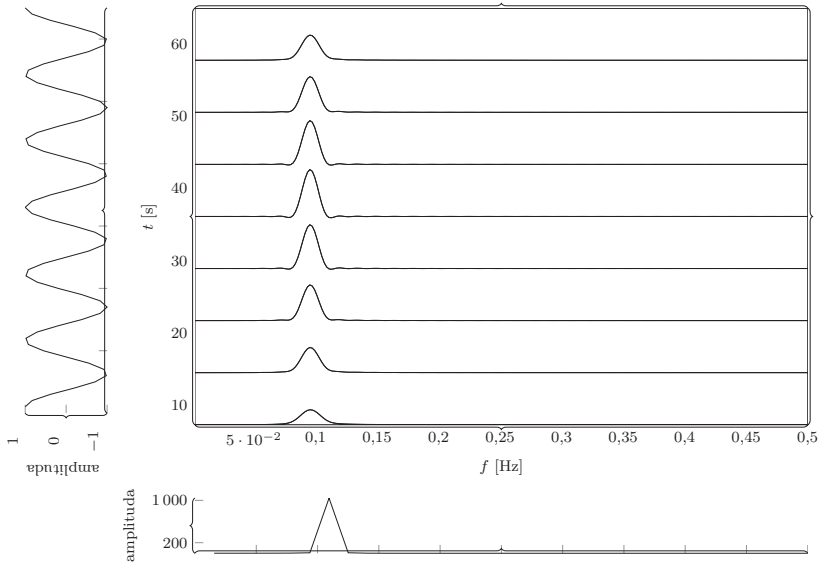
Rys. 3.13. Przykład reprezentacji czasowo-częstotliwościowej sygnału

Cechą charakterystyczną reprezentacji czasowo-częstotliwościowej jest to, że umożliwia opis zachodzących w sygnale zmian zarówno w dziedzinie czasu, jak i częstotliwości. Zaletą tej reprezentacji jest to, że dla jedno- i wielokomponentowych sygnałów niestacjonarnych możliwe jest przedstawienie każdego z tych komponentów osobno. Ponadto możliwe jest wskazanie początku i końca każdego z nich. Dla sygnałów z rys. 3.6 oraz 3.7 efekt ten jest niewidoczny dla reprezentacji częstotliwościowej, co zostało przedstawione na rys. 3.11 oraz 3.12. Wykażemy teraz, że przedstawienie składowych dla sygnałów niestacjonarnych oraz wskazanie początków i końców każdego z komponentów jest możliwe z wykorzystaniem reprezentacji czasowo-częstotliwościowej.

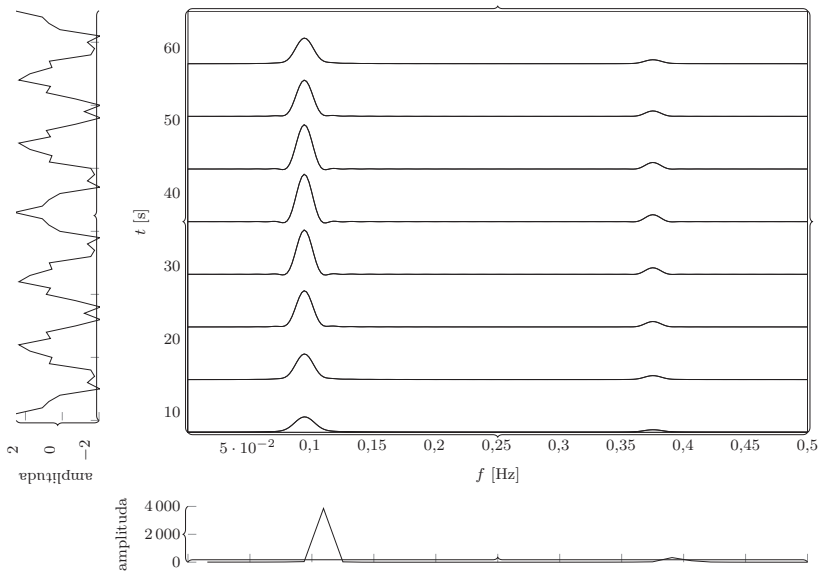
Klasyką metodą służącą do czasowo-częstotliwościowego opisu sygnału jest krótkoczasowa transformata Fouriera (ang. *Short-Time Fourier Transform*, STFT). STFT jest wyznaczana (dla przypadku ciągłego) na podstawie następującej zależności:

$$S_{\text{STFT}}(t, f) = \int_{-\infty}^{\infty} s(\tau)w(\tau - t)e^{-j2\pi f\tau} d\tau \quad (3.3)$$

3.3. Metody reprezentacji sygnałów

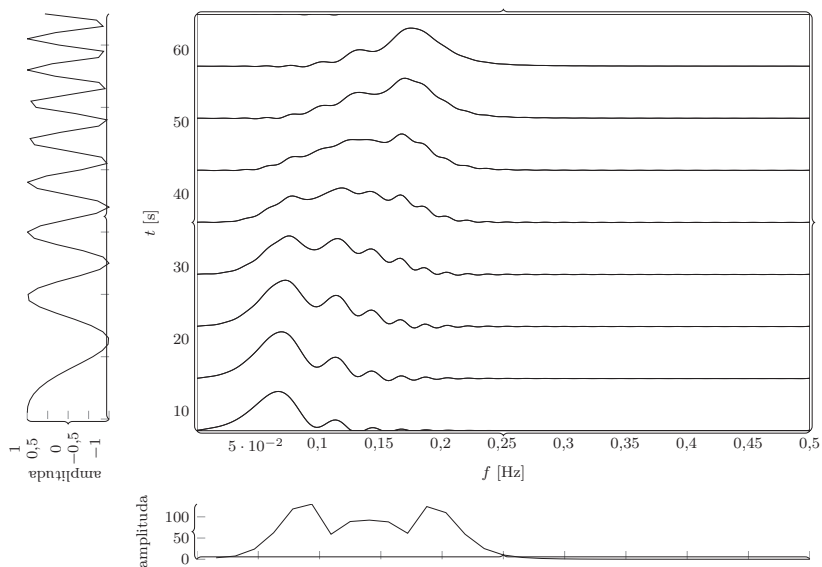


Rys. 3.14. Reprezentacja sygnału sinusoidalnego w dziedzinie czasowo-częstotliwościowej

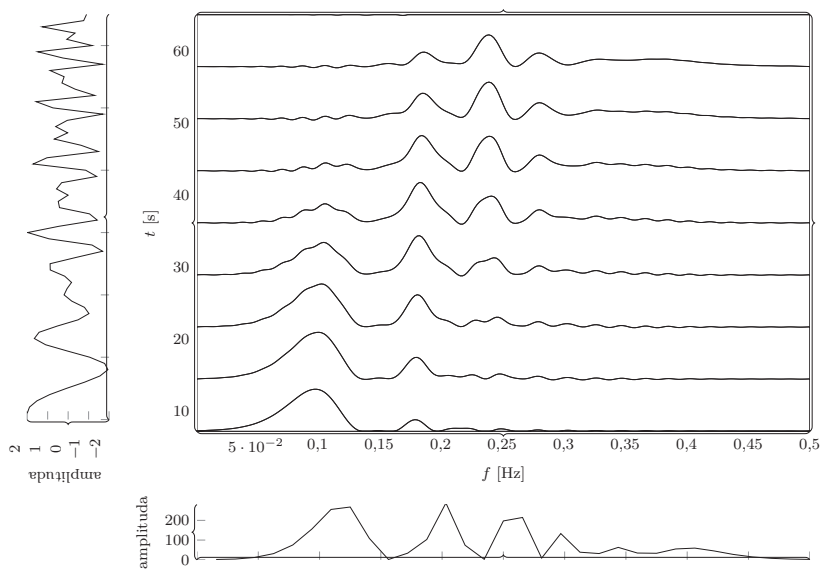


Rys. 3.15. Reprezentacja sygnału będącego złożeniem dwóch sygnałów sinusoidalnych w dziedzinie czasowo-częstotliwościowej

3. Sygnały, pomiary, szумы



Rys. 3.16. Reprezentacja sygnału o modulowanej częstotliwości w dziedzinie czasowo-częstotliwościowej



Rys. 3.17. Reprezentacja sygnału o modulowanej częstotliwości i amplitudzie w dziedzinie czasowo-częstotliwościowej

gdzie τ to przesunięcie na osi czasu, $w(t)$ jest oknem czasowym o ustalonej szerokości. W przypadku STFT popularnymi oknami czasowymi są okna Hanna, Hamminga oraz Blackmana. Na rysunkach 3.14–3.17 przedstawiono reprezentacje czasowo-częstotliwościowe sygnałów bazujących na krótkoczasowej transformacji Fouriera, których reprezentacje czasowe i częstotliwościowe przedstawiono odpowiednio na rys. 3.4–3.7 oraz 3.9–3.12. Prezentacja zawartości poszczególnych składowych częstotliwości sygnału w czasie, która została przedstawiona na rys. 3.14–3.17, gdzie częstotliwość jest na osi odciętych, a czas na osi rzędnych, umożliwia ocenę czy w analizowanym sygnale zachodzą zmiany. Z wykresów wynika (por. rys. 3.14 oraz 3.15), że w przypadku sygnałów sinusoidalnego i sinusoidalnego o dwóch komponentach ich moc widma nie ulega zmianie. Inaczej jest w przypadku sygnałów przedstawionych na rys. 3.6 oraz 3.7. Wykresy w dziedzinie czasowo-częstotliwościowej pokazują zmiany w mocy widma tych sygnałów (por. rys. 3.16, 3.17).

Innym przykładem transformacji, która umożliwia czasowo-częstotliwościową reprezentację sygnałów, jest **transformacja falkowa**. W metodzie tej dokonuje się przekształcenia wartości sygnału z dziedziny czasu do układu współrzędnych czas-skala, co umożliwia analizę zmian sygnału w dziedzinie częstotliwości w funkcji czasu. Efektem transformacji falkowej jest zbiór sygnałów elementarnych, które nazywane są falkami. Transformacja ta wyznaczana jest z następującej zależności:

$$S_{\text{WVT}}(\mathbf{a}, \mathbf{b}) = \frac{1}{\sqrt{\mathbf{a}}} \int_{-\infty}^{\infty} s(t) \Psi\left(\frac{t - \mathbf{b}}{\mathbf{a}}\right) dt \quad (3.4)$$

gdzie \mathbf{a} to parametr skali czasu, \mathbf{b} reprezentuje wartość przesunięcia falki, Ψ oznacza funkcję falkową.

Falka podstawowa (macierzysta) stanowi podstawę do wyznaczenia pozostałych falek przez zmianę parametrów skali czasu i przesunięcia falki. Zbudowany w ten sposób zbiór falek nazywany jest rodziną falek, który tworzy jądro transformacji i stanowi podstawę do dekompozycji wybranego sygnału.

W tej części pracy przedstawiono transformacje STFT oraz falkową. Transformacje Gabora oraz Hilberta–Huanga zostaną omówione w dalszej części tego opracowania.

3.4. Pomiary

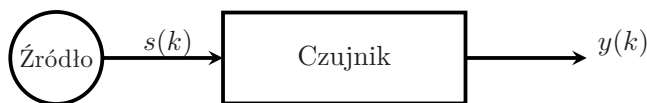
Przejdźmy teraz do omówienia zagadnień związanych z podstawowymi układami pomiarowymi oraz modelami tych układów. Rozdział ten częściowo opiera się na materiale opublikowanym w pracy [223].

3. Sygnały, pomiary, szumy

Rozważmy na początek prosty przypadek, tj. pewne źródło generuje sygnał $s(t)$, który mierzony jest przez czujnik pomiarowy. Wartość zmierzoną oznaczmy przez $y(t)$. W idealnej sytuacji pomiarowej nie istnieją żadne zjawiska mające niekorzystny wpływ na obserwowany obiekt lub na urządzenie pomiarowe. Sytuacja ta została zilustrowana na rys. 3.18. Model matematyczny dla tego przypadku można zapisać w formie:

$$y(k) = h(s(k)) \quad (3.5)$$

gdzie h jest pewną funkcją opisującą zależność pomiędzy wielkością mierzoną a pomiarem, z kolei k jest dyskretną chwilą czasu.



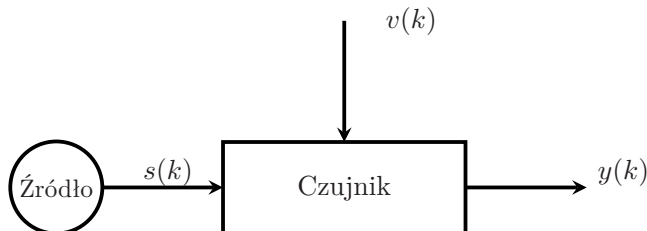
Rys. 3.18. Schemat prostego układu pomiarowego

Przedstawiona na rysunku 3.18 sytuacja jest trudna do osiągnięcia w praktyce. W rzeczywistości pomiary uzyskane z czujnika pomiarowego zawierają szumy. W przypadku konieczności uwzględniania ich, przedstawiony na rys. 3.18 model należy rozszerzyć do postaci takiej jak na rys. 3.19.

Niepożądane składowe w sygnałach pomiarowych dzielimy na szumy wewnętrzne i szumy zewnętrzne (tj. zakłócenia) [155]. Uwzględniając tę uwagę, model z rys. 3.19, po uwzględnieniu zakłóceń pomiarowych, przyjmuje następującą postać:

$$y(k) = h_v(s(k), v(k)) \quad (3.6)$$

gdzie $v(k)$ jest szumem zewnętrznym, a h_v reprezentuje model pomiaru. Szumem zewnętrznym są wszelkie niepożądane składniki występujące w pomiarach i mające swoje źródło poza obiektem lub poza układem pomiarowym.

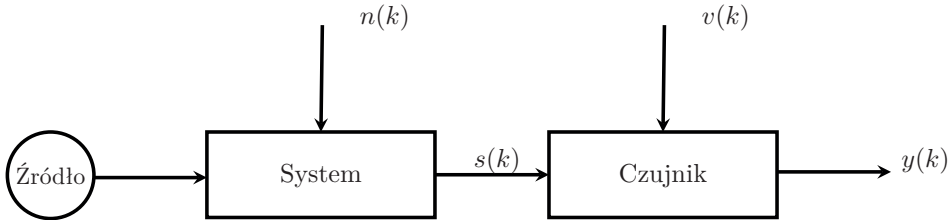


Rys. 3.19. Schemat prostego układu pomiarowego z zakłóceniami pomiarowymi

Schemat układu pomiarowego przedstawiony na rys. 3.19 jest właściwy dla przypadku obiektu statycznego. W przypadku obiektów dynamicznych konieczne jest uwzględnienie szumów wewnętrznych (procesowych). Wówczas model (3.6) uogólniamy do następującej formuły:

$$\begin{aligned} s(k+1) &= f_n(s(k), n(k)) \\ y(k) &= h_v(s(k), v(k)) \end{aligned} \quad (3.7)$$

gdzie $n(k)$ reprezentuje szum wewnętrzny (procesowy), natomiast f_n jest funkcją opisującą model procesu. Ilustrację przypadku opisanego zależnością (3.7) przedstawiono na rys. 3.20.



Rys. 3.20. Schemat prostego układu pomiarowego z zakłóceniami pomiarowymi i procesowymi

W przypadku obiektu niestacjonarnego równanie (3.7) uogólniamy z kolei do następującej postaci:

$$\begin{aligned} s(k+1) &= f_n(s(k), n(k), k) \\ y(k) &= h_v(s(k), v(k), k) \end{aligned} \quad (3.8)$$

gdzie zarówno f_n , jak i h_v są funkcjami zależnymi od czasu.

W przypadku szumu wewnętrznego (procesowego) zazwyczaj przyjmuje się, że jest to biały szum *gaussowski* [199]. W przypadku szumu zewnętrznego (zakłócenia) przyjęcie, zwłaszcza w zastosowaniach praktycznych, że szum ten ma naturę szumu białego *gaussowskiego*, może okazać się uproszczeniem [157]. Efektami przyjęcia takiego założenia, co do typu zakłóceń, mogą być np. niezadowalające wyniki odsumowania sygnałów pomiarowych lub ich modelowania. Jeśli wyniki działania algorytmów przetwarzania sygnałów, w których przyjęto założenie o *gaussowskim* typie szumów pomiarowych, dają niezadowalające wyniki, to wówczas należy rozważyć zastosowanie *niegaussowskich* modeli szumów w projektowanych algorytmach.

3. Sygnały, pomiary, szumy

Oprócz funkcji opisującej w zaproponowanym modelu wpływ szumów, wprowadzona została również zmienna opisująca typ szumów pomiarowych. Szumy te są jednym z głównych elementów wpływających na ograniczone możliwości układów pomiarowych. Dlatego modelowanie ich jest jednym z najważniejszych zagadnień zarówno praktycznych, jak i teoretycznych. W następnym rozdziale zostaną omówione niektóre z typów szumów, ich charakterystyki oraz metody ich modelowania.

3.5. Szumy

Szumy wewnętrzne (procesowe) związane są bezpośrednio z obserwowanym obiektem. Powstają one w obiektach, a ich wpływ na obiekt jest trwały. Oznacza to, że oddziałują zarówno na aktualny, jak i na przyszły stan obiektu (por. zależn. (3.7)). Trwały wpływ tego typu szumów na obiekt widoczny jest w jakości pomiarów [307].

Niedokładności w pomiarach mogą być skutkiem działania czynników zewnętrznych, takich jak: pole magnetyczne, temperatura, wilgoć lub być związane z przeprowadzaniem pomiaru, np. urządzeniem pomiarowym lub kanałem transmisji danych. Niedokładności w tym przypadku związane są z szumami, które dla odróżnienia nazywamy szumami zewnętrznymi lub zakłóceniami. W dalszej części pracy będziemy używać obu tych określeń zamiennie.

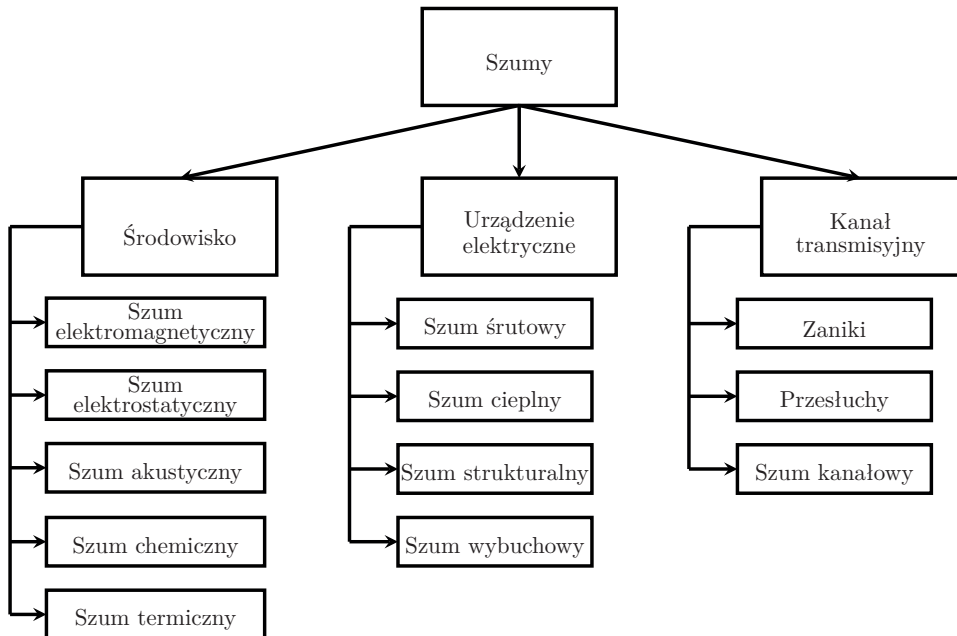
Niepożądane składniki w sygnałach mają negatywny wpływ na jakość ich przetwarzania. Stąd, w celu poprawy efektów przetwarzania sygnałów pomiarowych, istotne stają się zagadnienia związane z modelowaniem oraz budową charakterystyk szumów i zakłóceń. Ich modelowanie umożliwia z kolei konstrukcję uproszczonych opisów niepożądanych składowych w sygnałach pomiarowych. W efekcie możliwe jest przygotowanie opisu szumów i zakłóceń z wykorzystaniem kilku lub kilkunastu parametrów.

Z kolei opracowane modele, wraz z ustalonymi wartościami parametrów tych modeli, umożliwiają przygotowanie charakterystyk niepożądanych składowych w sygnałach pomiarowych. Wówczas ich znajomość ułatwia projektowanie metod usuwania niepożądanych składowych z pomiarów. Przypadkiem idealnym, choć trudnym do spełnienia w rzeczywistości, jest całkowite usunięcie szumów z pomiarów.

Źródeł powstawania szumów, zarówno tych procesowych, jak i zewnętrznych, jest wiele i są one różnorodne. W pracy tej skupimy się jedynie na fizycznych przyczynach ich powstawania. Na rysunku 3.21 zaproponowano klasyfikację głównych źródeł szumów i szumów pomiarowych, które mają swoje źródło w zjawiskach fizycznych. Zgodnie z zaproponowanym podziałem, niepożądane składniki pomiarów

mogą być związane z wpływem **środowiska** (otoczenia), **urządzeń elektronicznych** (np. elektroniczne urządzenia pomiarowe) oraz **kanałem transmisyjnym**. Do najważniejszych zakłóceń środowiskowych zaliczamy zakłócenia elektromagnetyczne, które pojawiają się głównie w zakresie częstotliwości radiowych (od kilku kHz do kilku GHz) oraz zakłócenia elektrostatyczne pojawiające się w efekcie obecności ładunków elektrycznych. Do innych zakłóceń zaliczamy zakłócenia akustyczne, które związane są z ruchem oraz wibrowaniem urządzeń i/lub obiektów pomiarowych oraz zjawiskiem odbicia fali dźwiękowej od przeszkód. Niepożądane składniki sygnałów pomiarowych mogą być wywołane również czynnikami chemicznymi oraz wpływem temperatury otoczenia.

Część z wymienionych zjawisk środowiskowych w otoczeniu obiektu może być wyeliminowana lub w znaczący sposób stłumiona, gdy pomiary dokonywane są w kontrolowanych warunkach laboratoryjnych. Niestety w przypadku gdy pomiary dokonywane są w warunkach innych niż laboratoryjne, zmniejszanie wpływu otoczenia jest ograniczone. W tym przypadku rozwiązaniem może okazać się modelowanie wpływu środowiska na sygnały pomiarowe. Przykładem zastosowania tego podejścia jest modelowanie wpływu natężenia pola magnetycznego na działanie magnetometru [375].



Rys. 3.21. Klasyfikacja zakłóceń

3. Sygnały, pomiary, szумы

Kolejna kategoria zakłóceń związana jest ze zjawiskami zachodzącymi w urządzeniach elektronicznych czy elektrycznych podczas przepływu prądu. W kontekście sieci czujnikowych typu BSN opartych na układach elektronicznych jest to jedno z istotniejszych zagadnień, ponieważ w czujnikach pomiarowych wykonanych w technologii MEMS szумы pomiarowe są głównie wywołane przez zjawiska powiązane z przepływem prądu i towarzyszące temu przepływowi przypadkowe fluktuacje ładunków elektrycznych. W urządzeniach elektronicznych mogą pojawić się m.in. szумы śrutowe, które wywołane są przepływem prądu w układach elektronicznych i powiązaniem z nim procesem generowania ładunków elektrycznych. Z kolei szумы ciepłne, związane są z chaotycznymi ruchami wolnych elektronów w rezystancjach układów elektronicznych. Efektem tych chaotycznych ruchów wolnych elektronów jest ciepło, którego wydzielanie skutkuje powstaniem tego szumu [231, 291, 445]. Należy zaznaczyć, że szum ten występuje we wszystkich materiałach i elementach stratnych. Kolejny typ szumu, tj. szum strukturalny, ma związek z niedokładnościami procesu technologicznego wytwarzania urządzeń elektronicznych, również tych wykonanych w technologii MEMS. Przykładem szumu strukturalnego jest szum lawinowy (ang. *avalanche noise*), który występuje w układach półprzewodnikowych i związany jest z przejściem nośników przez złącze $p-n$ [155]. Innym przykładem szumu strukturalnego jest szum wybuchowy (ang. *burst noise*). Również i ten typ szumu powodowany jest przez niedokładności procesu technologicznego wytwarzania urządzeń elektronicznych. Szum ten występuje w zakresie częstotliwości poniżej 100 Hz. Cechą charakterystyczną szumu wybuchowego jest to, że chwila pojawienia się tego szumu oraz czas jego trwania są przypadkowe.

Wśród szumów generowanych w kanale transmisyjnym podczas przesyłania danych wyróżniamy m.in. zaniki w kanale transmisyjnym. Zaniki powodowane są przez nakładanie się na siebie sygnałów przesyłanych w tym samym kanale transmisyjnym i są przyczyną tłumienia sygnału [313]. Kolejnym przykładem niekorzystnych efektów występujących w kanale transmisyjnym są przesłuchy, czyli efekt nakładania się sygnału przesyłanego w jednym kanale na sygnał przesyłany w innym kanale. Przesłuchy mogą pojawić się na przykład podczas wykonywania pomiarów z wykorzystaniem czujnika podłączonego do sieci energetycznej. Takie pomiary mogą być obciążone niechcianymi składowymi, których częstotliwość jest zbliżona do częstotliwości prądu w sieci energetycznej. Z kolei szum kanałowy związany jest z szumami pojawiającymi się w kanale transmisyjnym.

Szумы można również klasyfikować ze względu na ich charakterystykę częstotliwościową. W tym przypadku szумы dzielimy na:

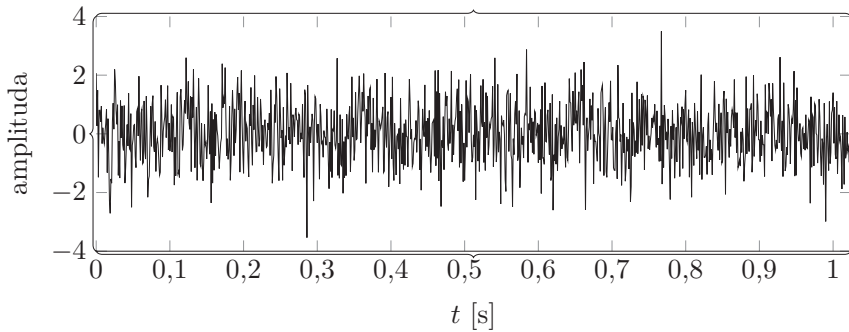
- biały szum,
- pasmowo ograniczony biały szum,
- szумы wąskopasmowe,

- szum kolorowy,
- szum impulsowy.

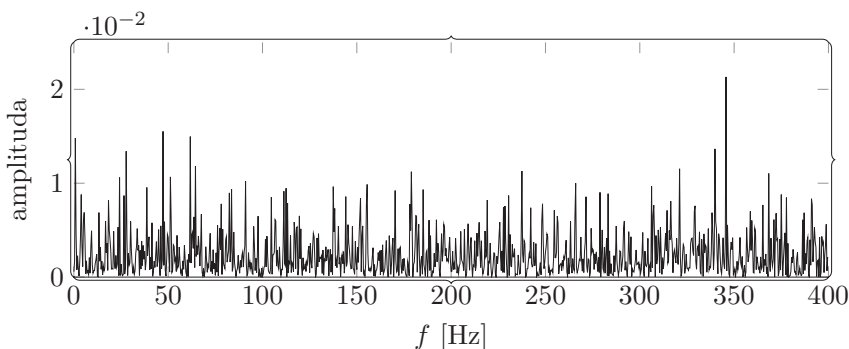
W dalszej części podrozdziału omówiono wybrane klasy wymienionych szumów.

3.5.1. Szum biały

Szum biały jest szumem charakteryzującym się płaskim widmem. Oznacza to, że jego intensywność jest równomierna w całym paśmie częstotliwości, tj. wszystkie częstotliwości mają jednakową moc. Warto również zaznaczyć, że szum biały cechuje się brakiem pamięci. To znaczy, że obecny stan obiektu generującego szum biały nie zależy od jego przeszłych stanów. Próbkę szumu białego przedstawiono na rys. 3.22, a na rys. 3.23 jego reprezentację częstotliwościową w postaci widma.



Rys. 3.22. Reprezentacja czasowa szumu białego



Rys. 3.23. Reprezentacja częstotliwościowa szumu białego

3. Sygnały, pomiary, szumy

Szczególnym przypadkiem szumu białego jest szum **gaussowski**, czyli funkcja rozkładu prawdopodobieństwa wartości w dziedzinie czasu jest funkcją *gaussowską*. Jeśli jednocześnie szum ten dodaje się do sygnału, to wówczas taki szum o rozkładzie opisanym funkcją *gaussowską* nazywamy **addytywnym białym szumem gaussowskim** (ang. *Additive White Gaussian Noise*, AWGN).

Należy w tym miejscu zaznaczyć, że jeśli funkcja rozkładu prawdopodobieństwa jest inna niż *gaussowska*, np. jest to rozkład równomierny, Poissona lub Cauchiego, to taki szum nazywamy szumem **białym niegaussowskim**.

3.5.2. Szum kolorowy

W praktycznych problemach przetwarzania sygnałów założenie o tym, że niepożądane składowe w pomiarach mają charakterystykę szumu białego, może okazać się niewystarczające. Jeśli uzyskane wyniki obróbki pomiarów nie dają zadowalających wyników, jednym z rozwiązań, które należy rozważyć, jest przyjęcie założenia o skorelowaniu szumów pomiarowych. Przyjęcie tego założenia wiąże się z opracowaniem lub wykorzystaniem algorytmów przetwarzania sygnałów, w których uwzględnia się niepożądane składowe o takich cechach.

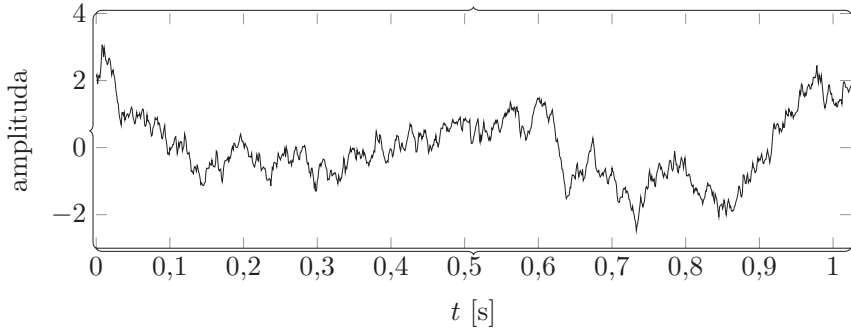
Pojęcie szumu skorelowanego (inaczej szumu kolorowego) związane jest z szumem, którego charakterystyka widmowa mocy znacząco różni się od charakterystyki widmowej szumu białego. W przypadku szumów kolorowych rozkład widmowy mocy nie jest równomierny w całym paśmie. Wyróżnia się kilka typów szumów kolorowych, a ich nazwa związana jest z tym, w jakim zakresie częstotliwości skupiona jest energia danego szumu. Wśród szumów kolorowych wyróżniamy szum:

- czerwony (Brown),
- różowy,
- niebieski,
- fioletowy.

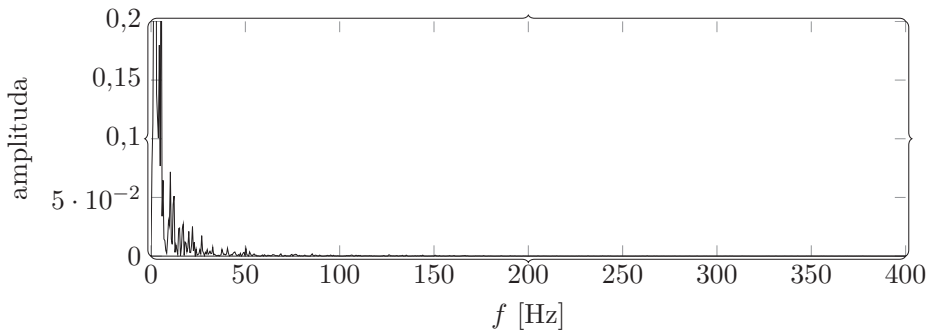
Szum czerwony (Brown)

Szum czerwony (lub szum Browna) jest szumem powiązany z chaotycznymi ruchami Browna. Szumy Browna w rzeczywistych układach wywoływane są przez fluktuacje termodynamiczne. W przypadku tego szumu jego energia skupiona jest w zakresie niskich częstotliwości. Cechą charakterystyczną szumu czerwonego jest to, że prędkość spadku jego widma jest odwrotnie proporcjonalna do kwadratu częstotliwości (tj. gęstość widmowa mocy jest proporcjonalna do $1/f^2$).

Uzyskanie szumu czerwonego jest prostą operacją, która sprowadza się do całkowania szumu białego. Przykład szumu czerwonego przedstawiono na rys. 3.24, a na rys. 3.25 jego reprezentację częstotliwościową.



Rys. 3.24. Reprezentacja czasowa szumu czerwonego



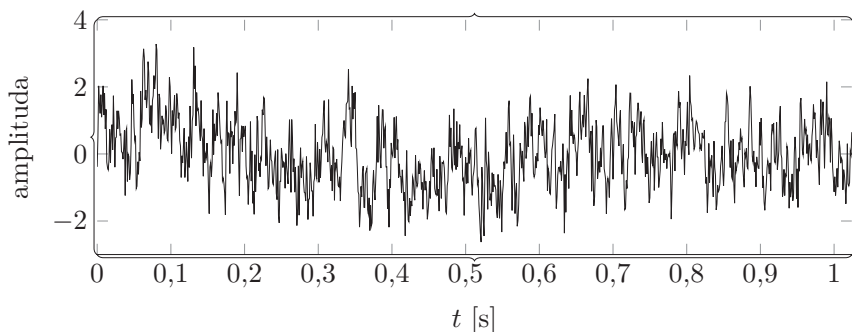
Rys. 3.25. Reprezentacja częstotliwościowa szumu czerwonego

Szum różowy

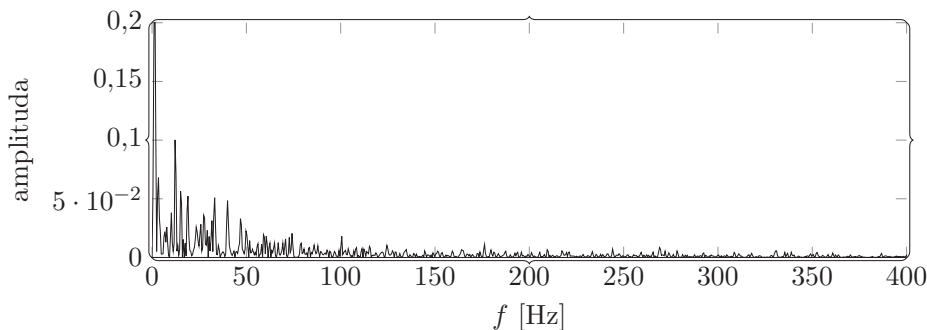
Szum różowy charakteryzuje się widmową gęstością mocy odwrotnie proporcjonalną do częstotliwości, tj. $1/f$. Zależność ta w pełnym zakresie częstotliwości spełniona jest jedynie teoretycznie. W praktyce zależność $1/f$ jest możliwa do spełnienia jedynie w ograniczonym zakresie częstotliwości. Wynika to z tego, że jeśli górna granica częstotliwości byłaby nieskończenie wielka, to wówczas również energia takiego szumu musiałaby być nieskończenie wielka.

3. Sygnały, pomiary, szумы

Podobnie jak to było dla szumu Browna, energia widmowa szumu $1/f$ skupiona jest w zakresie niskich częstotliwości. Widmowa gęstość mocy tego szumu opada 3 dB na oktawę. Próbkę szumu różowego przedstawiono na rys. 3.26. Z kolei na rys. 3.27 jego reprezentację częstotliwościową.



Rys. 3.26. Reprezentacja czasowa szumu różowego



Rys. 3.27. Reprezentacja częstotliwościowa szumu różowego

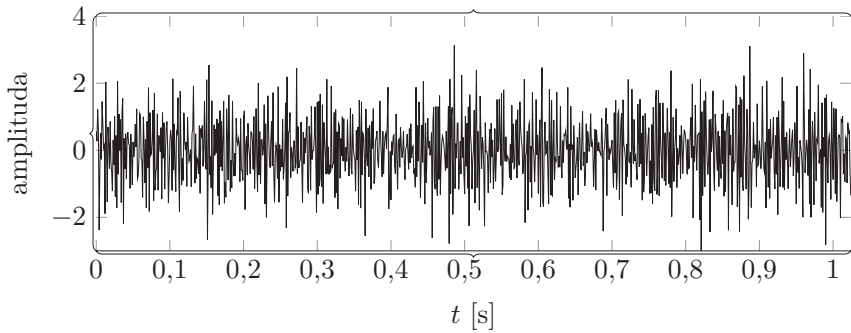
Szum typu $1/f$, pomimo jego powszechności, nie został do tej pory w pełni opisany. Nie ma żadnych metod generowania szumu mającego takie właściwości. W przypadku szumu różowego nie jest możliwe jego otrzymanie przez operację całkowania szumu białego (por. z szumem Browna). Pewne przybliżenie szumu $1/f$ można uzyskać, wykorzystując formalne uogólnienie ruchów Browna, tj. bazując na ułamkowym ruchu Browna (ang. *fractional Brownian motion*, fBm) oraz generowanym podczas tego ruchu, tj. ułamkowym szumie gaussowskim (ang.

fractional Gaussian noise, fGn) [486]. Szum różowy jest przykładem szumu typu *niegaussowskiego* [34, 113, 200, 418].

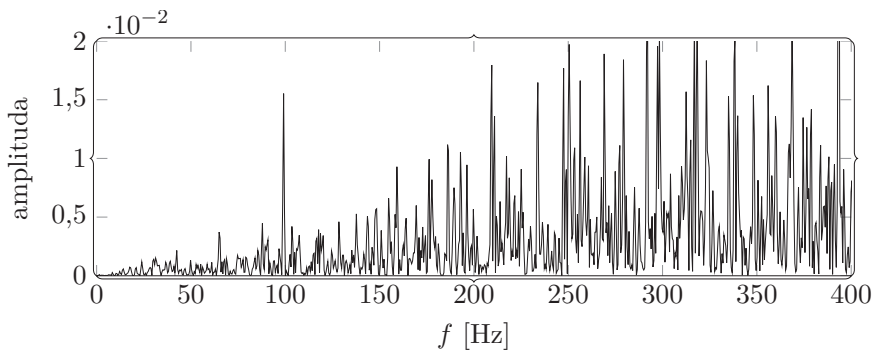
Szum niebieski

Kolejnym rodzajem szumu, który rozważymy, jest szum niebieski. W przypadku tego szumu jego energia skupiona jest w zakresie wysokich częstotliwości. Widmowa gęstość mocy jest proporcjonalna do jego częstotliwości f i rośnie 3 dB na oktawę. Oznacza to, że jego widmowa gęstość mocy rośnie wraz ze wzrostem jego częstotliwości.

Szum niebieski można traktować jako odwrotność szumu różowego, czyli również jest proporcjonalny do częstotliwości szumu z tą różnicą, że WGM rośnie, a nie maleje jak to jest w przypadku szumu $1/f$. Próbkę szumu niebieskiego przedstawiono na rys. 3.28, a jego reprezentację częstotliwościową na rys. 3.29.



Rys. 3.28. Reprezentacja czasowa szumu niebieskiego

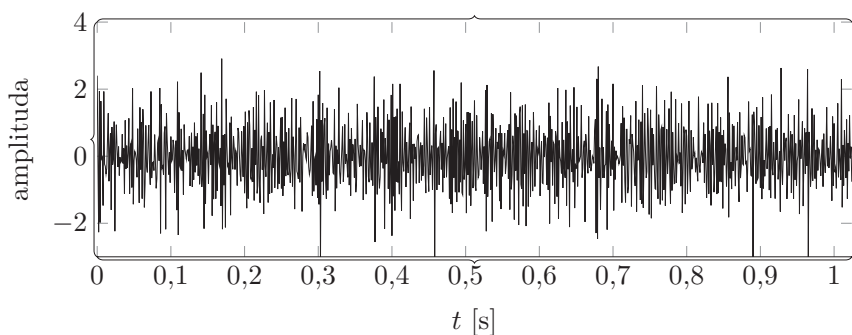


Rys. 3.29. Reprezentacja częstotliwościowa szumu niebieskiego

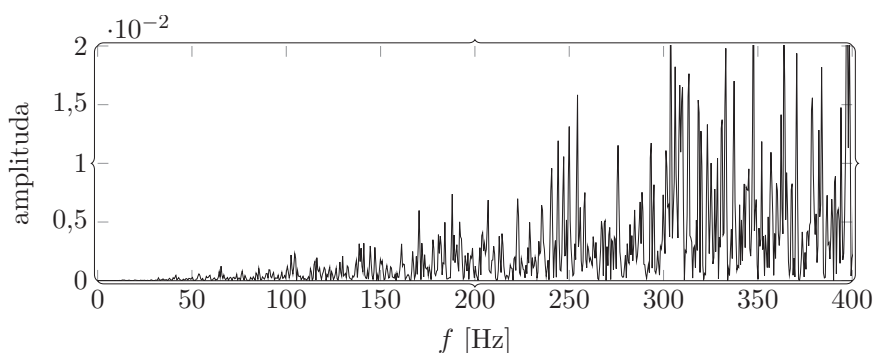
Szum fioletowy

Kolejnym typem szumu, którego energia skupiona jest w zakresie wysokich częstotliwości, jest szum fioletowy, tylko że energia ta skupiona jest jeszcze bardziej w zakresie wysokich częstotliwości niż na przykład w przypadku szumu niebieskiego. Szum fioletowy może być traktowany jako odwrotność szumu czerwonego, czyli jego energia jest proporcjonalna do kwadratu jego częstotliwości, tj. f^2 , z tym że w tym przypadku energia ta rośnie, a nie maleje jak w przypadku szumu czerwonego, 6 dB na oktawę.

Szum fioletowy może zostać wygenerowany przez zróżniczkowanie szumu białego. Przykładowy przebieg tego szumu przedstawiono na rys. 3.30. Z kolei na rys. 3.31 zilustrowano jego reprezentację częstotliwościową.



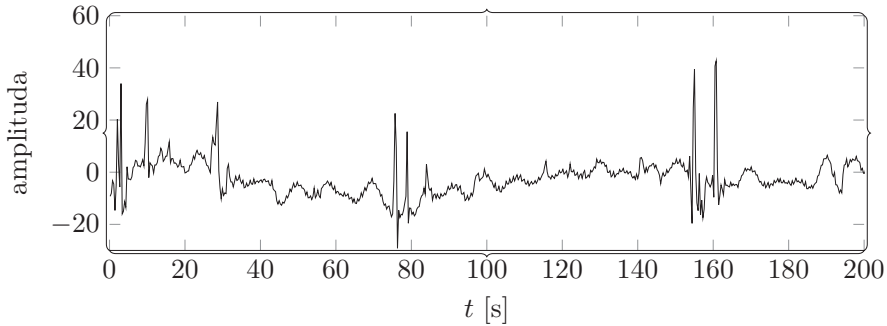
Rys. 3.30. Reprezentacja czasowa szumu fioletowego



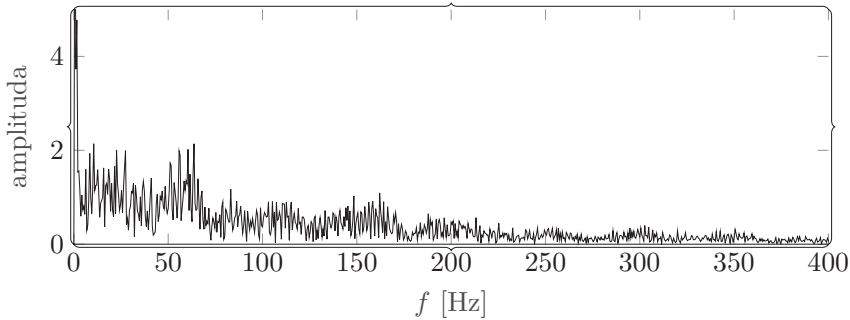
Rys. 3.31. Reprezentacja częstotliwościowa szumu fioletowego

3.5.3. Szum impulsowy

Szum impulsowy charakteryzuje się tym, że w sygnale pomiarowym pojawiają się krótkotrwałe impulsy o dużej amplitudzie. Częstotliwość ich występowania jest przypadkowa, podobnie jak wartość amplitudy. Szum impulsowy jest kolejnym przykładem szumu *niegaussowskiego*. Reprezentację tego szumu w dziedzinie czasu przedstawiono na rys. 3.32, natomiast na rys. 3.33 przedstawiono jego reprezentację częstotliwościową.



Rys. 3.32. Reprezentacja czasowa szumu impulsowego



Rys. 3.33. Reprezentacja częstotliwościowa szumu impulsowego

3.6. Miary jakości sygnałów pomiarowych

Istotną kwestią w przetwarzaniu sygnałów pomiarowych jest ich jakość. W literaturze zaproponowano wiele miar oceny jakości sygnałów. Najczęściej stosowaną

3. Sygnały, pomiary, szumy

miarą jest współczynnik SNR (ang. *Signal-To-Noise Ratio*), który jest stosunkiem mocy sygnału do mocy szumów. Wielkość SNR definiujemy w następujący sposób:

$$SNR = 10 \log_{10} \left(\frac{\sum_{k=1}^K (y(k))^2}{\sum_{k=1}^K (y(k) - s(k))^2} \right) \quad (3.9)$$

gdzie $y(k)$ jest sygnałem pomiarowym, a $s(k)$ to sygnał użyteczny. Dla współczynnika SNR jednostką jest decybel.

Inną miarą jakości sygnału pomiarowego jest wariant współczynnika SNR , tj. $PSNR$. Współczynnik $PSNR$ (ang. *Peak Signal-To-Noise Ratio*) wyrażany jest zależnością:

$$PSNR = 10 \log_{10} \left(\frac{(\max(y(k)))^2}{\sum_{k=1}^K (y(k) - s(k))^2} \right), \quad k = 1, 2, \dots, K \quad (3.10)$$

gdzie \max jest funkcją wyznaczającą maksymalną wartość w sygnale $y(k)$. Również i dla $PSNR$ jednostką jest decybel.

Wymieńmy jeszcze dwie inne miary wykorzystywane do oceny jakości sygnału pomiarowego. Jedną z nich jest współczynnik $RMSE$ (ang. *Root Mean Square Error*):

$$RMSE = \sqrt{\frac{1}{K} \sum_{k=1}^K (y(k) - s(k))^2} \quad (3.11)$$

oraz MAE (ang. *Mean Absolute Error*):

$$MAE = \frac{1}{K} \sum_{k=1}^K |y(k) - s(k)| \quad (3.12)$$

natomiast K w zależnościach (3.9)–(3.12) reprezentuje liczbę wszystkich dostępnych pomiarów.

4. Czujniki pomiarowe

*Informacje z zewnątrz są
przekształcane za pośrednictwem
wewnętrznych urządzeń bez względu
na to, czy chodzi o ustrój
organiczny, czy nieorganiczny.
Informacje te są przekształcane
w formy nadające się do
wykorzystania w dalszych stadiach
zachowania organizmu lub maszyny.*

Norbert Wiener

4.1. Wstęp

W ostatnich latach obserwuje się postęp w dziedzinie konstrukcji miniaturowych urządzeń takich, jak czujniki pomiarowe czy urządzenia wykonawcze. Elementy te stały się nie tylko typowymi składnikami profesjonalnych systemów automatyki przemysłowej czy urządzeń medycznych, ale zyskały również popularność w sprzęcie powszechnego użytku. Typowym przykładem z ostatnich lat jest tzw. inteligentny telefon (ang. *smartphone*). W wielu nowoczesnych telefonach można znaleźć takie elementy jak miniaturowe czujniki przyspieszenia lub prędkości kątowej. Popularność tego typu rozwiązań związana jest z niską ceną, ich niewielkim rozmiarem oraz niewielką wagą przy jednocześnie dużych możliwościach obliczeniowych i pomiarowych.

Wyraźny postęp w dziedzinie pomiarów i urządzeń pomiarowych związany jest z upowszechnieniem technologii, które umożliwiają produkcję czujników o takich cechach jak te, które wymieniono. Wytwarzane urządzenia o opisanych cechach bazują na technologii systemów mikroelektromechanicznych (MEMS). Technologia MEMS umożliwia produkcję m.in. urządzeń, które łączą systemy mikroelektroniczne z systemami mikromechanicznymi. Atutem urządzeń pomiarowych wykonanych w tej technologii jest to, że umożliwiają one wstępne przetworzenie sygnałów pomiarowych przez działające w tych urządzeniach algorytmy. Efektem połączenia

4. Czujniki pomiarowe

nowoczesnych technologii MEMS z algorytmami przetwarzania sygnałów są inteligentne czujniki pomiarowe.

Dodajmy, że alternatywą dla technologii MEMS jest technologia MOEMS (ang. *Micro-Opto-Electro-Mechanical Systems*). Technologia MOEMS wyróżnia się między innymi tym, że jako sygnały sterujące wykorzystuje się sygnały optyczne. Urządzenia wykonane w tej technologii charakteryzują się również tym, że zużywają mniej energii niż elementy wykonane w technologii MEMS, są również od nich tańsze i mniej waży. Zastosowania urządzeń wykonanych w tej technologii są zbliżone do tych, do których wykorzystuje się urządzenia bazujące na technologii MEMS. Wraz z rozwojem technologii MOEMS obszary jej zastosowania mogą się poszerzyć ze względu na unikatowe, w porównaniu do technologii MEMS, właściwości. Cechy technologii MOEMS ułatwiają budowę bardziej złożonych systemów, których zapotrzebowanie na energię jest znacznie mniejsze.

Czujniki są urządzeniami służącymi do konwersji sygnałów z otoczenia lub sił działających na czujnik na sygnał elektryczny. W przypadku układów wykonanych w technologii MOEMS sygnałem po konwersji będzie sygnał optyczny. Urządzenia pomiarowe wykonane w technologiach MEMS lub MOEMS można klasyfikować według różnych kryteriów. Jednym z nich jest podział ze względu na typ zjawiska fizycznego wykorzystywanego do przekształcania wielkości mierzonych na wielkości elektryczne lub optyczne. W tym przypadku urządzenia te dzielimy na:

- optyczne,
- akustyczne,
- elektrostatyczne,
- magnetyczne,
- indukcyjne
- i inne.

Wśród czujników pomiarowych można wyróżnić dwie klasy czujników, tj. czujniki parametryczne i czujniki generacyjne [96]. Czujniki parametryczne charakteryzują się tym, że pod wpływem wielkości mierzonej zmianie ulega parametr czujnika, np. rezystancja, indukcyjność lub pojemność. Sposób działania tych czujników wymaga doprowadzenia do nich dodatkowej energii. Do tej grupy zaliczamy czujniki [96]:

- tensometryczne,
- pojemnościowe,
- indukcyjnościowe,
- magnetorezystancyjne,
- fotorezystancyjne.

Z kolei czujniki generacyjne to czujniki, które pod wpływem wielkości mierzonej wytwarzają sygnał elektryczny. Do tych czujników zaliczamy, m.in.:

- termopary,
- czujniki piezoelektryczne,
- fotoogniwa.

Drugą grupą urządzeń wykonanych w technologiach MEMS i MOEMS stanowią urządzenia wykonawcze (tzw. aktuatory). Aktuatory przetwarzają sygnały sterujące (elektryczne lub optyczne) na wielkości mechaniczne. Wielkościami fizycznymi w większości przypadków są siły lub celowe ruchy. Do prawidłowego działania urządzeń wykonawczych wymagane jest dostarczenie do nich dodatkowego zasilania. Aktuatory można podzielić ze względu na metodę przetwarzania sygnałów sterujących na wielkości fizyczne:

- mechaniczne (ang. *mechanical actuator*),
- magnetyczne (ang. *magnetic actuator*),
- termiczne (ang. *thermal actuator*),
- elektrostatyczne (ang. *electrostatic actuators*).

Elementy wykonawcze można podzielić również ze względu na typ wytwarzanego celowego ruchu:

- translacyjne,
- rotacyjne.

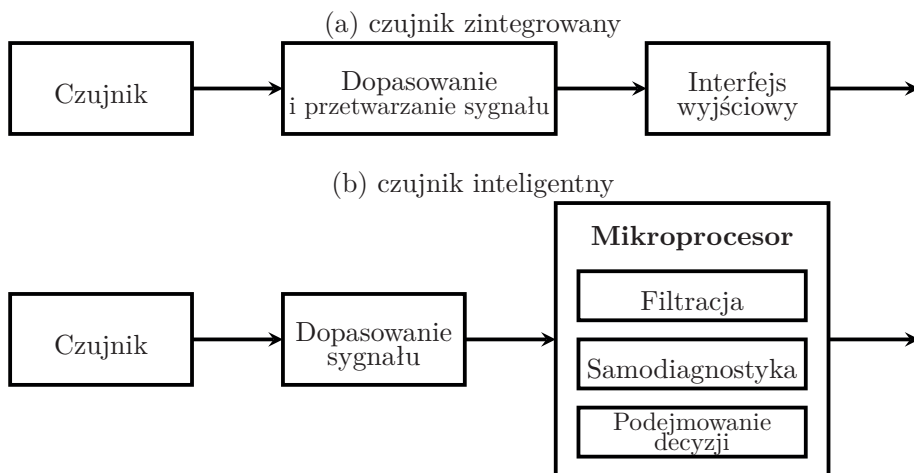
Podstawowe elementy wykonawcze, które dostępne są na rynku to mikropompy, mikrośiłowniki, mikrochwytaki oraz mikrozawory [181, 372, 440].

Zarówno czujniki pomiarowe, jak i urządzenia wykonawcze wykonane w technologiach MEMS lub MOEMS, można podzielić ze względu na stopień ich integracji [96] na:

- czujniki proste,
- czujniki zintegrowane,
- czujniki inteligentne.

Czujniki proste to czujniki, które dokonują jedynie konwersji wielkości mierzonej na wielkość elektryczną lub optyczną. Czujniki zintegrowane mają wbudowane dodatkowe elementy, takie jak wzmacniacze, filtry czy przetworniki analogowo-cyfrowe. Zgodnie z [173] czujniki inteligentne to urządzenia dostarczające funkcjonalności, które wykraczają poza funkcjonalności dostarczane przez czujniki zintegrowane. Cechy tych czujników upraszczają ich integrację z innymi urządzeniami pracującymi w środowisku sieciowym. Urządzenia te charakteryzują się również autonomicznością i zdolnością do adaptacji. Czujniki inteligentne charakteryzują się większym stopniem integracji w porównaniu do czujników zintegrowanych. Przekłada się to na większą liczbę dodatkowych elementów przetwarzających dane pomiarowe. Do tych elementów zaliczamy mikroprocesory oraz specjalizowane układy scalone. Na rysunku 4.1 przedstawiono schematy czujników: zintegrowanego (a) i inteligentnego (b).

4. Czujniki pomiarowe



Rys. 4.1. Schematy czujników pomiarowych: a) zintegrowanego i b) inteligentnego

Zakres zastosowań urządzeń wykonanych w technologiach MEMS i MOEMS jest bardzo szeroki. Wymieńmy niektóre z nich:

- przemysł samochodowy,
- telekomunikacja,
- medycyna i sport,
- przemysł militarny,
- rozrywka.

W różnych obszarach producenci oferują wyspecjalizowane produkty wykonane w omawianych technologiach. W obszarze medycyny i sportu wśród najważniejszych producentów wyróżniamy:

- Shimmer,
- Xsense,
- Zephyr Technology.

Do najpopularniejszych czujników dostępnych na rynku, które oferowane są przez podanych producentów, zaliczamy:

- czujniki przyśpieszenia (akcelerometry),
- czujniki żyroskopowe (żyroskopy),
- czujniki pola magnetycznego (magnetometry).

W dalszej części rozdziału omówione zostaną zagadnienia związane z modelowaniem czujników przyśpieszenia, żyroskopowych oraz natężenia pola magnetycznego. Przedstawione zostaną również podstawowe modele pomiarowe dla tych czujników oraz ich charakterystyki techniczne. Charakterystyki techniczne urządzeń zostaną omówione na przykładzie czujników produkowanych przez firmę Shimmer.

4.2. Czujnik przyśpieszenia

Czujnik przyśpieszenia (akcelerometr) jest urządzeniem umożliwiającym pomiar przyśpieszenia na skutek ruchu. Należy zaznaczyć, że w pomiarach tych czujników jest również obecna, oprócz składowej przyśpieszenia związanej z ruchem, składowa związana z przyśpieszeniem grawitacyjnym. Ponieważ obecność w pomiarach drugiej z tych składowych może w niektórych zadaniach sprawiać problemy obliczeniowe, budując model pomiarowy dla czujnika przyśpieszenia, należy wziąć ten fakt pod uwagę:

$$\mathbf{a} = \mathbf{a}_g + \mathbf{a}_F \quad (4.1)$$

gdzie \mathbf{a}_g reprezentuje składową związaną z przyśpieszeniem ziemskim, a \mathbf{a}_F jest składową związaną z siłą wymuszającą ruch obiektu.

Dokładny pomiar przyśpieszenia \mathbf{a} w praktyce nie jest możliwy ze względu na wpływające na pomiar szumy. Oznaczmy więc pomiar przyśpieszenia następująco:

$$\mathbf{y}_a = \mathbf{a} + \mathbf{v} \quad (4.2)$$

gdzie \mathbf{a} jest prawdziwym, ale nieznanym wektorem przyśpieszenia, z kolei \mathbf{v} to szum w sygnale pomiarowym.

Pomiar przyśpieszenia wykonywany jest zwykle w trzech osiach oznaczanych typowo jako x , y , z . Czujniki takie nazywamy czujnikami trójosiowymi. Poszczególne składowe wektora pomiaru przyśpieszenia oznaczamy następująco:

$$\mathbf{y}_a = [y_{a_x} \quad y_{a_y} \quad y_{a_z}]^T \quad (4.3)$$

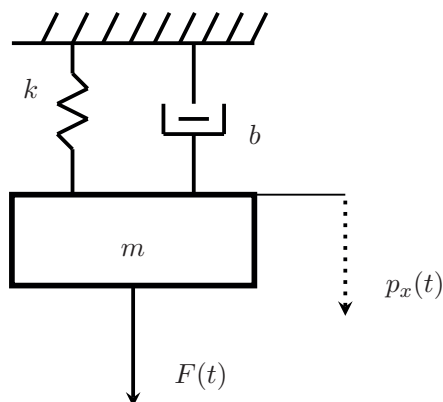
Na rysunku 4.2 przedstawiono uproszczony model czujnika przyśpieszenia. Głównym elementem układu jest masa przymocowana do sprężyny (charakteryzowanej przez współczynnik sprężystości k) oraz elementu tłumiącego (współczynnik tarcia b). W wyniku przemieszczenia się masy m , element sprężysty ulega rozciągnięciu lub ściśnięciu. Zmiana ta powoduje zmianę napięcia wyjściowego w układzie elektrycznym czujnika. Napięcie to jest następnie przetwarzane w czujniku w celu wyznaczenia wartości przyśpieszenia.

Przedstawiony na rys. 4.2 model można zapisać w postaci następującego równania różniczkowego:

$$m \frac{d^2}{dt^2} p_x(t) + b \frac{d}{dt} p_x(t) + k p_x(t) = F(t) = m a_x(t) \quad (4.4)$$

gdzie $p_x(t)$ jest przemieszczeniem się ciężarka w osi x pod wpływem działającej siły $F(t)$, $a_x(t)$ reprezentuje przyśpieszenie poruszającego się ciężarka w osi x . Przed-

4. Czujniki pomiarowe



Rys. 4.2. Mechaniczny model czujnika przyśpieszenia

stawione równanie dynamiki jest prawdziwe dla każdej z osi (x, y, z) pomiarowych czujnika przyśpieszenia.

Analizując równanie (4.4), zauważamy, że w stanie ustalonym zmienna Δp_x jest proporcjonalna do działającej w układzie siły:

$$\Delta p_x \propto F(t) = ma_x(t) \quad (4.5)$$

Przedstawione równanie (4.4) jest równaniem oscylatora. Równanie (4.4) można zapisać w formie następującego równania w przestrzeni stanu:

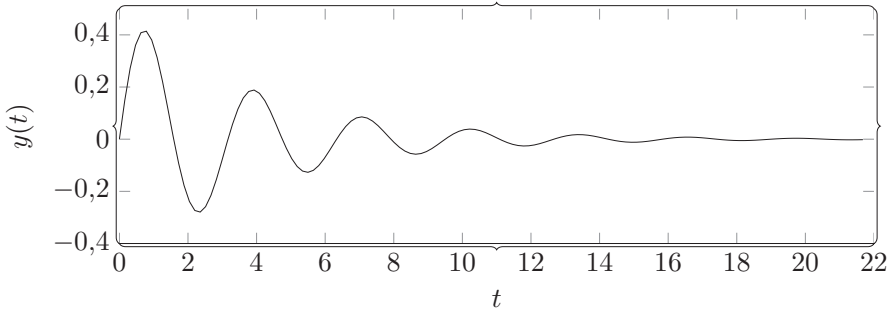
$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & \frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} F(t) \quad (4.6)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (4.7)$$

gdzie x_1 oraz x_2 są składowymi następującego wektora stanu:

$$\mathbf{x} = [x_1 \quad x_2]^T \quad (4.8)$$

Na rysunku 4.3 przedstawiono przykładową odpowiedź impulsową systemu opisanego równaniem (4.4).



Rys. 4.3. Odpowiedź impulsowa czujnika przyspieszenia

4.2.1. Zakłócenia w czujnikach przyspieszenia

Przedstawiony w podrozdziale 4.2 opis działania czujnika przyspieszenia nie uwzględniał wpływu szumów zewnętrznych na pomiar przyspieszenia. Rzeczywiste pomiary tej wielkości, oprócz interesującego nas sygnału, zawierają również niepożądane składowe. Omówione zostaną teraz zagadnienia związane z metodami analizy zakłóceń w czujnikach przyspieszenia. Rozdział ten częściowo opiera się na materiale opublikowanym w pracy [223].

Szumy w układach elektronicznych, również w tych wykonanych w technologii MEMS, mogą być związane zarówno ze zjawiskami zachodzącymi wewnątrz urządzenia, jak i z czynnikami zewnętrznymi. Głównymi przyczynami niedokładności pomiarowych w czujnikach przyspieszenia wykonanych w technologii MEMS są [22, 111, 145, 445]:

- nieortogonalność osi pomiarowych,
- fluktuacje przepływającego prądu,
- efekty cieplne związane z przepływem elektronów,
- wpływ temperatury otoczenia oraz samopodgrzewanie się czujnika.

Błędy w pomiarach przyspieszenia mogą być powodowane na przykład przez nieortogonalność osi pomiarowych czujnika, tj. osi x , y , z . W tym przypadku, możliwe jest kompensowanie niedokładności powodowane przez nieortogonalność osi pomiarowych czy też błąd systematyczny w pomiarach, dokonując procedury kalibracji czujnika.

Innym czynnikiem wpływającym na jakość pomiarów pochodzących z czujników wykonanych w technologii MEMS są fluktuacje przepływającego prądu. Fluktuacje te związane są głównie z efektami powierzchniowymi oraz defektami siatki krystalicznej.

Efekty cieplne występujące w układach MEMS są powodowane głównie przez ruch elektronów. Ich chaotyczne ruchy powodują wydzielanie energii cieplnej.

4. Czujniki pomiarowe

Wydzielanie energii cieplnej powoduje z kolei powstawanie szumu, który wpływa niekorzystnie na jakość pomiarów. Szum ten ma charakterystykę zbliżoną do szumu białego (jest to szum typu VRW, tj. *Velocity Random Walk*).

Wpływ temperatury na jakość pomiarów może mieć różne przyczyny. Jedną z nich jest otoczenie, w którym wykonywane są pomiary. Inna z przyczyn związana jest ze zjawiskiem samopodgrzewania się czujników wykorzystywanych w pomiarach. Zjawiska wywołane zmianami temperatury sprawiają, że w sygnałach pomiarowych mogą pojawić się szумы, których usunięcie jest trudnym zadaniem.

Część z wymienionych czynników wpływających na jakość działania czujnika przyspieszenia może być przedmiotem modelowania, a uzyskane wyniki mogą zostać wykorzystane do eliminacji ich negatywnego wpływu na wynik pomiaru. W tabeli 4.1 dokonano podsumowania głównych czynników wpływających na dokładność czujnika przyspieszenia [445].

Tabela 4.1. Typy błędów w czujnikach przyspieszenia

| Typ błędu | Opis |
|---|--|
| Błąd systematyczny (ang. <i>bias</i>) | Obciążenie jest mierzone w pomiarze przyspieszenia |
| Szum biały | Związany z ruchami Browna elektronów |
| Gleick2012 Nieortogonalność | Błąd deterministyczny spowodowany nieprostokątnym ustawieniem osi x , y , z czujnika |
| Szum BI (ang. <i>Bias Instability</i>) | Fluktuacja błędu systematycznego |

4.2.2. Model pomiarowy czujnika przyspieszenia

Bazując na wynikach analizy źródeł błędów dla czujnika przyspieszenia oraz na pracach [345, 350, 358], zaproponujemy następujący model pomiarowy dla czujnika przyspieszenia:

$$\bar{\mathbf{y}}_a = \mathbf{a} + \underbrace{\mathbf{v}_{a_{bs}} + \mathbf{v}_{a_{wh}}}_{\mathbf{v}} = \mathbf{R}_{a,b} (\mathbf{a}_g + \mathbf{a}_F) + \mathbf{v}_{a_{bs}} + \mathbf{v}_{a_{wh}} \quad (4.9)$$

gdzie \mathbf{a} jest prawdziwym, ale nieznanym przyspieszeniem (zawierającym składową związaną z wymuszeniem i z przyspieszeniem ziemskim), $\mathbf{v}_{a_{bs}}$ to niskoczęstotliwościowe (wolnozmiennie) zakłócenie (np. dryft), $\mathbf{v}_{a_{wh}}$ reprezentuje biały szum, \mathbf{a}_F jest przyspieszeniem związanym z siłą wymuszającą F , \mathbf{a}_g to składowa związana z pomiarem przyspieszenia ziemskiego, natomiast $\mathbf{R}_{a,b}$ to macierz rotacji prze-

kształcenia pomiędzy układem czujnika $\{B\}$ (indeks b) a układem odniesienia $\{A\}$ (patrz rozdz. 9.2).

Obecny w pomiarach dryft jest powiązany z szumem wywołanym głównie przez błąd systematyczny i jego fluktuacje. Z kolei biały szum obserwowany w pomiarach przyspieszenia związany jest z chaotycznym ruchem przepływających elektronów (szum typu VRW). Z zależności (4.9) wynika, że niebędący w ruchu akcelerometr mierzy w jednej z osi przyspieszenie ziemskie.

4.2.3. Charakterystyka czujnika przyspieszenia

W tabeli 4.2 przedstawiono parametry czujnika przyspieszenia będącego częścią platformy firmy Shimmer.

Tabela 4.2. Parametry dla czujnika przyspieszenia [367]

| Parametr | Wielkość |
|---------------------------|--|
| Zakres | $\pm 1,5$ [g]– $\pm 6,0$ [g] |
| Czułość | 800 [mV/g] (dla $\pm 1,5$ [g]); 206 [mV/g] (dla $\pm 6,0$ [g]) |
| Częstotliwość próbkowania | 5–1024 [Hz] |

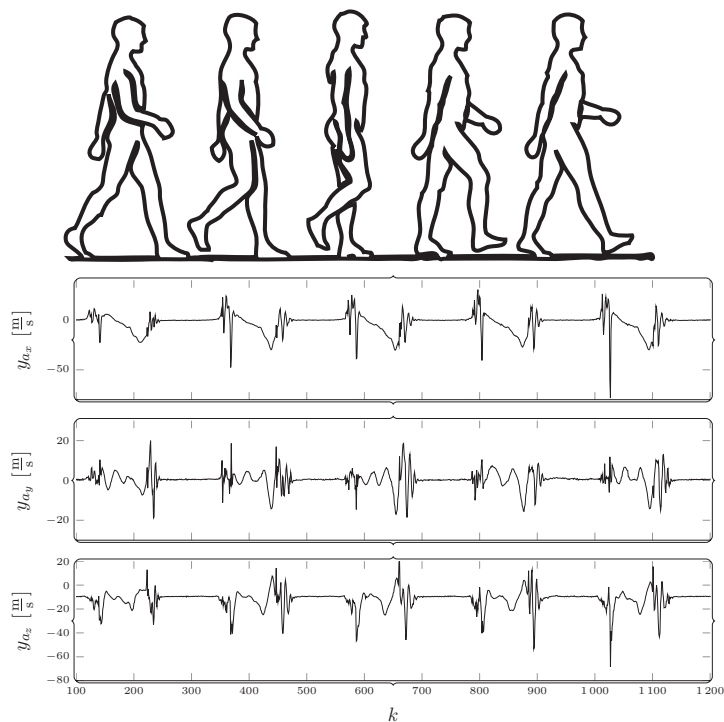
Czujniki przyspieszenia, pomimo swoich zalet i dużych możliwości wykorzystania danych pomiarowych pozyskanych podczas ich zastosowania, mają również swoje ograniczenia. Jednym z ograniczeń tych czujników jest brak możliwości określenia orientacji czujnika w przestrzeni, w przypadku gdy mierzony obiekt porusza się [228]. W takim przypadku akcelerometr mierzy nie tylko składową przyspieszenia związaną z przyspieszeniem ziemskim, ale również składową przyspieszenia związaną z ruchem obiektu (4.1). Brak możliwości odseparowania składowej związanej z przyspieszeniem ziemskim od składowej związanej z ruchem obiektu uniemożliwia estymację orientacji poruszającego się obiektu z wykorzystaniem jedynie danych z czujnika przyspieszenia.

Jedno z podejść, które pozwala odseparować opisane składowe w pomiarach przyspieszenia, bazuje na wykorzystaniu danych pomiarowych z dodatkowego czujnika. Tym czujnikiem jest żyroskop. Zastosowanie tego czujnika ułatwia wyznaczenie orientacji obiektu w przestrzeni, w przypadku gdy jest on w ruchu.

4.2.4. Przykładowe pomiary z wykorzystaniem czujnika przyspieszenia

Na rysunku 4.4 przedstawiono przykładowe pomiary, które wykonano z wykorzystaniem czujnika przyspieszenia dla osoby chodzącej. W przedstawionym

4. Czujniki pomiarowe



Rys. 4.4. Przykładowy pomiar przyśpieszenia podczas chodu człowieka. Sposób przymocowania czujnika przedstawiono na rys. 4.5



Rys. 4.5. Sposób przymocowania czujnika przyśpieszenia (przykład z podrozdz. 4.2.4)

przykładzie czujnik przyśpieszenia przymocowany został do jednego z butów (rys. 4.5), który nosiła podczas badania poruszająca się osoba.

4.3. Żyroskop

Żyroskopy są urządzeniami służącymi do pomiaru prędkości kątowej. Podobnie jak czujniki przyśpieszenia, również żyroskopy reagują na siły działające na to urządzenie pomiarowe. Czujniki żyroskopowe dzielimy na:

- kierunkowe,
- prędkościowe.

Żyroskopy kierunkowe służą do pomiaru orientacji obiektu w przestrzeni. Natomiast żyroskopy prędkościowe służą do pomiaru prędkości kątowych. W technologii MEMS wykonuje się najczęściej żyroskopy prędkościowe. W urządzeniach tych pomiaru prędkości kątowych dokonuje się dla każdej z osi obrotów. Typowy żyroskop umożliwia pomiar prędkości kątowych wokół osi x , y i z . Takie żyroskopy nazywane żyroskopami trójosiowymi. Pomiar z czujnika żyroskopowego będziemy oznaczać następująco:

$$\mathbf{y}_\omega = [y_{\omega_x} \quad y_{\omega_y} \quad y_{\omega_z}]^T \quad (4.10)$$

Żyroskopy prędkościowe wykonane w technologii MEMS do pomiarów wykorzystują efekt Coriolisa. Efekt Coriolisa wywoływany jest przez siłę Coriolisa. Siłę tę opisujemy zależnością:

$$F_c = 2m (\boldsymbol{\omega} \times \mathbf{v}) \quad (4.11)$$

gdzie $\boldsymbol{\omega}$ to prędkość kątowa obracającego się obiektu wokół zadanej osi, \mathbf{v} reprezentuje prędkość poruszającego się obiektu, a m oznacza jego masę. Na rysunku 4.6 zilustrowano efekt Coriolisa.

Podobnie jak dla czujnika przyśpieszenia, również dla prędkościowego czujnika żyroskopowego można przedstawić uproszczony model mechaniczny (patrz rys. 4.7).

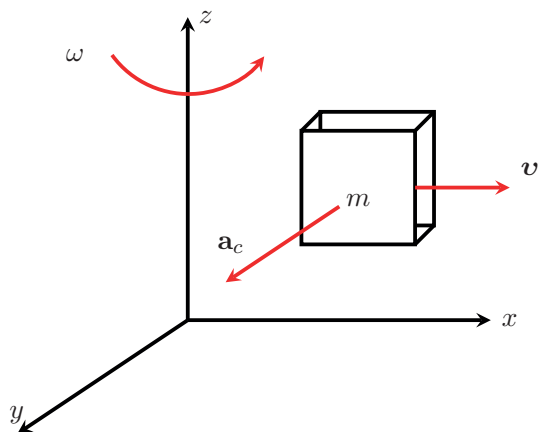
Głównymi elementami tego czujnika jest masa m , do której przymocowane są sprężyny o współczynnikach sprężystości k_1 i k_2 oraz układy tłumiące o współczynnikach b_1 i b_2 . Typowo czujnik żyroskopowy może dokonywać pomiarów w trzech osiach. W przedstawionym na rys. 4.7 schemacie ograniczono się do dwóch osi.

W żyroskopach opartych na tzw. efekcie Coriolisa wyróżniamy dwa główne układy, tj. układ wzbudzania drgań oraz układ służący do pomiaru analogowego sygnału napięciowego, który jest proporcjonalny do tych drgań. Budowa żyroskopu kierunkowego jest odzwierciedlona w modelu idealnego żyroskopu, który składa się z dwóch następujących części:

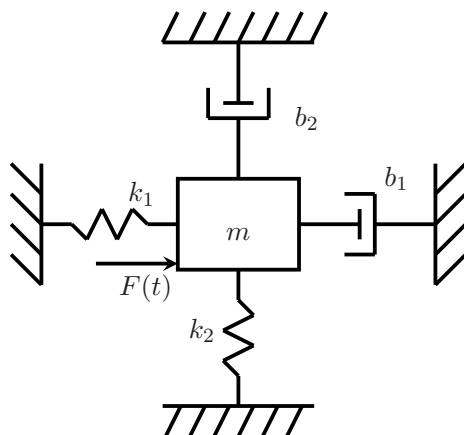
$$m \frac{d^2}{dt^2} p_x(t) + b_1 \frac{d}{dt} p_x(t) + k_1 p_x(t) = F(t) \quad (4.12a)$$

$$m \frac{d^2}{dt^2} p_y(t) + b_2 \frac{d}{dt} p_y(t) + k_2 p_y(t) = F_c(t) \quad (4.12b)$$

4. Czujniki pomiarowe



Rys. 4.6. Ilustracja efektu *Coriolisa* [121]



Rys. 4.7. Mechaniczny model czujnika żyroskopowego (dwuosiowego)

gdzie m oznacza masę elementu, b_1 , b_2 są współczynnikami tłumienia, k_1 , k_2 reprezentuje współczynniki sprężystości, $F_c(t)$ jest siłą Coriolisa (4.11); $p_x(t)$ jest przemieszczeniem w osi x , a $p_y(t)$ oznacza przesunięcie w osi y .

Równania (4.12a) oraz (4.12b) można przedstawić w postaci układu następujących równań stanu i wyjścia:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & \frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{1}{m} & 0 \\ 0 & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} F(t) \\ F_c(t) \end{bmatrix} \quad (4.13)$$

$$y(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} \quad (4.14)$$

gdzie x_1 , x_2 , x_3 oraz x_4 są składowymi wektora stanu:

$$\mathbf{x} = [x_1 \quad x_2 \quad x_3 \quad x_4]^T \quad (4.15)$$

4.3.1. Zakłócenia w czujnikach żyroskopowych

Podobnie jak jest to w przypadku czujnika przyśpieszenia, również pomiary żyroskopowe zawierają pewne niepożądane składowe. Część z tych niepożądanych składowych można wyeliminować na etapie kalibracji urządzania pomiarowego. Inne zakłócenia należy eliminować przez stosowanie metod bazujących na modelowaniu matematycznym lub algorytmach przetwarzania sygnałów.

Nasze rozważanie dotyczące pomiarów żyroskopowych rozpoczniemy od metod wykorzystywanych podczas kalibracji czujników. Jedną z przyczyn niedokładności w pomiarach żyroskopowych, które można kompensować podczas kalibracji, jest nieortogonalność osi pomiarowych. Brak kalibracji czujnika może doprowadzić do błędów w pomiarach, które będą się kumulowały. Może to doprowadzić do niedokładnych pomiarów prędkości kątowych.

Kalibracja czujników żyroskopowych jest procedurą dwuetapową. W pierwszym etapie kalibracja wykonywana jest na nieruchomym czujniku, natomiast w drugim etapie kalibracji dokonuje się, gdy czujnik jest w ruchu. Ruch odbywa się wokół wszystkich osi pomiarowych czujnika. Drugi z etapów wymaga precyzji i jest trudny do precyzyjnego wykonania bez odpowiedniego oprzyrządowania.

Oprócz nieortogonalności osi pomiarowych, które mają wpływ na obniżenie jakości pomiarów prędkości kątowych, innymi źródłami niedokładności w pomiarach żyroskopowych wykonanych w technologii MEMS są [22, 111, 145, 445]:

4. Czujniki pomiarowe

- efekty cieplne, związane z przepływem elektronów,
- fluktuacje przepływającego prądu,
- wpływ temperatury otoczenia oraz samopodgrzewanie się czujnika.

Efekty cieplne w czujniku prędkości kątowej mają te same konsekwencje, co w przypadku akcelerometru. Błędy pomiarowe związane są zarówno ze zmianą temperatury otoczenia, jak i z efektem samopodgrzewania czujnika. I podobnie jak w przypadku poprzednio omawianego czujnika, zależność pomiędzy temperaturą otoczenia a błędem systematycznym jest zależnością silnie nieliniową.

Fluktuacje przepływającego prądu wpływają na zmiany błędu systematycznego w pomiarach żyroskopowych. Składowe tego typu modelowane są z wykorzystaniem chaotycznych ruchów Browna, tj. błędzenia przypadkowego. Jest to jednak jedynie przybliżenie, które w wielu praktycznych zastosowaniach może okazać się niewystarczające. W tabeli 4.3 przedstawiono podsumowanie zjawisk, które mają wpływ na błędy w pomiarach żyroskopowych.

Tabela 4.3. Typy błędów w żyroskopie

| Typ błędu | Opis |
|---|--|
| Błąd systematyczny (ang. <i>bias</i>) | Obciążenie jest mierzone w pomiarze przyśpieszenia. |
| Szum biały | Związany z ruchami Browna elektronów. |
| Nieortogonalność | Błąd deterministyczny, spowodowany nieprostokątnym ustawieniem osi x , y , z czujnika. |
| Szum BI (ang. <i>Bias Instability</i>) | Fluktuacja błędu systematycznego |

4.3.2. Model pomiarowy czujnika żyroskopowego

Bazując na analizie przeprowadzonej w poprzednim podrozdziale oraz m.in. w pracach [66, 345, 358], proponujemy następujący model pomiaru dla czujnika żyroskopowego:

$$\bar{\mathbf{y}}_{\omega} = \boldsymbol{\omega} + \mathbf{v}_{\omega_{bs}} + \mathbf{v}_{\omega_{wh}} \quad (4.16)$$

gdzie $\boldsymbol{\omega}$ jest prawdziwą, ale nieznaną prędkością kątową obiektu, $\mathbf{v}_{\omega_{bs}}$ to niskoczęstotliwościowy składnik o cechach niestacjonarnych (np. dryft), $\mathbf{v}_{\omega_{wh}}$ reprezentuje biały szum. Dryft w czujnikach żyroskopowych powodowany jest głównie przez błąd systematyczny i jego fluktuacje.

4.3.3. Charakterystyka

Czujniki żyroskopowe produkowane są przez wielu producentów układów elektronicznych. W ofercie dostępne są czujniki o różnych parametrach, które dostosowane są do wymagań odbiorcy. W tabeli 4.4 przedstawiono parametry czujnika żyroskopowego oferowanego przez firmę Shimmer, który bazuje na układzie MPU-9150 producenta InvenSense.

Tabela 4.4. Parametry dla czujnika żyroskopowego [367]

| Parametr | Wielkość |
|---------------------------|-------------------|
| Zakres | ± 500 [deg/s] |
| Czułość | 2 [mV/(deg/s)] |
| Częstotliwość próbkowania | 5–1024 [Hz] |

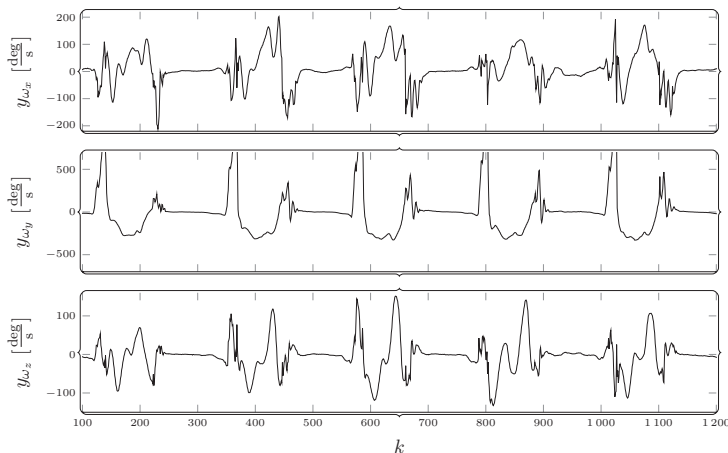
Oprócz prędkości kątowych możliwe jest również wyznaczenie wartości kąta obrotu przez wykonanie numerycznej operacji całkowania pomiarów żyroskopowych. Niestety wyniki uzyskane tą metodą są obciążone błędem. Jest to spowodowane niedokładnością pomiarową czujników żyroskopowych oraz kumulacją tych błędów podczas wykonywania operacji całkowania numerycznego. Wykonywanie operacji całkowania numerycznego dla danych pomiarowych z szumem pomiarowym może skutkować pojawieniem się w wynikach niskoczęstotliwościowych składowych, które są trudne do usunięcia.

Jedno z podejść ułatwiających ograniczenie niekorzystnego wpływu niskoczęstotliwościowych składowych bazuje na wykorzystaniu w obliczeniach pomiarów z czujnika pola magnetycznego. Zagadnienie to zostanie szczegółowo omówione w kolejnym rozdziale.

4.3.4. Przykładowe pomiary z wykorzystaniem czujnika żyroskopowego

Na rysunku 4.8 przedstawiono przykładowy pomiar wykonany z wykorzystaniem czujnika żyroskopowego dla osoby będącej w ruchu. Dane zbierano podczas tego samego eksperymentu pomiarowego, który opisano w podrozdziale 4.2.4. Analizując pomiary, można zauważyć, że zawierają one niepożądane składowe, które powinny zostać usunięte przed przystąpieniem do dalszych obliczeń.

4. Czujniki pomiarowe



Rys. 4.8. Przykładowy pomiar prędkości kątowej podczas chodu człowieka (kontynuacja przykładu z rys. 4.4 oraz 4.5)

4.4. Czujnik pola magnetycznego (magnetometr)

Czujniki magnetometryczne służą do pomiaru natężenia i/lub kierunku pola magnetycznego Ziemi w ich otoczeniu. Tak jak czujniki przyśpieszenia i żyroskopy, również i te urządzenia pomiarowe wykonywane są w technologii MEMS. Najczęściej spotykane rozwiązania magnetometrów to czujniki trójosiowe. W odróżnieniu od omówionych już czujników przyśpieszenia i żyroskopów, wykorzystanie danych dostarczanych przez te urządzenia pomiarowe wymaga większego nakładu obliczeń. Jest to związane z tym, że pomiary z magnetometru różnią się w zależności od tego, w jakim miejscu na Ziemi są wykonywane. Czujniki pola magnetycznego dzielimy na czujniki do pomiaru wielkości [233]:

- natężenia i kierunku pola magnetycznego (ang. *vector magnetometer*),
- natężenia pola magnetycznego (ang. *total field magnetometer*).

Kierunek pola magnetycznego określa inklinacja i deklinacja magnetyczna. Inklinacja magnetyczna (nachylenie magnetyczne) jest kątem zawartym pomiędzy wektorem natężenia ziemskiego pola magnetycznego a płaszczyzną poziomą [443]. Deklinacją magnetyczną (odchyleniem magnetycznym) jest kąt zawarty pomiędzy południkiem geograficznym a południkiem magnetycznym w miejscu obserwacji [443]. Przykładowo wartości tych wielkości dla Wrocławia wynoszą odpowiednio $\approx 4^\circ$ i $\approx 66^\circ$ [312].

4.4. Czujnik pola magnetycznego (magnetometr)

Typowe oznaczenie wektora pomiarów uzyskanych z trójosiowego czujnika magnetometrycznego to:

$$\mathbf{y}_m = [y_{m_x} \quad y_{m_y} \quad y_{m_z}]^T \quad (4.17)$$

Składowe wektora natężenia pola magnetycznego, zmierzone z wykorzystaniem trójosiowego magnetometru, mają następującą postać:

$$\mathbf{y}_m = \mathbf{R}_{a,b} \begin{bmatrix} \cos(\tilde{h}) \\ 0 \\ -\sin(\tilde{h}) \end{bmatrix} \quad (4.18)$$

gdzie $\mathbf{R}_{a,b}$ jest macierzą rotacji, \tilde{h} to wartość nachylenia magnetycznego w danej lokalizacji. Macierz rotacji $\mathbf{R}_{a,b}$ opisuje zależność pomiędzy układem czujnika i układem globalnym. Bazowy układ współrzędnych jest tak zorientowany, że pierwsza z osi zgodna jest z magnetyczną północą (jest to kierunek wskazywany przez kompas magnetyczny), a trzecia oś skierowana jest do góry [378].

Projektanci czujników pola magnetycznego wykorzystują wiele znanych zjawisk fizycznych. W zależności od konstrukcji urządzenia oraz zjawiska fizycznego, na którego bazie wykonywany jest pomiar, czujniki pola magnetycznego dzielimy na czujniki pól słabych, średnich i silnych. Najczęściej spotykane czujniki oparte są na zjawisku magnetorezystancji, na zjawisku Halla oraz wykorzystujące siłę Lorentza.

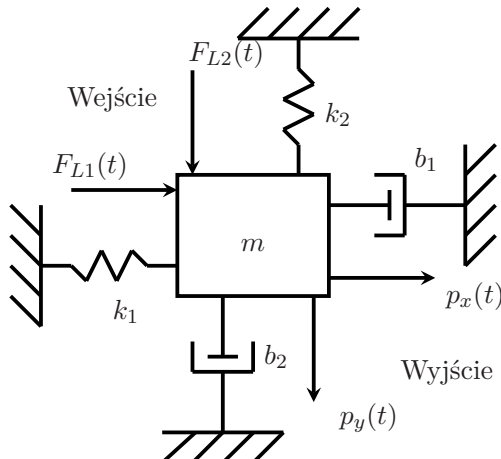
Magnetorezystancja dotyczy zmiany rezystancji materiału przewodzącego pod wpływem pola magnetycznego. Zjawisko Halla natomiast dotyczy powstawania różnicy potencjałów w przewodniku po umieszczeniu go w polu magnetycznym prostopadłym do przepływu prądu. Siła Lorentza jest siłą działającą na cząstkę z ładunkiem elektrycznym, poruszającą się w polu elektromagnetycznym.

Najpopularniejsze czujniki magnetometryczne bazują na zjawisku magnetorezystancji lub efekcie Halla. Do budowy czujników magnetorezystancyjnych wykorzystywane są specjalistyczne materiały magnetyczne [54]. Ponadto w czujnikach tego typu występuje niekorzystne zjawisko, jakim jest efekt histerezy. Efekt ten wymusza konieczność okresowego rozmagnesowania urządzeń zbudowanych w tej technologii. Z kolei czujniki oparte na efekcie Halla nie mają wad charakterystycznych dla czujników opartych na zjawisku magnetorezystancji, niestety ich czułość jest znacznie mniejsza. Ponadto zapotrzebowanie na energię w tego typu czujnikach jest znacznie wyższe.

Alternatywą dla opisanych typów czujników są, zyskujące dużą popularność w ostatnich latach, czujniki wykorzystujące siłę Lorentza. Charakteryzują się małym poborem energii oraz dużą czułością. Ponadto nie wymagają do budo-

4. Czujniki pomiarowe

wy specjalnych materiałów magnetycznych. Uproszczony model czujnika pola magnetycznego wykorzystującego siłę Lorentza przedstawiono na rys. 4.9 [242].



Rys. 4.9. Elektromechaniczny model czujnika pola magnetycznego bazującego na sile Lorentza [242]

Schemat magnetometru wykorzystującego siłę Lorentza jest podobny do schematu żyroskopu. W opisanym przypadku model czujnika magnetometrycznego ma postać [54]:

$$m \frac{d^2}{dt^2} p_x(t) + b_1 \frac{d}{dt} p_x(t) + k_1 p_x(t) + F_{el1,1}(t) - F_{el1,2}(t) = F_{L1}(t) \quad (4.19a)$$

$$m \frac{d^2}{dt^2} p_y(t) + b_2 \frac{d}{dt} p_y(t) + k_2 p_y(t) + F_{el2,1}(t) - F_{el2,2}(t) = F_{L2}(t) \quad (4.19b)$$

gdzie $F_{L1}(t)$, $F_{L2}(t)$ są siłami, które powstają w polu magnetycznym (siły Lorentza), $F_{eli,j}(t)$ siły elektrostatyczne związane z pracującymi w układzie czujnika kondensatorami. Siły te wyznaczone są z następujących zależności:

$$F_{eli,1}(t) = -\frac{1}{2} \epsilon_{0m} A_m \frac{(V_{i,1} - V_m)^2}{(L_m + y_1)^2} \quad (4.20a)$$

$$F_{eli,2}(t) = \frac{1}{2} \epsilon_{0m} A_m \frac{(V_m - V_{i,2})^2}{(L_m - y_2)^2} \quad (4.20b)$$

gdzie ϵ_{0m} to stała dielektryczna próżni, $V_{i,j}$ są potencjałami stałych okładek kondensatora (statorów), V_m to potencjał poruszającego się ciężarka, L_m to

4.4. Czujnik pola magnetycznego (magnetometr)

odległość pomiędzy elektrodami w stanie spoczynku, A_m to całkowita powierzchnia elektrody.

4.4.1. Zakłócenia w czujnikach pola magnetycznego

Podobnie jak w przypadku czujników przyśpieszenia i żyroskopowych, na dokładność czujnika pola magnetycznego wpływają szумы, nieliniowości oraz dryft. Dodatkowymi źródłami zakłóceń dla czujnika pola magnetycznego są materiały magnetyczne, które znajdują się na stałe w otoczeniu czujnika pola magnetycznego. Mogą one wywoływać efekty trwałego namagnesowania lub przyczynić się do deformacji pola magnetycznego. Zakłócenia tego typu mogą być kompensowane po wykorzystaniu dedykowanych algorytmów [133].

Drugi typ zakłóceń związany jest również z obecnością elementów zakłócających, ale elementy te nie są na stałe obecne w otoczeniu czujników. Również te zakłócenia mogą być modelowane, a uzyskane wyniki wykorzystano do kompensacji ich niekorzystnego wpływu na działanie czujników natężenia pola magnetycznego [347].

4.4.2. Model pomiarowy czujnika pola magnetycznego

Wykorzystując wyniki analizy z poprzedniego rozdziału oraz z prac [205, 338, 345, 347, 358], wprowadźmy następujący model dla czujnika natężenia pola magnetycznego:

$$\mathbf{y}_m = \mathbf{R}_{a,b}\mathbf{h} + \mathbf{v}_{m_{bs}} + \mathbf{v}_{m_{wh}} \quad (4.21)$$

gdzie \mathbf{h} jest prawdziwym, ale nieznanym natężeniem pola magnetycznego, $\mathbf{v}_{m_{bs}}$ to niskoczęstotliwościowy składnik o naturze niestacjonarnej (np. dryft), $\mathbf{v}_{m_{wh}}$ reprezentuje biały szum.

4.4.3. Charakterystyka

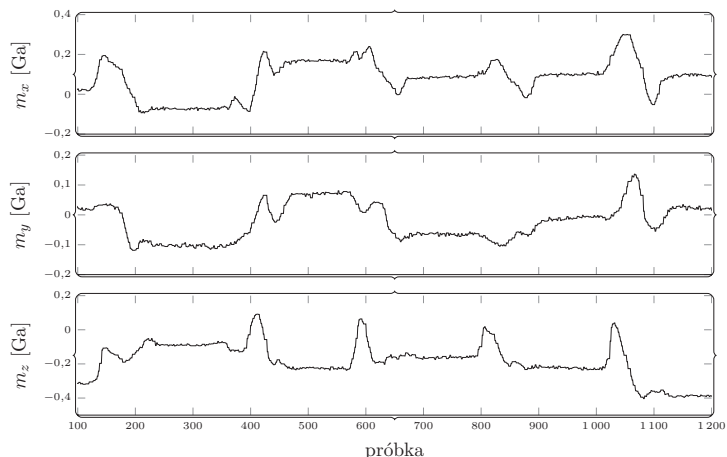
W tabeli 4.5 przedstawiono parametry techniczne czujnika pola magnetycznego dostępnego dla platformy Shimmer.

Tabela 4.5. Parametry dla czujnika pola magnetycznego [367]

| Parametr | Wielkość |
|---------------------------|----------------|
| Zakres | $\pm 4,5$ [Ga] |
| Rozdzielczość | 7 [mGa] |
| Częstotliwość próbkowania | 5–1024 [Hz] |

4.4.4. Przykładowe pomiary z wykorzystaniem czujnika magnetometrycznego

Na koniec rozdziału przedstawimy przykładowy pomiar wykonany z wykorzystaniem magnetometru (rys. 4.10) dla chodzącej osoby. Pomiar wykonano w trakcie eksperymentu wykonanego podczas badania opisanego w podrozdziale 4.2.4. Dodanie pomiarów z magnetometru poprawia wyniki szacowania orientacji obiektu w przestrzeni. Teoretycznie prawidłowe wyniki powinny być możliwe do uzyskania z wykorzystaniem jedynie pomiarów z czujników przyspieszenia i prędkości kątowych. Ze względu na błędy pomiarowe obecne w danych, w praktyce jest to bardzo trudne. Dodanie pomiarów natężenia pola magnetycznego znacząco wpływa na poprawę jakości estymacji orientacji obiektu w przestrzeni. Dlatego bardzo wielu producentów decyduje się na dodawanie do akcelerometru i żyroskopu również magnetometru.



Rys. 4.10. Przykładowy pomiar natężenia pola magnetycznego podczas chodu człowieka (kontynuacja przykładu z rys. 4.4, 4.5 oraz 4.8)

4.5. Metody analizy zakłóceń w czujnikach pomiarowych

W dalszej części tego rozdziału omówione zostaną zagadnienia związane z metodami analizy zakłóceń w czujnikach pomiarowych. Rozdział ten częściowo bazuje na materiale opublikowanym w pracy [223].

4.5. Metody analizy zakłóceń w czujnikach pomiarowych

Dane pomiarowe pozyskane z omówionych we wcześniejszych podrozdziałach czujników zawierają pewne niepożądane składowe, które powinny zostać usunięte zanim zostaną przetworzone na kolejnych poziomach systemu fuzji danych. Przyczyną powstania tych niepożądanych składowych w sygnałach pomiarowych są zjawiska zachodzące w czujnikach podczas ich działania (np. wzrost temperatury spowodowany m.in. przez przepływający prąd) oraz w ich otoczeniu (np. zakłócenia powodowane przez elementy będące w pobliżu czujników pola magnetycznego). Jak już to zaznaczono, jednym ze sposobów na usunięcie niepożądanych składowych z pomiarów jest ich modelowanie. Ważne jest, by w modelach tych uwzględniono zjawiska mające wpływ na dokładność uzyskiwanych pomiarów.

W wielu aplikacjach konstrukcja modeli szumów jest koniecznością, ponieważ producenci podają jedynie podstawowe informacje na temat produkowanych urządzeń. A poprawę dokładności zebranych danych pomiarowych można uzyskać jedynie przez rozpoznanie szumów wpływających na działanie czujnika, ich modelowanie oraz usunięcie z pomiarów.

Podstawowe metody, które znalazły zastosowanie w analizie szumów w czujnikach przyśpieszenia, żyroskopowych oraz pola magnetycznego (tzn. czujniki MARG), to metody bazujące na estymacji **widmowej gęstości mocy** (ang. *Power Spectral Density*, PSD) oraz **wariancji Allana**.

Widmowa gęstość mocy (WGM) jest klasyczną metodą do opisu struktury częstotliwościowej analizowanego sygnału przez wyznaczenie widma wartości średniokwadratowej sygnału [33]. Wyznaczenie tej struktury ułatwia pozyskanie istotnych informacji o składnikach widma sygnału, co może być przydatne podczas opracowania algorytmów ich przetwarzania.

Widmowa gęstość mocy ma swoje ograniczenia, co przekłada się na ograniczony zakres jej stosowania w zagadnieniach praktycznych. Ograniczenia te związane są głównie ze sposobem wyznaczania WGM. Metoda ta przeznaczona jest głównie do analizy sygnałów stacjonarnych. Wciąż jednak pozostaje jednym z najpopularniejszych narzędzi do analizy sygnałów w dziedzinie częstotliwości [33]. Pomimo ograniczeń tego podejścia metoda ta jest wykorzystywana m.in. do analizy sygnałów z czujników inercyjnych, tj. akcelerometru i żyroskopu [175, 327] oraz z czujników natężenia pola magnetycznego [205, 277, 365].

Alternatywą dla opisanej metody jest podejście bazujące na **wariancji Allana**. Wariancja Allana bazuje na dwupróbkowej statystyce, która wyznaczana jest jako odchylenie standardowe różnic wartości średnich wielkości mierzonych w sąsiadujących ze sobą przedziałach czasowych [10]. Pierwotnie metody bazujące na wariancji Allana były wykorzystywane do mierzenia stabilności oscylatorów oraz zegarów [11], ponieważ metoda ta pozwala na charakterystykę fluktuacji fazy i częstości analizowanego sygnału. Obecnie podejście to wykorzystywane jest m.in. do konstrukcji metod wspomagających rozpoznawanie typów zakłóceń oraz ich

4. Czujniki pomiarowe

oceny dla urządzeń pomiarowych. Wariancja Allana znalazła również zastosowanie w analizie danych pomiarowych zbieranych z czujników typu IMU oraz MARG.

Zaznaczmy, że podejście do analizy sygnałów pomiarowych bazujące na wariancji Allana charakteryzuje się tym, że dane z czujników analizowane są w dziedzinie czasu.

Metody wykorzystujące wariancję Allana składają się z dwóch etapów, na pierwszym z nich wylicza się wartość wariancji Allana dla analizowanego sygnału, na drugim z etapów dokonywana jest estymacja wartości błędów. Warto zaznaczyć, że wyznaczona wartość wariancji Allana zależy od długości przedziału czasowego, na którym analiza jest przeprowadzona. Stąd bardzo istotnym etapem metod wykorzystujących wariancję Allana jest dobór odpowiedniego przedziału czasowego.

Widmowa gęstość mocy

Widmową gęstość mocy definiuje się jako transformację Fouriera funkcji autokorelacji. Niech

$$Y(f) = \int_{-\infty}^{\infty} y(t)e^{-j2\pi ft} dt \quad (4.22)$$

będzie transformacją Fouriera sygnału pomiarowego $y(t)$. Zdefiniujmy również funkcję autokorelacji:

$$R_{yy}(\tau) = y(\tau) * y(-\tau) = \int_{-\infty}^{\infty} y(t)y(t + \tau)dt \quad (4.23)$$

gdzie $*$ jest splotem funkcji. Dla sygnałów stacjonarnych widmowa gęstość mocy związana jest z funkcją autokorelacji przekształceniem Fouriera:

$$S_{yy}(f) = \int_{-\infty}^{\infty} R_{yy}(\tau)e^{-j2\pi f\tau} d\tau \quad (4.24)$$

Widmowa gęstość mocy wyraża miarę mocy średniej sygnału przypadającej na jednostkę częstotliwości. Inną, istotną cechą gęstości widmowej mocy jest jej związek z funkcją autokorelacji, co widać w zależności (4.24). Głównym celem estymacji gęstości widmowej mocy sygnału jest analiza jego struktury częstotliwościowej.

Wariancja Allana

Alternatywą dla WGM jest metoda analizy zakłóceń bazująca na wariancji Allana. W metodzie tej na pierwszym etapie dzielimy analizowany sygnał pomiarowy $y(k)$, $k = 1, 2, \dots, K$ na N_p przedziałów. Każdy z wyznaczonych przedziałów

4.5. Metody analizy zakłóceń w czujnikach pomiarowych

ma długość czasu T . Następnie wyznaczamy wartość średnią dla każdego z tych przedziałów, korzystając z następującej zależności:

$$\hat{y}(n_p) = \frac{1}{T} \sum_{i=1}^T y((n_p - 1)T + i), \quad n_p = 1, 2, \dots, N_p \quad (4.25)$$

Uśrednianie powtarzamy dla pozostałych przedziałów. Po wyznaczeniu wszystkich wartości przechodzimy do drugiego etapu, w którym wyznaczamy wariancję Allana z zależności [12, 148, 240, 120]:

$$\sigma_{Av}^2(T) = \frac{1}{2(T-1)} \sum_{n_p=1}^{T-1} (\hat{y}(n_p + 1) - \hat{y}(n_p))^2 \quad (4.26)$$

Szczegóły dotyczące wykorzystania algorytmów bazujących na wariancji Allana w zastosowaniu do czujników inercyjnych przedstawiono m.in. w pracach [175, 423].

4.5.1. Reprezentacja szumu z wykorzystaniem wariancji Allana

W tej części zostanie opisana metoda bazująca na wariancji Allana do charakteryzowania najważniejszych typów szumów, które pojawiają się w pomiarach wykonanych czujnikami IMU oraz MARG. Wyróżnijmy następujące szумы w tych czujnikach [120, 256]:

- szum kwantyzacji (ang. *Quantization Noise*),
- szum ARW i VRW (odpowiednio ang. *Angular Random Walk* oraz *Velocity Random Walk*),
- szum BI (ang. *Bias Instability*),
- szum RRW (ang. *Rate Random Walk*),
- szum RR (ang. *Rate Ramp*).

W kolejnych podpunktach omówimy wymienione szумы oraz wskażemy, w jaki sposób mogą być one scharakteryzowane z wykorzystaniem wariancji Allana.

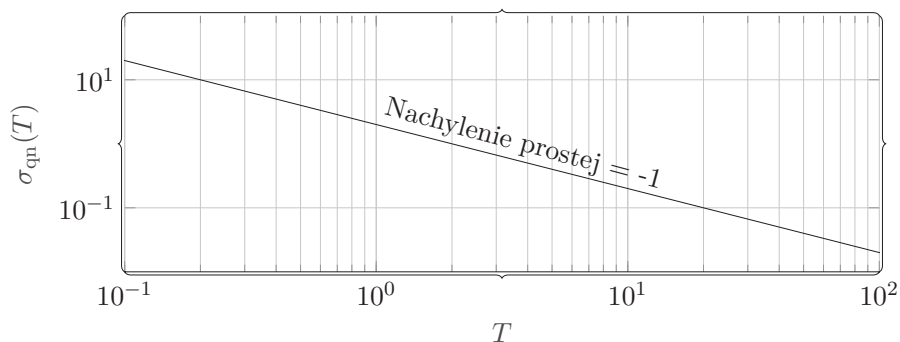
Szum kwantyzacji

Szum kwantyzacji (czasami nazywany również błędem kwantyzacji) związany jest z operacją kwantyzacji sygnału pomiarowego w przetwornikach A/C czujników pomiarowych. Szum ten można wyrazić, korzystając z wariancji Allana w następujący sposób [256, 120]:

$$\sigma_{qn}^2(T) = \frac{3Q_{qn}^2}{T^2} \quad (4.27)$$

gdzie Q_{qn} reprezentuje pewien współczynnik dla szumu kwantyzacji.

4. Czujniki pomiarowe



Rys. 4.11. Szum kwantyzacji wyrażony wariacją Allana

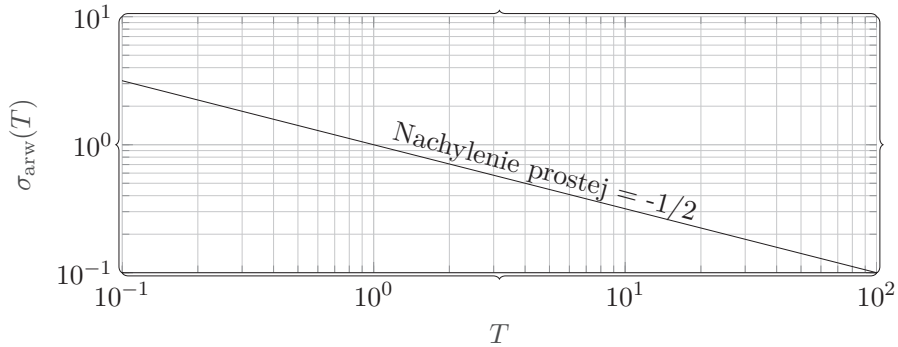
Na rysunku 4.11 przedstawiono wykres w skali *log-log* przedstawiający szum σ_{qn}^2 w zależności od czasu T . Charakterystyczne jest to, że prosta dla tego szumu w skali *log-log* ma nachylenie -1 . Warto również podkreślić, że z wykresu możliwe jest odczytanie amplitudy szumu σ_{qn}^2 w punkcie czasu $T = \sqrt{3}$.

Interesującym zagadnieniem jest sposób modelowania tego szumu. Szum kwantyzacji może być modelowany jako szum addytywny tylko w sytuacji, gdy spełnione są pewne warunki techniczne pomiaru [182]. Pierwszy z tych warunków dotyczy odpowiednio wysokiej rozdzielczości kwantyzacji. Tylko w przypadku, gdy rozdzielczość zastosowanego przetwornika A/C spełnia ten warunek, możliwe jest uzyskanie addytywności szumu kwantyzacji. Jest to spowodowane tym, że operacja kwantyzacji jest operacją nieliniową i tylko przy odpowiednio wysokiej rozdzielczości przetwornika A/C możliwe jest uzyskanie pożądanych właściwości szumu kwantyzacji. Z kolei jeśli błędy w przetworniku A/C miałyby mieć naturę szumu białego, to nie wystarczy, by rozdzielczość kwantyzacji była odpowiednio wysoka. Musi być ona również właściwie dobrana [182].

Szумы ARW i VRW

Inne szумы, które mogą zawierać pomiary z czujników IMU oraz MARG, to szумы typu ARW i VRW. W szumach tych dominują składowe o wysokich częstotliwościach, co oznacza m.in., że czas korelacji tych szumów jest bardzo krótki. Również w przypadku szumów ARW i VRW możemy posłużyć się wariacją Allana [256, 120] do ich opisu. Odpowiednia zależność ma postać:

$$\sigma_{arw}^2(T) = \frac{Q_{arw}^2}{T}. \quad (4.28)$$



Rys. 4.12. Szum ARW (VRW) wyrażony wariacją Allana

Na rysunku 4.12 przedstawiono przykładowy wykres ilustrujący wariację Allana dla szumów ARW i VRW. W tym przypadku nachylenie prostej na wykresie *log-log* wynosi $-1/2$ (rys. 4.12). Natomiast wartość parametru σ_{arw}^2 dla tego szumu może zostać wyznaczona bezpośrednio z wykresu dla $T = 1$.

Szum BI

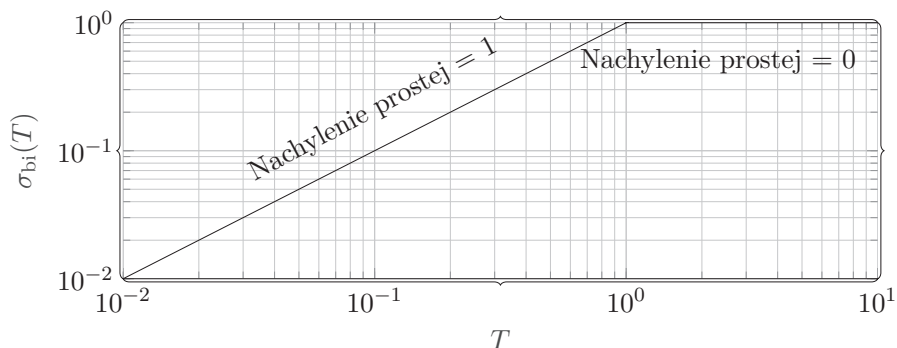
Głównymi źródłami generującymi szum tego typu są przepływ prądu w obwodach elektronicznych oraz drgania wywołane przez inne elementy wchodzące w skład czujników pomiarowych [256, 120]. Szum ten charakteryzuje się tym, że dominują w nim składowe o niskich częstotliwościach, które przyczyniają się do powstawania fluktuacji błędu systematycznego. Szum BI generowany jest w wyniku działania pola elektrycznego. Jest to tzw. szum migotliwy. Scharakteryzować ten szum można, bazując na wariacji Allana. Odpowiednia zależność ma następującą postać:

$$\sigma_{bi}^2(T) = \frac{2Q_{bi}^2}{\pi} \left[\ln 2 - \frac{\sin(\pi f_0 T)^3}{2(\pi f_0 T)^2} \left(\sin(\pi f_0 T) + \right. \right. \\ \left. \left. + 4\pi f_0 T \cos(\pi f_0 T) + C_{bi}(2\pi f_0 T) - C_{bi}(4\pi f_0 T) \right) \right] \quad (4.29)$$

gdzie f_0 jest częstotliwością odcięcia, Q_{bi} to współczynnik charakteryzujący szum typu BI, C_{bi} oznacza funkcję kosinus całkowity.

Na rysunku 4.13 przedstawiono wykres wariacji Allana dla szumu BI. Wykres ten różni się od dwóch wcześniej omówionych, ponieważ składa się z dwóch prostych. Jedna z nich na wykresie *log-log* ma nachylenie 1, a druga ma nachylenie równe 0.

4. Czujniki pomiarowe



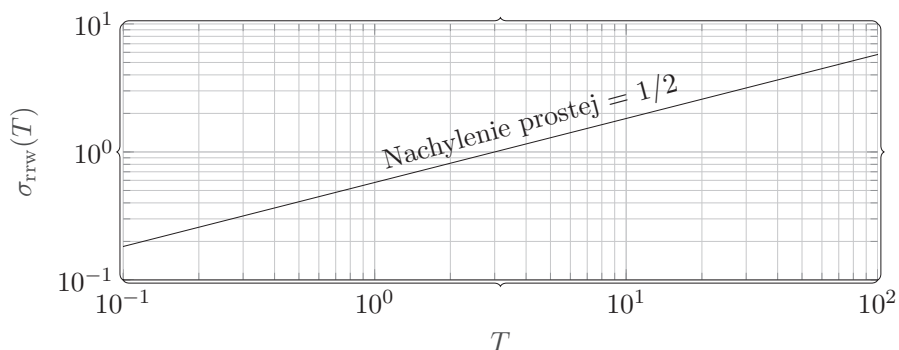
Rys. 4.13. Szum BI wyrażony wariancją Allana

Szum RRW

Z kolei szum RRW jest szumem, w którym przeważają składowe o niskich częstotliwościach. W czujnikach typu IMU oraz MARG szum ten może mieć niekorzystny wpływ na pomiary przyspieszenia (ang. *Velocity Rate Random Walk*, VRRW) i prędkości kątowej (ang. *Angular Rate Random Walk*, ARRW). Bazując na wariancji Allana, szum ten można opisać następującą zależnością:

$$\sigma_{rrw}^2(T) = \frac{Q_{rrw}^2 T}{3} \quad (4.30)$$

gdzie Q_{rrw} jest współczynnikiem charakterystycznym dla szumu typu RRW.



Rys. 4.14. Szum RRW wyrażony wariancją Allana

4.5. Metody analizy zakłóceń w czujnikach pomiarowych

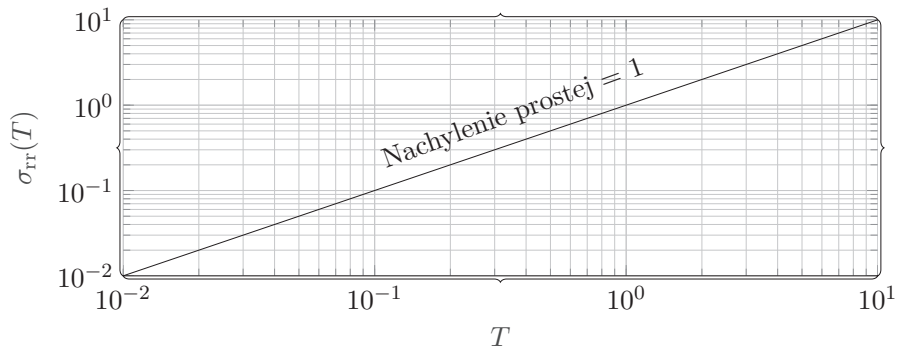
Wykres wariancji Allana dla szumu RRW zilustrowano na rys. 4.14. Wykres w skali *log-log* dla szumu tego typu ma nachylenie 1/2. Wykres ten umożliwia również odczyt amplitudy szumu RRW dla $T = 3$.

Szum RR

Szum RR związany z błędem systematycznym wykonanych pomiarów. Jedną z przyczyn powstawania błędów systematycznych jest niedokładność czujnika pomiarowego. Szum RR również może zostać scharakteryzowany na bazie wariancji Allana z wykorzystaniem następującej formuły:

$$\sigma_{rr}^2(T) = \frac{Q_{rr}^2 T^2}{2} \quad (4.31)$$

gdzie Q_{rr} jest współczynnikiem właściwym dla szumu typu RR.



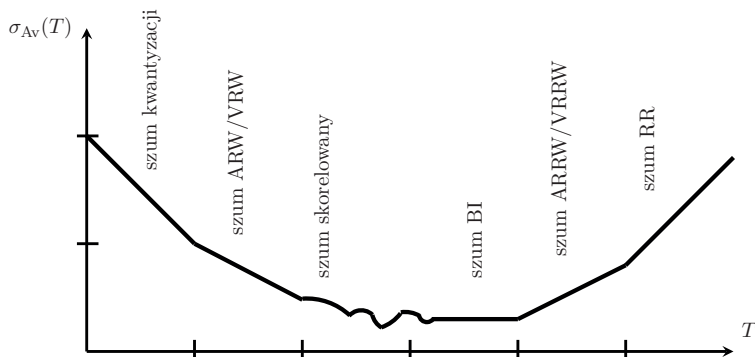
Rys. 4.15. Szum RR wyrażony wariancją Allana

Na wykresie (rys. 4.15) w skali *log-log* szum RR reprezentowany jest przez prostą o nachyleniu 1. Z kolei wartość amplitudy szumu RR wyznaczana jest dla $T = \sqrt{2}$ dla wykresu w skali *log-log*.

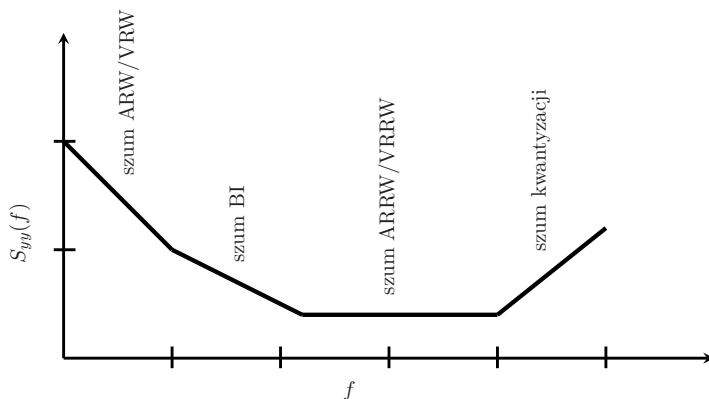
4.5.2. Podsumowanie metod modelowania zakłóceń w czujnikach pomiarowych

Analiza szumów pomiarowych z wykorzystaniem takich narzędzi, jak wariancja Allana nie jest zadaniem trywialnym. W przypadku, gdyby pomiary zawierały tylko jeden rodzaj szumu, wówczas wykresy je reprezentujące miałyby postać taką jak ta, która została przedstawiona na rys. 4.11–4.15. W rzeczywistości analizowane pomiary mogą być superpozycją wielu różnych szumów, co sprawi, że i wykres

4. Czujniki pomiarowe



Rys. 4.16. Ilustracja błędów opisanych wariancją Allana [174]



Rys. 4.17. Ilustracja błędów opisanych gęstością widmową mocy [174]

końcowy będzie różny od tych, które omówiono dla każdego z rodzajów szumów. Na rysunku 4.16 przedstawiono wykres, który jest wypadkową dla różnych typów szumów, które mogą pojawić się w pomiarach. Z wykresu wynika, że możliwe jest odróżnienie typu szumów na podstawie wykresu wariancji Allana, który utworzono na bazie odpowiednich sygnałów pomiarowych.

Z kolei jednoznaczne odróżnienie typów szumów może być przydatne w ocenie tego, jakie szумы występują w analizowanych pomiarach, oraz które z nich w nich dominują. Wykres ten może również posłużyć do oceny jakości algorytmów służących do usuwania z sygnałów pomiarowych niepożądanych składowych.

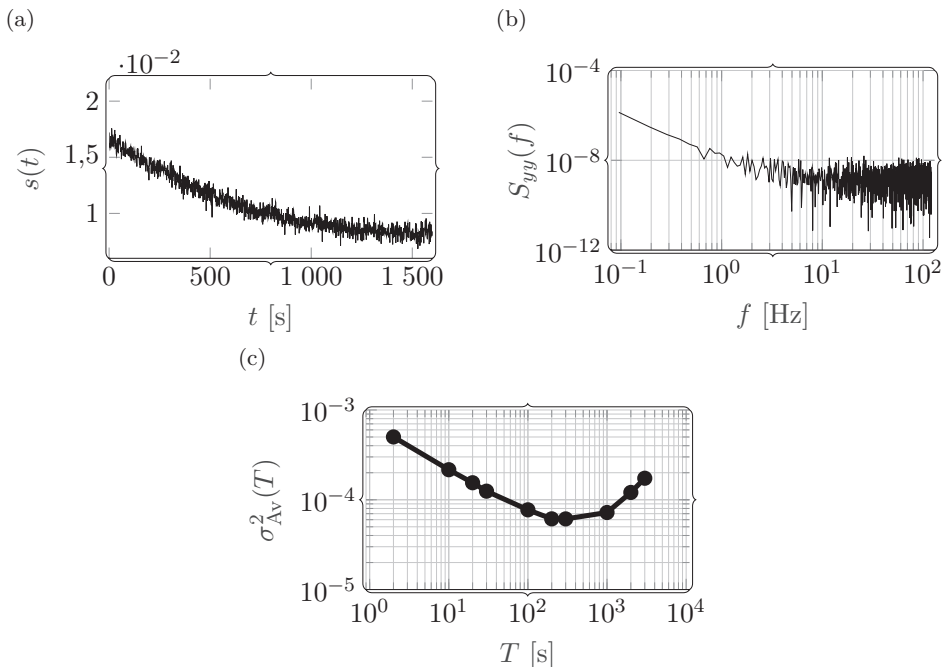
Dla porównania na rys. 4.17 przedstawiono wykres przedstawiający zależność pomiędzy gęstością widmową mocy a częstotliwością. Porównując wykresy za-

uważamy, że analiza szumów w pomiarach z wykorzystaniem wariacji Allana wykazała znacznie bardziej precyzyjne określenie, jakiego typu szumy zawarte są w pomiarach. W przypadku gęstości widmowej mocy informacja o szumach jest znacznie skromniejsza.

4.5.3. Zastosowanie wariacji Allana oraz gęstości widmowej mocy do analizy pomiarów z czujników inercyjnych i pola magnetycznego

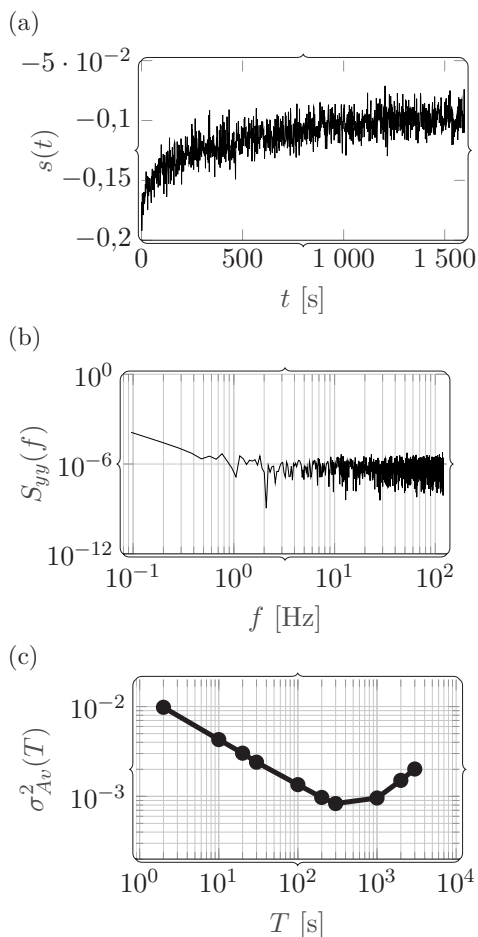
Na koniec tego rozdziału przedstawimy przykłady wykorzystania omówionej metody bazującej na wariacji Allana do analizy sygnałów z czujników wykonywanych w technologii MEMS. W badaniach wykorzystaliśmy dane z czujników przyspieszenia (rys. 4.18), żyroskopowych (rys. 4.19) oraz magnetometrów (rys. 4.20). Dla porównania przedstawiono również analogiczną analizę z wykorzystaniem gęstości widmowej mocy. Wyniki analizy danych pomiarowych prezentujemy na wykresach zamieszczonych na rys. 4.18–4.20.

Na wykresach reprezentujących wariacje Allana dla pomiarów żyroskopowych (dla każdej z osi) obserwujemy prostą o nachyleniu $-1/2$, $1/2$, 1 oraz krótki



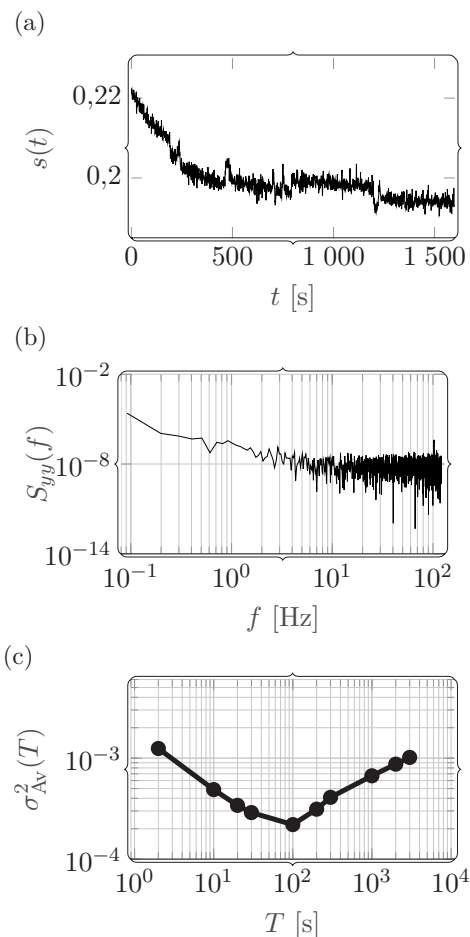
Rys. 4.18. Analiza pomiaru z czujnika żyroskopowego dla osi x :
 a) pomiar, b) widmowa gęstość mocy, c) wariacja Allana

4. Czujniki pomiarowe



Rys. 4.19. Analiza pomiaru z czujnika przyśpieszenia dla osi x :

- a) pomiar, b) widmowa gęstość mocy, c) wariancja Allana



Rys. 4.20. Analiza pomiaru z czujnika pola magnetycznego dla osi x :

- a) pomiar, b) widmowa gęstość mocy, c) wariancja Allana

odcinek prostej o nachyleniu 0. Wynika z tego, że w pomiarach z żyroskopu obecne są szumy BI, ARRW/VRRW, RR oraz szum kwantyzacji. Warto zauważyć, że przeprowadzenie podobnej analizy dla wykresów gęstości widmowej mocy jest znacznie trudniejsze, w porównaniu do tej wykonanej na bazie wariancji Allana.

Podobną analizę można przeprowadzić dla pomiarów przyśpieszenia i natężenia pola magnetycznego. Analiza ta umożliwi określenie typów szumów w danych pomiarowych, co ułatwi dobór odpowiednich metod przetwarzania danych pomia-

4.5. Metody analizy zakłóceń w czujnikach pomiarowych

rowych (np. do modelowania i usuwania niepożądanych składowych z sygnałów pomiarowych). W tej części pracy skupiliśmy się na modelowaniu pomiarów z czujników typu IMU lub MARG. Jednakże warto dodać, że obszar zastosowań wariacji Allana nie ogranicza się tylko do danych pomiarowych pochodzących z tych czujników. Podejście to jest również stosowane do modelowania szumów w innych typach czujników pomiarowych [120].

5. Przetwarzanie wstępne

Doskonałość jest osiągnięta nie wtedy, gdy nie można nic już dodać, ale wtedy, gdy nie można już nic zabrać.

Antoine de Saint-Exupéry

5.1. Wstęp

W rozdziale omówiono podstawowe zagadnienia związane ze wstępnym przetwarzaniem danych pomiarowych. Typowe zadania, które należy rozważyć na tym etapie to:

- normalizacja i standaryzacja,
- usuwanie składowych o niskich częstotliwościach (ang. *detrending*, *high-pass filtering*),
- usuwanie składowych o wysokich częstotliwościach (ang. *denoising*, *low-pass filtering*).

Wymienione zadania, będące częścią procesu wstępnego przetwarzania danych pomiarowych, są niezbędnymi składowymi tego procesu. Zaznaczmy, że problematyka wstępnego przetwarzania danych pomiarowych ma duże znaczenie praktyczne. Właściwie przeprowadzone wstępne przetwarzanie danych pomiarowych wpływa na jakość uzyskanych wyników.

Oprócz zagadnień związanych z normalizacją i standaryzacją danych pomiarowych większość tego rozdziału zostanie poświęcona metodom rozwiązania zadania usuwania składowych o niskich i wysokich częstotliwościach z danych pomiarowych. W rozdziale tym szczególną uwagę skupiono na metodach usuwania niepożądanych składowych, które wykorzystują modelowanie matematyczne. Dokonałiśmy przeglądu podejść zarówno takich, które są przeznaczone dla problemów liniowych, jak i dla problemów nieliniowych. Część tego rozdziału została również poświęcona na omówienie metod, które są przeznaczone dla sygnałów niestacjonarnych.

W rozdziale tym omówiono nie tylko klasyczne metody, które usuwają niepożądane składowe z sygnałów po ich analizie w dziedzinie czasu czy częstotliwości. Rozwa-

5. Przetwarzanie wstępne

żyliśmy również podejścia wykorzystujące reprezentację czasowo-częstotliwościową. Omówiliśmy również podstawowe zagadnienia związane z nowym podejściem do analizy sygnałów, jak i do usuwania niepożądanych składowych, które wykorzystują rzadką reprezentację sygnału (ang. *sparse signal representation*).

5.2. Normalizacja i standaryzacja sygnałów pomiarowych

Sygnały pomiarowe pochodzące z rozmaitych czujników mogą przyjmować wartości z różnych zakresów, co może mieć niekorzystny wpływ na wyniki ich przetwarzania. Dlatego w wielu przypadkach może okazać się, że konieczna będzie normalizacja lub standaryzacja danych pomiarowych. Normalizacja i standaryzacja należą do dwóch podstawowych operacji ujednociających. Rozpocznijmy nasze rozważania od operacji **normalizacji**. Normalizacja to przeskalowanie sygnałów pomiarowych z zakresu oryginalnego do zakresu docelowego. Zwykle po wykonaniu tej operacji, pomiary przyjmują wartości z zakresów $\langle 0, 1 \rangle$ lub $\langle -1, 1 \rangle$. Opisaną transformację danych pomiarowych wykonuje się zgodnie ze wzorem:

$$y_{\text{norm}}(k) = \frac{y(k) - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}, \quad k = 1, 2, \dots, K \quad (5.1)$$

gdzie y_{max} jest maksymalną wartością występującą w zbiorze pomiarowym, y_{min} reprezentuje z kolei najmniejszą wartość w tym zbiorze. Normalizacja danych pomiarowych jest niezbędna w wielu metodach przetwarzania danych, ponieważ wiele z tych metod wymaga, by dane te mieściły się w pewnym zakresie. Główną wadą tego podejścia jest to, że wyniki uzyskiwane z jej wykorzystaniem, w przypadku gdy w pomiarach występują wartości odstające, mogą okazać się niezadowolające. Niska jakość danych po znormalizowaniu może wpłynąć na pogorszenie efektów ich przetwarzania na kolejnych etapach.

Druga z wymienionych operacji transformujących to **standaryzacja**. Standaryzacja polega na przeskalowaniu wartości wszystkich próbek sygnału pomiarowego względem wartości średniej oraz odchylenia standardowego zgodnie z następującą zależnością:

$$y_{\text{stand}}(k) = \frac{y(k) - \hat{y}}{y_{\sigma}}, \quad \text{dla } k = 1, 2, \dots, K \quad (5.2)$$

gdzie

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K y(k) \quad (5.3a)$$

$$y_\sigma = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (y(k) - \hat{y})^2} \quad (5.3b)$$

W wyniku standaryzacji sygnałów pomiarowych wartość średnia sygnału pomiarowego będzie wynosić 0, a odchylenie standardowe będzie równe 1.

5.3. Usuwanie szumów o niskich częstotliwościach

Na etapie wstępnego przetwarzania danych pomiarowych należy również usunąć z danych pomiarowych szumy. Szumy w pomiarach mogą być skupione w zakresie niskich lub wysokich częstotliwości, mogą być również ulokowane w konkretnym paśmie (np. zakłócenia związane z siecią energetyczną). Na początek rozważmy problem usuwania z pomiarów szumów z zakresu niskich częstotliwości. Przyjmijmy, że dysponujemy zakłóconymi pomiarami $y(k)$ pewnego sygnału $s(k)$:

$$y(k) = s(k) + v_L(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.4)$$

gdzie $v_L(k)$ jest pewnym zakłóceniem zawierającym sygnał ze składowymi o niskich częstotliwościach. Oznaczmy przez \mathcal{F}_L algorytm estymacji szumów, które zawierają składowe o niskich częstotliwościach:

$$\tilde{v}_L(k) = \mathcal{F}_L(y(k)), \quad \text{dla } k = 1, 2, \dots, K \quad (5.5)$$

Sformułowanie zadania

Zadanie usuwania szumów o niskich składowych sprowadza się do zadania estymacji szumu $\tilde{v}_L(k)$ takiego, w którym:

$$v_L(k) \approx \tilde{v}_L(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.6)$$

gdzie

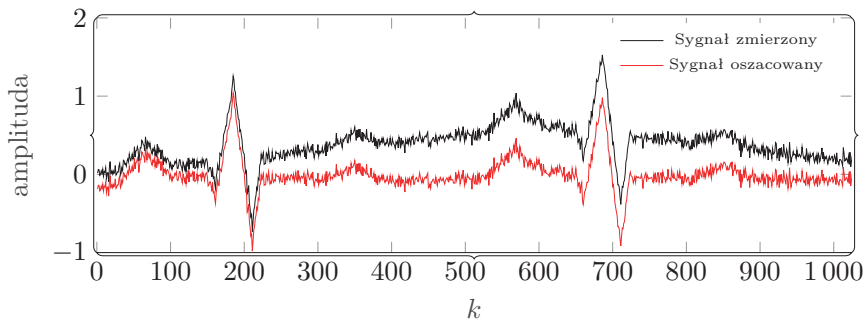
$$\tilde{v}_L(k) = \min_{\mathcal{F}_L} Q(\mathcal{F}_L(y(k))) \quad (5.7a)$$

$$Q(\mathcal{F}_L) = \sum_{k=0}^K (s(k) - \mathcal{F}_L(y(k)))^2 \quad (5.7b)$$

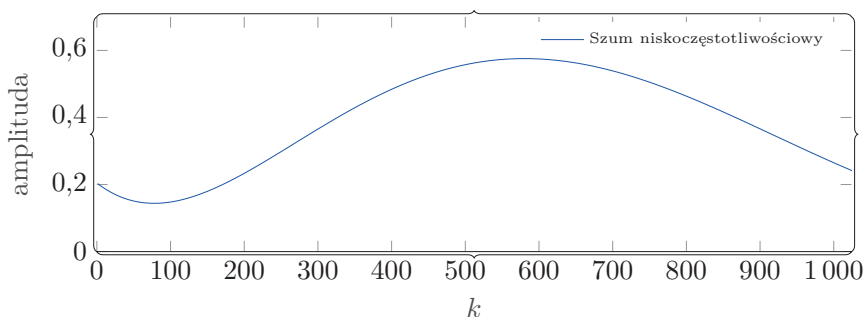
Ostatecznie, biorąc pod uwagę (5.4), usuwanie sygnału ze składowymi o niskich częstotliwościach wykonujemy przez:

$$\tilde{s}(k) = y(k) - \tilde{v}_L(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.8)$$

5. Przetwarzanie wstępne



Rys. 5.1. Ilustracja zadania usuwania szumów o niskich częstotliwościach



Rys. 5.2. Sygnał zawierający składowe o niskiej częstotliwości dla przykładu z rys. 5.1

Ilustrację zadania usuwania z sygnału pomiarowego szumu zawierającego składowe o niskich częstotliwościach przedstawiono na rys. 5.1. Z kolei wynik estymacji szumu o niskich częstotliwościach przedstawiono na rys. 5.2.

5.4. Usuwanie szumów o wysokich częstotliwościach

Rozważmy teraz przez analogię zadanie usuwania szumów o wysokich częstotliwościach. Niech $y(k)$ będą pomiarami pewnego sygnału $s(k)$:

$$y(k) = s(k) + v_H(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.9)$$

gdzie $v_H(k)$ jest szumem o wysokich częstotliwościach. Niech z kolei \mathcal{F}_H będzie algorytmem estymacji szumu wysokoczęstotliwościowego:

$$\tilde{v}_H(k) = \mathcal{F}_H(y(k)), \quad \text{dla } k = 1, 2, \dots, K \quad (5.10)$$

Sformułowanie problemu

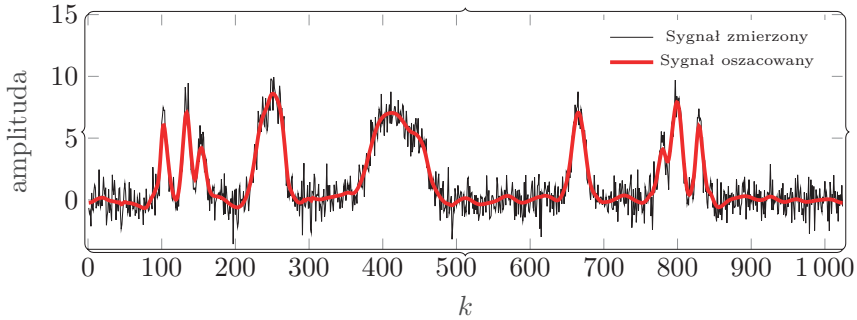
Zadanie usuwania szumów o wysokich częstotliwościach sprowadza się do zadania estymacji sygnału $\hat{v}_H(k)$ takiego, że:

$$v_H(k) \approx \tilde{v}_H(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.11)$$

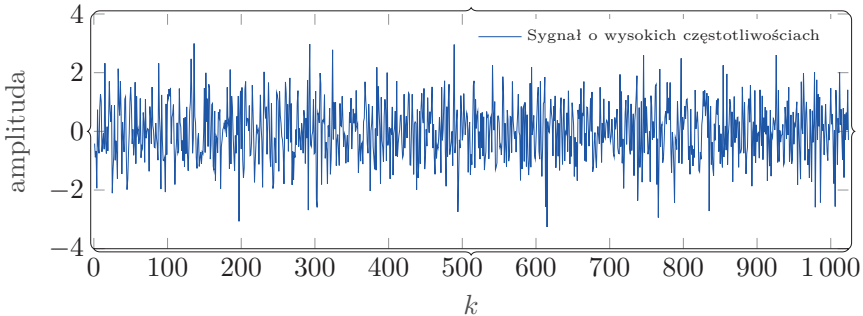
gdzie

$$\tilde{v}_H(k) = \min_{\mathcal{F}_H} Q(\mathcal{F}_H(y(k))) \quad (5.12a)$$

$$Q(\mathcal{F}_H) = \sum_{k=0}^K (s(k) - \mathcal{F}_H(y(k)))^2 \quad (5.12b)$$



Rys. 5.3. Ilustracja zadania usuwania szumów o wysokich częstotliwościach



Rys. 5.4. Sygnał zawierający składowe o wysokich częstotliwościach wykorzystany w przykładzie z rys. 5.3

5. Przetwarzanie wstępne

Ostatecznie, biorąc pod uwagę (5.4), usuwanie sygnału ze składowymi o wysokich częstotliwościach wykonujemy przez

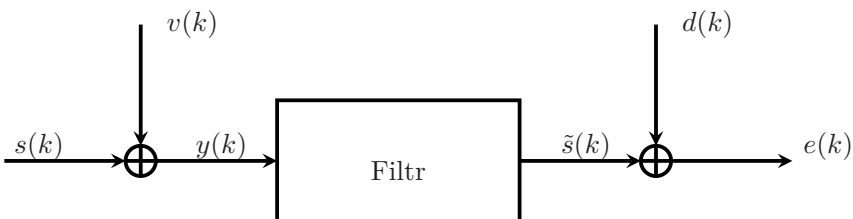
$$\tilde{s}(k) = y(k) - \tilde{v}_H(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.13)$$

Wynik usunięcia z sygnału pomiarowego szumu wysokoczęstotliwościowego przedstawiono na rys. 5.3, natomiast na rys. 5.4 przedstawiono estymowany szum, który zawarty był w danych pomiarowych.

5.5. Filtr Wienera

Po sformułowaniu zadań usuwania niepożądanych składowych o niskich i wysokich częstotliwościach omówimy teraz problem projektowania algorytmów do rozwiązywania tych zadań. Rozważania rozpoczniemy od klasy optymalnych filtrów liniowych, tj. filtrów Wienera.

Na początek rozważmy zadanie filtracji optymalnej, które przedstawiono na rys. 5.5. W filtrze tym możemy wyróżnić sygnał wejściowy $y(k)$, który jest sygnałem z szumami. Na wyjściu filtra optymalnego możemy wyróżnić sygnał estymowany $\tilde{s}(k)$, błąd filtracji $e(k)$ oraz sygnał odniesienia $d(k)$. Zadaniem filtra optymalnego jest przekształcenie sygnału pomiarowego podanego na wejście $y(k)$ na sygnał wyjściowy $\tilde{s}(k)$ w taki sposób, by sygnał estymowany możliwie najmniej różnił się od sygnału odniesienia $d(k)$, tj. błąd filtracji $e(k)$ powinien być możliwie najmniejszy.



Rys. 5.5. Ilustracja zadania estymacji optymalnej

Sformalizujmy przedstawiony opis. Na rysunku 5.5 wielkość $e(k)$ oznacza błąd oszacowania, który wyznaczamy z zależności:

$$e(k) = d(k) - \tilde{s}(k), \quad \text{dla } k = 1, 2, \dots, K \quad (5.14)$$

Zadanie filtracji optymalnej polega na estymacji sygnału $s(k)$ na podstawie zaszumionych obserwacji $y(k)$ tak, że:

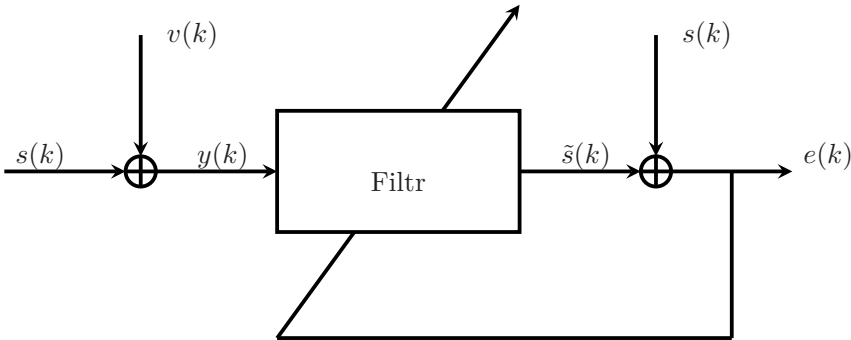
$$\tilde{s}(k) = \min_{\mathcal{F}} Q(\mathcal{F}), \quad \text{dla } k = 1, 2, \dots, K \quad (5.15)$$

gdzie

$$Q(\mathcal{F}) = \sum_{k=1}^K (d(k) - \mathcal{F}(y(k)))^2 = \sum_{k=1}^K (d(k) - \tilde{s}(k))^2 = \sum_{k=1}^K (e(k))^2 \quad (5.16)$$

Cechą filtra optymalnego jest to, że jego parametry raz dobrane nie mogą być zmieniane. W praktyce takie rozwiązanie ma swoje ograniczenia, które związane są ze specyfiką rzeczywistych pomiarów. Alternatywne sformułowanie zadania filtracji bazuje na filtracji adaptacyjnej. Zadanie filtracji optymalnej formułuje się tak, że parametry filtra nie są dobrane raz, ale mogą się zmieniać w czasie. Na bieżąco wyznaczone parametry decydują o charakterystyce filtra i jego skuteczności. Zauważmy, że w przypadku, gdy sygnały $y(k)$ oraz $d(k)$ są sygnałami o niezmiennych w czasie charakterystykach (tj. sygnałami stacjonarnymi), to oczywiście raz wyznaczone parametry pozostają stałe w czasie.

Przykładem filtra adaptacyjnego jest filtr Wienera. W filtrze Wienera podobnie jak w filtrze optymalnym wyróżniamy sygnał odniesienia $d(k)$ oraz błąd filtracji $e(k)$. Celem projektowym dla filtra Wienera jest takie dobranie jego parametrów, by błąd filtracji był możliwie najmniejszy. Zadanie to zostało przedstawione na ogólnym schemacie filtra Wienera na rys. 5.6.



Rys. 5.6. Filtr Wienera

Zaznaczmy, że dla filtra Wienera rozwiązanie jest optymalne przy spełnieniu następujących warunków:

- sygnał wyjściowy jest liniową kombinacją sygnałów wejściowych,
- dane wejściowe są sygnałami stacjonarnymi.

5. Przetwarzanie wstępne

Projektowanie filtra Wienera wymaga podjęcia kilku decyzji projektowych. Dwie najważniejsze z nich to wybór kryterium oceny jakości filtracji oraz czas trwania odpowiedzi filtra. Kryterium oceny jakości może bazować np. na kryterium średniokwadratowym (ang. *Mean-Square Error*, MSE) lub na wartości oczekiwanej wartości bezwzględnej błędu filtracji (ang. *expectation of the absolute value of the estimation error*) [158].

W przypadku wyboru kryterium średniokwadratowego podczas projektowania filtra Wienera, problem estymacji sprowadza się do następującego zadania optymalizacji:

$$\tilde{s}(k) = \min_{\mathcal{F}} Q(\mathcal{F}) \quad (5.17)$$

gdzie \mathcal{F} jest filtrem o poszukiwanych parametrach, a $Q(\mathcal{F})$ jest wskaźnikiem jakości w postaci:

$$Q(\mathcal{F}) = E \left[(d(k) - \tilde{s}(k))^2 \right] = E \left[e^2(k) \right], \quad \text{dla } k = 1, 2, \dots, K \quad (5.18)$$

gdzie $\hat{s}(k)$ to oszacowanie nieznanego sygnału, który będzie zbliżony do sygnału referencyjnego $d(k)$.

Z zależności (5.18) wynika, że jakość filtra Wienera zależy od właściwości przetwarzanego sygnału. Jeśli te właściwości będą niezmiennie w czasie, to również parametry tego filtra nie będą zmieniały się w czasie [484].

Oprócz wyboru kryterium oceny jakości, należy również ustalić czas trwania odpowiedzi filtra. Decyzja ta wiąże się z wyborem struktury modelu. Dwie podstawowe struktury, które można wykorzystać do konstrukcji filtra Wienera, to model o:

- skończonej odpowiedzi impulsowej (ang. *Finite Impulse Response*, FIR),
- nieskończonej odpowiedzi impulsowej (ang. *Infinite Impulse Response*, IIR).

Model o skończonej odpowiedzi impulsowej

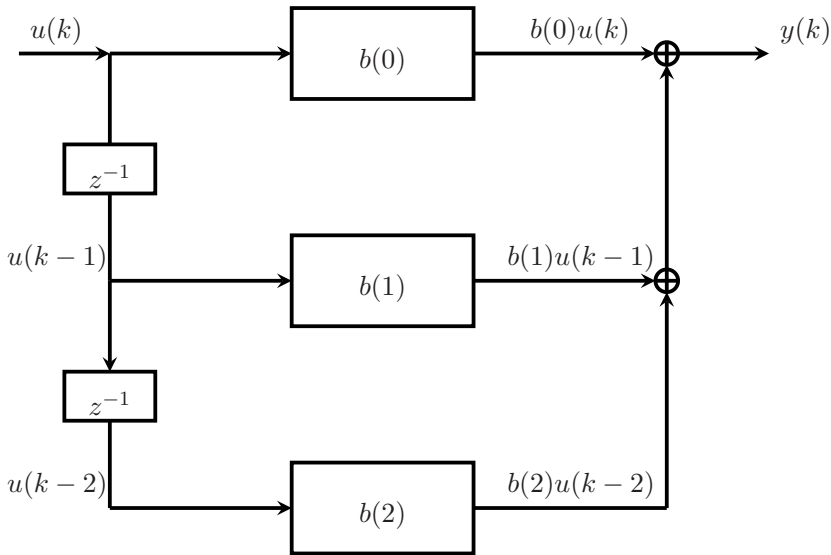
Model o skończonej odpowiedzi impulsowej jest modelem bez sprzężenia zwrotnego, a każda próbka sygnału wyjściowego jest średnią ważoną od kilku do kilkudziesięciu próbek sygnału wejściowego. Zaletami filtrów FIR jest prosta budowa, która przekłada się na prostotę ich projektowania. Ważne jest również to, że filtry te są zawsze stabilne i nie zniekształcają przetwarzanego sygnału [484].

Wadą filtrów FIR jest z kolei to, że ich złożoność obliczeniowa jest bardzo duża. Wydajne filtry o skończonej odpowiedzi impulsowej wymagają dużej liczby parametrów, co przekłada się na liczbę wymaganych obliczeń. Warto jednak zaznaczyć, że postęp w obszarze układów elektronicznych sprawia, że zdolności obliczeniowe współczesnych jednostek są bardzo duże, co powoduje, że konstruktorzy wracają do filtrów FIR, pomimo ich dużej złożoności obliczeniowej.

Ogólny model struktury FIR zapisujemy w formie następującej zależności:

$$y(k) = b_0u(k) + b_1u(k-1) + \dots + b_{M-1}u(k-M) = \sum_{m=0}^M b_mu(k-m) \quad (5.19)$$

gdzie $u(k)$ jest sygnałem wejściowym, $y(k)$ jest sygnałem wyjściowym, b_m to zestaw parametrów modelu, a M określa liczbę tych parametrów. Przykładową strukturę modelu FIR przedstawiono na rys. 5.7.



Rys. 5.7. Struktura modelu o skończonej odpowiedzi impulsowej

Struktura FIR charakteryzuje się tym, że M przeszłych próbek wraz z próbką bieżącą jest sumowana z pewnymi wagami (tutaj b_m). Cechą charakterystyczną modelu o skończonej odpowiedzi impulsowej jest to, że dla pobudzenia impulsowego, (tj. deltą Diraca) $u(k) = \delta(k)$, jego odpowiedź jest zbieżna.

Rozważmy jeszcze jeden przypadek. Jeżeli na wejście FIR podamy sekwencje impulsów, których odległości są niewielkie, to odpowiedzi modelu będą na siebie nachodzić. Dochodzi wówczas do konwolucji (splotu) odpowiedzi impulsowej FIR z sygnałem wejściowym.

Model o nieskończonej odpowiedzi impulsowej

Przeciwnie niż to było w przypadku modelu FIR, w modelu o nieskończonej odpowiedzi impulsowej do szacowania bieżącej wartości wyjściowej wykorzystuje

5. Przetwarzanie wstępne

się nie tylko bieżące wartości wejściowe, ale również przeszłe wartości wyjściowe. Cechą charakterystyczną tej struktury jest to, że ich częścią jest sprzężenie zwrotne oraz linia opóźniająca na ich wyjściu. Zaletą modeli IIR jest to, że filtry bazujące na tej strukturze mają bardzo dobre (tj. strome) charakterystyki amplitudowo-częstotliwościowe o stosunkowo niedużej liczbie parametrów wykorzystanych do ich budowy [484]. Wynika z tego, że złożoność tego modelu może być znacznie mniejsza niż jest w przypadku filtra FIR.

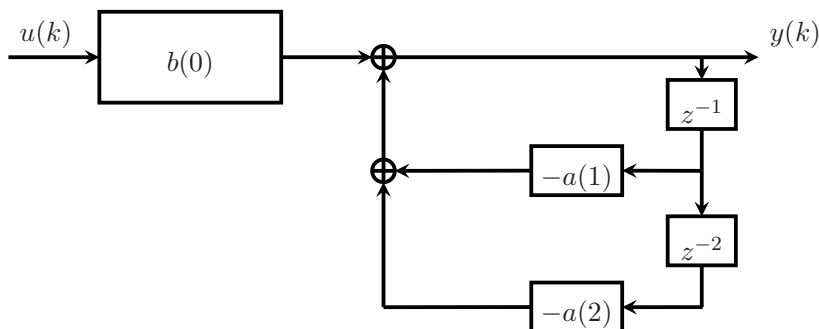
Przedstawmy teraz ogólne równanie dla struktury IIR [258]:

$$\begin{aligned} \bar{y}(k) &= b_0 u(k) - a_0 y(k-1) - a_1 y(k-2) - \dots - a_M y(k-M-1) = \\ &= b_0 u(k) - \sum_{m=0}^M a_m y(k-m-1) \end{aligned} \quad (5.20)$$

gdzie $u(k)$ reprezentuje sygnał wejściowy, $y(k)$ jest sygnałem wyjściowym, b_0 oraz a_0, a_1, \dots, a_M to zestaw parametrów modelu. Zmienna M określa liczbę parametrów a_m . Na rysunku 5.8 przedstawiono przykładowy schemat dla struktury IIR.

To, który model spośród omówionych wybierzemy do budowy filtra Wienera, zależy głównie od względów praktycznych [158]. W modelach typu FIR do wyznaczenia wartości sygnału w chwili $y(k)$ wykorzystuje się jedynie bieżącą próbkę sygnału wejściowego oraz próbki poprzednie tego sygnału. Warto zauważyć, że dysponując skończonym zbiorem próbek sygnału wejściowego, zawsze możemy wyznaczyć sygnał wyjściowy.

Wspomnieliśmy już o tym, że jedną z cech odróżniających model IIR od modelu FIR jest występowanie w tym pierwszym sprzężenia zwrotnego. Oznacza to, że w przypadku modelu IIR wartość wyjściowa jest zależna nie tylko od sygnału wejściowego



Rys. 5.8. Struktura modelu o nieskończonej odpowiedzi impulsowej

wego, ale również od poprzednich wartości sygnału wyjściowego. O takim modelu mówimy, że ma pamięć. Zwykle struktury z pamięcią przewyższają te, które pamięci nie mają. Warto jednak zauważyć, że w pewnych przypadkach pamięć modelu może mieć niekorzystny wpływ na jakość filtra bazującego na takiej strukturze.

W rozwiązaniach bazujących na modelu IIR mogą pojawić się niekorzystne efekty związane z przetwarzaniem przeszłych wartości wyjściowych. Takim niekorzystnym efektem jest możliwość pojawienia się na wyjściu takiego modelu sygnału oscylacyjnego [258]. Jest to niekorzystna cecha modeli FIR, która musi być brana pod uwagę podczas projektowania filtrów bazujących na tej strukturze. Również sprzężenie zwrotne, które jest charakterystyczne dla filtrów IIR, może sprawić, że układy te będą niestabilne, tj. będą się wzbudzać [484].

Inną wadą tego rozwiązania jest to, że filtry IIR budowane na bazie modelu o nieskończonej odpowiedzi impulsowej wprowadzają zniekształcenia do przetwarzanego sygnału. Zniekształcenia te związane są z tym, że w filtrach tych czas opóźnienia poszczególnych jego częstotliwości na ich wyjściach jest różny [484].

Zaletą tego modelu jest natomiast to, że do uzyskania sygnału o pożądanym cechach wymagana jest zazwyczaj mniejsza liczba parametrów w porównaniu do modelu typu FIR. Z tej przyczyny, pomimo wad struktury typu IIR, jest ona często brana pod uwagę na etapie projektowania filtrów.

Przedstawimy teraz dwie wersje filtra Wienera, z których jedna została oparta na modelu typu FIR, a druga na modelu typu IIR.

Filtr Wienera o skończonej odpowiedzi impulsowej

Koncepcję filtra Wienera można wykorzystać do konstrukcji adaptacyjnego filtra Wienera o skończonej odpowiedzi impulsowej, którego wagi b_m będą dobierane tak, by minimalizować błąd średniokwadratowy (5.18). Niech błąd filtracji będzie wyznaczany z zależności:

$$e(k) = y(k) - \sum_{m=0}^M b_m y(k-m) \quad (5.21)$$

gdzie b_m są składowymi wektora $\mathbf{b} = [b_0 \ b_1 \ \dots \ b_M]^T$. Parametry modelu FIR (5.19) można wyznaczyć, korzystając z numerycznych metod optymalizacji. W tym celu należy ustalić wskaźnik jakości estymacji. Na przykład wskaźnik jakości może mieć następującą postać:

$$Q(\mathbf{b}) = \sum_{k=1}^K e^2(k) \quad (5.22)$$

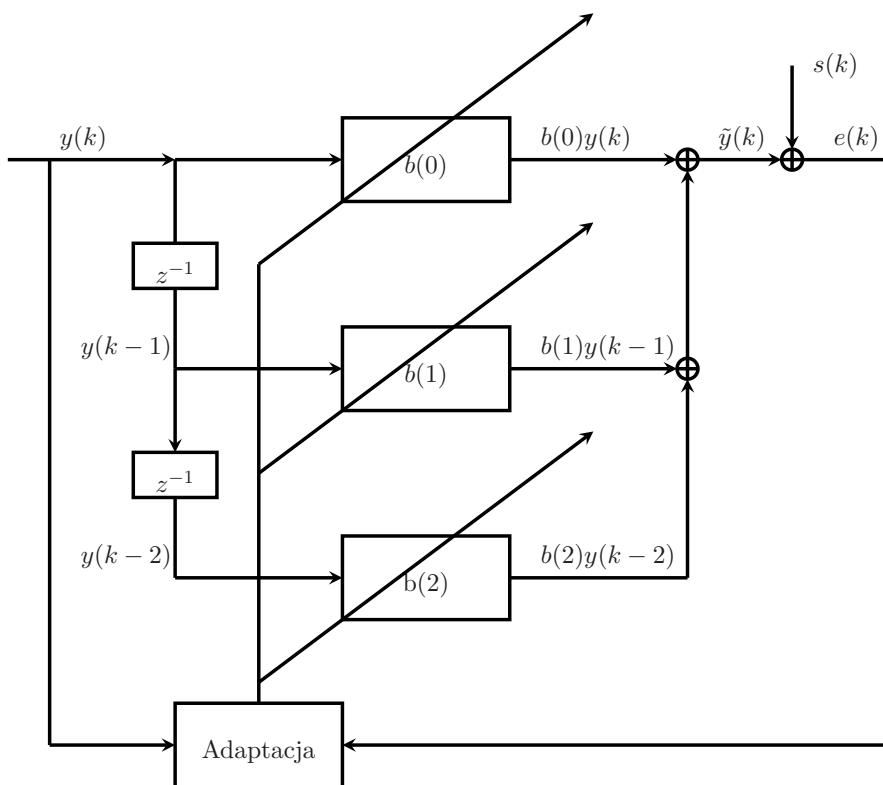
gdzie $e(k)$ jest błędem filtracji zdefiniowanym jak w (5.21). Numeryczne metody optymalizacji to jedno z możliwych podejść, które można wykorzystać do projek-

5. Przetwarzanie wstępne

towania filtrów Wienera. W niektórych przypadkach jest to konieczność, ponieważ nie zawsze znane są funkcje rozkładu gęstości prawdopodobieństwa obserwowanych zmiennych losowych, które muszą być znane podczas projektowania algorytmów dla filtra Wienera [484]. Numeryczna formuła do wyznaczania wartości wektora wag filtra FIR ma następującą ogólną postać:

$$\mathbf{b}^{(i+1)} = \mathbf{b}^{(i)} + \eta \nabla_{\mathbf{b}} Q(\mathbf{b}^{(i)}) \quad (5.23)$$

gdzie i jest indeksem i -tej iteracji, η współczynnikiem prędkości zbieżności. Filtr adaptacyjny FIR przedstawiono na rys. 5.9.



Rys. 5.9. Filtr Wienera o skończonej odpowiedzi impulsowej

Formuła (5.23) do iteracyjnego wyznaczania parametrów filtra Wienera przedstawia ogólne podejście do ich estymacji. Dwa najważniejsze podejścia wykorzystywane do rozwiązania tego zadania nazywane są filtrami LMS (ang. *Least Mean Squares*) oraz RLS (ang. *Recursive Least Squares*). Filtr LMS charakteryzuje się

prostotą w implementacji i małą złożonością obliczeniową. Natomiast główną jego wadą jest wolna zbieżność [484].

Alternatywą dla filtra LMS jest filtr RLS. W przeciwieństwie do filtra LMS, filtr RLS jest filtrem z pamięcią, który dodatkowo charakteryzuje się dużą zbieżnością. Złożoność tego filtra jest znacznie wyższa niż w przypadku filtra LMS. Wadą tego podejścia jest z kolei to, że jest ono czułe na błędy zaokrągleń. Oznacza to, że algorytm RLS jest algorytmem potencjalnie niestabilnym numerycznie.

Filtr Wienera o nieskończonej odpowiedzi impulsowej

Analogicznie do filtra Wienera, bazującego na modelu FIR, możliwe jest zaprojektowanie filtra wykorzystującego model IIR. W tym przypadku błąd filtracji wyznaczamy z zależności:

$$e(k) = y(k) - \left(b_0 y(k) - \sum_{m=0}^M a_m \bar{y}(k - m - 1) \right) \quad (5.24)$$

Współczynniki filtra IIR zapisujemy w formie następującego wektora parametrów $\mathbf{h} = [b_0 \ a_0 \ a_1 \ \dots \ a_M]$. Na bazie zależności (5.24) możemy również zaproponować następujący wskaźnik:

$$Q(\mathbf{h}) = \sum_{k=1}^K e^2(k) \quad (5.25)$$

Również i w tym przypadku do estymacji parametrów \mathbf{h} modelu wykorzystujemy jeden ze znanych schematów numerycznej optymalizacji. Ogólna postać takiego schematu ma postać:

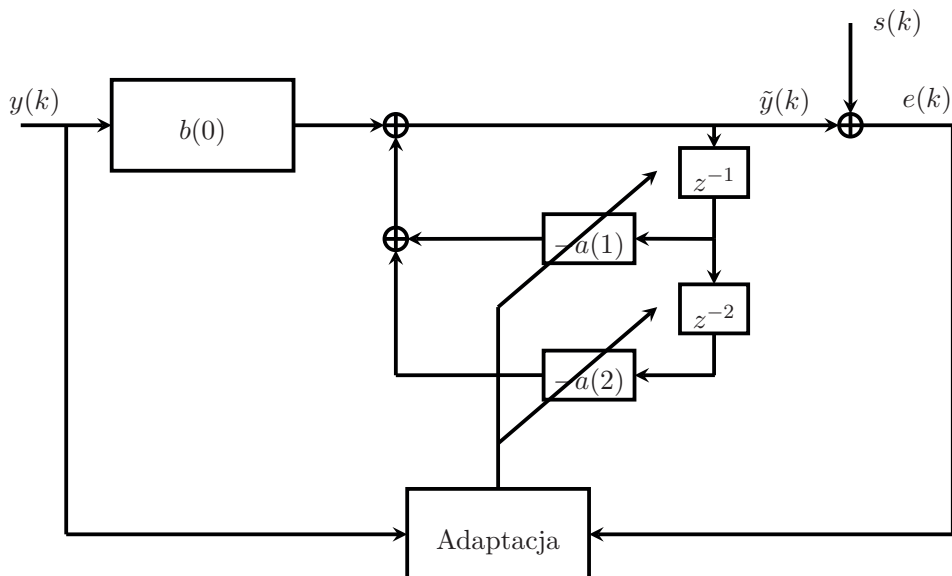
$$\mathbf{h}^{(i+1)} = \mathbf{h}^{(i)} + \eta \nabla_{\mathbf{h}} Q(\mathbf{h}^{(i)}) \quad (5.26)$$

Również w przypadku filtra Wienera, bazując na modelu IIR, możliwe jest zaprojektowanie wersji algorytmu LMS i RLS. Z kolei na rys. 5.10 przedstawiono przykładową strukturę adaptacyjną filtra Wienera wykorzystującego model IIR.

5.6. Metody przetwarzania wstępnego

W poprzednim podrozdziale omówiliśmy zagadnienia związane z filtracją adaptacyjną na przykładzie filtra Wienera. Wskazaliśmy, że filtr Wienera w przypadku, gdy przetwarzane sygnały są sygnałami stacjonarnymi i liniowymi, daje optymalne rozwiązanie zadania filtracji. W rzeczywistych zadaniach sygnały stacjonarne i liniowe nie występują, przez co zadanie projektowania filtra Wienera jest opisem

5. Przetwarzanie wstępne



Rys. 5.10. Filtr Wienera o nieskończonej odpowiedzi impulsowej

pewnej idealnej sytuacji, ponieważ w praktyce większość dostępnych sygnałów to sygnały niestacjonarne.

Opracowano wiele metod przetwarzania sygnałów pomiarowych, które są przystosowane do przetwarzania sygnałów niestacjonarnych. Niestety nie ma jednej uznanej metody, która jest rekomendowana do wszystkich typów sygnałów. Zadaniem projektanta algorytmów przetwarzania sygnałów jest dobór odpowiedniego algorytmu dostosowanego do konkretnych sygnałów pomiarowych.

W ogólności możemy wyróżnić następujące podejścia do przetwarzania sygnałów bazujące na [417]:

- transformacjach sygnałów,
- modelowaniu matematycznym,
- wnioskowaniu statystycznym,
- sztucznej inteligencji.

Podejście bazujące na **transformacjach sygnałów** charakteryzuje się tym, że na początkowym etapie konieczna jest dekompozycja sygnału pomiarowego na zbiór prostych sygnałów. W przypadku np. transformacji Fouriera zbiór ten składa się z sygnałów elementarnych (sinusów i kosinusów). W przypadku innych, np. transformacji Hilberta–Huanga, nie są to sygnały elementarne tak jak w przypadku transformacji Fouriera, ale ich przebiegi są mniej złożone w porównaniu do sygnału

oryginalnego. Oprócz wymienionych, możemy jeszcze wyróżnić transformację falkową oraz transformację Gabora.

Drugie z podejść, bazujące na **modelach matematycznych**, należy do grupy metod parametrycznych. Zaletą tego podejścia jest to, że, konstruuując model na potrzeby przetwarzania sygnałów, możliwe jest uwzględnienie dodatkowych informacji o przetwarzanych sygnałach, które mogą usprawnić proces ich przetwarzania. Wadą tego podejścia jest to, że algorytmy bazujące na modelach matematycznych wykazują dużą czułość na zmiany w przetwarzanych sygnałach [417]. Przykładem może być sytuacja, w której dokonano błędnego założenia, np. co do typu szumów w pomiarach. Jeżeli w konstrukcji algorytmu bazującego na modelu przyjęliśmy, że sygnał pomiarowy zawiera biały szum, należy się spodziewać pogorszenia jakości przetwarzania w przypadku, gdy szum będzie miał inną charakterystykę niż biały szum.

Na przetwarzanie sygnałów pomiarowych można spojrzeć również jak na zadanie **analizy procesów losowych**. Obecność szumów w sygnale sprawia, że precyzyjne ustalenie przyszłych jego wartości nie jest możliwe. Można natomiast wyznaczyć jego charakterystyki statystyczne takie, jak np. wartość średnia. W takim przypadku do przetwarzania sygnałów można wykorzystać m.in. wnioskowanie bayesowskie [355, 417]. Podejście to łączy informacje z przetwarzanego sygnału z wiedzą *a priori* dotyczącą rozkładów parametrów związanych z sygnałem pomiarowym.

Sztuczna inteligencja (SI) w ostatnich latach zyskała dużą popularność, co przelożyło się m.in. na dużą liczbę nowych algorytmów, które znalazły zastosowanie w wielu obszarach związanych również z przetwarzaniem sygnałów [179, 225, 421]. W ostatnich latach dochodzi również do prób wykorzystania metod sztucznej inteligencji w przetwarzaniu sygnałów. Stosowanie SI w przetwarzaniu sygnałów umożliwia rozwój metod nieliniowego przetwarzania sygnałów, jak również systemów adaptacyjnych czy inteligentnych. O potrzebie wykorzystania metod sztucznej inteligencji w przetwarzaniu sygnałów może również świadczyć organizowane od kilku lat pod patronatem organizacji IEEE warsztaty *IEEE Workshop on Machine Learning for Signal Processing*.

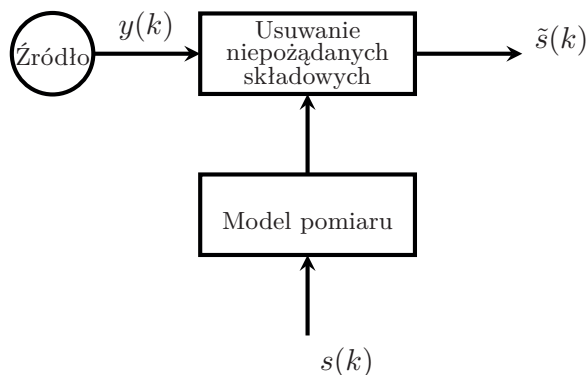
W następnym rozdziale omówimy metody przetwarzania wstępnego, które bazują na modelach. Zaproponujemy również oryginalną metodę, która łączy podejście bazujące na modelowaniu matematycznym z podejściem opartym na transformacjach sygnałów.

5.7. Podstawy przetwarzania wstępnego z modelem

Na rysunku 3.18 przedstawiono prosty schemat układu pomiarowego. Dysponując sygnałem pomiarowym, możemy, korzystając z metod modelowania matematycznego, usunąć z niego niepożądane składowe. Metody bazujące na modelowaniu możemy

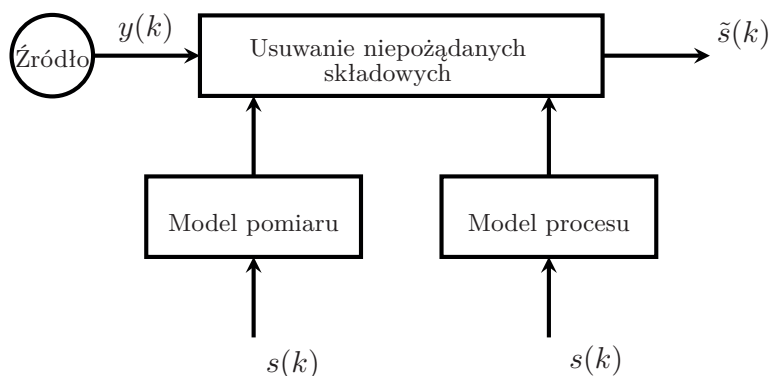
5. Przetwarzanie wstępne

podzielić na takie, w których wykorzystuje się model środowiska pomiarowego lub takie, w których wykorzystujemy dwa modele. Jeden z tych modeli to model środowiska pomiarowego, a drugi to model procesu, który generuje obserwowany sygnał. Przypadki te zostały przedstawione na rys. 5.11 oraz 5.12.



Rys. 5.11. Filtracja z modelem wykorzystująca model o środowisku pomiarowym

Uwzględniając w projektowanym algorytmie model pomiaru, możliwa jest poprawa końcowego wyniku estymacji sygnału pomiarowego $s(k)$. W tym przypadku poprawę uzyskujemy przez uzupełnienie modelu o informację związaną z wykonywanym pomiarem. Oprócz wiedzy o pomiarze, możliwe jest uwzględnienie informacji o procesie, z którym powiązany jest sygnał pomiarowy. Przypadek ten został z kolei zilustrowany na rys. 5.12.



Rys. 5.12. Filtracja z modelem wykorzystująca model o środowisku pomiarowym i model procesu

Do konstrukcji algorytmów wstępnego przetwarzania sygnałów pomiarowych mogą zostać wykorzystane zarówno modele liniowe, jak i nieliniowe. Mogą to być również modele, w których istnieje możliwość uwzględnienia niestacjonarności w analizowanych sygnałach. Oczywiście, można budować modele, które łączą w sobie kilka z wymienionych cech.

Na koniec tego podrozdziału wyróżnijmy kluczowe kroki procesu przetwarzania sygnałów, w których wykorzystywane są modele matematyczne. Pierwszy z nich dotyczy wyboru modelu, który zostanie wykorzystany do konstrukcji algorytmu przetwarzania sygnałów. Wybór powinien opierać się na wcześniej przeprowadzonej analizie sygnałów i szumów pomiarowych. Drugi etap związany jest z estymacją parametrów wybranego modelu. Zadanie estymacji rozwiązuje się z wykorzystaniem danych pomiarowych. Ostatni krok to estymacja sygnału z wykorzystaniem opracowanego modelu i na podstawie danych pomiarowych.

W kolejnych podrozdziałach dokonano przeglądu modeli, które mogą zostać zastosowane do konstrukcji algorytmów usuwania niepożądanych składowych z sygnałów pomiarowych, jak również do odtwarzania użytecznego sygnału z pomiarów.

5.7.1. Modele liniowe

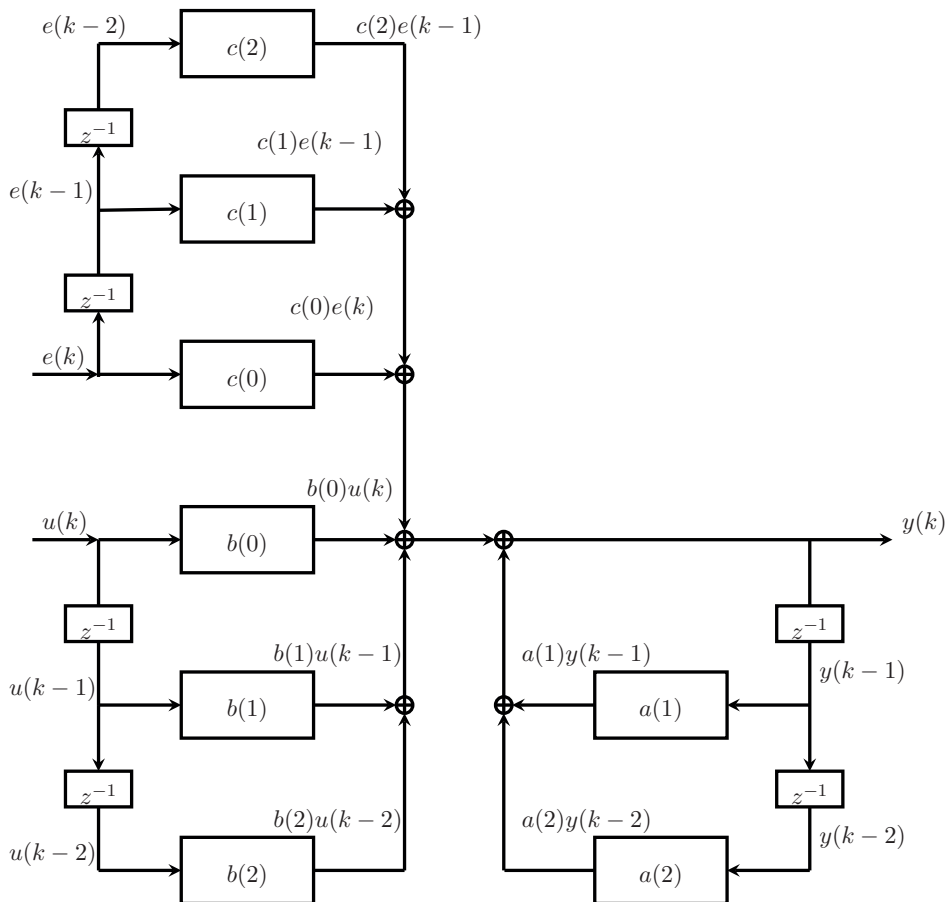
Modele z rodziny ARMAX (ang. *AutoRegressive Moving Average eXogenous model*) stanowią szeroką klasę modeli wykorzystywanych do modelowania sygnałów pomiarowych. Specjalnymi przypadkami modelu ARMAX są modele AR (ang. *AutoRegressive Model*), MA (ang. *Moving Average model*), ARMA (ang. *AutoRegressive Moving Average model*) oraz ARX (ang. *AutoRegressive eXogenous Model*). Model MA w nomenklaturze przetwarzania sygnałów nazywany jest modelem FIR (patrz podrozdz. 5.5). Natomiast model AR wykorzystywany na potrzeby przetwarzania sygnałów nazywany jest modelem IIR (patrz również podrozdz. 5.5).

Rozpocznijmy rozważania od najbardziej ogólnego z wymienionych modeli, tj. modelu ARMAX. Równanie opisujące ten model ma następującą postać:

$$\bar{y}(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_{M_a} y(k-M_a) = b_1 u(k-1) + b_2 u(k-2) + \dots + b_{M_b} u(k-M_b) + c_1 e(k-1) + c_2 e(k-2) + \dots + c_{M_c} e(k-M_c) \quad (5.27)$$

gdzie $u(k)$ jest wejściem, $y(k)$ reprezentuje wyjście, a $e(k)$ to błąd. Wielkości M_a , M_b oraz M_c określają liczbę parametrów modelu, tj. a_i , b_i i c_i . Na rysunku 5.13 przedstawiono schemat struktury ARMAX (pol. Model Autoregresywny z Ruchomą Średnią i zewnętrznym wejściem).

5. Przetwarzanie wstępne



Rys. 5.13. Schemat modelu ARMAX

Szczególnym przypadkiem modelu ARMAX jest model AR (pol. Model Autoregresywny). Zależność matematyczna służąca do opisu modelu AR ma z kolei postać:

$$\bar{y}(k) + a_1y(k - 1) + a_2y(k - 2) + \dots + a_{M_a}y(k - M_a) = e(k) \quad (5.28)$$

Szczególnym przypadkiem modelu ARMAX jest również model MA (pol. Model Ruchomej Średniej). Równanie opisujące ten model zapisujemy następująco:

$$\bar{y}(k) + c_1e(k - 1) + c_2e(k - 2) + \dots + c_{M_c}e(k - M_c) = e(k) \quad (5.29)$$

5.7. Podstawy przetwarzania wstępnego z modelem

Jeśli połączymy modele (5.28) i (5.29), to otrzymamy model ARMA (pol. Model Autoregresywny z Ruchomą Średnią), którego równanie ma postać:

$$\begin{aligned} \bar{y}(k) + a_1y(k-1) + a_2y(k-2) + \dots + a_{M_a}y(k-M_a) = \\ c_1e(k-1) + c_2e(k-2) + \dots + c_{M_c}e(k-M_c) \end{aligned} \quad (5.30)$$

Kolejnym szczególnym przypadkiem modelu ARMAX jest model ARX (pol. Model Autoregresywny z Zewnętrznym Wejściem), którego formalizm matematyczny jest następujący:

$$\begin{aligned} \bar{y}(k) + a_1y(k-1) + a_2y(k-2) + \dots + a_{M_a}y(k-M_a) = \\ = b_1u(k-1) + b_2u(k-2) + \dots + b_{M_b}u(k-M_b) + e(k) \end{aligned} \quad (5.31)$$

Mając ustalony model z rodziny ARMAX jako element składowy algorytmu przetwarzania sygnałów pomiarowych należy następnie ustalić strukturę modelu oraz wyznaczyć związane z nią parametry. W przypadku omówionej klasy modeli ustalenie struktury wiąże się z doбором odpowiednich stopni wielomianów oraz wartości opóźnień.

5.7.2. Modele odcinkowo-liniowe

Modele z rodziny ARMAX są klasycznymi modelami wykorzystywanymi w przypadku filtracji sygnałów stacjonarnych z zakłóceniami o charakterystyce białego szumu. Modele tej klasy nie pozwalają na uzyskania zadowalających wyników w przypadku, gdy filtracji podlegają sygnały pomiarowe zakłócone szumem kolorowym lub *niegaussowskim*.

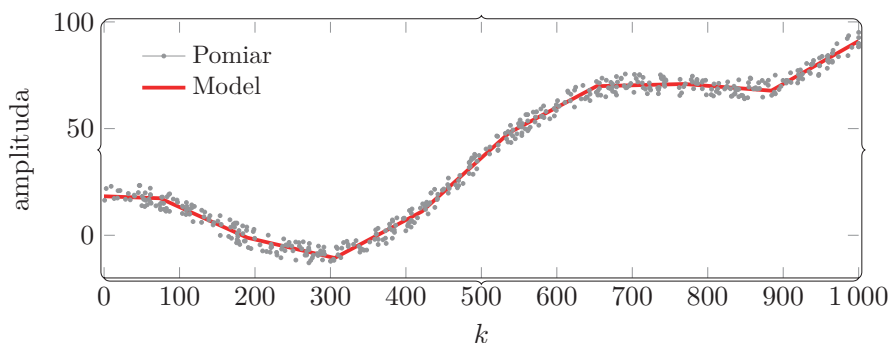
Jednym z podejść, które można zastosować do rozwiązania zadania usuwania niepożądanych składowych z pomiarów jest wykorzystanie modeli odcinkowo-liniowych. Istotą tego podejścia jest estymacja sygnałów, która oparta jest na modelach liniowych. Metoda ta ułatwia estymowanie sygnałów, dla których zastosowanie metod liniowych nie daje zadowalających wyników. Przy czym warto podkreślić, że, stosując to podejście, zgadzamy się na mniejszą dokładność odcinkowo-liniowego przybliżenia sygnału. Wadą tego podejścia jest z kolei to, że nie każdy sygnał można estymować zestawem modeli liniowych z zadowalającą dokładnością. Opracowując model odcinkowo-liniowy, należy rozwiązać kilka dodatkowych problemów. Jednym z nich jest metoda podziału sygnału na obszary, które będą przybliżane modelami liniowymi. Jest to istotne zagadnienie, ponieważ ma wpływ na końcowy efekt modelowania. Kolejną wadą prezentowanego podejścia jest to, że do uzyskania gładkich przejść pomiędzy kolejnymi regionami potrzeba wielu modeli liniowych.

5. Przetwarzanie wstępne

Ogólna postać modelu lokalnie liniowego do aproksymacji nieliniowego systemu ma następującą postać:

$$\bar{y}(k) = a_0^{(i)} + \sum_{m_a=1}^{M_a} a_{m_a}^{(i)} y(k - m_a) + \sum_{m_b=1}^{M_b} b_{m_b}^{(i)} u(k - m_b) \quad (5.32)$$

gdzie $i = 1, 2, \dots, \mathcal{P}$ są punktami pracy. Punktem pracy nazywamy punkt, w którym charakterystyka systemu nieliniowego przybliżana jest modelem liniowym. Liczba punktów pracy jest wyznaczana na podstawie danych pomiarowych i ma ona wpływ na jakość modelowania. Wadą tego podejścia jest natomiast to, że do dokładnego modelowania może być konieczne wykorzystanie dużej liczby modeli. Może mieć to również wpływ na złożoność obliczeniową algorytmu estymacji parametrów modelu. Złożoność obliczeniowa algorytmu estymacji parametrów może zależeć od tego, czy konieczne jest wyznaczenie punktów pracy, w których dane pomiarowe będą przybliżane modelami lokalnymi.



Rys. 5.14. Przykład modelowania z wykorzystaniem modelu odcinkowo-liniowego

Na rysunku 5.14 przedstawiono przykład modelowania sygnału pomiarowego z wykorzystaniem modelu odcinkowo-liniowego.

5.7.3. Szereg Volterry

Szereg Volterry jest jednym z najstarszych podejść do modelowania systemów nieliniowych [41], które z powodzeniem wykorzystuje się m.in. do modelowania sygnałów. Formalnie szereg Volterry jest uogólnieniem klasycznego systemu liniowego. Szereg Volterry buduje się z elementów, które są sumą ważoną iloczynów sygnałów wyjściowych. Elementy te nazywa się operatorami Volterry, a ich liczba określa rząd szeregu. W przypadku szeregu pierwszego rzędu model składa się

z jednego operatora Voltery i jest odpowiednikiem operacji splotu dla systemów liniowych. Równanie opisujące szereg Voltery pierwszego rzędu ma postać [176, 306]:

$$\bar{y}(k) = k_{\text{imp}0} + k_{\text{imp}1}(1)u(k-1) + k_{\text{imp}1}(2)u(k-2) + \dots + k_{\text{imp}1}(M)u(k-M) = k_{\text{imp}0} + \sum_{m=1}^M k_{\text{imp}1}(m)u(k-m) \quad (5.33)$$

gdzie $k_{\text{imp}i}$ reprezentuje odpowiedź impulsową.

Warto również zauważyć, że przedstawiony model w postaci szeregu Voltery pierwszego rzędu (5.33) ma podobną konstrukcję do modelu FIR. Dlatego model ten nazywany jest modelem NFIR (ang. *Nonlinear* FIR).

W praktycznych zadaniach częściej spotyka się szeregi Voltery wyższych rzędów. Ogólna postać równania (5.33) opisującego szereg Voltery M_q -tego stopnia można uogólnić do postaci:

$$\bar{y}(k) = k_{\text{imp}0} + \sum_{j=1}^{M_q} \left(\sum_{m_1=0}^{M-1} \dots \sum_{m_j=0}^{M-1} \mathbf{k}_{\text{imp}j}(m_1, \dots, m_j) u(k-m_1) \dots u(k-m_j) \right) \quad (5.34)$$

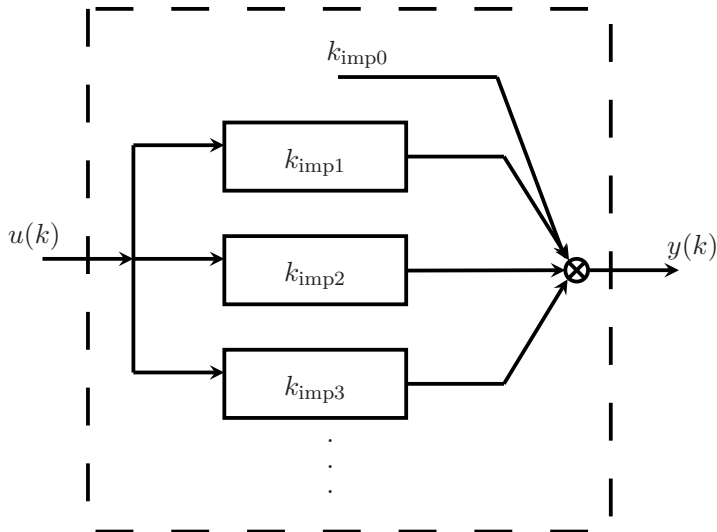
gdzie M_q jest stopniem modelu, M reprezentuje rząd szeregu Voltery, a $\mathbf{k}_{\text{imp}j}$ to wektor zawierający odpowiedzi impulsowe $k_{\text{imp}j}$ dla $j = 1, \dots, M_q$. Schemat blokowy modelu bazującego na szeregu Voltery przedstawiono na rys. 5.15.

Jedną z wad szeregu Voltery jest jego duża złożoność obliczeniowa, która wyklucza stosowanie go w niektórych praktycznych zadaniach [110]. Analizując zależność (5.34), zauważymy, że liczba wymaganych obliczeń wzrasta wraz ze wzrostem rzędu modelu. Dlatego zakres stosowalności tego modelu jest ograniczony, zwłaszcza w przypadku metod przetwarzania sygnałów projektowanych na potrzeby systemów czasu rzeczywistego.

5.7.4. Modele blokowo zorientowane

Modelowanie blokowo zorientowane to kolejne podejście do modelowania systemów nieliniowych. Modele te charakteryzują się tym, że budowane są na bazie liniowych modeli dynamicznych oraz nieliniowych modeli statycznych. Dwa podstawowe modele w tej grupie to modele Hammersteina i Wienera. W modelu Wienera jego pierwszy człon jest liniowym modelem dynamicznym, a drugi nieliniowym modelem statycznym (rys. 5.16). W przypadku modelu Hammersteina kolejność

5. Przetwarzanie wstępne



Rys. 5.15. Schemat blokowy modelu bazującego na szeregu Volterry

członów jest odwrotna, tj. nieliniowy model statyczny poprzedza liniowy model dynamiczny, który jest modelem wyjściowym (rys. 5.17).

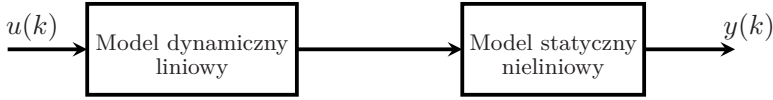
W modelach Wienera i Hammersteina oddziela się część liniową od nieliniowej, co znacząco upraszcza analizę i konstrukcję tych modeli.

Model Wienera

Schemat modelu Wienera przedstawiono na rys. 5.16. Natomiast zależność opisująca relację pomiędzy wejściem i wyjściem w tym modelu ma postać [318, 438]:

$$\begin{aligned} \bar{y}(k) &= \sum_{j=0}^{M_q} b_j \left(\sum_{m=0}^{M-1} k_{\text{imp1}}(m) u(k-m) \right)^j = \\ &= \sum_{j=0}^{M_q} b_j \left(\sum_{m_1=0}^{M-1} \cdots \sum_{m_j=0}^{M-1} k_{\text{imp1}}(m_1) \cdots k_{\text{imp1}}(m_j) u(k-m_1) \cdots u(k-m_j) \right) \end{aligned} \quad (5.35)$$

gdzie b_j jest zestawem dodatnich parametrów modelu, M_q jest stopniem modelu, M reprezentuje jego rząd, a k_{imp1} to odpowiedź impulsowa systemu dynamicznego liniowego.



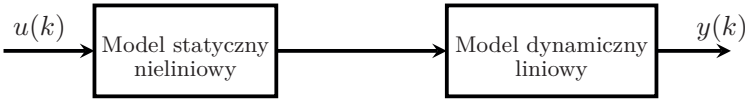
Rys. 5.16. Model Wienera

Model Hammersteina

W modelu Hammersteina kolejność bloków jest przeciwna do tej, jaka została opisana dla modelu Wienera, czyli pierwszy blok jest modelem statycznym nieliniowym, a drugi z tych bloków to model dynamiczny liniowy. Schemat blokowy tego modelu przedstawiono na rys. 5.17. Z kolei zależność (5.36) jest formalnym zapisem tego modelu [438]:

$$\bar{y}(k) = \sum_{m=0}^{M-1} k_{\text{imp1}}(m) \left(\sum_{j=0}^{M_q} b_j u^j(k-m) \right) \quad (5.36)$$

gdzie M_q jest stopniem modelu, M reprezentuje jego rząd, k_{imp1} to odpowiedź impulsowa systemu dynamicznego liniowego, a b_j jest zestawem parametrów modelu o dodatnich wartościach.



Rys. 5.17. Model Hammersteina

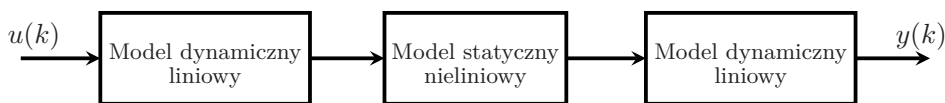
Model Wienera–Hammersteina

Szczególnym przypadkiem modelu Wienera i Hammersteina jest model Wienera–Hammersteina. Na rysunku 5.18 przedstawiono ogólną strukturę tego modelu, natomiast zależność (5.37) opisuje jego model matematyczny:

$$\bar{y}(k) = \sum_{m_2=0}^{M-1} k_{\text{imp2}}(m_2) \sum_{j=0}^{M_q} b_j \left(\sum_{m_1=0}^{M-1} k_{\text{imp1}} u(k-m_2-m_1) \right)^j \quad (5.37)$$

gdzie M_q jest stopniem modelu, M reprezentuje jego rząd, k_{imp1} to odpowiedź impulsowa systemu dynamicznego liniowego wyjściowego, k_{imp2} to odpowiedź impulsowa systemu dynamicznego liniowego wejściowego, a b_j jest zestawem parametrów modelu o dodatnich wartościach.

5. Przetwarzanie wstępne



Rys. 5.18. Model Wienera–Hammersteina

5.7.5. Modele NARMAX

Uogólnieniem przedstawionych w rozdziale 5.7.1 modeli z rodziny ARMAX są modele NARMAX (ang. *Nonlinear AutoRegressive Moving Average eXogenous model*) [236]. Struktura modeli z rodziny NARMAX jest bardzo ogólna, a do ich konstrukcji można wykorzystać inne modele [41].

Na przykład, szczególnym przypadkiem modelu NARMAX jest model NARX (ang. *Nonlinear AutoRegressive eXogenous model*). Jest to nieliniowy model autoregresywny z zewnętrznym wejściem, którego równanie ma następującą postać [41]:

$$\begin{aligned} \bar{y}(k) = g_{\text{nx}} & \left(y(k-1) + y(k-2) + \dots + y(k-M_y) \right. \\ & \left. u(k-d_u) + u(k-d_u-1) + \dots + u(k-d_u-M_u) \right. \\ & \left. e(k-1) + e(k-2) + \dots + e(k-M_e) \right) + e(k) \end{aligned} \quad (5.38)$$

gdzie $y(k)$, $u(k)$, $e(k)$ reprezentują odpowiednio wyjście, wejście i błąd, M_y , M_u , M_e reprezentuje liczbę poprzednich wartości wyjść, wejść i błędów, g_{nx} jest pewną nieliniową funkcją, a wielkość d_u jest opóźnieniem.

Zauważmy, że w zależności (5.38) występuje pewna nieliniowa funkcja g_{nx} . Funkcja ta w modelach NARMAX może mieć różną postać. Stąd w kolejnych podpunktach przedstawiono wybrane modele NARMAX, w których ich część nieliniowa bazuje m.in. na sieciach neuronowych oraz transformacji falkowej. Omówiono również dwa nowe podejścia wykorzystujące transformację Hilberta–Huanga oraz uczenie głębokie.

Model NARX bazujący na sieci neuronowej

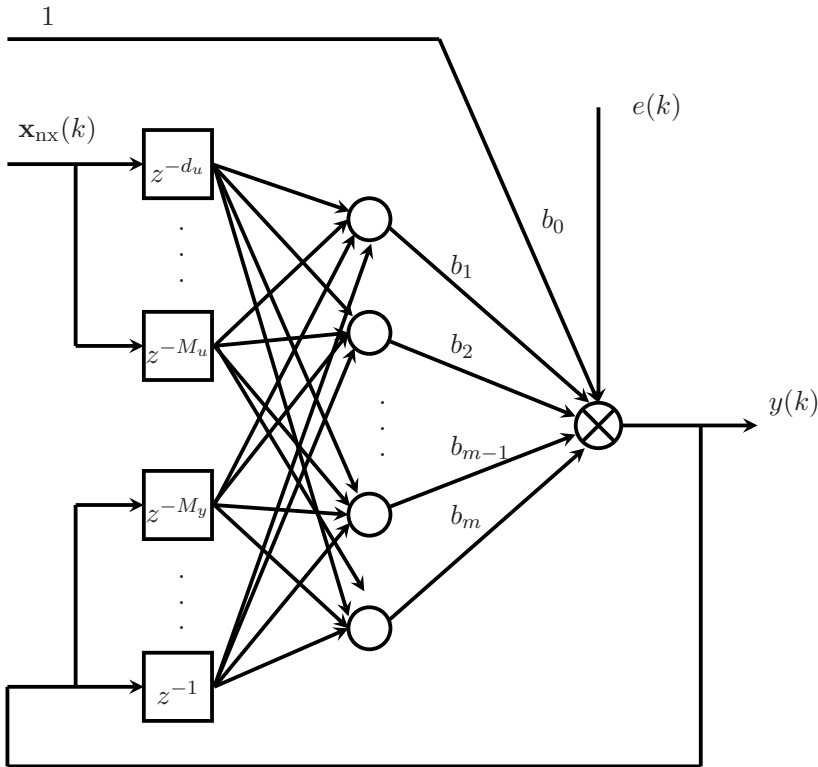
Do budowy modeli NARMAX można wykorzystać np. sieć neuronową. Sieć neuronowa jest już klasycznym podejściem do modelowania systemów nieliniowych, które można wykorzystać również do modelowania sygnałów. W podpunkcie tym przedstawimy też sposób wykorzystania sieci neuronowej do budowy modelu NARX. Model taki bazujący na sieci neuronowej ma postać:

$$\bar{y}(k) = g_{\text{nx}}(\mathbf{x}_{\text{nx}}(k)) = b_0 + \sum_{m=1}^M b_m \Phi_m(\mathbf{x}_{\text{nx}}(k)) + e(k) \quad (5.39)$$

gdzie Φ_m jest **funkcją aktywacji**, $e(k)$ reprezentuje błąd, b_m zestaw parametrów modelu, a $\mathbf{x}_{\text{nx}}(k)$ to wektor zawierający zarówno sygnały pomiarowe z wejścia, jak i wyjścia:

$$\mathbf{x}_{\text{nx}}(k) = [y(k-1) \quad y(k-2) \cdots y(k-M_y) \quad u(k-1) \quad u(k-2) \cdots u(k-M_u)]^T \quad (5.40)$$

Na rysunku 5.19 przedstawiono schemat modelu NARX z wykorzystaniem sieci neuronowej.



Rys. 5.19. Model NARX bazujący na rekurencyjnej sieci neuronowej

Model NARMAX z wykorzystaniem uczenia głębokiego

Do budowy modeli NARMAX można również wykorzystać nowoczesne metody uczenia maszynowego, w tym uczenie głębokie. Cechą charakterystyczną uczenia głębokiego jest to, że wykorzystywane w nim modele mają strukturę hierarchiczną złożoną z wielu nieliniowych warstw.

5. Przetwarzanie wstępne

Przykładem modelu wykorzystywanego w uczeniu głębokim jest model typu DLN (ang. *Deep Learning Networks*). Struktura tego modelu bazuje na strukturze perceptronu wielowarstwowego (ang. *Multilayer Perceptron*). Głębokość sieci DLN jest określana przez liczbę warstw ukrytych. Typowa sieć DLN ma przynajmniej dwie warstwy ukryte i mniej neuronów niż typowa sieć MLP [35, 161].

Niestety, dodając kolejne warstwy w sieci DLN, jej złożoność obliczeniowa wzrasta, co ma negatywny wpływ m.in. na prędkość procesu uczenia. Wzrost złożoności sieci DLN wpływa, z jednej strony na wydłużenie się czasu uczenia sieci, z drugiej natomiast sprawia, że zwiększa się zapotrzebowanie sieci na kolejne przykłady uczące. Sprawia to, że dla sieci takich wymagane są bardzo duże zbiory danych.

Typowa metoda uczenia wielowarstwowego perceptronu bazuje na wstecznej propagacji błędów. Niestety wadą tego podejścia jest duża złożoność obliczeniowa, która wpływa niekorzystnie na szybkość zbieżności uczenia. Algorytmy bazujące na wstecznej propagacji mają również tendencję do utykania w minimach lokalnych [178] oraz są wrażliwe na początkowy dobór wag [215]. W literaturze zaproponowano wiele technik, które kompensują niekorzystny wpływ tych zjawisk. Przykładem jest uzupełnienie procesu uczenia o algorytm nieliniowego grupowania [250] lub przez dodanie *momentum* [276]. W drugim przypadku mówimy o tzw. metodzie wstecznej propagacji błędów z *momentum* (ang. *Momentum Error Back Propagation*).

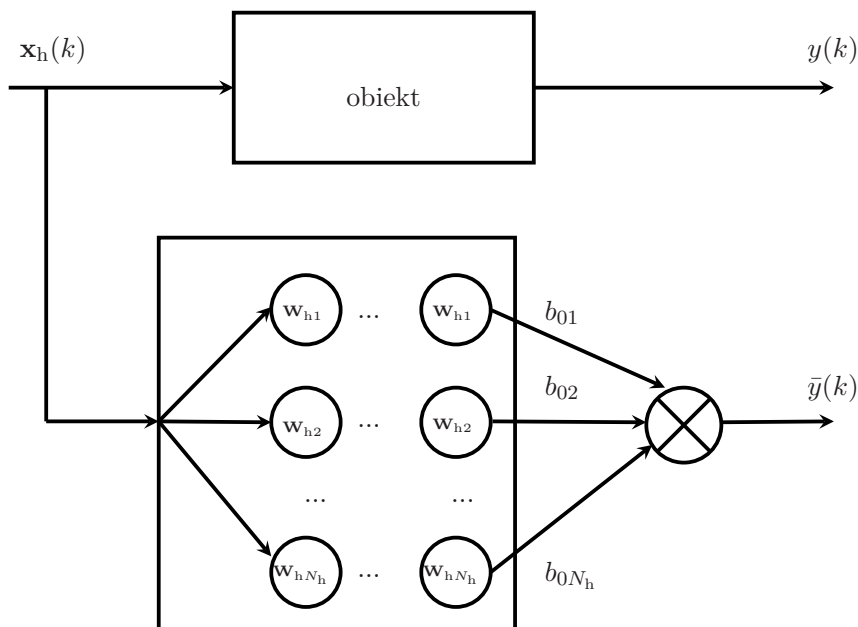
Niestety żadna z wymienionych metod nie eliminuje głównego problemu związanego z minimami lokalnymi, który powodowany jest przez niewłaściwy dobór początkowych wartości wag sieci. Przykładem podejścia do rozwiązania problemu niewłaściwego doboru początkowych wartości wag sieci MLP jest właśnie podejście nazywane uczeniem głębokim (ang. *deep learning*).

Z naszego punktu widzenia istotne jest to, że również metody uczenia głębokiego można wykorzystać do konstrukcji modeli typu NARMAX. W pracy [348] zaproponowano podejście do budowy szczególnego przypadku modelu NARMAX, tj. modelu NARMA z wykorzystaniem uczenia głębokiego na potrzeby zadania usuwania niepożądanych składowych z sygnałów pomiarowych. Warto podkreślić, że uczenie głębokie może być wykorzystane w tym zadaniu jedynie na początkowym etapie konstrukcji modelu.

Przedstawimy teraz model bazujący na omówionym w rozdziale 5.7.5 modelu NARMA (5.38) oraz uczeniu głębokim. Na bazie tego modelu można zaproponować następującą sieć wielowarstwową [348]:

$$\bar{y}(k) = \mathbf{B}_h \Phi_{hN_h} \left(\mathbf{W}_{hN_h} \Phi_{hN_h-1} \left(\dots \mathbf{W}_{h3} \Phi_{h2} \left(\mathbf{W}_{h2} \Phi_{h1} \left(\mathbf{W}_{h1} \mathbf{x}_h(k) + \right. \right. \right. \right. \\ \left. \left. \left. \left. + b_{01} \right) + b_{02} \right) + b_{03} + \dots + b_{0N_h-1} \right) + b_{0N_h} \right) \right) \quad (5.41)$$

gdzie N_h to liczba ukrytych warstw sieci, \mathbf{W}_{hi} reprezentuje macierz wag w każdej i -tej warstwie, b_{0i} to dodatkowa waga o stałych wartościach w każdej i -tej warstwie (ang. *bias*), $\mathbf{B}_h = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_{N_h}]^\top$ jest macierzą wag wyjściowych, N_h liczba węzłów w warstwie ukrytej, $\mathbf{x}_h(k)$ jest wektorem (patrz (5.40)), natomiast Φ_{hi} jest funkcją aktywacji, gdzie $i = 1, 2, \dots, N_h$. Za funkcję aktywacji można przyjąć np. funkcję sigmoidalną [458]. Na rysunku 5.20 przedstawiono przykład sieci wielowarstwowej.



Rys. 5.20. Model NARMA wykorzystujący głęboką sieć neuronową

5.7.6. Model TVARMAX

W podrozdziale 5.7.1 omówiliśmy klasę liniowych modeli autoregresyjnych oraz ruchomej średniej. Z kolei w podrozdziale 5.7.5 przedstawiliśmy uogólnienie tych modeli, tj. modeli NARMAX wraz z przykładami ich wykorzystania. W tym podrozdziale dokonamy uogólnienia modeli ARMAX i NARMAX do klasy modeli niestacjonarnych odpowiednio TVARMAX (ang. *Time-Varying ARMAX*) i TVNARMAX (ang. *Time-Varying NARMAX*).

Klasyczne podejście do modelowania obiektów niestacjonarnych bazuje na algorytmach rekurencyjnych. Za przykład możemy wymienić np. rekurencyjną

5. Przetwarzanie wstępne

metodę najmniejszych kwadratów lub filtr Kalmana. Do konstrukcji liniowych modeli systemów niestacjonarnych stosowane jest również podejście wykorzystujące funkcje bazowe [245, 485]. Przykładowo, uogólnieniem modelu AR dla systemów niestacjonarnych jest model TVAR (ang. *Time-Varying AR*), którego struktura jest następująca:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + e(k) \quad (5.42)$$

gdzie M_a określa liczbę przeszłych wyjść. Z kolei uogólnieniem modelu ARX jest model TVARX (ang. *Time-Varying ARX*). Równanie dla tego modelu ma formę:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + \sum_{m_b=1}^{M_b} b_{m_b}(k) u(k - m_b) + e(k) \quad (5.43)$$

gdzie M_a , M_b określa, odpowiednio, liczbę przeszłych wyjść i wejść. Uogólnieniem modelu ARMA dla systemów niestacjonarnych jest natomiast model TVARMA (ang. *Time-Varying ARMA*) o następującej zależności między wejściami a wyjściami:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + \sum_{m_c=1}^{M_c} c_{m_c}(k) e(k - m_c) \quad (5.44)$$

gdzie $e(k)$ jest błędem, a M_c określa liczbę przeszłych wartości wyliczonych błędów oszacowania. Dokonajmy jeszcze uogólnienia modelu ARMAX dla przypadku niestacjonarnego. Otrzymujemy wówczas model TVARMAX (ang. *Time-Varying ARMAX*) postaci:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + \sum_{m_b=1}^{M_b} b_{m_b}(k) u(k - m_b) + \sum_{m_c=1}^{M_c} c_{m_c}(k) e(k - m_c) \quad (5.45)$$

Oprócz wspomnianych klasycznych metod konstrukcji modeli dla systemów niestacjonarnych bazujących na ruchomej średniej i autoregresji, do konstrukcji modeli można również wykorzystać funkcje bazowe. Parametry modeli takich klas modeli jak TVAR czy TVARMA mogą zostać przedstawione w formie sumy ważonej funkcji bazowych. Przykładowo, dla modelu TVAR zależność na wyznaczenie wartości parametrów a_{m_a} modelu (5.42) ma postać [159, 245]:

$$a_{m_a}(k) = \sum_{m_d=1}^{M_d} a_{m_a, m_d} \Gamma_{m_d}(k) \quad (5.46)$$

5.7. Podstawy przetwarzania wstępnego z modelem

gdzie $m_d = 1, 2, \dots, M_d$, M_d jest liczbą funkcji bazowych, Γ jest pewną funkcją bazową. Jeśli modelem bazowym byłby model TVARX, to oprócz zależności dla parametrów a_{m_a} dodatkowo należy określić zależność do wyliczania parametrów b_{m_b} , tj.

$$b_{m_b}(k) = \sum_{m_d=1}^{M_d} b_{m_b, m_d} \Gamma_{m_d}(k) \quad (5.47)$$

Za funkcje bazowe można wykorzystać również wyniki transformacji Fouriera, dyskretnej transformacji kosinusowej lub falkowej. Na przykład uwzględniając (5.46), zależność (5.42) zapiszemy w formie:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} \sum_{m_d=1}^{M_d} a_{m_a, m_d} \Gamma_{m_d}(k) y(k - m_a) + e(k) \quad (5.48)$$

natomiast, biorąc pod uwagę (5.46) oraz (5.47), zależność (5.43) zapisujemy w formie:

$$\begin{aligned} \bar{y}(k) = & \sum_{m_a=1}^{M_a} \sum_{m_d=1}^{M_d} a_{m_a, m_d} \Gamma_{m_d}(k) y(k - m_a) + \\ & + \sum_{m_b=1}^{M_b} \sum_{m_d=1}^{M_d} b_{m_b, m_d} \Gamma_{m_d}(k) y(k - m_b) + e(k) \end{aligned} \quad (5.49)$$

5.7.7. Modele z rzadką reprezentacją sygnału

Przejdziemy teraz do omówienia nowego podejścia do modelowania sygnałów, które bazuje na rzadkiej reprezentacji. Zanim przejdziemy do omówienia szczegółów tej metody, zwróćmy uwagę na pewną kwestię. W praktycznych zadaniach, w których wykorzystywane są dane pomiarowe, jednym z istotnych zagadnień jest estymacja sygnału na podstawie danych pomiarowych. W typowych przypadkach akwizycja danych pomiarowych jest liniowa. Wynika z tego, że estymacja sygnału na podstawie danych pomiarowych sprowadza się do rozwiązania układu równań liniowych [55]. Stąd, niech związek pomiędzy pomiarem \mathbf{y} a sygnałem \mathbf{s} ma postać:

$$\mathbf{y} = \Xi \mathbf{s} \quad (5.50)$$

gdzie \mathbf{y} reprezentuje wektor pomiaru taki, że $\mathbf{y} \in \mathbb{R}^K$, z kolei \mathbf{s} , to nieznaną wektor sygnału, dla którego $\mathbf{s} \in \mathbb{R}^{K_s}$, a Ξ jest macierzą pomiarów taką, że $\Xi \in \mathbb{R}^{K \times K_s}$. Chcąc estymować sygnał na podstawie danych pomiarowych, należy rozwiązać zadany układ równań (5.50). By było to możliwe, liczba próbek pomiarowych musi być co najmniej równa długości estymowanego sygnału. Jeśli liczba pomiarów

5. Przetwarzanie wstępne

będzie mniejsza niż długość sygnału, to wówczas nie jest możliwe jednoznaczne rozwiązanie układu równań (5.50). Uwaga ta jest również podstawą twierdzenia Shannona–Kotelnikowa.

Zastosowanie twierdzenia Shannona–Kotelnikowa w praktyce w wielu przypadkach prowadzi do tego, że liczba zebranych danych jest nadmiarowa. Przykładem jest powszechność kompresji obrazów i sygnałów dźwiękowych, która umożliwia usunięcie części z zebranych danych.

Jeśli rzeczywiście liczba zbieranych podczas akwizycji danych jest nadmiarowa, to zasadne jest pytanie o to, czy konieczne jest gromadzenie tak dużej liczby danych tylko po to, by na późniejszych etapach część z nich usunąć. W tej sytuacji istotnym problemem staje się zagadnienie ustalenia takiej reprezentacji sygnału, która umożliwi jego estymację na podstawie mniejszej liczby pomiarów bez widocznej straty jakości. Problem ten staje się bardzo istotny w przypadku danych pozyskiwanych podczas bardzo kosztownych pomiarów lub wówczas, gdy dane te muszą być po ich zebraniu przesłane przez sieć, np. sieć BSN.

Rozwiązaniem jest **rzadka reprezentacja sygnałów**. Rzadki sygnał (ang. *sparse signal*) charakteryzuje się tym, że ma on niewiele składowych różnych od zera [471]. Podstawą modelowania z wykorzystaniem takich sygnałów jest ustalenie konkretnej reprezentacji, w której sygnał ten będzie miał cechy sygnału rzadkiego. Reprezentację taką nazywamy reprezentacją rzadką (ang. *sparse representation*). Niestety wiele rzeczywistych sygnałów nie wykazuje cech rzadkości, dlatego sygnały takie można reprezentować w innych bazach [211]. Stąd poszukiwanie rzadkiej reprezentacji dla sygnału pomiarowego sprowadza się do ustalenia pewnej bazy, w której to sygnał będzie miał cechy sygnału rzadkiego. Wśród najpopularniejszych baz wyróżnić możemy bazy wykorzystujące transformację Fouriera i falkową.

Sformalizujmy teraz pojęcie rzadkości sygnału. Niech \mathbf{s} będzie sygnałem o K_s próbkach. Jeśli przez $K_{s \neq 0}$ określimy liczbę próbek w sygnale, których wartości są różne od zera, to sygnał rzadki spełnia nierówność:

$$K_{s \neq 0} \ll K_s \quad (5.51)$$

W przypadku gdy zachodzi (5.51), nie jest możliwe jednoznaczne rozwiązanie układu równań (5.50). W tym przypadku układ równań jest silnie niedookreślony i nie ma jednoznacznego rozwiązania. Istnieje jednak metoda umożliwiająca rozwiązanie tego układu równań. W metodzie tej, przed rozpoczęciem poszukiwania rozwiązania, do tego układu równań dodawany jest pewien warunek. Celem dołączenia tego warunku jest narzucenie wymagania, by znalezione rozwiązanie spełniało możliwie najlepiej cechy rzadkości, tj. takiego, które ma największą liczbę składowych zerowych.

5.7. Podstawy przetwarzania wstępnego z modelem

Rozwiązanie dla tak sformułowanego problemu może zostać znalezione przez rozwiązanie następującego zadania optymalizacji [471]:

$$\tilde{\mathbf{s}} = \min_{\mathbf{s}} \quad \|\mathbf{s}\|_0 \quad (5.52)$$

p.o. $\mathbf{y} = \Xi \mathbf{s}$

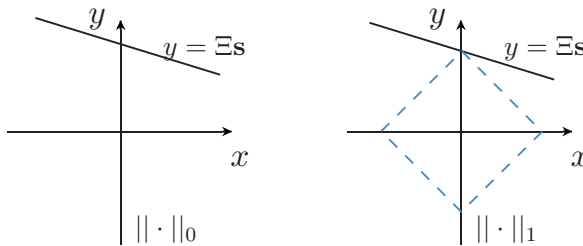
gdzie $\|\cdot\|_0$ reprezentuje tzw. normę zerową wektora. Niestety algorytm ten charakteryzuje się niestabilnością numeryczną oraz wymaga rozwiązania złożonego problemu kombinatorycznego.

Alternatywą dla podejścia (5.52) jest algorytm, który bazuje na normie pierwszej. W tym przypadku, bazując na tej normie, zadanie optymalizacyjne ma następującą postać [471]:

$$\tilde{\mathbf{s}} = \min_{\mathbf{s}} \quad \|\mathbf{s}\|_1 \quad (5.53)$$

p.o. $\mathbf{y} = \Xi \mathbf{s}$

gdzie $\|\cdot\|_1$ reprezentuje normę pierwszą wektora. Algorytm ten nie tylko nie ma wad metody wykorzystującej normę zerową wektora, ale też sprowadza to zadanie do wypukłego problemu optymalizacyjnego. Ilustracja geometrii rozwiązania dla omówionych norm została przedstawiona na rys. 5.21.



Rys. 5.21. Ilustracja geometrii rozwiązania dla przykładowych norm

Rzadka reprezentacja jest nowym podejściem, które coraz częściej wykorzystywane jest w przetwarzaniu sygnałów. Na przykład w pracy [209] przedstawiono podejście bazujące na rzadkiej reprezentacji do modelowania niskoczęstotliwościowych zakłóceń w sygnałach.

Omówimy teraz przykład wykorzystania tego podejścia w modelowaniu sygnałów.

Model ARMAX z wykorzystaniem rzadkiej reprezentacji

W pracy [368] zaproponowano metodę konstrukcji modelu bazującego na modelach ARMAX oraz reprezentacji rzadkiej. Jednocześnie zaproponowano metodę

5. Przetwarzanie wstępne

rekonstrukcji sygnałów rzadkich, która wykorzystuje metody optymalizacji rzadkiej. Przyjmijmy, że sygnał pomiarowy zawiera następujące składowe:

$$\bar{y}(k) = s(k) + v(k) + \underbrace{\kappa_0(k) + \kappa_1(k) + \kappa_2(k)}_{\kappa(k)} \quad (5.54)$$

gdzie $s(k)$ jest szukanym, nieznanym sygnałem, $v(k)$ oznacza biały szum, $\kappa(k)$ reprezentuje zakłócenie o następujących składowych $\kappa_0(k)$ (zakłócenie szpilkowe), $\kappa_1(k)$ (zakłócenie odcinkowo-liniowe), $\kappa_2(k)$ (odcinkowo-liniowy trend). Zakłócenia $\kappa(k)$ są zakłóceniami rzadkimi, które spełniają następujące zależności:

$$\xi_0(k) = \kappa_0(k), \quad k = 1, 2, \dots, K \quad (5.55a)$$

$$\xi_1(k) = \kappa_1(k) - \kappa_1(k-1), \quad k = 2, 3, \dots, K \quad (5.55b)$$

$$\xi_2(k) = \kappa_2(k+1) - 2\kappa_2(k) + \kappa_2(k-1), \quad k = 2, 3, \dots, K-1 \quad (5.55c)$$

Przedstawimy teraz zaproponowany w pracy [368] model sygnału pomiarowego, który bazuje na omówionym już modelu ARMAX (patrz (5.27)):

$$\bar{y}(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_{M_a} y(k-M_a) = b_1 u(k-1) + b_2 u(k-2) + \dots + b_{M_b} u(k-M_b) + c_1 e(k-1) + c_2 e(k-2) + \dots + c_{M_c} e(k-M_c) + \kappa(k) \quad (5.56)$$

gdzie a_i , b_i oraz c_i są parametrami modelu. Przedstawione podejście wymaga zastosowania odpowiednich metod estymacji parametrów modelu (5.56) [368]. Typowe podejście w tym przypadku wykorzystuje algorytmy optymalizacji rzadkiej (ang. *sparse optimization*).

5.7.8. Modele w dziedzinie czasowo-częstotliwościowej

W podrozdziale omówiono podejście do konstrukcji modeli, które bazuje na strukturach typu NARMAX oraz na reprezentacji sygnałów w dziedzinie czasowo-częstotliwościowej. Zaprezentowane zostaną dwa rozwiązania, tj. w pierwszym z nich wykorzystano transformację falkową. Drugie z nich jest nowym podejściem bazującym na transformacji Hilberta–Huanga i metodzie EMD.

Modele bazujące na transformacji falkowej

W rozdziale 3 wskazaliśmy, że transformacja falkowa jest jednym z przykładów przekształcenia pozwalającą na jednoczesną reprezentację sygnału w dziedzinie czasu i skali. Ciągła transformata falkowa pewnego sygnału $s(t)$ wyznaczana jest

5.7. Podstawy przetwarzania wstępnego z modelem

z następującej zależności:

$$S_{\text{WVT}}(\mathbf{a}, \mathbf{b}) = \frac{1}{\sqrt{\mathbf{a}}} \int_{-\infty}^{\infty} s(t) \Psi\left(\frac{t - \mathbf{b}}{\mathbf{a}}\right) dt \quad (5.57)$$

Efektom działania tego przekształcenia jest zestaw sygnałów elementarnych. Sygnały te są sygnałami oscylacyjnymi o różnych czasach trwania i o różnym widmie. Zależność (5.57) reprezentuje filtrację sygnału $s(t)$ przez sygnał analizujący $\Psi(t)$, który został przeskalowany w dziedzinie czasu (parametr) [484]. Dzięki specyficznym właściwościom tych sygnałów możliwe jest ich wykorzystanie np. w modelach typu TVARMAX [245].

Transformację falkową można zdyskretyzować, próbując odpowiednie wartości skali (\mathbf{a}) oraz przesunięcia (\mathbf{b}). W przypadku wykorzystania skali diadycznej (ang. *dyadic scalling*) rodzina falek przyjmuje następującą postać:

$$\Psi_{j,m}(t) = \frac{1}{\sqrt{2^j}} \Psi\left(\frac{t - m2^j}{2^j}\right) = 2^{-j/2} \Psi(2^{-j}t - m) \quad (5.58)$$

gdzie j reprezentuje wielkość parametryzującą skalę, a m jest przesunięciem w czasie. Rodzina falek (5.58) tworzy diadyczną ortonormalną bazę falkową.

Ostatecznie dyskretna transformacja falkowa (ang. *Discrete Wavelet Transform*, DWT) przyjmuje postać:

$$d_{j,m} = \int_{-\infty}^{\infty} s(t) \Psi_{j,m} dt \quad (5.59)$$

Z kolei odwrotna transformacja falkowa (ang. *Inverse Discrete Wavelet Transform*, IDWT) jest sumą iloczynów wartości $d_{j,m}$ oraz odpowiednich funkcji bazowych $\Psi_{j,m}$, która obejmuje wszystkie poziomy i przesunięcia:

$$s(k) = \sum_{j=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} d_{j,m} \Psi_{j,m}(t) \quad (5.60)$$

Z funkcją falkową można powiązać funkcję skalującą \mathcal{R} . Funkcję skalującą definiujemy, korzystając z następującej zależności:

$$\mathcal{R}_{j_0,m}(t) = 2^{j_0/2} \mathcal{R}(2^{j_0}t - m) \quad (5.61)$$

gdzie j_0 jest liczbą całkowitą.

5. Przetwarzanie wstępne

Uwzględniając (5.61), zdekomponowany z wykorzystaniem transformacji falkowej sygnał $s(k)$ można przedstawić jako [245]:

$$s(k) = \sum_{m=-\infty}^{\infty} c_{j_0,m} \mathcal{R}_{j_0,m}(t) + \sum_{j=j_0}^{\infty} \sum_{m=-\infty}^{\infty} d_{j,m} \Psi_{j,m}(t) \quad (5.62)$$

gdzie $c_{j_0,m}$ to współczynniki otrzymywane na drodze filtracji dolnoprzepustowej, natomiast $d_{j,m}$ na drodze filtracji górnoprzepustowej [484].

Model WTVAR

Przedstawimy teraz metodę konstrukcji modelu bazującego na transformacji falkowej i modelu TVAR. Dla tego przypadku model TVAR ma następującą postać (por. (5.42)):

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + e(n) \quad (5.63)$$

Uwzględniając (5.62), zależność (5.63) przyjmuje następującą postać [245]:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} \sum_{m_w=1}^{M_w} c_{m_a,m_w} y_{m_w}(k - m_a) + e(n) \quad (5.64)$$

gdzie M_w jest liczbą funkcji bazowych. Modele typu (5.64), dla odróżnienia od modeli TVAR, nazwiemy modelami WTVAR (ang. *wavelet TVAR*).

Model WTVARX

Przez analogię przedstawiony zostanie teraz model będący połączeniem modelu TVARX i transformacji falkowej [245]. Przypomnijmy (por. (5.43)), że opis matematyczny dla modelu TVARX ma postać:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k) y(k - m_a) + \sum_{m_b=1}^{M_b} b_{m_b}(k) u(k - m_b) \quad (5.65)$$

Uwzględniając tym razem zależność (5.62) w modelu (5.65), otrzymamy następujący model:

$$\bar{y}(k) = \sum_{m_a=1}^{M_a} \sum_{m_w=1}^{M_w} c_{m_a,m_w} y_{m_w}(k - m_a) + \sum_{m_b=1}^{M_b} \sum_{m_u=1}^{M_u} d_{m_b,m_u} u_{m_u}(k - m_b) + e(k) \quad (5.66)$$

Modele typu (5.66) z kolei nazwiemy modelami WTVARX (ang. *wavelet TVARX*). Jedną z zalet przedstawionych modeli WTVAR oraz WTVARX jest to, że umożliwiają one modelowanie sygnałów o bardzo stromych zboczach [245]. Podejście to wpływa również na poprawę jakości modelowania sygnałów niestacjonarnych, która jest możliwa do uzyskania z wykorzystaniem typowych modeli TVAR oraz TVARX. Z kolei w pracy [244] wskazano na dobre właściwości uogólniające modeli WTVAR i WTVARX.

Modele bazujące na transformacji Hilberta–Huanga

Do konstrukcji modeli z rodziny NARMAX można również wykorzystać transformację Hilberta–Huanga (HHT). Transformacja HHT jest stosunkowo nową koncepcją, która cieszy się w ostatnich latach dużym zainteresowaniem. Głównymi elementami transformacji HHT są algorytm EMD (ang. *Empirical Mode Decomposition*) oraz transformacja Hilberta. Algorytm EMD służy do dekompozycji sygnału na zestaw sygnałów nazywanych modami lub funkcjami składowymi IMF (ang. *Intrinsic Mode Functions*). Z kolei transformację Hilberta wykorzystuje się do wyznaczania chwilowych częstotliwości i amplitud mod zdekomponowanego sygnału.

Jedną z cech transformacji HHT jest zdolność do separowania składowych sygnałów, które nakładają się na siebie w dziedzinie czasu i częstotliwości. Inną, wyróżniającą cechą tego podejścia jest to, że pozwala na analizę sygnałów niestacjonarnych.

Zasadniczą częścią transformacji Hilberta–Huanga jest algorytm EMD, który dokonuje adaptacyjnej dekompozycji sygnału na funkcje składowe, co upodabnia go np. do transformacji falkowej. W tabeli 5.1 zestawiono transformację HHT z transformacjami Fouriera i falkową.

Tabela 5.1. Porównanie transformacji Fouriera, falkowej i Hilberta–Huanga

| | Transformacja Fouriera | Transformacja falkowa | Transformacja Hilberta–Huanga |
|------------------|----------------------------------|--------------------------|----------------------------------|
| Baza | <i>a priori</i> | <i>a priori</i> | adaptacyjna |
| Prezentacja | energia, czas i częstotliwość | energia, czas i skala | energia, czas i częstotliwość |
| Nieliniowość | nie | nie | tak |
| Niestacjonarność | nie | tak | tak |

Algorytm EMD

Metoda EMD została pierwotnie zaproponowana w pracy [169]. Bazując na tej pracy, przedstawimy podstawy tego podejścia. Niech algorytm EMD będzie reprezentowany przez \mathcal{E} , wówczas dekompozycję sygnału $s(k)$ z wykorzystaniem metody EMD oznaczymy jako:

$$\mathbf{H}_{\text{imf}N_{\text{imf}}}(k) = \mathcal{E}(s(k)), \quad k = 1, 2, \dots, K \quad (5.67)$$

gdzie

$$\mathbf{H}_{\text{imf}N_{\text{imf}}}(k) = \{h_1(k) \ h_2(k) \ \dots \ h_{N_{\text{imf}}}(k) \ r(k)\}, \quad k = 1, 2, \dots, K \quad (5.68)$$

$s(k) \in \mathbb{R}^K$, $h_n(k) \in \mathbb{R}^K$ jest n -tą modą, $r(k) \in \mathbb{R}^K$ jest sygnałem rezydualnym (resztowym), k jest indeksem czasu, K to liczba wszystkich pomiarów, a N_{imf} to liczba wyznaczonych mod.

Sygnał $r(k)$ jest sygnałem wolnozmiennym. Natomiast wyznaczone funkcje składowe IMF są sygnałami nieokresowymi o lokalnej średniej równej zero [169].

Wyznaczone mody są względem siebie prawie ortogonalne [169], a ich zbiór jest pełny. Sygnały IMF wyznaczone dla sygnału pierwotnego $s(k)$ spełniają:

$$s(k) = \sum_{n=1}^{N_{\text{imf}}} \mathbf{H}_{\text{imf}n}(k) + r(k) \quad (5.69)$$

Sygnał resztowy ($r(k)$) można traktować jako funkcję IMF. Wówczas zależność (5.69) można zapisać w postaci:

$$s(k) = \sum_{n=1}^{N_{\text{imf}}+1} \mathbf{H}_{\text{imf}n}(k) \quad (5.70)$$

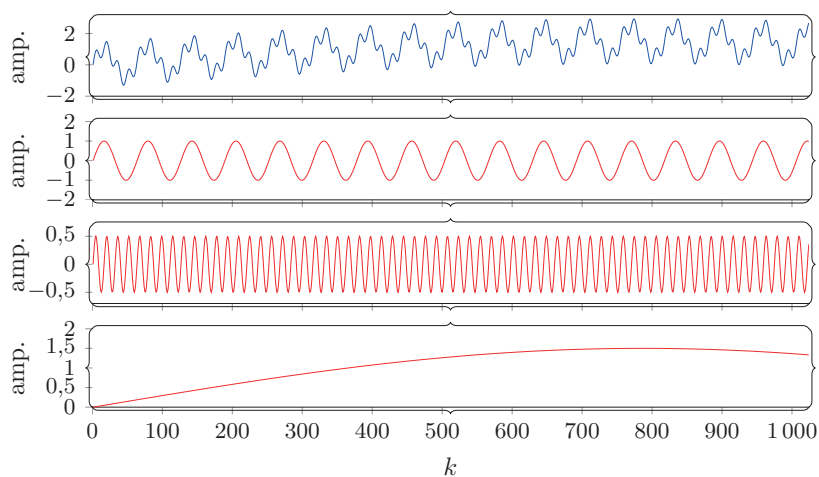
Zależność (5.69) będziemy oznaczać jako:

$$s(k) = \mathcal{E}^{-1}(\mathbf{H}_{\text{imf}N_{\text{imf}}}(k)) \quad (5.71)$$

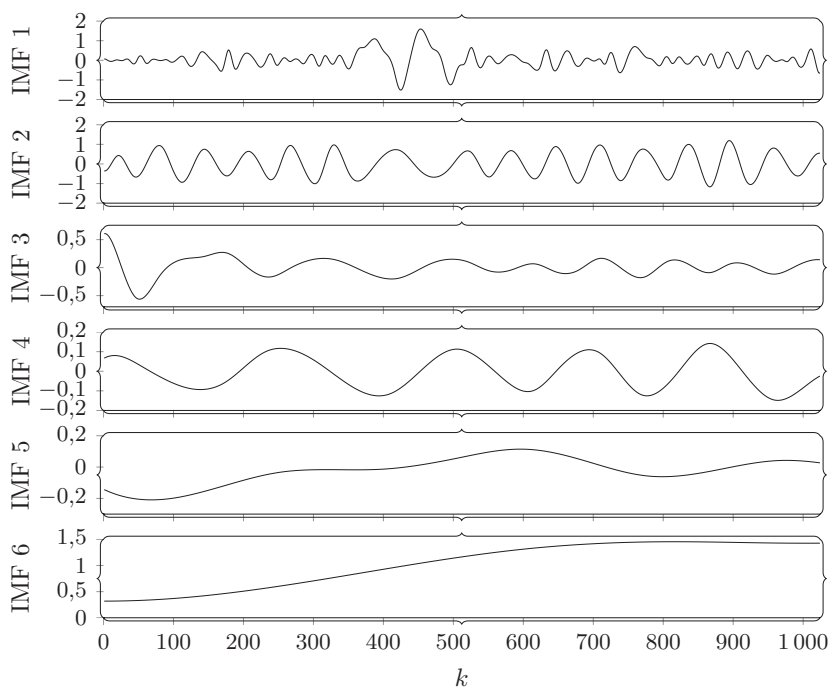
Na rysunku 5.23 przedstawiono wynik dekompozycji syntetycznego sygnału (rys. 5.22) na mody IMF.

Cechą charakterystyczną algorytmu EMD jest to, że bazuje on jedynie na analizowanym sygnale pomiarowym. Oznacza to, że w przeciwieństwie do np. transformacji Fouriera lub falkowej metoda EMD nie wymaga żadnych wstępnych założeń odnośnie do analizowanego sygnału. Na przykład w przypadku transformacji Fouriera analizowany sygnał musi być sygnałem okresowym i stacjonarnym. W przeciwnym przypadku uzyskane wyniki mogą być niezadowolające.

5.7. Podstawy przetwarzania wstępnego z modelem

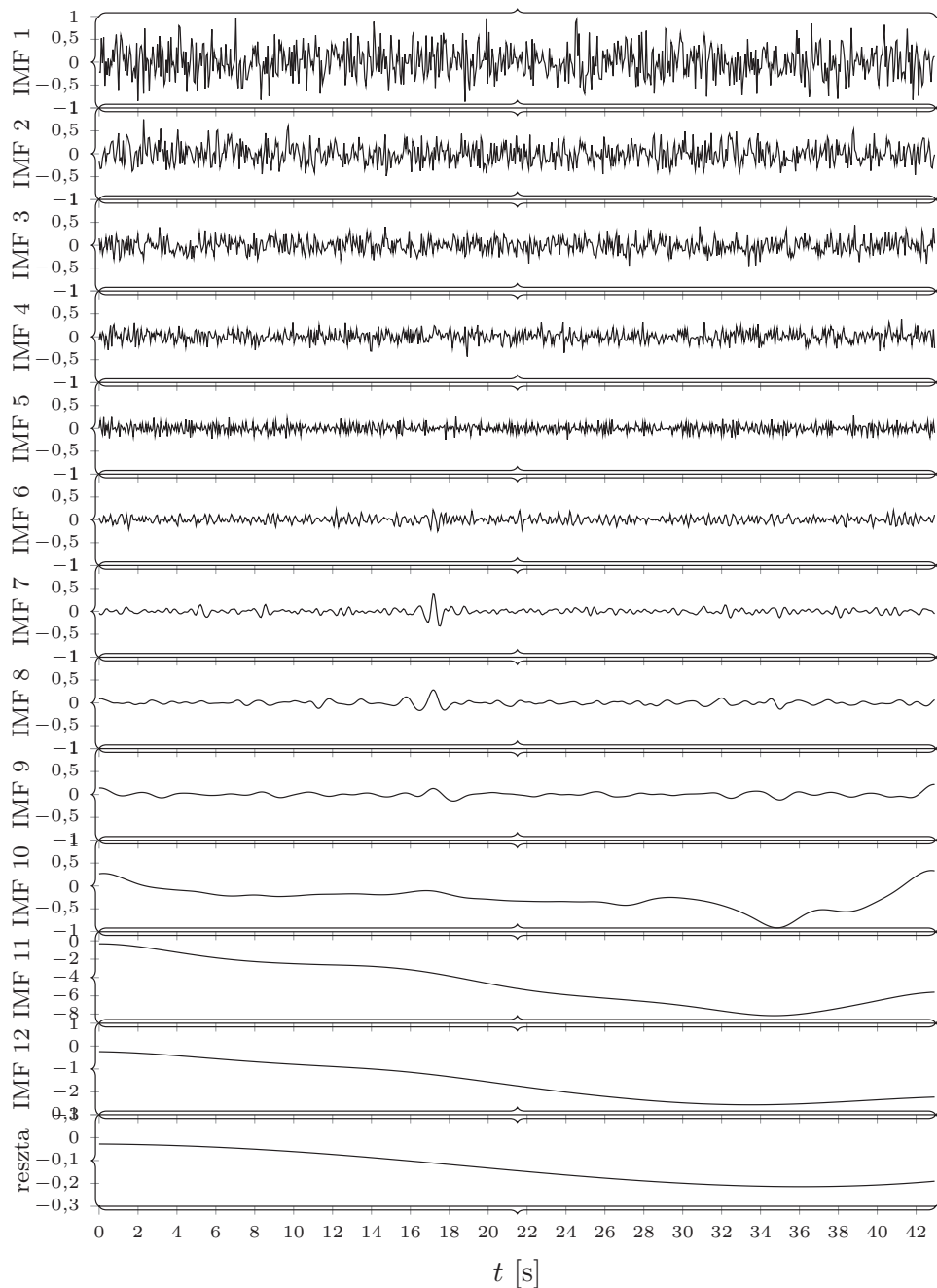


Rys. 5.22. Przykładowy sygnał syntetyczny (niebieska linia) oraz jego składowe (czerwone linie)



Rys. 5.23. Mody wyznaczone dla sygnału syntetycznego (rys. 5.22) wyznaczone algorytmem EMD

5. Przetwarzanie wstępne



Rys. 5.24. Przykład dekompozycji sygnału na mody (funkcje IMF)

Wymienione cechy algorytmu EMD sprawiają, że jest to narzędzie przydatne w przetwarzaniu sygnałów niestacjonarnych [169].

Dla porównania na rysunku 5.24 przedstawiono efekt dekompozycji sygnału pomiarowego wykonanej z wykorzystaniem algorytmu EMD.

Na rysunku 5.25 przedstawiono schemat blokowy algorytmu EMD. Natomiast na rys. 5.26 przedstawiono jeden z etapów wyznaczania funkcji składowej sygnału pomiarowego.

Jedną z wad algorytmu EMD jest efekt nazywany mieszaniem funkcji (ang. *mode mixing*). Efekt ten oznacza (a), że IMF składa się z sygnałów o różnych modach, (b), że sygnał o tej samej skali zawiera się w różnych funkcjach IMF [269, 449].

W funkcjach wyznaczonych podczas dekompozycji sygnału algorytmem EMD może też wystąpić zjawisko *aliasingu*. Skutkiem jest to, że widma funkcji składowych IMF mogą się na siebie nakładać. Przyczyną powstania tego zjawiska jest niewłaściwie dobrana częstotliwość próbkowania.

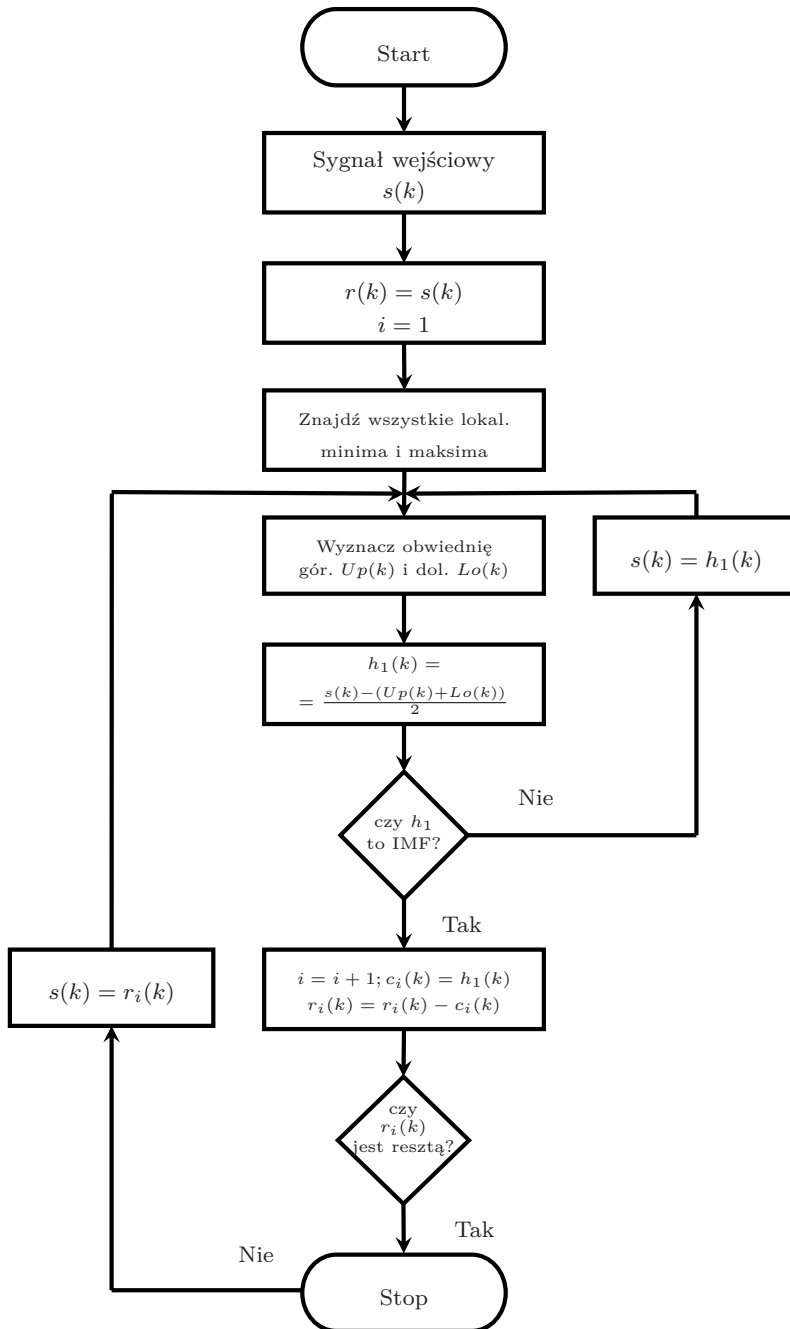
Innym niekorzystnym fenomenem występującym podczas dekompozycji sygnałów z wykorzystaniem algorytmu EMD jest zjawisko nazywane (ang. *end effect artifacts*). Efekt ten jest spowodowany przez niewłaściwie oszacowaną liczbą ekstremów w sygnale, co wpływa na niepoprawne wyznaczenie obwiedni sygnału.

EEMD

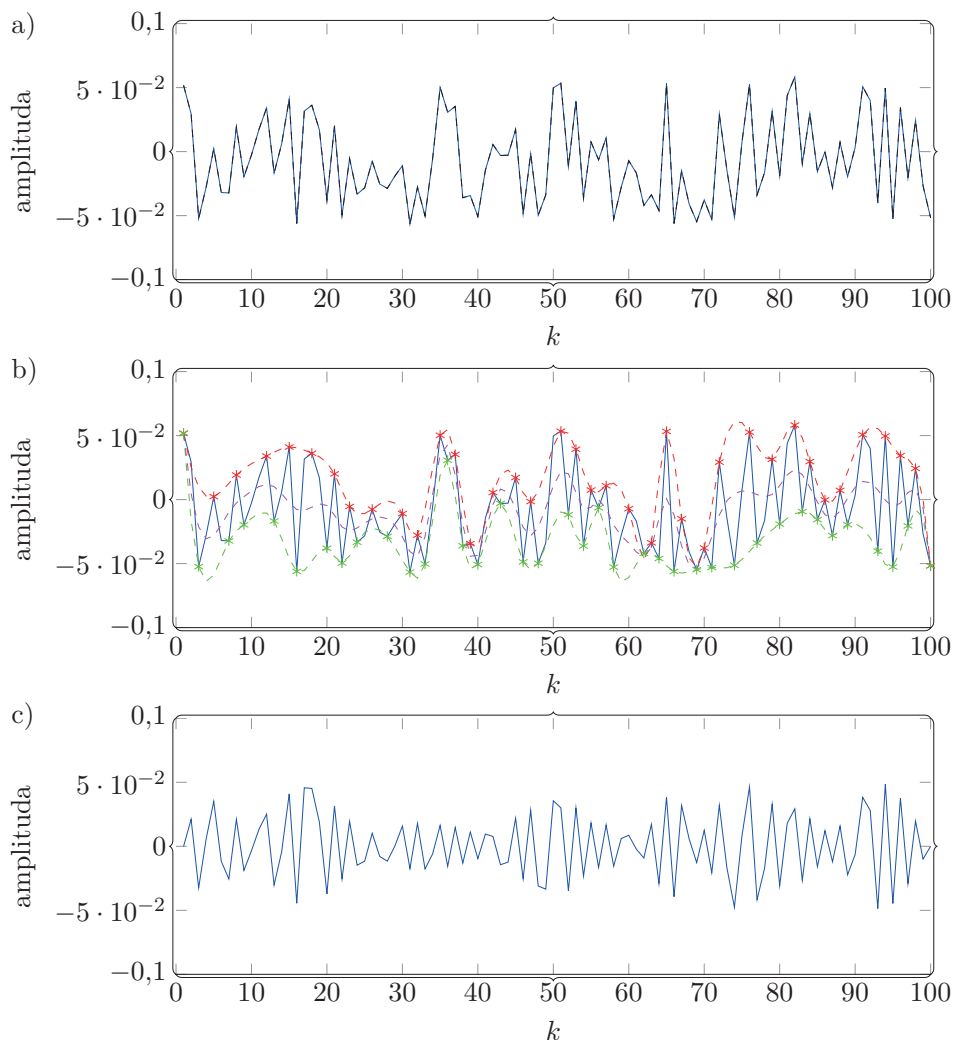
Oryginalny algorytm EMD ma pewne wady, które zostały omówione w poprzednim punkcie. W celu usunięcia części z nich zaproponowano pewne warianty tej metody. Jednym z nich jest algorytm EEMD (ang. *Ensemble Empirical Mode Decomposition*) [449]. W metodzie tej zmodyfikowano pierwszą i ostatnią fazę dekompozycji sygnału (w porównaniu do oryginalnej metody EMD). W pierwszej fazie do analizowanego sygnału dodanych jest kilka realizacji białego szumu. W wyniku otrzymujemy kilka zestawów funkcji IMF, wyznaczonych na podstawie oryginalnego sygnału wraz z jego wariantami. Z kolei modyfikacja w ostatniej fazie związana jest z tym, że ostateczny zestaw funkcji składowych IMF jest uśrednieniem z tych wielu zestawów wygenerowanych funkcji składowych IMF. Takie podejście do zadania dekompozycji sprawia, że algorytm EEMD jest bardziej odporny na zjawisko mieszania funkcji składowych (mod).

Wadą tego podejścia jest natomiast to, że wygenerowany zestaw funkcji składowych IMF może w tym przypadku nie być zupełny. Co więcej, zastosowanie uśredniania sprawia, że dla różnych realizacji można uzyskać inną liczbę funkcji IMF. Może to prowadzić do pewnych trudności na etapie rekonstrukcji oryginalnego sygnału.

5. Przetwarzanie wstępne



Rys. 5.25. Schemat blokowy algorytmu EMD



Rys. 5.26. Ilustracja działania algorytmu EMD: a) sygnał oryginalny, b) wyznaczenie obwiedni dolnej i górnej, c) wynikowa moda

CEEMD

Kolejną modyfikacją algorytmu EMD to algorytm o nazwie komplementarny EEMD (ang. *complementary EEMD*, CEEMD). Struktura algorytmu CEEMD jest podobna do struktury algorytmu EEMD. Różnica związana jest ze sposobem generowania szumu białego. W przypadku tej metody szum generowany w parach, co sprawia, że otrzymujemy dwa zestawy funkcji IMF.

5. Przetwarzanie wstępne

Metoda CEEMD usuwa podstawowe problemy związane z algorytmem EMD, tj. mieszania funkcji składowych oraz aliasing. Jednakże w algorytmie CEEMD pojawiają się inne niekorzystne zjawiska, które w niektórych zadaniach dekompozycji sygnałów pomiarowych mogą sprawiać problemy. Na przykład metoda CEEMD jest metodą znacznie bardziej wymagającą obliczeniowo w porównaniu do algorytmów EMD czy EEMD [477]. Jest to związane ze zwiększeniem liczby operacji koniecznych do wykonania na etapie przesiewania. Inną wadą algorytmu CEEMD jest to, że możliwe jest wygenerowanie funkcji, które nie będą miały właściwości funkcji składowych IMF [477].

CEEMDAN

W pracy [408] zaproponowano algorytm CEEMDAN, który usuwa część niekorzystnych cech algorytmu CEEMD. W porównaniu do wcześniej omówionych metod, wynik dekompozycji sygnału z wykorzystaniem algorytmu CEEMDAN jest zbiorem zupełnym. Mniejsza jest też wymagana liczba obliczeń dla tej metody. W wyniku tych uproszczeń czas wykonywania metody znacznie się skrócił.

Charakterystykę algorytmu EMD zakończmy następującą uwagą. Wyznaczone przez ten algorytm funkcje składowe IMF mogą zostać wykorzystane do konstrukcji zespołu filtrów. Zespół filtrów składa się z wielu filtrów pasmowo-przepustowych, które umożliwiają dekompozycję sygnału na składowe należące do różnych przedziałów częstotliwości. Podczas konstrukcji zespołu filtrów ważne jest to, by pokrywały pełny zakres częstotliwości sygnału wejściowego [126, 448].

Dwa główne zadania, które mogą zostać rozwiązane z wykorzystaniem zespołu filtrów, bazującego na algorytmie EMD, jest zadanie usuwania składowych o niskich i wysokich częstotliwościach z sygnału pomiarowego

Model ETVAR

W pracy [393] zaproponowano model niestacjonarnej autoregresji TVAR (por. (5.42)) bazujący na algorytmie EMD. Zaproponowaną w pracy modyfikację modelu TVAR wykorzystującą algorytm EMD nazwiemy ETVAR (ang. EMD TVAR). Wiemy, że w wyniku dekompozycji sygnału pomiarowego $y(k)$ bazującej na metodzie EMD otrzymujemy zbiór funkcji składowych IMF (5.69). Każdą z wyznaczonych funkcji IMF w przedstawionym podejściu będziemy reprezentować następującym modelem TVAR:

$$\bar{h}_n(k) = \sum_{m_a=1}^{M_a} a_{m_a}(k)h_n(k - m_a) + e(k), \quad k = 1, 2, \dots, K \quad (5.72)$$

gdzie $a_{m_a}(k)$ reprezentuje M_a niestacjonarnych parametrów modelu każdej z funkcji IMF. Do reprezentacji tych parametrów w omawianej pracy zaproponowano

następującą zależność:

$$a_{m_a}(k) = \sum_{m_d=1}^{M_d} a_{m_a, m_d} \Gamma_{m_d}(k) \quad (5.73)$$

gdzie a_{m_a, m_d} są wagami, a Γ_{m_d} reprezentuje funkcje bazowe. Do aproksymacji parametrów a_{m_a, m_d} można wykorzystać np. wielomiany Legendre'a lub DPSS (ang. *Discrete Prolate Spheroidal Sequences*) [393]. Uwzględniając zależności (5.73), (5.72) oraz (5.69) otrzymujemy model ETVAR w następującej postaci:

$$\begin{aligned} \bar{y}_n(k) &= \sum_{n=1}^{N_{\text{imf}}} \sum_{m_a=1}^{M_a} \left(\sum_{m_d=1}^{M_d} a_{m_a, m_d} \Gamma_{m_d}(k) h_n(k - m_a) \right) + e(k) = \\ &= \sum_{n=1}^{N_{\text{imf}}} \sum_{m_a=1}^{M_a} \sum_{m_d=1}^{M_d} a_{m_a, m_d} \left(\Gamma_{m_d}(k) h_n(k - m_a) \right) + e(k), \quad k = 1, 2, \dots, K \end{aligned} \quad (5.74)$$

Modele z rodziny TVARMAX, wykorzystujące transformację Hilberta–Huanga do estymacji sygnału, osiągają bardzo dobre wyniki zwłaszcza dla sygnałów niestacjonarnych. Istotny wpływ na jakość wyników estymacji ma wybór wariantu algorytmu EMD.

5.7.9. Modele w przestrzeni stanów

Do przetwarzania wstępnego sygnałów można wykorzystać również **modele w przestrzeni stanu**. Modele te składają się z zestawu równań stanu (dynamiki) oraz równań wyjścia (pomiarów). Modele w przestrzeni stanu zawierają zmienne stanu, funkcję opisującą dynamikę oraz wejścia systemu. Z kolei równanie wyjścia wiąże wyjście ze stanem oraz wejściami modelowanego systemu. Postać modelu w przestrzeni stanów ma formę:

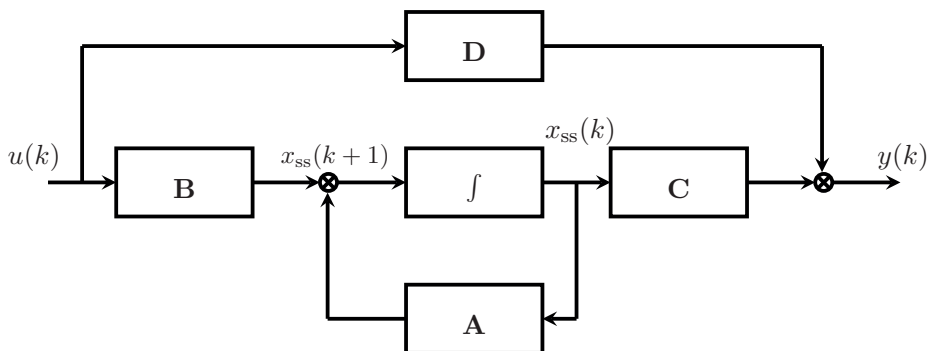
$$\begin{aligned} \mathbf{x}_{\text{ss}}(k+1) &= f_{\text{ss}}(\mathbf{x}_{\text{ss}}(k), \mathbf{u}(k)) + \mathbf{n}(k) \\ \mathbf{y}(k) &= g_{\text{ss}}(\mathbf{x}_{\text{ss}}(k), \mathbf{u}(k)) + \mathbf{v}(k) \end{aligned} \quad (5.75)$$

gdzie N_{ss} jest liczbą zmiennych stanu, reprezentuje wektor $\mathbf{x}_{\text{ss}}(k) = [x_1(k), x_2(k), \dots, x_{N_{\text{ss}}}(k-1)]^T$. Równanie (5.75) jest ogólnym opisem nieliniowych zależności w przestrzeni stanów. Przedstawmy jeszcze szczególny przypadek tego modelu, tj. model dynamiki dla systemu liniowego, który zapisujemy następująco:

$$\begin{aligned} \mathbf{x}_{\text{ss}}(k+1) &= \mathbf{A}\mathbf{x}_{\text{ss}}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \quad (5.76)$$

5. Przetwarzanie wstępne

gdzie \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} to macierze, odpowiednio, stanu, wejść, wyjść i przenoszenia. Schemat modelu systemu liniowego w przestrzeni stanu przedstawiono na rys. 5.27.



Rys. 5.27. Schemat modelu liniowego w przestrzeni stanu

5.8. Modelowanie szumów pomiarowych

W podrozdziale tym przedstawimy zadanie estymacji szumów w sygnałach pomiarowych z wykorzystaniem wybranych modeli omówionych w poprzednim podrozdziale. Podstawowy model sygnału pomiarowego, który będzie wykorzystywany w tej części, ma następującą postać:

$$y(k) = s(k) + v(k), \quad k = 1, 2, \dots, K \quad (5.77)$$

gdzie $s(k)$ jest nieznanym sygnałem, a $v(k)$ reprezentuje zakłócenie.

Z rozdziału 3 wiemy, że sygnał pomiarowy może zawierać np. składowe wolnozmiennne (niskoczęstotliwościowe) oraz szybkozmiennne (wysokoczęstotliwościowe). Uwzględniając to, zmodyfikujemy model (5.77) do postaci:

$$y(k) = s(k) + v_L(k) + v_H(k), \quad k = 1, 2, \dots, K \quad (5.78)$$

gdzie $v_L(k)$ jest zakłóceniem zawierającym składowe o niskich częstotliwościach, a $v_H(k)$ to zakłócenia zawierające składowe o wysokich częstotliwościach.

5.8.1. Modelowanie z wykorzystaniem transformacji falkowej

Jednym z przykładów wykorzystania transformacji falkowej (por. podrozdz. 5.7.8) jest modelowanie szumów w sygnałach pomiarowych; niech model sygnału pomiarowego będzie taki jak (5.77). Załóżmy, że dysponujemy też pomiarami sygnału $y(k)$, $k = 1, 2, \dots, K$. Zastosujemy teraz do sygnału $y(k)$ dyskretną

transformację falkową (por. (5.59)):

$$d_{j,m} = \int_{-\infty}^{\infty} y(t)\Psi_{j,m}(t)dt \quad (5.79)$$

gdzie $d_{j,m}$ reprezentują współczynniki falkowe. Dla uproszczenia zapiszmy dyskretną transformację falkową (5.79) jako [109, 217]:

$$\mathbf{d} = \mathbf{W}_{\text{WT}}\mathbf{y} \quad (5.80)$$

gdzie \mathbf{d} reprezentuje współczynniki falkowe, \mathbf{W}_{WT} jest macierzą ortonormalną spełniającą warunek Parsevala (ang. *Parseval condition*), tj. $\mathbf{W}_{\text{WT}}^T\mathbf{W}_{\text{WT}} = \mathbf{I}$, a \mathbf{y} jest wektorową reprezentacją sygnału pomiarowego $y(k)$. Stosując odwrotną operację do (5.80), tj.

$$\mathbf{y} = \mathbf{W}_{\text{WT}}^T\mathbf{d} \quad (5.81)$$

otrzymamy oryginalną postać sygnału pomiarowego.

W pracy [112] udowodniono, że optymalną estymację sygnału $s(k)$ można uzyskać, wykorzystując **operator eliminacji progowej** (ang. *thresholding*). Wy różniamy dwa główne operatory eliminacji progowej dla metod wykorzystujących transformację falkową (\mathfrak{T} oznacza próg):

$$\tilde{\mathbf{d}} = \begin{cases} \mathbf{d}, & |d| > \mathfrak{T} \\ 0, & |d| \leq \mathfrak{T} \end{cases} \quad (5.82)$$

operator eliminacji progowej twardej (ang. *hard thresholding*) oraz

$$\tilde{\mathbf{d}} = \begin{cases} \text{sgn}(d)(|d| - \mathfrak{T}), & |d| > \mathfrak{T} \\ 0, & |d| \leq \mathfrak{T} \end{cases} \quad (5.83)$$

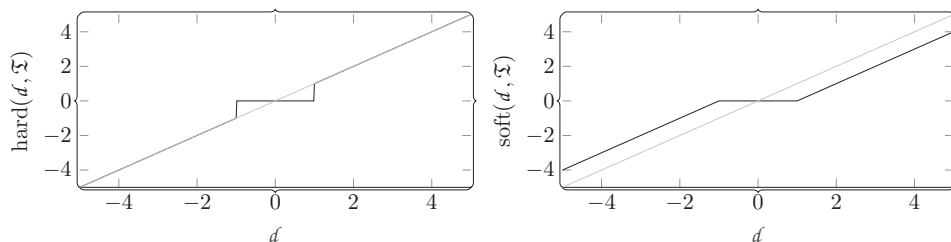
operator eliminacji progowej miękkiej (ang. *soft thresholding*). Operator eliminacji progowej twardej oznaczamy jako $\text{hard}(\mathbf{d}, \mathfrak{T})$, a operator eliminacji progowej miękkiej niech ma następujące oznaczenie $\text{soft}(\mathbf{d}, \mathfrak{T})$. Na rysunku 5.28 przedstawiono zarówno operator eliminacji progowej twardej, jak i miękkiej.

Po zastosowaniu operatora eliminacji progowej twardej lub miękkiej możliwa jest estymacja sygnału $s(k)$. W tym celu należy zastosować następującą formułę:

$$\tilde{\mathbf{s}} = \mathbf{W}_{\text{WT}}^T\tilde{\mathbf{d}}, \quad (5.84)$$

gdzie $\tilde{\mathbf{s}}$ jest wektorową reprezentacją sygnału $\tilde{s}(k)$.

5. Przetwarzanie wstępne



Rys. 5.28. Operator progowy twardy (z lewej) i miękki (z prawej)

5.8.2. Modelowanie z wykorzystaniem empirycznej dekompozycji mod (EMD)

Jedną z cech funkcji składowych IMF wyznaczanych przez dekompozycję oryginalnego sygnału $s(k)$ z wykorzystaniem algorytmu EMD jest to, że w modach o niskich indeksach skupione są szумы o wysokich częstotliwościach. Z kolei w modach o wysokich indeksach skupione są szумы o niskich częstotliwościach. Fakt ten wykorzystuje się do projektowania algorytmów filtracji bazujących na algorytmie EMD.

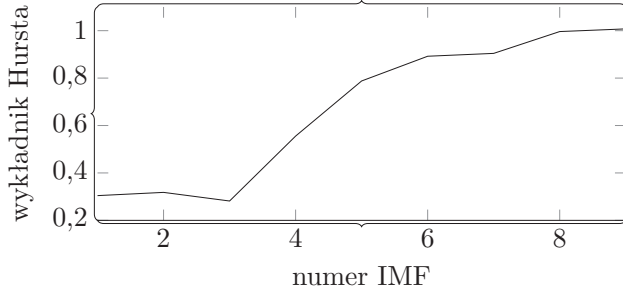
Filtracja przez częściową rekonstrukcję

Podstawowa metoda usuwania szumów z sygnałów pomiarowych, w której wykorzystuje się algorytm EMD, bazuje na omówionej właściwości funkcji składowych IMF. W metodzie tej estymacji sygnału dokonuje się z wykorzystaniem tylko części z wyznaczonych funkcji składowych IMF, tj.

$$\tilde{s}(k) = \sum_{n=n_{th}}^{N_{imf}+1} h_n(k) \quad (5.85)$$

gdzie $1 < n_{th} \leq N_{imf} + 1$ jest indeksem modu, który oddziela składowe o niskich częstotliwościach od składowych zawierających wysokie częstotliwości. W literaturze opisano wiele metod do wyznaczania progu n_{th} . Na przykład w pracach [216] oraz [455] zaproponowano podejście porównujące funkcję rozkładu prawdopodobieństwa analizowanego sygnału z funkcjami rozkładu prawdopodobieństwa każdej z mod. Z kolei w pracy [290] zaproponowano rozwiązanie bazujące na porównaniu energii każdej z mod i na tej podstawie określanie wartości n_{th} . Interesujące podejście opisano w pracy [284], gdzie wartość progu n_{th} wyznaczana jest na bazie beztrendowej analizy fluktuacyjnej (ang. *Detrended Fluctuation Analysis*, DFA). Istota tego podejścia została zilustrowana na rys. 5.29.

W metodzie bazującej na algorytmie DFA wartość n_{th} wyznaczana jest na podstawie wartości wykładnika Hursta [29] i przyjętej arbitralnie wartości progu.



Rys. 5.29. Wartość wykładnika Hursta dla każdej z mod

Eliminacja progowa

Przedstawimy teraz alternatywną metodę usuwania szumów z sygnału pomiarowego. Podejście to oryginalnie zostało opracowane na potrzeby algorytmów usuwania szumów z sygnałów pomiarowych bazujących na transformacji falkowej. Metoda ta nazywa się eliminacją progową (ang. *thresholding*) (por. 5.8.1). W podejściu tym usuwanie niepożądanych składowych z sygnałów pomiarowych wykonuje się we wszystkich wyznaczonych modach. Dwa główne podejścia wykorzystujące operator eliminacji progowej to:

$$\tilde{s}(k) = \sum_{n=1}^{N_{\text{imf}}+1} \tilde{h}_n(k) \quad (5.86)$$

oraz [455]

$$\tilde{s}(k) = \sum_{n=n_{th1}}^{n_{th2}} \tilde{h}_n(k) + \sum_{n=n_{th2}+1}^{N_{\text{imf}}} \tilde{h}_n(k) \quad (5.87)$$

gdzie

$$\tilde{h}_n(k) = \begin{cases} h_n(k) & |h_n(k)| > \mathfrak{T}_n \\ 0 & |h_n(k)| \leq \mathfrak{T}_n \end{cases} \quad (5.88)$$

dla **operatora eliminacji progowej twardej** (ang. *hard thresholding*) oraz

$$\tilde{h}_n(k) = \begin{cases} \text{sgn}(h_n(k)) (|h_n(k)| - \mathfrak{T}_n) & |h_n(k)| > \mathfrak{T}_n \\ 0 & |h_n(k)| \leq \mathfrak{T}_n \end{cases} \quad (5.89)$$

dla **operatora eliminacji progowej miękkiej** (ang. *soft thresholding*), gdzie $k = 1, 2, \dots, K$.

5. Przetwarzanie wstępne

W zależności (5.87) wprowadzono dwa dodatkowe parametry, tj. k_{th1} i k_{th2} . Są to parametry pomocnicze, dla których algorytm wyznaczania ich wartości zaproponowano np. w [217].

Bezpośrednie wykorzystanie filtracji progowej twardej lub miękkiej do funkcji IMF zwykle prowadzi do niezadowolających wyników [217, 455]. Dlatego w przypadku funkcji składowych IMF stosuje się operator eliminacji progowej w przedziałach (ang. *interval thresholding*).

Niech $z_n(m) = [z_n(m) \ z_n(m+1)]$ będzie przedziałem pomiędzy dwoma sąsiadującymi punktami przejścia analizowanego sygnału przez zero (w omawianym problemie są to funkcje IMF). Wówczas zależności (5.88) oraz (5.89) przyjmą postać [455]:

$$\tilde{h}_n(z_n(m)) = \begin{cases} h_n(z_n(m)) & |h_n(r_n(m))| > \mathfrak{T}_n \\ 0 & |h_n(r_n(m))| \leq \mathfrak{T}_n \end{cases} \quad (5.90)$$

dla **filtracji progowej twardej** oraz

$$\tilde{h}_n(z_n(m)) = \begin{cases} \operatorname{sgn}(h_n(z_n(m))) \frac{|h_n(z_n(m))| - \mathfrak{T}_n}{h_n(z_n(m))} & |h_n(r_n(m))| > \mathfrak{T}_n \\ 0 & |h_n(r_n(m))| \leq \mathfrak{T}_n \end{cases} \quad (5.91)$$

dla **filtracji progowej miękkiej**. Wielkość $r_n(m)$ oznacza indeks czasu, w którym n -ta moda przyjmuje wartość maksymalną w danym przedziale $z_n(m)$.

Wielkość $h_n(z_n(m))$ wskazuje na punkt ekstremum w zadanym przedziale. Problem wyboru progę \mathfrak{T} omówiono m.in. w pracach [217, 455].

5.8.3. Modelowanie z wykorzystaniem rzadkiej reprezentacji

W podrozdziale 5.7.7 omówiono problem wykorzystania w modelowaniu rzadkiej reprezentacji sygnałów. W tym podrozdziale zostanie z kolei przedstawiony problem usuwania szumów o niskich i wysokich częstotliwościach z wykorzystaniem podejścia bazującego na rzadkiej reprezentacji sygnałów (ang. *sparse representation*). Rozważmy na początek rozszerzony model sygnału pomiarowego w postaci:

$$y(k) = s(k) + \kappa(k) + v(k), \quad k = 1, 2, \dots, K \quad (5.92)$$

gdzie $v(k)$ jest zakłóceniem o właściwościach białego szumu, $\kappa(k)$ jest sygnałem rzadkim (por. (5.54)), a $s(k)$ to nieznany sygnał, który należy oszacować na podstawie pomiarów $y(k)$.

Zauważmy, że, bazując na modelu pomiaru (5.92), w celu oszacowania sygnału $s(k)$ należy estymować zarówno sygnał $\kappa(k)$, jak i $v(k)$. Wynika z tego, że:

$$\tilde{s}(k) = y(k) - \tilde{\kappa}(k) - \tilde{v}(k) \quad (5.93)$$

Ponieważ przyjęto, że $v(k)$ jest białym szumem, do jego usunięcia można więc wykorzystać jedną z podstawowych metod filtracji sygnałów. Problemem jest natomiast oszacowanie sygnału rzadkiego $\kappa(k)$.

Sygnał rzadki może przybliżyć różnicami pierwszego rzędu, tj. $\kappa(k) - \kappa(k-1)$ (por. (5.55)). Zadanie estymacji sygnału $\kappa(k)$ sformułujemy jako następujące zadanie optymalizacji:

$$\tilde{\kappa}(k) = \min_{\kappa} \begin{array}{l} \|\kappa(k) - \kappa(k-1)\|_1 \\ \text{p.o.} \quad \|y(k) - \bar{y}(k)\|_2^2 \leq K\sigma^2 \end{array} \quad (5.94)$$

Z kolei problem (5.94) może zostać przeformułowany do postaci:

$$\tilde{\kappa}(k) = \min_{\kappa} \left\{ \frac{1}{2} \|y(k) - \bar{y}(k)\|_2^2 + \lambda \|\mathbf{D}\kappa(k)\|_1 \right\} \quad (5.95)$$

gdzie

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \end{bmatrix} \quad (5.96)$$

a λ to mnożnik Lagrange'a.

Rozwiązaniem zadania (5.96) jest algorytm TVD (ang. *Total Variation Denoising*, TVD):

$$\tilde{\kappa}(k) = \text{tvd}(y(k), \lambda) \quad k = 1, 2, \dots, K \quad (5.97)$$

Model FLSA

W przypadku, gdy zarówno sygnał $s(k)$ oraz jego pochodna są sygnałami rzadkimi, to zadanie (5.95) sprowadza się do następującej postaci:

$$\tilde{\kappa}(k) = \min_{\kappa} \left\{ \frac{1}{2} \|y(k) - \bar{y}(k)\|_2^2 + \lambda_0 \|\kappa(k)\|_1 + \lambda_1 \|\mathbf{D}\kappa(k)\|_1 \right\} \quad (5.98)$$

gdzie λ_0 oraz λ_1 są mnożnikami Lagrange'a.

5. Przetwarzanie wstępne

Rozwiązaniem tego zadania jest algorytm:

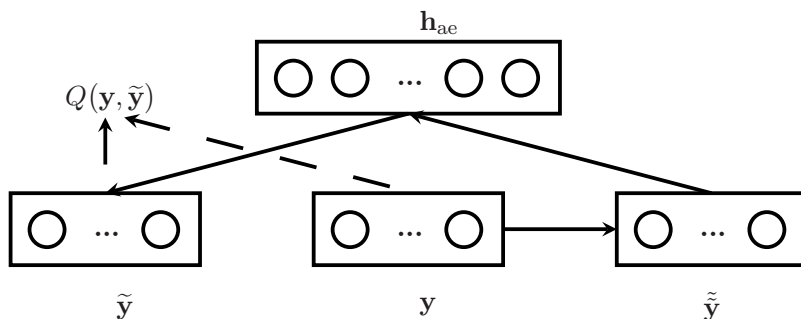
$$\tilde{\kappa}(k) = \text{soft}(\text{tvd}(y(k), \lambda_1), \lambda_0) \quad k = 1, 2, \dots, K \quad (5.99)$$

Problem (5.98) w literaturze nazywany jest modelem FLSA (ang. *Fused Lasso Signal Approximator*). Funkcja soft jest operatorem eliminacji progowej miękkiej (por. podrozdz. 5.8.1).

5.8.4. Modelowanie z wykorzystaniem uczenia głębokiego

Przedstawimy teraz inne, również nowe, podejście do modelowania szumów bazujące na łącznym podejściu, w którym wykorzystano metody uczenia głębokiego oraz transformację falkową. Przedstawione podejście zostało zaproponowane w pracy [451]. W pracy zastosowano głęboką sieć (ang. *Deep Neural Network*, DNN) bazującą na **autoenkoderach odsumiających** (ang. *denoising autoencoder*). Sieć ta została wykorzystana w omawianym podejściu do wspomaganie metody usuwania szumów z pomiarów, która bazuje na transformacji falkowej. Podejście do usuwania niepożądanych składowych z sygnałów pomiarowych zostało omówione w podrozdziale 5.8.1. Z kolei zagadnienia związane z wykorzystaniem sieci głębokich w zadaniach modelowania omówiono w podrozdziale 5.7.5.

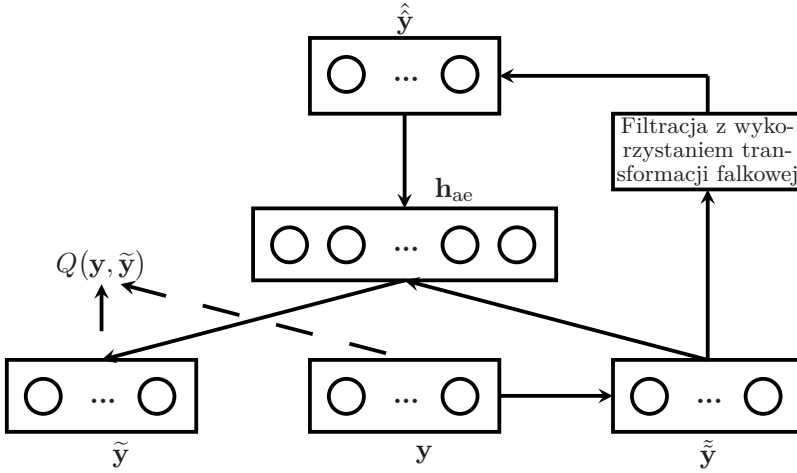
Architektura autoenkodera odsumiającego została zilustrowana na rys. 5.30. Szczegóły dotyczące tej metody zamieszczono m.in. w [141, 450].



Rys. 5.30. Architektura autoenkodera odsumiającego

Bazując na architekturze autoenkodera odsumiającego (rys. 5.30), na rys. 5.31 przedstawiono jego zmodyfikowaną wersję, w której jedną z głównych części jest algorytm transformacji falkowej. Przyjmijmy następujący model sygnału pomiarowego:

$$y(k) = s(k) + v(k), \quad k = 1, 2, \dots, K \quad (5.100)$$



Rys. 5.31. Zmodyfikowana architektura autoenkodera odzsumiającego z wykorzystaniem transformacji falkowej

Niech teraz K -elementowy sygnał $y(k)$ będzie reprezentowany przez wektor, tj. $\mathbf{y} \in \mathbb{R}^K$. Wprowadźmy również funkcję \mathcal{T} , która:

$$\tilde{\mathbf{y}} = \mathcal{T}(\mathbf{y}) \tag{5.101}$$

gdzie $\mathcal{T} : \mathbb{R}^K \rightarrow \mathbb{R}^K$. Głównym zadaniem funkcji \mathcal{T} jest zaszumienie sygnału pomiarowego $y(k)$. Kolejnym krokiem w zaproponowanym rozwiązaniu jest wykonanie dyskretnej transformacji falkowej na sygnale, który reprezentowany jest przez wektor $\tilde{\mathbf{y}}$ (por. (5.101)), w celu wyznaczenia zestawu współczynników falkowych:

$$\mathbf{d} = \mathbf{W}_{\text{WT}} \tilde{\mathbf{y}} \tag{5.102}$$

i zastosowanie operatora eliminacji progowej miękkiej:

$$\tilde{\mathbf{d}} = \begin{cases} \text{sgn}(d) (|d| - \mathfrak{T}), & |d| > \mathfrak{T} \\ 0, & |d| \leq \mathfrak{T} \end{cases} \tag{5.103}$$

W następnym kroku wykonujemy operację odtwarzania sygnału na podstawie zmodyfikowanych współczynników transformacji falkowej, korzystając z następującej zależności (por. (5.81)):

$$\hat{\mathbf{y}} = \mathbf{W}_{\text{WT}}^T \tilde{\mathbf{d}} \tag{5.104}$$

5. Przetwarzanie wstępne

Kolejny etap to wygenerowanie wektora cech, na podstawie sygnału $\hat{\mathbf{y}}$, z zależności:

$$\mathbf{h}_{ae} = \mathcal{G}_{\text{kod}}(\mathbf{W}_{aeh} \cdot \hat{\mathbf{y}} + \mathbf{b}_h) \quad (5.105)$$

gdzie \mathbf{W}_{aeh} to macierz wag dla warstw ukrytych, \mathbf{b}_h jest wektorem wyrazów wolnych (ang. *bias*) dla warstw ukrytych, a \mathcal{G}_{kod} jest funkcją kodującą.

W kolejnym kroku należy odtworzyć, na podstawie wyznaczonych cech (5.105), sygnał wejściowy:

$$\tilde{\mathbf{y}} = \mathcal{G}_{\text{dek}}(\mathbf{W}_{aeo} \cdot \mathbf{h}_{ae} + \mathbf{b}_o) \quad (5.106)$$

gdzie $\tilde{\mathbf{y}}$ jest estymowanym na podstawie wygenerowanych cech \mathbf{h}_{ae} sygnałem wejściowym $y(k)$, \mathbf{W}_{aeo} jest macierzą wag dla warstw wyjściowych, a \mathbf{b}_o jest wektorem obciążeń dla tej warstwy, natomiast \mathcal{G}_{dek} reprezentuje funkcję dekodującą.

Parametry autoenkodera odszumiającego wyznaczymy, rozwiązując następujące zadanie optymalizacyjne:

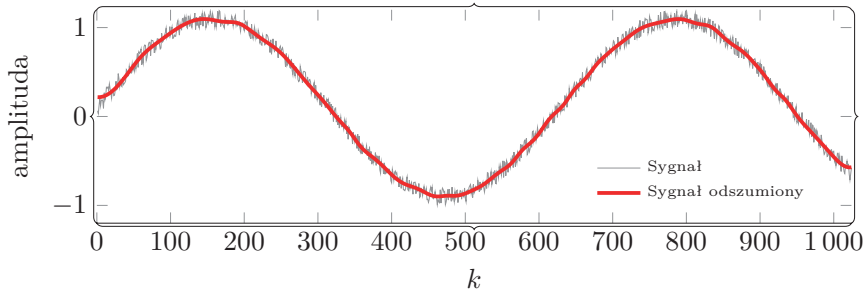
$$\begin{aligned} \{\mathbf{W}_{aeh}^*, \mathbf{W}_{aeo}^*, \mathbf{b}_h^*, \mathbf{b}_o^*\} &= \min_{\mathbf{W}_{aeh}, \mathbf{W}_{aeo}, \mathbf{b}_h, \mathbf{b}_o} Q(\mathbf{y}, \tilde{\mathbf{y}}) = \\ &= Q(\mathbf{y}, \mathcal{G}_{\text{dek}}(\mathbf{W}_{aeo} \cdot \mathcal{G}_{\text{kod}}(\mathbf{W}_{aeh} \cdot \hat{\mathbf{y}} + \mathbf{b}_h) + \mathbf{b}_o)) \end{aligned} \quad (5.107)$$

5.9. Usuwanie składowych o wysokich częstotliwościach

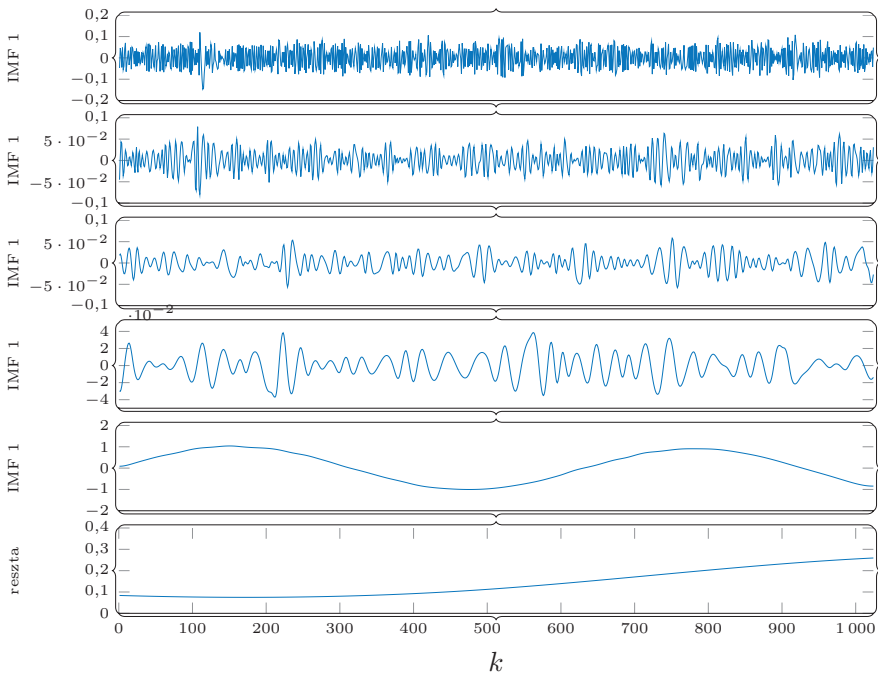
Przedstawimy teraz przykłady zastosowania omówionych metod do usuwania niepożądanych składowych z sygnałów pomiarowych. Pierwszy przykład dotyczy zadania usuwania z sygnału składowych o wysokich częstotliwościach. Przedstawiona metoda będzie bazowała na algorytmie EMD. W przykładzie rozważmy dwa sygnały testowe. Pierwszy z nich ma przebieg taki jak funkcja *sinus* (rys. 5.32). Drugim z przykładowych sygnałów jest sygnał *dopplerowski*, którego przykładowy przebieg przedstawiono na rys. 5.35. Zarówno w jednym, jak i w drugim przypadku do sygnałów został dodany biały szum *gaussowski*.

W metodzie bazującej na algorytmie EMD pierwszy etap związany jest z dekompozycją oryginalnego sygnału pomiarowego ($y(k)$) na zbiór mod i sygnał rezydualny (por. z zależn. (5.67) oraz (5.68)). Dla pierwszego z przykładowych sygnałów efekt dekompozycji sygnału z wykorzystaniem algorytmu EMD przedstawiono na rys. 5.33. Po wyznaczeniu mod kolejnym krokiem w omówionym podejściu bazującym na wyborze odpowiednich mod jest ustalenie, które mody zawierają sygnał, a które składowe niepożądane (szumy szybkozmienne). W tym celu wykorzystaliśmy metodę opartą na estymacji wykładnika Hursta (por. zależn. (5.85)).

5.9. Usuwanie składowych o wysokich częstotliwościach



Rys. 5.32. Wynik usunięcia zakłóceń o wysokich częstotliwościach z sygnału *sinus*

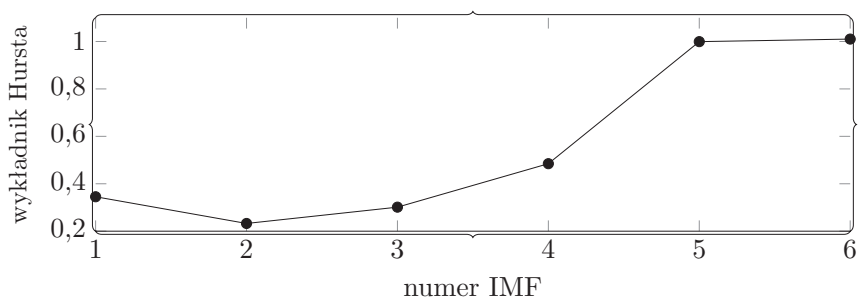


Rys. 5.33. Wynik dekompozycji sygnału na kolejne funkcje IMF dla pierwszego z sygnałów (funkcja *sinus*)

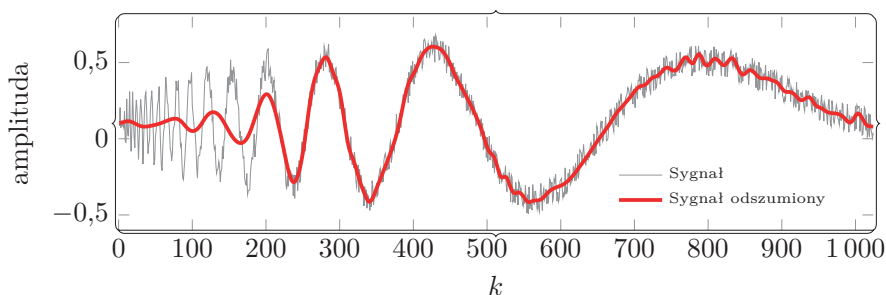
Wyznaczenie wykładnika Hursta dla każdej z mod umożliwia konstrukcję wykresu, którego przykład przedstawiono na rys. 5.34.

Wybierając wyznaczone funkcje IMF, które będą wykorzystane do odtworzenia sygnału, dokonujemy analizy wykładników Hursta. Do estymacji sygnału wybieramy te mody, które mają wysoką wartość wykładnika Hursta. Wartość progu (n_{th}),

5. Przetwarzanie wstępne



Rys. 5.34. Wartość wykładnika dla kolejnych funkcji IMF dla pierwszego z sygnałów (funkcja *sinus*)



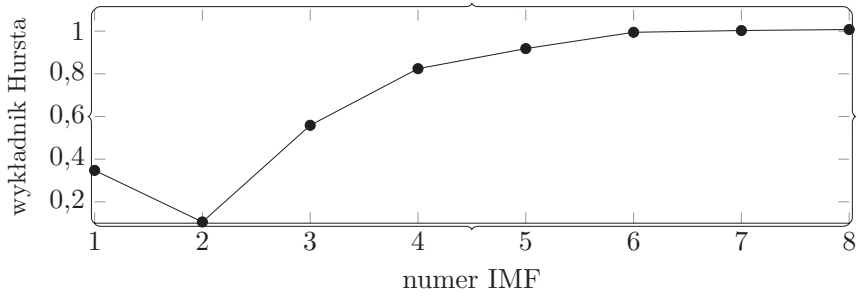
Rys. 5.35. Wynik usunięcia zakłóceń o wysokich częstotliwościach z sygnału dopplerowskiego

który oddzieli mody zawierające sygnał od tych zawierających szumy należy dobrać eksperymentalnie. Po wybraniu odpowiednich mod zastosowano zależność (5.85) w celu estymacji sygnału $s(k)$. Wynik dla tego przykładu przedstawiono na rys. 5.32.

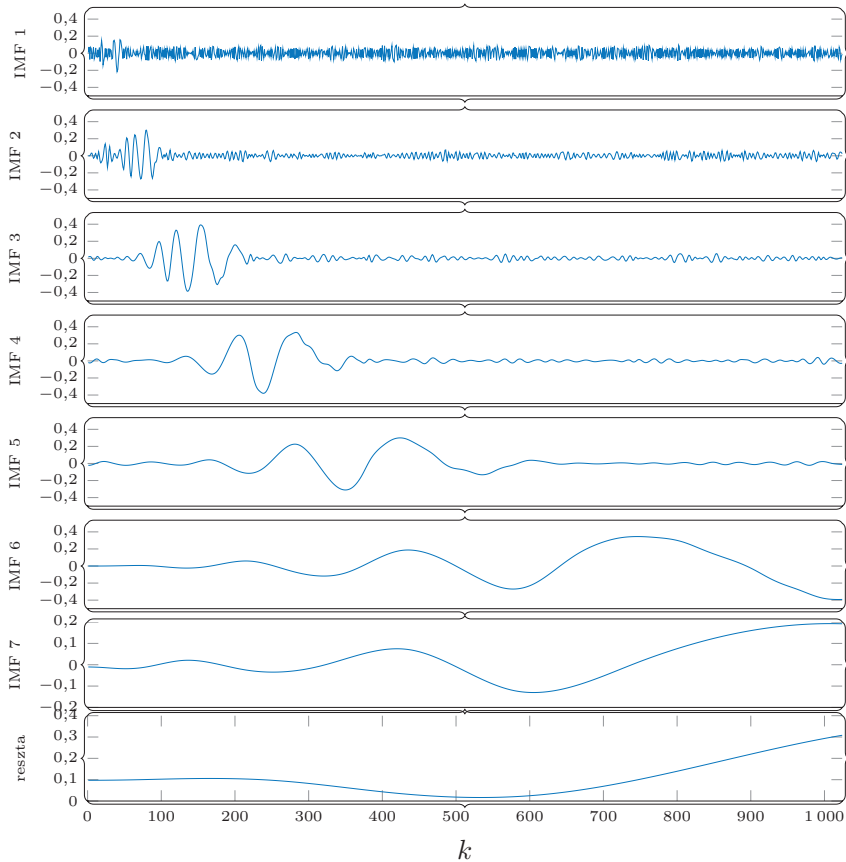
Opisane postępowanie zastosowaliśmy również do drugiego z sygnałów (rys. 5.35). Wynik dekompozycji sygnału pomiarowego został przedstawiony na rys. 5.37, natomiast wartość wykładnika Hursta dla każdej z wyznaczonych mod przedstawiono na rys. 5.36. Po dobraniu odpowiedniej wartości progu dla wykładnika Hursta wyznaczono wartość n_{th} , a następnie wykorzystano zależność (5.85) w celu estymacji sygnału $s(k)$ z pomiarów $y(k)$.

Analizując otrzymane wyniki usunięcia składowych o wysokich częstotliwościach z przetworzonego sygnału doplerowskiego zauważamy, że jakość estymacji oryginalnego sygnału jest bardzo słaba, gdy częstotliwość sygnału jest wysoka. Estymacja poprawia się, gdy częstotliwość sygnału dopplerowskiego zmniejsza się.

5.9. Usuwanie składowych o wysokich częstotliwościach



Rys. 5.36. Wartość wykładnika dla kolejnych funkcji IMF dla sygnału dopplerowskiego



Rys. 5.37. Wynik dekompozycji sygnału na kolejne funkcje IMF dla sygnału dopplerowskiego

6. Poziom podstawowy: fuzja danych

System nerwowy i maszyna automatyczna są fundamentalnie do siebie podobne w tym, że podejmują decyzje na bazie decyzji podjętych w przeszłości.

Norbert Wiener

6.1. Wstęp

Poprzedni rozdział został poświęcony problematyce wstępnego przetwarzania danych pomiarowych. Zagadnienie to jest istotne głównie z tego powodu, że wstępne przetworzenie danych ułatwia przygotować zebrane dane do ich przetwarzania na kolejnych etapach będących właściwymi etapami fuzji danych.

W kolejnych trzech rozdziałach omówione zostaną zagadnienia dotyczące fuzji danych strumieniowych. Rozłożenie zagadnień fuzji danych strumieniowych na trzy rozdziały ma bezpośredni związek z jednym z typowych podejść do jej podziału, tj. podziału ze względu na poziom abstrakcji. Dlatego kolejne trzy rozdziały będą dotyczyły zagadnień związanych z fuzją danych (poziom 1), fuzją właściwości (poziom 2) oraz fuzją decyzji (poziom 3). Każdy z tych poziomów został szczegółowo scharakteryzowany w rozdziale 2.

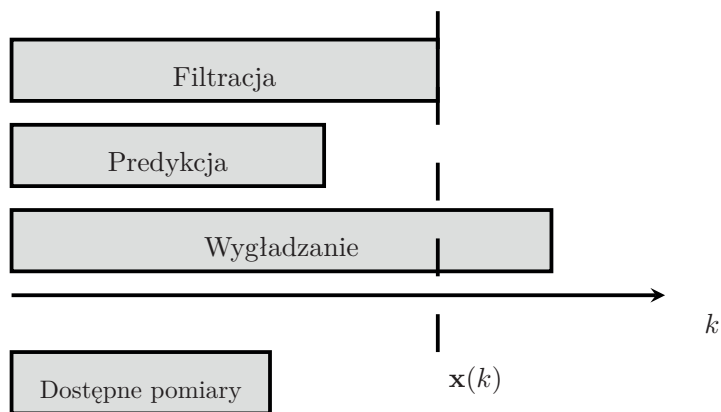
Niniejszy rozdział poświęcony został pierwszemu z wymienionych typów fuzji, tj. fuzji danych. Fuzja danych należy do metod bezpośrednich (ang. *direct fusion*). W literaturze ten rodzaj fuzji danych nazywany jest również fuzją na poziomie podstawowym (ang. *low-level fusion*). Cechą charakterystyczną fuzji danych na tym poziomie jest to, że fuzji mogą podlegać jedynie **dane jednomodalne**. Jednomodalność danych oznacza, że są to dane związane z pomiarami wykonywanymi dla tego samego zjawiska. Przy czym do pomiarów tego zjawiska mogą być wykorzystane czujniki różnego typu. Przykładem tego są dane z czujników przyspieszenia i żyroskopowych wykorzystywane do estymacji położenia obiektu w przestrzeni [337].

6. Poziom podstawowy: fuzja danych

Głównym zadaniem fuzji danych na poziomie podstawowym jest **estymacja stanu**. Estymacja stanu ma na celu odtworzenie na podstawie zakłóconych sygnałów pomiarowych pewnego przybliżenia przebiegu prawdziwego sygnału [281]. Zauważmy, że z zadaniem estymacji stanu związanych jest wiele dodatkowych zadań. Do najważniejszych z nich należą zagadnienia związane z modelowaniem matematycznym procesów dynamicznych, układów pomiarowych oraz procesów akwizycji danych. Inne, równie istotne, zagadnienia, które należy rozważyć na tym poziomie, związane są z metodami doboru kryteriów i oceny jakości estymacji oraz algorytmami przetwarzania zakłóconych sygnałów pomiarowych.

6.2. Estymacja stanu

W ogólności estymacja stanu dotyczy trzech zagadnień, tj. **filtracji optymalnej** (filtracja w trybie na bieżąco), **predykcji** i **wygładzania** (filtracja w trybie wsadowym). Zasadnicza różnica pomiędzy tymi zadaniami związana jest z relacją pomiędzy danymi pomiarowymi a stanem. Na rysunku 6.1 zilustrowano różnicę pomiędzy tymi zadaniami.



Rys. 6.1. Trzy typy problemów estymacji stanu

W pierwszym przypadku, tj. w zadaniu filtracji, estymacji stanu dokonujemy na podstawie wszystkich dostępnych pomiarów. Oznacza to, że dokonujemy estymacji **stanu bieżącego**. Z kolei zadanie predykcji polega na estymacji **przyszłych stanów** tzn. tych, dla których nie dysponujemy żadnymi pomiarami. Z kolei wygładzanie związane jest z estymacją **przeszłych stanów**.

Jak można zauważyć, istotnymi dla tego rozdziału pojęciami będą pojęcia **stanu** oraz **zmiennych stanu**. Wprowadźmy teraz definicję stanu [191, 317]:

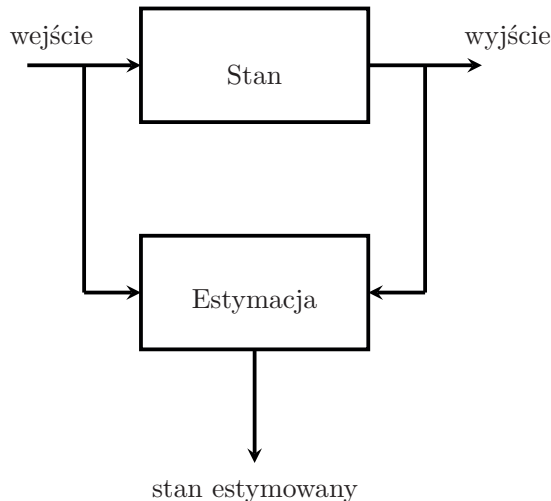
Definicja 6.2.1 (Stan). *Stan to najmniejszy zbiór wielkości określających w pełni skutki przeszłych oddziaływań na system, który jest wystarczający do wyznaczenia zachowania systemu w przyszłości dla znanych wymuszeń oraz dla znanego opisu systemu. Stan możemy przedstawić w formie następującego wektora:*

$$\mathbf{x}(k) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\mathcal{D}_x} \end{bmatrix} \quad (6.1)$$

gdzie \mathcal{D}_x to liczba składowych wektora, a k jest indeksem pomiaru. Składowe wektora stanu nazywamy zmiennymi stanu.

Zmienne stanu określają \mathcal{D}_x -wymiarową **przestrzeń stanu**. Punkt w przestrzeni stanu dla chwili k nazywa się **punktem bieżącym**. Stan systemu może zostać odtworzony na podstawie danych pomiarowych. W tym celu należy rozwiązać zadanie estymacji [281].

Definicja 6.2.2 (Estymacja). *Estymacja polega na określeniu, na podstawie zakłóconych pomiarów, przybliżenia wartości pewnej nieznannej wielkości.*



Rys. 6.2. Ilustracja zadania estymacji

6. Poziom podstawowy: fuzja danych

Przedstawimy teraz ogólny schemat postępowania dla zadania estymacji stanu obiektu. Niech stan, tak jak ustaliliśmy, będzie oznaczony przez $\mathbf{x}(k)$. Wejście oznaczmy natomiast przez $\mathbf{u}(k)$:

$$\mathbf{u}(k) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{\mathcal{D}_u} \end{bmatrix} \quad (6.2)$$

gdzie \mathcal{D}_u oznacza liczbę składowych wektora. Wyjście systemu (rys. 6.2) oznaczmy z kolei przez $\mathbf{y}(k)$:

$$\mathbf{y}(k) = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\mathcal{D}_y} \end{bmatrix} \quad (6.3)$$

gdzie \mathcal{D}_y jest liczbą składowych tego wektora. Podsumowując, dysponując zakłóconymi pomiarami sygnałów wejściowych $\mathbf{u}(k)$ oraz wyjściowych $\mathbf{y}(k)$, estymujemy stan obiektu $\mathbf{x}(k)$.

Zdefiniujmy jeszcze czym jest optymalny estymator stanu [134].

Definicja 6.2.3 (Optymalny estymator stanu). *Optymalny estymator stanu to algorytm przetwarzający dane pomiarowe w celu oszacowania stanu obiektu, w którym wykorzystywana jest wiedza o systemie oraz dynamice toru pomiarowego z jednoczesnym założeniem statystyk dotyczących tak szumów systemowych (tj. dynamicznych), jak i pomiarowych oraz przyjętych warunków początkowych.*

6.2.1. Model w przestrzeni stanu

Pierwszym krokiem w zadaniu estymacji jest przygotowanie odpowiednich modeli matematycznych. Konstrukcja modeli dla konkretnego zadania, poza prostymi przypadkami, jest zadaniem trudnym. Jest to spowodowane tym, że rzeczywiste obiekty są złożone, a wiedza na ich temat może okazać się niewystarczająca.

Podstawowym modelem wykorzystywanym w estymacji stanu jest model w przestrzeni stanów (por. z (5.75)). W modelu tym wyróżniamy dwa główne elementy, tj. **model systemu** (model dynamiki systemu) (ang. *system dynamic model*) oraz **model układu pomiarowego** (ang. *measurement model*).

Model systemu

Przyjmijmy, że znamy całą historię analizowanego obiektu, tj. znamy wszystkie poprzednie wartości zmiennej stanu $\mathbf{x}(k)$, tj. dysponujemy następującym zestawem pomiarów $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)$. Teraz, korzystając z przeszłych danych pomiarowych zmiennej stanu $\mathbf{x}(k)$, wyznaczmy kolejną wartość tej zmiennej, tj. $\mathbf{x}(K+1)$. Ogólną postać modelu systemu można zapisać w postaci (por. (3.8)):

$$\mathbf{x}(k+1) = f_n(\mathbf{x}(k), \mathbf{n}(k), k) \quad (6.4)$$

Zadanie wyznaczania następnych wartości zmiennych stanu \mathbf{x} można rozwiązać m.in. w modelu probabilistycznym. W tym celu przyjmijmy, że stan systemu jest zmienną losową. Po tym ustaleniu kolejny krok to wyznaczenie rozkładu warunkowego w następującej formie $P(\mathbf{x}(K+1)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$. Zaznaczmy, że znajomość tego rozkładu umożliwiłaby wyznaczenie następnych wartości zmiennej stanu, tj. $\mathbf{x}(k+1)$. Niestety w przypadku rzeczywistych problemów i danych pomiarowych wyznaczenie tego rozkładu warunkowego jest zadaniem trudnym, a w niektórych przypadkach może okazać się niemożliwe.

Przedstawione zadanie można natomiast rozwiązać, jeśli zostanie przyjęty tzw. **warunek Markowa** (ang. *Markov condition*):

$$P(\mathbf{x}(K+1)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)) = P(\mathbf{x}(K+1)|\mathbf{x}(K)) \quad (6.5)$$

Prawdopodobieństwo warunkowe (6.5) nazywane jest **prawdopodobieństwem przejścia** (ang. *transition probability density*). Z kolei z przyjętego warunku Markowa wynika, że do wyznaczenia przyszłych wartości zmiennej stanu $\mathbf{x}(k)$ nie potrzebna jest znajomość jej wszystkich przeszłych wartości. Po przyjęciu tego warunku obliczenia bieżącej wartości wektora stanu $\mathbf{x}(k+1)$ wystarczy oprzeć na wartości z chwili bezpośrednio ją poprzedzającej.

Znając prawdopodobieństwo przejścia (6.5) oraz początkowy rozkład $P(\mathbf{x}(1))$, możliwe jest wyznaczenie rozkładu prawdopodobieństwa z następującej zależności:

$$P(\mathbf{x}(k+1)) = \int_{\mathbf{x}(k)} P(\mathbf{x}(k+1)|\mathbf{x}(k))P(\mathbf{x}(k))d\mathbf{x} \quad (6.6)$$

Łączny rozkład prawdopodobieństwa dla sekwencji $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)$ wyznaczany jest z wykorzystaniem następującej zależności:

$$P(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)) = P(\mathbf{x}(0)) \prod_{k=1}^K P(\mathbf{x}(k)|\mathbf{x}(k-1)) \quad (6.7)$$

Model układu pomiarowego

Niestety w większości przypadków nie jest możliwe bezpośrednie obserwowanie stanów obiektu. Natomiast możliwe jest obserwowanie wyjść obiektu, które powiązane są ze zmiennymi stanu. Z kolei by możliwa była obserwacja stanu obiektu, konieczne jest ustalenie modelu dla układu pomiarowego. Model układu pomiarowego może być w pełni opisany przez zastosowanie warunkowego rozkładu prawdopodobieństwa, tj. $P(\mathbf{y}(K)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K); \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K-1))$. W modelu tym zmienne $\mathbf{y}(k)$ reprezentują pomiary wyjścia obserwowanego obiektu. Ogólna postać modelu układu pomiarowego ma następującą postać (patrz (3.8)):

$$\mathbf{y}(k) = h_v(\mathbf{x}(k), \mathbf{v}(k), k) \quad (6.8)$$

Przyjmując brak pamięci układu pomiarowego, probabilistyczny model dla tego układu ma następującą postać:

$$P(\mathbf{y}(K)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K); \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K-1)) = P(\mathbf{y}(K)|\mathbf{x}(K)) \quad (6.9)$$

Omówione modele, tj. modele systemu i układu pomiarowego można podsumować i przedstawić w postaci **probabilistycznego modelu w przestrzeni stanu**.

Definicja 6.2.4 (Probabilistyczny model w przestrzeni stanów). *Probabilistyczny model w przestrzeni stanów to model będący sekwencją następujących warunkowych rozkładów prawdopodobieństwa:*

$$\mathbf{x}(k+1) \sim P(\mathbf{x}(k+1)|\mathbf{x}(k)) \quad (6.10a)$$

$$\mathbf{y}(k) \sim P(\mathbf{y}(k)|\mathbf{x}(k)) \quad (6.10b)$$

dla $k = 1, 2, \dots, K$.

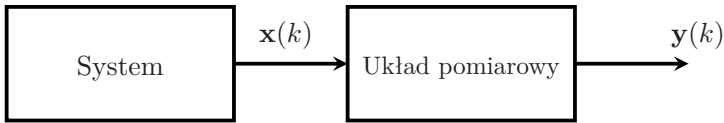
Zauważmy, że probabilistyczny model w przestrzeni stanów może być traktowany jako alternatywa opisu (5.75). Model wprowadzony wcześniej, tj. model (5.75), służy do estymacji wielkości skalarnej lub wektorowej. Z kolei model (6.10a) i (6.10b) wykorzystuje się do szacowania funkcji rozkładu prawdopodobieństwa zmiennej stanu. Model ten jest ogólny i w łatwy sposób może zostać zaadaptowany do różnych zadań modelowania.

6.2.2. Zadanie estymacji stanu

Dysponując probabilistycznym modelem w przestrzeni stanu (6.10a) i (6.10b) oraz pomiarami wyjścia $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K-1)$, sformułujemy, a następnie rozwiążemy zadanie estymacji stanu.

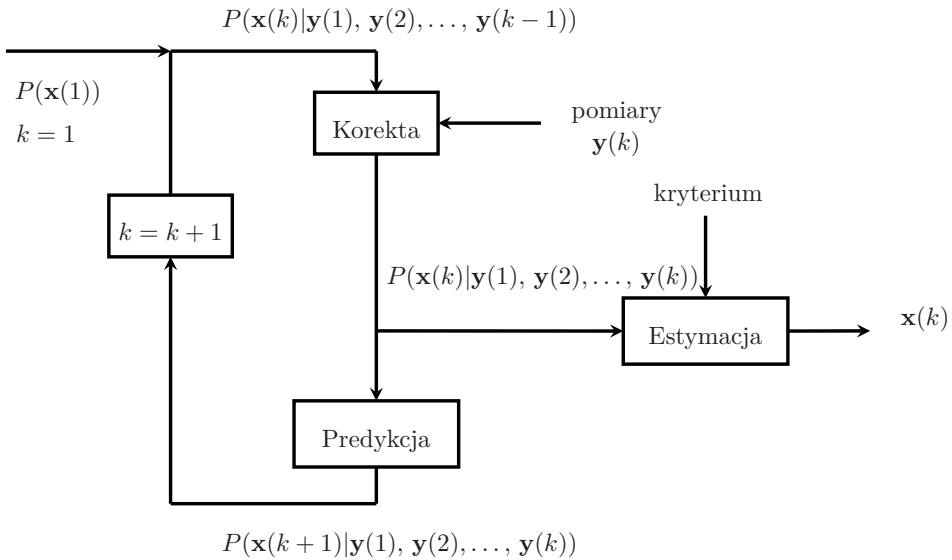
Sformułowanie zadania

Rozważmy pewien system, którego stan jest opisany zmienną losową. Dla tego systemu rozważmy zadanie estymacji wartości wektora stanu $\mathbf{x}(K)$, która nie jest pomiarowo dostępna. Załóżmy jednocześnie, że znany jest zbiór pomiarów $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)$, których wartości są powiązane ze zmienną stanu $\mathbf{x}(K)$ przez pewien układ pomiarowy (rys. 6.3). Zadaniem jest znalezienie wartości wektora $\mathbf{x}(K)$ z wykorzystaniem tych pomiarów, tj. przez ustalenie $P(\mathbf{x}(K)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K-1))$ [281].



Rys. 6.3. Schemat systemu i układu pomiarowego

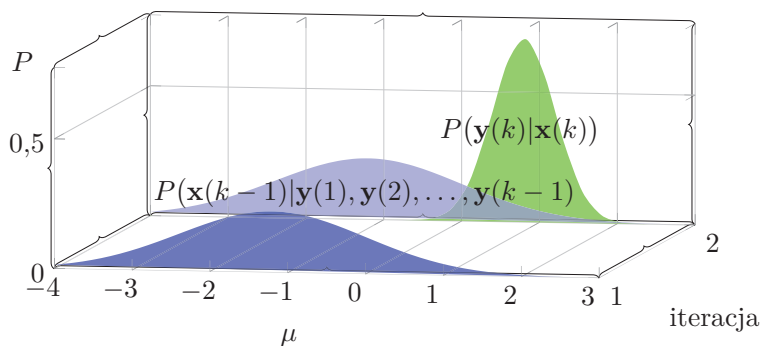
Rozwiązanie ogólne



Rys. 6.4. Ogólny schemat rozwiązania zadania predykcji (na podst. [416])

Na rysunku 6.4 przedstawiono schemat rozwiązania zadania estymacji wartości wektora stanu. Pierwszym etapem na przedstawionym schemacie jest **predykcja**. W kroku tym wiedza dotycząca wektora stanu $\mathbf{x}(k)$ jest ekstrapolowana na przyszłą

6. Poziom podstawowy: fuzja danych



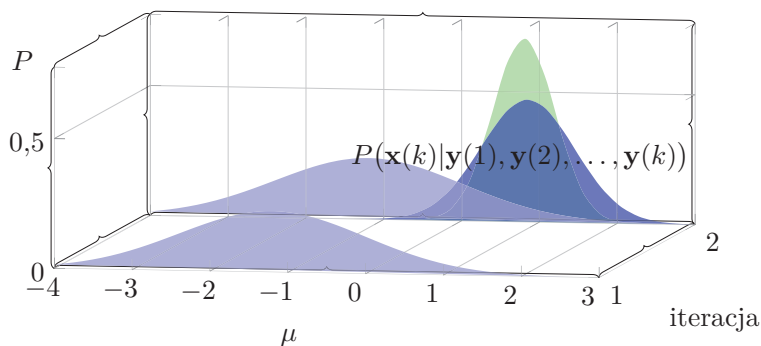
Rys. 6.5. Ilustracja kroku predykcji

wartość tego wektora, tj. $\mathbf{x}(k+1)$. Stąd, korzystając z *bayesowskiego* twierdzenia o prawdopodobieństwie warunkowym, otrzymujemy:

$$\begin{aligned}
 P(\mathbf{x}(k+1)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k)) &= \\
 &= \int_{\mathbf{x}(k)} P(\mathbf{x}(k+1)|\mathbf{x}(k))P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k))d\mathbf{x}(k) \quad (6.11)
 \end{aligned}$$

Krok predykcji został przedstawiony na rys. 6.5. Zmienna μ na wykresie oznacza wartość oczekiwaną wartości zmiennej stanu.

Drugim krokiem jest **korekcja**, w kroku tym wiedza pozyskana z pomiaru $\mathbf{y}(k)$ jest wykorzystana do poprawy funkcji opisującej warunkową gęstość prawdopodo-



Rys. 6.6. Ilustracja kroku korekcji

bieństwa:

$$P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k)) = \frac{1}{c} P(\mathbf{y}(k)|\mathbf{x}(k)) P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1)) \quad (6.12)$$

gdzie c jest następującą stałą:

$$c = \int_{\mathbf{y}(k)} P(\mathbf{x}(k)|\mathbf{x}(k)) P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1)) d\mathbf{x}(k) \quad (6.13)$$

Krok korekcji został przedstawiony na rys. 6.6.

6.2.3. Filtr Bayesa

Bazując na wynikach rozważań z poprzedniego podrozdziału przedstawimy teraz algorytm estymacji wartości wektora stanu (a dokładniej filtracji na bieżąco). Algorytm ten nazywany jest w **filtrze Bayesa** (patrz algorytm 1). Filtr Bayesa jest ogólnym podejściem do estymacji wartości wektora stanu. W podejściu tym wyróżniamy dwa etapy, tj. etap predykcji oraz etap korekcji. Celem predykcji w filtrze Bayesa jest oszacowanie przyszłej wartości stanu $\mathbf{x}(k+1)$ na podstawie stanu bieżącego, tj. $\mathbf{x}(k)$. Z kolei korekcja służy do poprawienia predykcji z etapu poprzedniego z wykorzystaniem aktualnego pomiaru $\mathbf{y}(k)$.

Zauważmy, że w przypadku ciągłego stanu przedstawiony algorytm estymacji wartości wektora stanu jest trudny do praktycznej realizacji. Głównym powodem trudności obliczeniowych jest konieczność estymacji wielowymiarowej funkcji rozkładu gęstości prawdopodobieństwa.

Podkreślmy jeszcze, że przedstawiony schemat może być wykorzystany w zadaniu estymacji wartości wektora stanu w przypadku dyskretnym. W celu przekształcenia algorytmu estymacji ciągłego wektora stanu na jego dyskretny odpowiednik konieczna jest zamiana w filtrze Bayesa (alg. 1), występujących we wzorach całek na sumy.

Filtr Bayesa, jako ogólne podejście do rozwiązania zadania estymacji stanu, stał się podstawą do konstrukcji filtrów, które znalazły szerokie zastosowanie w zagadnieniach praktycznych. Do najbardziej znanych filtrów wykorzystujących koncepcję filtra Bayesa należą **filtr Kalmana**, **rozszerzony filtr Kalmana**, **bezsładowy filtr Kalmana** oraz **filtr cząsteczkowy**.

6.3. Filtracja optymalna – stan ciągły

Opiszemy teraz zadanie estymacji stanu dla przypadku ciągłego. Do rozważań przyjmujemy model, który jest liniowy zarówno w części opisującej dynamikę

Algorytm 1: Filtr Bayesa

Wejście : zestaw pomiarów $(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$ **while** nie spełniono warunku stopu **do** **predykcja:**

$$\begin{aligned}
 P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1)) &= \\
 &= \int_{\mathbf{x}(k)} P(\mathbf{x}(k)|\mathbf{x}(k-1))P(\mathbf{x}(k-1)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1))d\mathbf{x}(k)
 \end{aligned} \tag{6.14}$$

korekcja:

$$\begin{aligned}
 P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k)) &= \\
 &= \frac{P(\mathbf{y}(k)|\mathbf{x}(k))p(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1))}{\int_{\mathbf{x}(k)} P(\mathbf{y}(k)|\mathbf{x}(k))P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1))d\mathbf{x}(k)}
 \end{aligned} \tag{6.15}$$

end**Wyjście :** sekwencja stanów $(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$

systemu, jak i w części dotyczącej układu pomiarowego. Model systemu i układu pomiarowego uzupełniamy o addytywny biały szum *gaussowski*.

Model systemu

Niech model dynamiki systemu liniowego z addytywnym białym szumem *gaussowskim* ma postać:

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{n}(k) \tag{6.16}$$

gdzie \mathbf{A} jest macierzą systemową, $\mathbf{x}(k)$ reprezentuje z kolei stan obiektu, $\mathbf{n}(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(k))$ jest szumem systemowym. Macierz $\mathbf{Q}(k)$ nazywana jest **macierzą kowariancji** szumu systemowego. Początkowe wartości macierzy kowariancji $\mathbf{Q}(k)$ należy ustalić przed rozpoczęciem obliczeń. Przyjmując założenie o białym szumie systemowym możemy przyjąć do dalszych obliczeń, że wartość oczekiwana zakłóceń jest równa:

$$E[\mathbf{n}(k)] = \mathbf{0} \tag{6.17}$$

a autokorelacja:

$$E[\mathbf{n}(i)\mathbf{n}^\top(j)] = \mathbf{Q}(i)\delta_K(i, j) \tag{6.18}$$

gdzie $\delta_K(i, j)$ jest deltą Kroneckera.

W celu oszacowania wartości wektora stanu $\mathbf{x}(k)$ w kolejnych chwilach należy wyliczyć jego estymaty *a priori* (zależn. (6.16)) oraz macierz kowariancji błędu procesu odtwarzania $\mathbf{P}(k)$:

$$\begin{aligned} E[\mathbf{x}(k+1)] &= \mathbf{A}(k)E[\mathbf{x}(k)] \\ \mathbf{P}(k+1) &= \mathbf{A}(k)\mathbf{P}(k)\mathbf{A}^\top(k) + \mathbf{Q}(k) \end{aligned} \quad (6.19)$$

Mając określony model opisujący dynamikę systemu z szumem *gaussowskim*, kolejnym etapem jest konstrukcja modelu układu pomiarowego.

Model układu pomiarowego

Przyjmijmy następujący model układu pomiarowego:

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k) \quad (6.20)$$

gdzie $\mathbf{v}(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}(k))$ jest zakłóceniem pomiarowym, macierz $\mathbf{R}(k)$ reprezentuje macierz kowariancji dla szumów pomiarowych, a $\mathbf{H}(k)$ to macierz pomiarowa. Zaznaczmy, że początkowe wartości macierzy kowariancji \mathbf{R} należy ustalić przed rozpoczęciem obliczeń.

Podobnie jak dla szumów systemowych, również w przypadku szumów pomiarowych (tj. zakłóceń) przyjmujemy, że mają one charakter addytywnego *gaussowskiego* szumu białego, co oznacza, że zachodzą dla niego:

$$\begin{aligned} E[\mathbf{v}(k)] &= \mathbf{0} \\ E[\mathbf{v}(i)\mathbf{v}^\top(j)] &= \mathbf{R}(i)\delta_K(i, j) \end{aligned} \quad (6.21)$$

Jednocześnie przyjmuje się, że szumy systemowe i pomiarowe nie są ze sobą skorelowane:

$$E[\mathbf{n}(i)\mathbf{v}^\top(j)] = 0 \quad (6.22)$$

6.3.1. Filtr Kalmana

Przedstawiony algorytm estymacji wartości wektora stanu można wpisać w schemat przedstawiony na rys. 6.4. Na schemacie tym wyróżnione są dwa elementy, tj. krok predykcji i krok korekcji. Bazując na tym schemacie można zaproponować algorytm postępowania, który w literaturze nazywany jest **filtrem Kalmana**. Ustalmy teraz następujący model w przestrzeni stanu:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}(k)\mathbf{x}(k) + \mathbf{n}(k) \\ \mathbf{y}(k) &= \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k) \end{aligned} \quad (6.23)$$

Algorytm 2: Filtr Kalmana

Wejście : $\mu = E[\mathbf{x}]$, \mathbf{Q} , \mathbf{R} , zestaw pomiarów $(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$ **while** nie spełniono warunku stopu **do** **predykcja:**

$$\mu^-(k) = \mathbf{A}(k-1)\mu(k) \quad (6.25a)$$

$$\mathbf{P}^-(k) = \mathbf{A}(k-1)\mathbf{P}(k-1)\mathbf{A}^\top(k-1) + \mathbf{Q}(k-1) \quad (6.25b)$$

korekcja:

$$\mathbf{e}(k) = \mathbf{y}(k) - \mathbf{H}(k)\mu^-(k) \quad (6.26a)$$

$$\mathbf{S}(k) = \mathbf{H}(k)\mathbf{P}^-(k)\mathbf{H}^\top(k) + \mathbf{R}(k) \quad (6.26b)$$

$$\mathbf{K}(k) = \mathbf{P}^-(k)\mathbf{H}^\top(k)\mathbf{S}^{-1}(k) \quad (6.26c)$$

$$\mu(k) = \mu^-(k) + \mathbf{K}(k)\mathbf{e}(k) \quad (6.26d)$$

$$\mathbf{P}(k) = \mathbf{P}^-(k) - \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}^\top(k) \quad (6.26e)$$

end**Wyjście :** sekwencja stanów $(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$

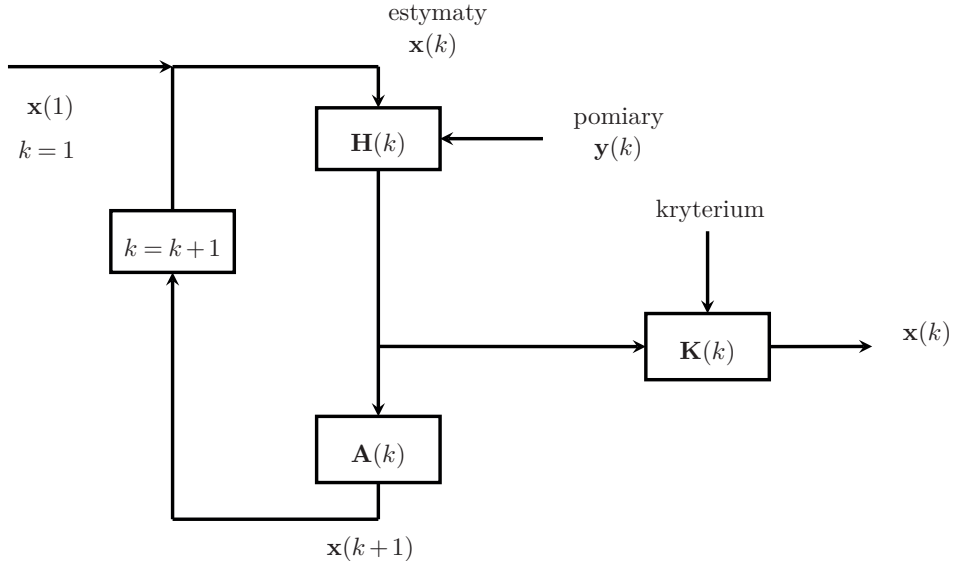
Niech wartość oczekiwaną dla rozkładu *gaussowskiego* reprezentuje:

$$\mu = E[\mathbf{x}] \quad (6.24)$$

Dysponując opisem dynamiki procesu oraz toru pomiarowego możliwe jest wprowadzenie zestawu równań do estymacji wartości wektora stanu $\mathbf{x}(k)$. Zestaw tych równań, nazwany filtrem Kalmana, zapisano w formie pseudokodu (patrz alg. 2). Znak minus w indeksie górnym przy μ oraz \mathbf{P} oznacza estymację *a priori* wartości wektora stanu.

W filtrze Kalmana, podobnie jak to było dla filtra Bayesa, wyróżniamy dwa etapy, tj. etap predykcji i etap korekcji. Z tymi etapami związane jest rekurencyjne obliczanie dwóch następujących macierzy \mathbf{S} oraz \mathbf{K} . Macierz \mathbf{S} jest **macierzą kowariancji procesu predykcji**, natomiast macierz \mathbf{K} nazywana jest **wzmocnieniem Kalmana** (ang. *Kalman gain*). Zmienna $\mathbf{e}(k)$ jest nazywana błędem predykcji. Obliczenia w filtrze Kalmana wykonywane są aż do uzyskania możliwie najlepszej wartości wzmocnienia \mathbf{K} . Na rysunku 6.7 przedstawiono przepływ informacji w filtrze Kalmana.

Zauważmy, że przedstawiony filtr Kalmana jest rekurencyjnym podejściem do wyznaczania minimalno-wariancyjnej estymaty wektora stanu systemu. Z wyprowadzeniem równań filtra Kalmana można zapoznać się m.in. w pracy [354].



Rys. 6.7. Przepływ informacji w filtrze Kalmana

Jeżeli szумы systemowe i pomiarowe, tak jak w naszych założeniach, będą addytywnymi szumami białymi, to wówczas filtr Kalmana jest najlepszym rozwiązaniem dla przypadku liniowego [370]. Natomiast gdy szумы systemowe i/lub pomiarowe są szumami kolorowymi, to wówczas, stosując filtr Kalmana, nie uzyskamy optymalnych rozwiązań. Podobna sytuacja występuje w przypadku, gdy zastosujemy opisany filtr Kalmana dla systemu nieliniowego. W kolejnych rozdziałach omówimy, w jaki sposób można rozwiązać problem w przypadku systemów nieliniowych.

Przykład 6.1 (Filtr Kalmana: estymacja stanu w \mathbb{R}^1). W przykładzie przedstawimy sposób konstrukcji filtra Kalmana dla zadania estymacji zmiennej stanu $x \in \mathbb{R}^1$. Przygotowany algorytm wykorzystamy do usuwania niepożądanych składowych o wysokich częstotliwościach z sygnału pomiarowego. Oznaczmy zestaw pomiarów przez $y(k)$, $k = 1, 2, \dots, K$. Założymy jednocześnie, że szумы procesowe $(n(k))$ i pomiarowe $(v(k))$ mają rozkład normalny, tj. $\mathcal{N}(\mu, \sigma)$ o ustalonej wartości oczekiwanej μ oraz odchyleniu standardowym σ . Model procesu i układu

6. Poziom podstawowy: fuzja danych

pomiarowego w przypadku jednowymiarowym ma postać:

$$\begin{aligned}x(k+1) &= x(k) + n(k) \\ y(k) &= x(k) + v(k)\end{aligned}\tag{6.27}$$

Zarówno pomiar, jak i stan w zaproponowanym modelu są wielkościami skalarnymi. Uwzględniając to, zapiszmy równania Kalmana w następującej formie:

$$\begin{aligned}\mu^-(k) &= \mu(k) \\ P^-(k) &= P(k-1) + Q(k-1)\end{aligned}\tag{6.28}$$

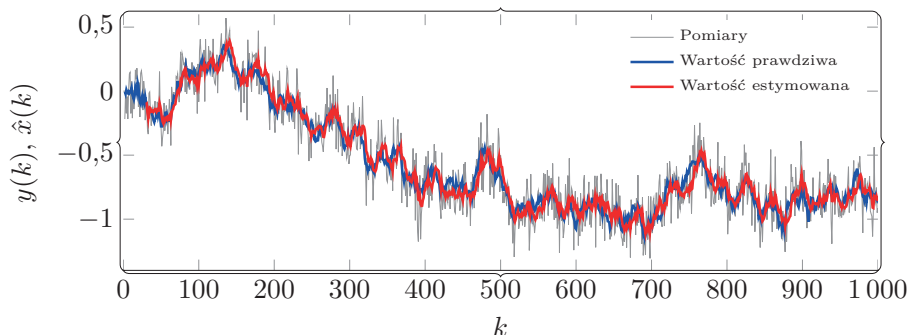
dla kroku predykcji oraz

$$\begin{aligned}e(k) &= y(k) - \mu^-(k) \\ S(k) &= P^-(k) + R(k) \\ K(k) &= \frac{P^-(k)}{S(k)}\end{aligned}\tag{6.29}$$

$$\mu(k) = \mu^-(k) + K(k)e(k)$$

$$P(k) = P^-(k) - K(k)S(k)K(k)$$

dla kroku korekcji. W kroku początkowym przyjmujemy: $A = 1$, $H = 1$ oraz $Q(0) = Q_0$, $R(0) = R_0$. Wynik zastosowania filtra Kalmana dla opisanych danych pomiarowych przedstawiono na rys. 6.8.



Rys. 6.8. Wynik działania filtra Kalmana (przykł. 6.1)

Przykład 6.2 (Filtr Kalmana: śledzenie obiektu na płaszczyźnie). W tym przykładzie rozpatrzono problem estymacji stanu obiektu poruszającego się na płaszczyźnie. Niech $\mathbf{y}(k)$, $k = 1, 2, \dots, K$ będzie zbiorem pomiarów prędkości poruszającego się po płaszczyźnie obiektu. Wektor $\mathbf{y}(k) = [y_x \ y_y]^\top$ jest dwuwymiarowy. Załóżmy również, że szумы procesowe ($\mathbf{n}(k)$) oraz szумы pomiarowe ($\mathbf{v}(k)$) mają rozkład normalny, tj. $\mathcal{N}(\mu, \sigma)$. Wówczas model procesu i pomiaru ma odpowiednio postać:

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{n}(k) \\ \mathbf{y}(k) &= \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)\end{aligned}\tag{6.30}$$

Dysponując pomiarami prędkości, formuła do wyznaczania położenia w kierunku osi x poruszającego się z tą prędkością obiektu na płaszczyźnie ma postać:

$$p_x(k+1) = p_x(k) + \mathbf{v}_x(k)\Delta t\tag{6.31}$$

gdzie $p_x(k)$ to bieżące położenie obiektu na płaszczyźnie wzdłuż osi x , Δt reprezentuje krok całkowania, a \mathbf{v}_x jest prędkością liniową z jaką obserwowany obiekt porusza się. Analogiczne równanie można zapisać dla osi y . Biorąc pod uwagę to, że obiekt porusza się z prędkością $\mathbf{v} = [\mathbf{v}_x, \mathbf{v}_y]^\top$ na płaszczyźnie o współrzędnych (p_x, p_y) , wektor stanu dla zadania będzie miał postać:

$$\mathbf{x}(k) = \begin{bmatrix} p_x \\ p_y \\ \mathbf{v}_x \\ \mathbf{v}_y \end{bmatrix}\tag{6.32}$$

Uwzględniając (6.31) oraz (6.32), model procesu przyjmuje postać:

$$\mathbf{x}(k+1) = \begin{bmatrix} p_x(k+1) \\ p_y(k+1) \\ \mathbf{v}_x(k+1) \\ \mathbf{v}_y(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} p_x(k) \\ p_y(k) \\ \mathbf{v}_x(k) \\ \mathbf{v}_y(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ n_x(k) \\ n_y(k) \end{bmatrix}\tag{6.33}$$

gdzie

$$\begin{aligned}\mathbf{v}_x(k+1) &= \mathbf{v}_x(k) + n_x(k) \\ \mathbf{v}_y(k+1) &= \mathbf{v}_y(k) + n_y(k)\end{aligned}\tag{6.34}$$

6. Poziom podstawowy: fuzja danych

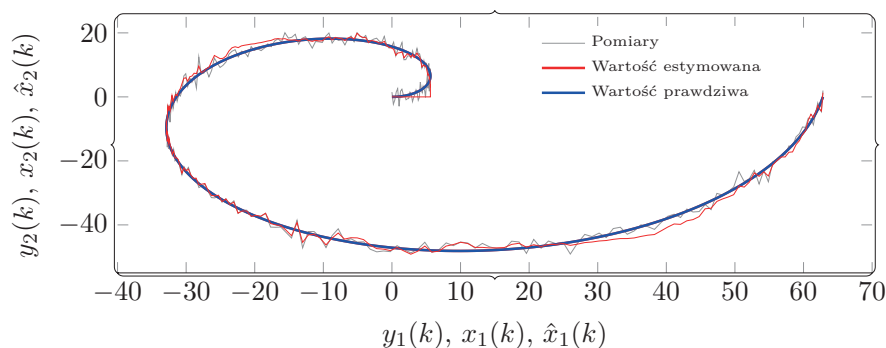
Znając model procesu dla zadania śledzenia obiektu na płaszczyźnie, kolejny krok to budowa modelu układu pomiarowego. W tym przypadku mamy:

$$y_x(k) = p_x(k) + v_x(k) \quad (6.35)$$

gdzie v_x jest szumem pomiarowym. Analogiczne równanie można zapisać dla drugiej z osi. Biorąc pod uwagę (6.32) oraz (6.35) otrzymujemy:

$$\mathbf{y}(k) = \begin{bmatrix} y_x(k) \\ y_y(k) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} p_x(k) \\ p_y(k) \\ \mathbf{v}_x(k) \\ \mathbf{v}_y(k) \end{bmatrix} + \begin{bmatrix} v_x(k) \\ v_y(k) \end{bmatrix} \quad (6.36)$$

Na rysunku 6.9 przedstawiono zbiór danych pomiarowych oraz wartości estymowane z wykorzystaniem opisanego w tym przykładzie filtra Kalmana. W celu porównania otrzymanych wyników szacowania wartości zmiennej stanu na rysunku tym zamieszczono również dane referencyjne (prawdziwe).



Rys. 6.9. Wynik zastosowania filtra Kalmana (przykl. 6.2)

6.3.2. Rozszerzony filtr Kalmana

Podstawowa wersja filtra Kalmana może być stosowana jedynie wówczas, gdy dane pochodzą z pomiarów systemu liniowego. Jeżeli obiekt ma nieliniową charakterystykę, to alternatywą dla niego jest rozszerzony filtr Kalmana (ang. *Extended Kalman Filter*, EKF). Cechą wspólną filtra EKF i podstawowego filtra Kalmana jest to, że do poprawnego ich działania wymagane jest to, by szum pomiarowy był addytywnym białym szumem *gaussowskim*. W praktycznych zadaniach warunek ten nie zawsze możliwy jest do spełnienia.

Na początek rozważmy ogólniejszy (w porównaniu do (6.23)) model w przestrzeni stanu, w którym zarówno model systemu, jak i model układu pomiarowego są nieliniowe:

$$\begin{aligned}\mathbf{x}(k+1) &= f(\mathbf{x}(k)) + \mathbf{n}(k) \\ \mathbf{y}(k) &= h(\mathbf{x}(k)) + \mathbf{v}(k)\end{aligned}\tag{6.37}$$

gdzie f jest nieliniowym opisem dynamiki systemu, z kolei h reprezentuje nieliniowy opis układu pomiarowego, $\mathbf{n}(k)$ i $\mathbf{v}(k)$ są szumami, które dotyczą odpowiednio systemu oraz układu pomiarowego (patrz rozdz. 3.4). W jednym i drugim przypadku szumy te mają naturę białego szumu *gaussowskiego*. Wprowadźmy również, tak jak to było w przypadku filtra Kalmana, macierze kowariancji \mathbf{Q} (dla procesu) i \mathbf{R} (dla układu pomiarowego).

Funkcje f oraz h (6.37) nie mogą być wprost wykorzystane w filtrze Kalmana. Natomiast możliwe jest, jeśli są to funkcje gładkie, ich rozwinięcie w szereg Taylora, oraz linearyzacja równań (6.37). Rozwinięcie funkcji f oraz h (pomijamy pochodne wyższego rzędu) ma postać:

$$\begin{aligned}f(\mathbf{x}) &\cong f(\bar{\mathbf{x}}) + \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\mathbf{x}} (\mathbf{x} - \bar{\mathbf{x}}) \\ h(\mathbf{x}) &\cong h(\bar{\mathbf{x}}) + \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\mathbf{x}} (\mathbf{x} - \bar{\mathbf{x}})\end{aligned}\tag{6.38}$$

gdzie $\left. \frac{\partial f}{\partial \bar{\mathbf{x}}} \right|_{\bar{\mathbf{x}}}$ oraz $\left. \frac{\partial h}{\partial \bar{\mathbf{x}}} \right|_{\bar{\mathbf{x}}}$ są **macierzami Jacobiego**, natomiast $\bar{\mathbf{x}}$ oznacza punkt pracy. Dokonując aproksymacji równań w przestrzeni stanu w otoczeniu pewnej wartości $\bar{\mathbf{x}}$ zmiennej $\mathbf{x}(k)$ oraz dokonując następującego podstawienia:

$$\begin{aligned}\mathbf{F}(\bar{\mathbf{x}}) &= \left. \frac{\partial f}{\partial \bar{\mathbf{x}}} \right|_{\bar{\mathbf{x}}} \\ \mathbf{H}(\bar{\mathbf{x}}) &= \left. \frac{\partial h}{\partial \bar{\mathbf{x}}} \right|_{\bar{\mathbf{x}}}\end{aligned}\tag{6.39}$$

otrzymamy liniowe przybliżenie równań (6.37):

$$\begin{aligned}\mathbf{x}(k+1) &\cong \mathbf{F}(\bar{\mathbf{x}})\mathbf{x}(k) + (\mathbf{I} - \mathbf{F}(\bar{\mathbf{x}}))\bar{\mathbf{x}} + \mathbf{n}(k) \\ \mathbf{y}(k) &\cong \mathbf{h}(\bar{\mathbf{x}}) + \mathbf{H}(\bar{\mathbf{x}})(\mathbf{x}(k) - \bar{\mathbf{x}}) + \mathbf{v}(k)\end{aligned}\tag{6.40}$$

gdzie \mathbf{I} reprezentuje macierz jednostkową.

Bazując na opisie w przestrzeni stanu (6.40), można wyprowadzić równania dla rozszerzonego filtra Kalmana. Równania tego algorytmu przedstawiono w formie pseudokodu (alg. 3).

Algorytm 3: Rozszerzony filtr Kalmana

W wejście : $\mu = E(\mathbf{x})$, macierze \mathbf{Q} , \mathbf{R} , zestaw pomiarów
 $(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$

while nie spełniono warunku stopu **do**

predykcja:

$$\mu^-(k) = f(\mathbf{x}(k-1)) \quad (6.41a)$$

$$\mathbf{P}^-(k) = \mathbf{F}(\mu(k))\mathbf{P}(k-1)\mathbf{F}^T(\mu(k-1)) + \mathbf{Q}(k-1) \quad (6.41b)$$

korekcja:

$$\mathbf{e}(k) = \mathbf{y}(k) - h(\mu^-(k)) \quad (6.42a)$$

$$\mathbf{S}(k) = \mathbf{H}(\mu^-(k))\mathbf{P}^-(k-1)\mathbf{H}^T(\mu^-(k)) + \mathbf{R}(k) \quad (6.42b)$$

$$\mathbf{K}(k) = \mathbf{P}^-(k)\mathbf{H}^T(\mu^-(k))\mathbf{S}^{-1}(k) \quad (6.42c)$$

$$\mu(k) = \mu^-(k) + \mathbf{K}(k)\mathbf{e}(k) \quad (6.42d)$$

$$\mathbf{P}(k) = \mathbf{P}^-(k) - \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}^T(k) \quad (6.42e)$$

end

Wyjście : sekwencja stanów $(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$

\mathbf{F} oraz \mathbf{H} reprezentują *macierze Jacobiego* odpowiednio dla funkcji f oraz h . Równania rozszerzonego filtra Kalmana szczegółowo wyprowadzono m.in. w pracy [354].

Zaletą filtra EKF jest możliwość estymacji wartości wektora stanu obiektów nieliniowych przez co rozszerzony filtr Kalmana umożliwia rozwiązanie zadania filtracji dla szerokiej klasy obiektów nieliniowych. Niestety w przypadku, gdy obiekt charakteryzuje się silną nieliniowością, to wówczas jakość wyników uzyskiwanych z wykorzystaniem filtra EKF może być niezadowalająca. Inną wadą tego rozwiązania jest to, że zarówno w części modelującej dynamikę systemu, jak i układ pomiarowy, szum ma charakterystykę addytywnego szumu białego. Pewną niedogodnością jest również wymaganie, by funkcje opisujące nieliniowości w równaniach stanu i wyjścia, tj. f oraz h były różniczkowalne.

Przykład 6.3 (Rozszerzony filtr Kalmana: estymacja sygnału sinusoidalnego). Rozważmy następujące zadanie estymacji amplitudy i częstotliwości sygnału sinusoidalnego:

$$\mathbf{y}(k) = A \sin(\omega k) \quad (6.43)$$

gdzie A jest amplitudą fali, ω reprezentuje pulsację, a $\mathbf{y}(k)$ to pomiar. Pomiar, którymi dysponujemy, zawierają niepożądane składowe w postaci szumów. Przed dalszymi obliczeniami wprowadźmy pomocniczą zmienną $\phi = \omega k$. Przyjmując teraz stałą wartość pulsacji, otrzymamy:

$$\begin{aligned}\dot{\phi} &= \omega \\ \dot{\omega} &= 0\end{aligned}\tag{6.44}$$

Zauważmy również, że dla stałej wartości amplitudy zachodzi:

$$\dot{A} = 0\tag{6.45}$$

Biorąc pod uwagę (6.44) oraz (6.45), możemy zapisać równanie stanu w formie:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\omega} \\ \dot{A} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \omega \\ A \end{bmatrix} + \mathbf{n}\tag{6.46}$$

gdzie \mathbf{n} reprezentuje szum.

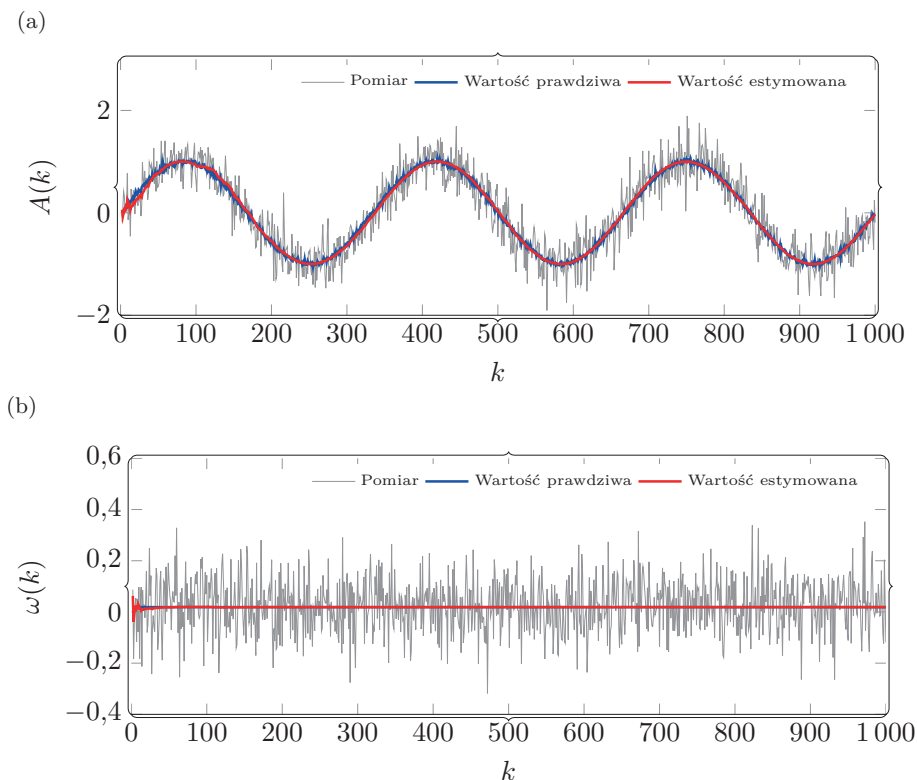
Wynik działania filtra EKF, którego podstawą jest model w przestrzeni stanu (6.46), przedstawiono na rys. 6.10.

6.3.3. Bezśladowy filtr Kalmana

Rozszerzony filtr Kalmana stosowany jest głównie w systemach, które charakteryzują się słabymi nieliniowościami. W przypadku systemów o silnych nieliniowościach stosowanie EKF może prowadzić do niezadowalających wyników. Alternatywą jest zastosowanie bezśladowego filtra Kalmana (ang. *Unscented Kalman Filter*, UKF), który bazuje na **przekształceniu bezśladowym**. Istotą filtra UKF jest to, że zamiast aproksymacji nieliniowej funkcji transformacji, estymuje się parametry rozkładu *gaussowskiego* dla każdej zmiennej wektora stanu. Pojedyncie to prowadzi do uproszczenia algorytmu filtracji, w porównaniu do filtra EKF, ponieważ przez zastosowanie transformacji bezśladowej nie ma konieczności prowadzenia obliczeń związanych z wyliczaniem pochodnych i jakobianów.

Przekształcenie bezśladowe (rys. 6.11) jest metodą wyznaczania parametrów zmiennej losowej poddanej nieliniowemu przekształceniu [189, 428]. Przekształcenia tego dokonuje się tylko w wybranych punktach, tj. w **punktach sigma** (ang. *sigma points*). Punkty te opisują prawdziwą wartość średnią i kowariancję zmiennej losowej. Ich liczba dobierana jest eksperymentalnie. Typowo do obliczeń przyjmuje się $2\mathcal{D}_x + 1$, gdzie \mathcal{D}_x to wymiar wektora stanu [354]. Punkty powinny zostać dobrane w taki sposób, by uzyskać możliwie najlepsze przybliżenie parametrów zmiennej losowej.

6. Poziom podstawowy: fuzja danych



Rys. 6.10. Wynik zastosowania rozszerzonego filtra Kalmana:

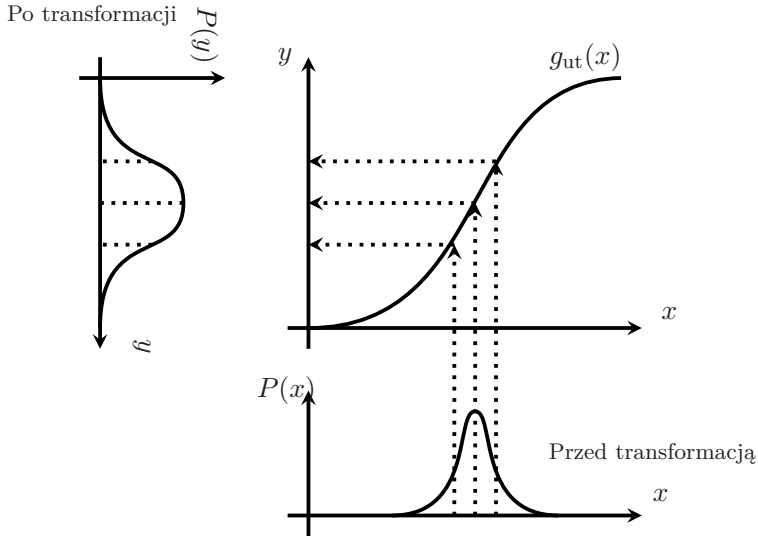
a) estymowana wartość amplitudy, b) estymowana wartość fazy. Dane z przykładu 6.3

Zastosowanie przekształcenia bezśladowego do filtra Kalmana upraszcza obliczenia i pozwala na stosowanie go w przypadku obiektów charakteryzujących się silnymi nieliniowościami. Podobnie jak dla filtrów Kalmana oraz EKF, również w przypadku filtra UKF wyróżniamy dwa etapy (tj. etap predykcji i korekcji), które powtarzane są cyklicznie. Ponieważ przekształcenie bezśladowe wykonywane jest w każdym kroku algorytmu oddzielnie dla każdego z etapów, to nakład obliczeń może być duży w porównaniu do filtra EKF.

Przyjmijmy do naszych rozważań następujący model:

$$\begin{aligned} \mathbf{x} &\approx \mathcal{N}(\boldsymbol{\mu}, \mathbf{P}) \\ \mathbf{y} &= g_{\text{ut}}(\mathbf{x}) \end{aligned} \tag{6.47}$$

gdzie g_{ut} jest nieliniową funkcją, $\boldsymbol{\mu}$ reprezentuje wartość oczekiwaną, a \mathbf{P} to macierz kowariancji, \mathbf{x} to wektor stanu, a \mathbf{y} stanowi wektor wyjściowy.



Rys. 6.11. Ilustracja działania przekształcenia bezśladowego

Koncepcja transformacji bezśladowej jest podobna do metody Monte Carlo [153], w której istotnym elementem jest wybór próbek. Różnica pomiędzy metodami polega na tym, że w metodzie Monte Carlo wybór próbek jest wyborem losowym, natomiast w transformacji bezśladowej odbywa się on deterministycznie. Deterministyczny sposób wyboru próbek w transformacji bezśladowej wpływa znacząco na liczbę wymaganych próbek w stosunku do metody Monte Carlo [208]. Dalej podano kolejne kroki algorytmu transformacji bezśladowej.

Algorytm transformacji bezśladowej

1. Wygeneruj $2\mathcal{D}_x + 1$ par punktów (średnia i kowariancja), tj. punkty sigma:

$$\mathcal{X}^{(0)} = \mu \quad (6.48a)$$

$$\mathcal{X}^{(i)} = \mu + \sqrt{\mathcal{D}_x + \lambda_{ut}} \left[\sqrt{\mathbf{P}} \right]_i \quad (6.48b)$$

$$\mathcal{X}^{(i+\mathcal{D}_x)} = \mu - \sqrt{\mathcal{D}_x + \lambda_{ut}} \left[\sqrt{\mathbf{P}} \right]_i \quad (6.48c)$$

dla $i = 1, 2, \dots, \mathcal{D}_x$

6. Poziom podstawowy: fuzja danych

gdzie $[\sqrt{\cdot}]_i$ jest i -tą kolumną macierzy \mathbf{P} , \mathcal{D}_x to wymiar wektora stanu \mathbf{x} , a λ_{ut} jest współczynnikiem wyliczanym z zależności:

$$\lambda_{\text{ut}} = \epsilon^2 (\mathcal{D}_x + \mathbf{f}) - \mathcal{D}_x \quad (6.49)$$

gdzie $\sqrt{\mathbf{P}}$ spełnia następujący warunek $\mathbf{P} = \sqrt{\mathbf{P}}\sqrt{\mathbf{P}}^\top$. Z kolei parametry ϵ oraz \mathbf{f} określają wartość λ_{ut} , która definiuje stopień skupienia punktów sigma wokół wartości średniej. Do wyznaczania pierwiastka kwadratowego z macierzy \mathbf{P} można wykorzystać jedną z metod numerycznych. Jedną z polecanych metod jest algorytm dekompozycji Choleskiego (ang. *Cholesky decomposition*) [115].

2. Wyznacz:

$$\mathcal{Y}^{(i)} = g_{\text{ut}} \left(\mathcal{X}^{(i)} \right) \quad (6.50)$$

dla $i = 0, 1, \dots, 2\mathcal{D}_x$.

3. Estymuj średnią, kowariancję oraz kowariancję krzyżową (ang. *cross-covariance*) dla wyznaczonych w drugim kroku punktów z następujących zależności:

$$\mu = \sum_{i=0}^{2\mathcal{D}_x} \mathcal{A}_{\mu i} \mathcal{Y}^{(i)} \quad (6.51a)$$

$$\mathbf{S} = \sum_{i=0}^{2\mathcal{D}_x} \mathcal{A}_{S i} \left(\mathcal{Y}^{(i)} - \mu \right) \left(\mathcal{Y}^{(i)} - \mu \right)^\top \quad (6.51b)$$

gdzie

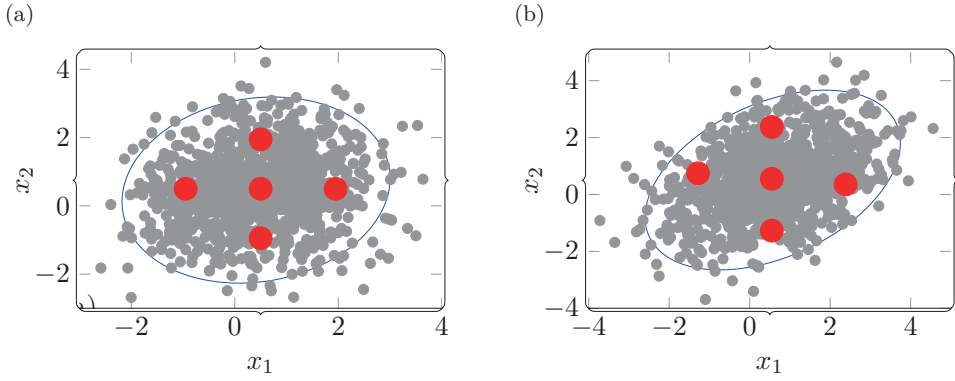
$$\mathcal{A}_{\mu 0} = \frac{\lambda_{\text{ut}}}{\mathcal{D}_x + \lambda_{\text{ut}}} \quad (6.52a)$$

$$\mathcal{A}_{S 0} = \frac{\lambda_{\text{ut}}}{\mathcal{D}_x + \lambda_{\text{ut}}} + \left(1 - \epsilon^2 + \mathbf{g} \right) \quad (6.52b)$$

$$\mathcal{A}_{\mu i} = \frac{1}{2(\mathcal{D}_x + \lambda_{\text{ut}})} \quad \text{dla } i = 1, 2, \dots, 2\mathcal{D}_x \quad (6.52c)$$

$$\mathcal{A}_{S i} = \frac{1}{2(\mathcal{D}_x + \lambda_{\text{ut}})} \quad \text{dla } i = 1, 2, \dots, 2\mathcal{D}_x \quad (6.52d)$$

są wagami, a \mathbf{g} jest parametrem określającym rozkład zmiennej stanu \mathbf{x} . Przykład działania transformacji bezśladowej przedstawiono na rys. 6.12.



Rys. 6.12. Przykład działania transformacji bezśladowej:
a) przed transformacją, b) po transformacji

Algorytm dla bezśladowego filtra Kalmana

Podamy teraz algorytm dla bezśladowego filtra Kalmana [190], którego ogólną koncepcję można przedstawić w formie następującej zależności:

$$P(\mathbf{x}(k)|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)) \simeq \mathcal{N}(\mathbf{x}(k)|\mu(k), \mathbf{P}(k)) \quad (6.53)$$

Zadaniem filtra jest oszacowanie rozkładu funkcji gęstości prawdopodobieństwa na podstawie wygenerowanych próbek z rozkładu normalnego. Kolejne kroki UKF przedstawiono w formie pseudokodu (alg. 4).

Algorytm 4: Bezśladowy filtr Kalmana

Wejście : $\mu = E(\mathbf{x})$, macierze \mathbf{Q} , \mathbf{R} , zestaw pomiarów $(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$
while nie spełniono warunku stopu **do**

predykcja:

Krok 1

$$\mathcal{X}^{(0)}(k-1) = \mu(k-1) \quad (6.54a)$$

$$\mathcal{X}^{(i)}(k-1) = \mu(k-1) + \sqrt{K + \lambda_{\text{ut}}} \left[\sqrt{\mathbf{P}(k-1)} \right]_i \quad (6.54b)$$

$$\mathcal{X}^{(i+\mathcal{D}_x)}(k-1) = \mu(k-1) - \sqrt{K + \lambda_{\text{ut}}} \left[\sqrt{\mathbf{P}(k-1)} \right]_i \quad (6.54c)$$

$$i = 1, 2, \dots, \mathcal{D}_x$$

Krok 2

$$\hat{\mathcal{Y}}^{(i)}(k) = f_{\text{ut}}(\mathcal{X}^{(i)}(k-1)) \quad (6.55)$$

$$i = 0, 1, \dots, 2\mathcal{D}_x$$

Krok 3

$$\mu^-(k) = \sum_{i=0}^{2\mathcal{D}_x} \mathcal{A}_{\mu i} \mathcal{Y}^{(i)}(k) \quad (6.56a)$$

$$\mathbf{P}^-(k) = \sum_{i=0}^{2\mathcal{D}_x} \mathcal{A}_{S i} (\mathcal{Y}^{(i)}(k) - \mu^-(k)) (\mathcal{Y}^{(i)}(k) - \mu^-(k))^{\text{T}} + \mathbf{Q}(k-1) \quad (6.56b)$$

korekcja:

Krok 1

$$\mathcal{X}^{-(0)}(k) = \mu(k)^- \quad (6.57a)$$

$$\mathcal{X}^{-(i)}(k) = \mu(k)^- + \sqrt{K + \lambda_{\text{ut}}} \left[\sqrt{\mathbf{P}(k)^-} \right]_i \quad (6.57b)$$

$$\mathcal{X}^{-(i+\mathcal{D}_x)}(k) = \mu(k)^- - \sqrt{K + \lambda_{\text{ut}}} \left[\sqrt{\mathbf{P}(k)^-} \right]_i \quad (6.57c)$$

$$i = 1, 2, \dots, \mathcal{D}_x$$

Krok 2

$$\hat{\mathcal{Y}}^{(i)}(k) = g_{\text{ut}}(\mathcal{X}^{-(i)}(k)), \quad (6.58)$$

$$i = 0, 1, \dots, 2\mathcal{D}_x$$

Algorytm 4 cd.

Krok 3

$$\mu(k) = \sum_{i=0}^{2D_x} \mathcal{A}_{\mu i} \hat{\mathcal{Y}}^{(i)}(k), \quad (6.59a)$$

$$\mathbf{S}(k) = \sum_{i=0}^{2D_x} \mathcal{A}_{S i} (\hat{\mathcal{Y}}^{(i)}(k) - \mu(k)) (\hat{\mathcal{Y}}^{(i)}(k) - \mu(k))^{\top} + \mathbf{R}(k), \quad (6.59b)$$

$$\mathbf{C}(k) = \sum_{i=0}^{2D_x} \mathcal{A}_{S i} (\mathcal{X}^{-(i)}(k) - \mu(k)^{-}) (\hat{\mathcal{Y}}^{(i)}(k) - \mu(k))^{\top} \quad (6.59c)$$

Krok 4

$$\mathbf{K}(k) = \mathbf{C}(k)\mathbf{S}(k)^{-1}, \quad (6.60a)$$

$$\mu(k) = \mu(k)^{-} + \mathbf{K}(k) (\mathbf{y}(k) - \mu(k)), \quad (6.60b)$$

$$\mathbf{P}(k) = \mathbf{P}(k)^{-} + \mathbf{K}(k)\mathbf{S}(k)\mathbf{K}(k)^{\top} \quad (6.60c)$$

end

Wyjście : sekwencja stanów $(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$

Przykład 6.4 (Estymacja kątów Eulera opisujących orientacje obiektów w \mathbb{R}^3). W przykładzie przedstawiono sposób wykorzystania filtra UKF do estymacji kątów Eulera. Kąty Eulera są jedną z metod reprezentacji orientacji obiektu w przestrzeni. Do estymacji tej orientacji wymagany jest odpowiedni model i dane pomiarowe. W omówionym przykładzie danymi pomiarowymi są dane z czujnika żyroskopowego, tj. prędkości kątowe. Oznaczmy zestaw danych pomiarowych jako $\mathbf{y}_{\omega}(k)$, $k = 1, 2, \dots, K$. Wektor prędkości kątowych ma trzy składowe, tj. $\mathbf{y}_{\omega} = [y_{\omega_x} \ y_{\omega_y} \ y_{\omega_z}]^{\top}$ (por. (4.10)).

Przystąpimy teraz do omówienia drugiego z elementów, tzn. modelu procesu. Zdefiniujmy teraz odpowiedni wektor stanu. Niech ma on następującą postać:

$$\mathbf{x}(k) = \begin{bmatrix} \varphi \\ \theta \\ \Psi \end{bmatrix} \quad (6.61)$$

6. Poziom podstawowy: fuzja danych

gdzie $\varphi(k)$, $\theta(k)$ oraz $\Psi(k)$ to kąty Eulera (patrz rozdz. 9.2). Model procesu i układu pomiarowego mają w tym przypadku postać:

$$\begin{aligned} \mathbf{x}(k+1) &= \begin{bmatrix} \varphi(k+1) \\ \theta(k+1) \\ \Psi(k+1) \end{bmatrix} = \\ &= \begin{bmatrix} 1 & \sin \varphi(k) \tan \theta(k) & \cos \varphi(k) \tan \theta(k) \\ 0 & \cos \varphi(k) & -\sin \varphi(k) \\ 0 & \sin \varphi(k) \sec \theta(k) & \cos \varphi(k) \sec \theta(k) \end{bmatrix} \begin{bmatrix} y_{\omega_x}(k) \\ y_{\omega_y}(k) \\ y_{\omega_z}(k) \end{bmatrix} + \mathbf{n}(k) \end{aligned} \quad (6.62a)$$

$$\mathbf{y}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \varphi(k) \\ \theta(k) \\ \Psi(k) \end{bmatrix} + \mathbf{v}(k) \quad (6.62b)$$

Po ustaleniu opisu w przestrzeni stanu kolejnym krokiem jest zastosowanie filtra UKF (patrz alg. 4). Wynik zastosowania filtra UKF dla danych z przykładu (6.4) przedstawiono na rys. 6.13.

6.3.4. Filtr Cząsteczkowy

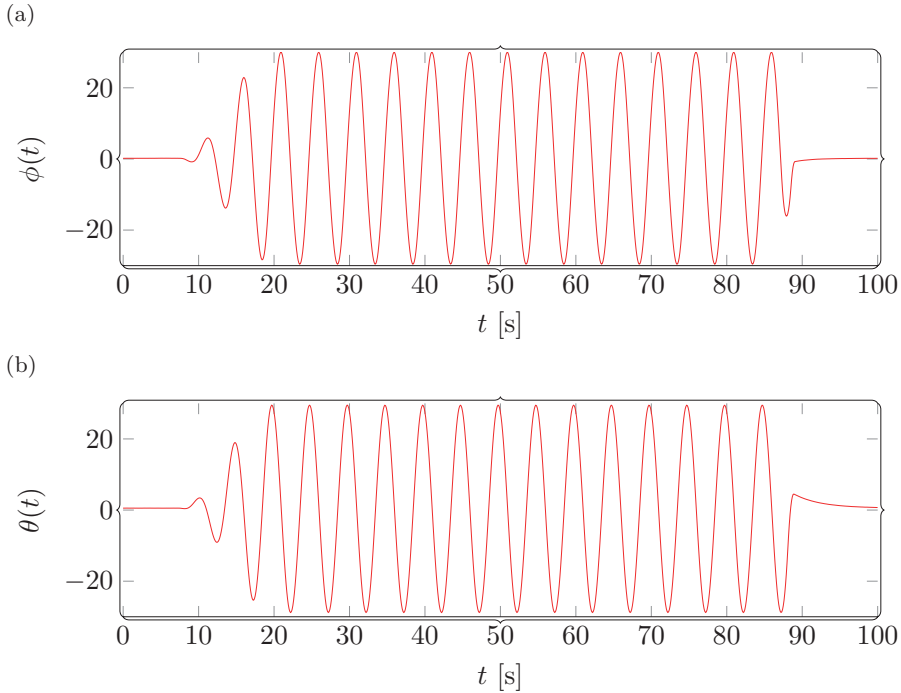
Przedstawione w poprzednich podrozdziałach podejścia przeznaczone są dla zadań, w których rozkłady prawdopodobieństwa są unimodalne. W przypadku wielomodalnych rozkładów prawdopodobieństwa omówione dotychczas metody mogą okazać się niewystarczające. Alternatywą dla przedstawionych rozwiązań jest na przykład, wykorzystujący koncept filtra Kalmana, filtr cząsteczkowy (ang. *Particle Filter*, PF). Filtr cząsteczkowy charakteryzuje się tym, że bazuje on na procedurze Monte Carlo. Procedura ta została w tym przypadku wykorzystana do rozwiązywania *bayesowskich* równań filtrujących (patrz równ. (6.11)–(6.13)). Rozpocznijmy od omówienia, w jaki sposób można wykorzystać metodę Monte Carlo w zadaniu wnioskowania i/lub filtrowania *bayesowskiego*.

Metoda Monte Carlo w zadaniu wnioskowania bayesowskiego

Problem wnioskowania (również filtrowania) bayesowskiego sprowadza się do następującego zadania:

$$E [g_{\text{pf}}(\mathbf{x}) | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)] = \int_{\mathbf{x}} g_{\text{pf}}(\mathbf{x}) P(\mathbf{x} | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)) d\mathbf{x} \quad (6.63)$$

gdzie g_{pf} jest arbitralnie dobraną funkcją. Do wyznaczenia całki w równaniu (6.63) można wykorzystać np. numeryczną metodę Monte Carlo. W przypadku idealnym



Rys. 6.13. Wynik estymacji kątów $\phi(k)$ (a) oraz $\theta(k)$ (b) z wykorzystaniem filtra UKF dla danych z zadania (6.4)

zadanie to sprowadza się do wygenerowania K_{pf} niezależnych statystycznie próbek, tj. $\mathbf{x}^{(i)} \sim P(\mathbf{x}|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$, $i = 1, 2, \dots, K_{\text{pf}}$ oraz do oszacowania wartości oczekiwanej z następującej zależności:

$$E[g_{\text{pf}}(\mathbf{x})|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)] \approx \frac{1}{K_{\text{pf}}} \sum_{i=1}^{K_{\text{pf}}} g_{\text{pf}}(\mathbf{x}^{(i)}) \quad (6.64)$$

W większości przypadków nie jest możliwe wygenerowanie próby bezpośrednio z rozkładu $\mathbf{x}^{(i)} \sim P(\mathbf{x}|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$. Jest to spowodowane tym, że postać analityczna tej zależności jest skomplikowana lub w ogóle nie istnieje. Jedną z popularniejszych metod otrzymywania próby z rozkładu $\mathbf{x}^{(i)} \sim P(\mathbf{x}|\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$ jest algorytm Gibbsa [53]. Wadą tego algorytmu jest niestety to, że wyznaczony rozkład prawdopodobieństwa musi być aktualizowany w chwili, gdy pojawią się nowe dane pomiarowe. Rozwiązaniem są metody sekwencyjne, których cechą charakterystyczną jest to, że bierze się w nich pod uwagę ciąg rozkładów.

6. Poziom podstawowy: fuzja danych

Innym podejściem do aproksymacji funkcji rozkładu gęstości prawdopodobieństwa jest podejście wykorzystujące funkcję ważności (ang. *importance sampling*). W metodzie tej dokonujemy dekompozycji tej funkcji do następującej postaci:

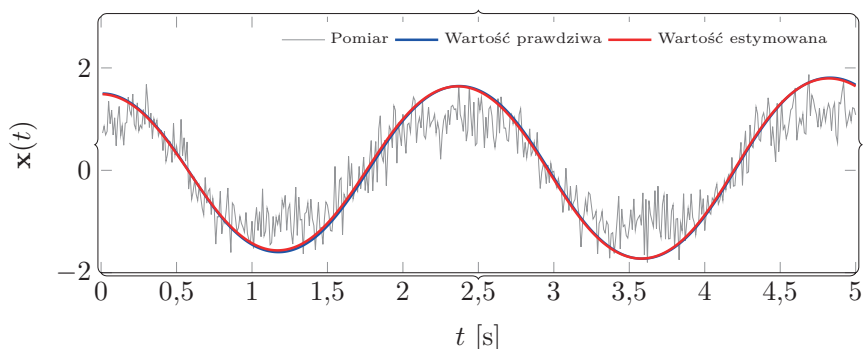
$$E [g_{\text{pf}}(\mathbf{x}) | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)] = \int_{\mathbf{x}} \left[g_{\text{pf}}(\mathbf{x}) \frac{P(\mathbf{y} | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))}{\mathcal{W}_{\text{pf}}(\mathbf{x} | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))} \right] \mathcal{W}_{\text{pf}}(\mathbf{x} | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K)) d\mathbf{x} \quad (6.65)$$

gdzie \mathcal{W}_{pf} jest funkcją ważności. W tej metodzie próbki generowane są z wykorzystaniem funkcji \mathcal{W}_{pf} , a następnie nadawane są im stosowne wagi. Wadą tego podejścia jest to, że zastosowana metoda może sprawiać problemy natury numerycznej.

Alternatywą dla metody funkcji ważności jest metoda sekwencyjnej funkcji ważności (ang. *sequential importance sampling*). Idea tego podejścia polega na generowaniu ciągu rozkładów $\mathcal{W}_{\text{pf}}(\mathbf{x} | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$ wraz z odpowiednimi wagami. Dysponując ciągiem rozkładów oraz odpowiednimi wagami, dokonujemy oszacowania wartości oczekiwanej z następującej zależności:

$$E [g_{\text{pf}}(\mathbf{x}) | \mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k)] \approx \sum_{i=1}^{K_{\text{pf}}} \mathbf{b}^{(i)} g(\mathbf{x}^{(i)}) \quad (6.66)$$

Z sekwencyjną funkcją ważności wiąże się problem degeneracji próby [53]. Efektem tego niekorzystnego zjawiska jest to, że po kilku iteracjach wszystkie (oprócz jednej) wagi przyjmują wartości bliskie zeru. W celu zapobiegania temu niekorzystnemu zjawisku do omówionej wcześniej metody wprowadza się dodatkowe losowanie (ang. *resampling*). Opracowano kilka metod wprowadzających dodatkowe losowanie, m.in. *resampling wielomianowy* (ang. *multinomial resampling*),



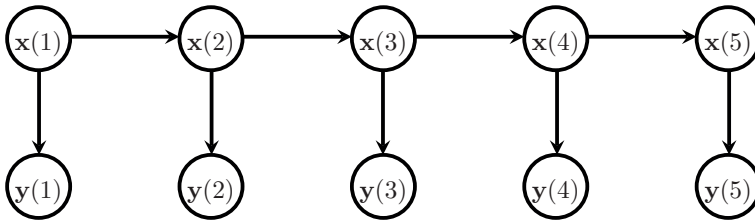
Rys. 6.14. Wyniki działania filtra cząsteczkowego (przykł. 6.5)

resampling resztowy (ang. *residual resampling*), *resampling systematyczny* (ang. *systematic resampling*).

Przykład 6.5 (Zastosowanie filtrów cząsteczkowych). *Przykład ilustruje wykorzystanie filtra cząsteczkowego do estymacji sygnału sinusoidalnego na podstawie zakłóconych danych pomiarowych. Dane pomiarowe oraz wynik przedstawiono na rys. 6.14. Przykład został opracowany na podstawie algorytmów przedstawionych w pracy [19].*

6.4. Filtracja optymalna – stan dyskretny

Rozdział ten poświęcony jest problemowi estymacji stanu dyskretnego. Stan dyskretny oznacza, że zmienne wektora stanu ($\mathbf{x}(k)$) mogą przyjmować wartości tylko z pewnego zbioru $\Omega = \{\zeta_1, \zeta_2, \dots, \zeta_{N_\Omega}\}$, gdzie N_Ω jest liczbą dyskretnych stanów. Wartości te nie są bezpośrednio mierzalne, mogą być jedynie obserwowane przez tzw. zmienne obserwowane $\mathbf{y}(k)$. Również i te zmienne mogą przyjmować jedynie wartości ze skończonego zbioru wartości $\Theta = \{\nu_1, \nu_2, \dots, \nu_{N_\Theta}\}$. Zmienna N_Θ reprezentuje liczbę wszystkich wartości, które może przyjmować zmienna obserwowana $\mathbf{y}(k)$. Opisana sytuacja została zilustrowana na rys. 6.15.



Rys. 6.15. Przykład Ukrytego Modelu Markowa

W przypadku gdy obserwacje przyjmują jedynie wartości dyskretne, zadanie filtracji polega na wyznaczeniu sekwencji stanów, przez które przechodzi obserwowany proces. Takie zadanie rozwiązywane jest na podstawie wcześniej zebranych danych pomiarowych. Jeśli obserwacje przyjmują wartości ze zbioru liczb rzeczywistych, to wykorzystanie opisanego podejścia jest również możliwe, ale wymaga dodatkowych działań. Te działania związane są z dyskretyzacją sygnału $\mathbf{y}(k)$. Popularną metodą dyskretyzacji danych pomiarowych jest metoda VQ (ang. *Vector Quantization*). Warto w tym miejscu zaznaczyć, że w wielu przypadkach zastosowanie dyskretyzacji może doprowadzić do utraty znaczącej informacji o pomiarach.

6. Poziom podstawowy: fuzja danych

Opisany model estymacji ukrytego dyskretnego stanu systemu na podstawie zbioru dyskretnych obserwacji nazywa się Ukrytym Modelem Markowa (ang. *Hidden Markov Model*, HMM).

6.4.1. Ukryte Modele Markowa

Ukryte Modele Markowa definiujemy za pomocą następujących wielkości:

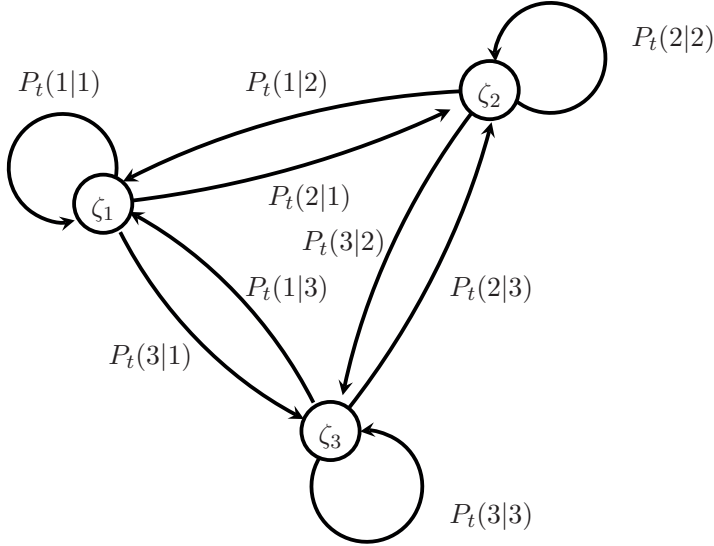
- skończony zbiór wartości ukrytych stanów $\Omega = \{\zeta_1, \zeta_2, \dots, \zeta_{N_\Omega}\}$ określa, jakie wartości mogą przyjmować ukryte stany $\mathbf{x}(k)$;
- skończony zbiór wartości obserwacji $\Theta = \{\nu_1, \nu_2, \dots, \nu_{N_\Theta}\}$ określa, jakie wartości mogą przyjmować wielkości obserwowane $\mathbf{y}(k)$;
- początkowy rozkład prawdopodobieństwa $P_0(\mathbf{x}(0))$ określa prawdopodobieństwo wyboru stanu k jako pierwszego stanu (ang. *initial state probability*);
- prawdopodobieństwo przejścia $P_t(\mathbf{x}(k)|\mathbf{x}(k-1))$ od stanu $\mathbf{x}(k-1)$ do stanu $\mathbf{x}(k)$ (ang. *state transition probability*);
- prawdopodobieństwo obserwacji $P_y(\mathbf{y}(k)|\mathbf{x}(k))$ dla poszczególnych stanów $\mathbf{x}(k)$ (ang. *observation probability*).

W modelach typu HMM przyjmuje się, że $P_t(\mathbf{x}(k)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(k-1))$ jest równe prawdopodobieństwu przejścia (warunek Markowa). Kolejne założenie dotyczy układu pomiarowego. W przypadku HMM przyjmujemy, że układ pomiarowy nie ma pamięci. W związku z tym zachodzi $P_y(\mathbf{y}(j)|\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(k)) = P_y(\mathbf{y}(j)|\mathbf{y}(j))$. Oznacza to, że $\mathbf{y}(j)$ zależy od stanu układu $\mathbf{x}(k)$ tylko w danej chwili.

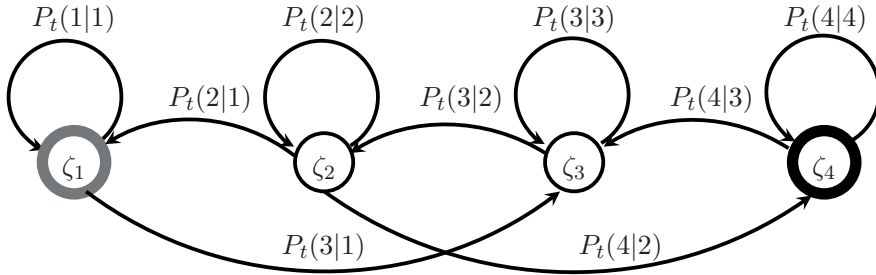
Ukryte Modele Markowa charakteryzują się **ergodycznością**. Oznacza to, że obserwacja jednej sekwencji stanów, np. $(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(\infty))$ jest wystarczająca do określenia prawdopodobieństwa przejścia dla wszystkich pozostałych sekwencji przejść pomiędzy stanami. Ukryty Model Markowa jest ergodyczny wówczas, gdy prawdopodobieństwa wszystkich stanów są niezerowe. W praktyce oznacza to, że wszystkie stany modelu są osiągalne z każdego innego stanu modelu. Przykład ergodycznego HMM przedstawiono na rys. 6.16.

Oddzielną klasę Ukrytych Modeli Markowa stanowi HMM bazujący na modelu Bakisa. W modelu tym stany są uporządkowane, a przejścia pomiędzy poszczególnymi stanami są możliwe od stanów o niższych indeksach do stanów o wyższych indeksach, tj. $P_t(\mathbf{x}(i)|\mathbf{x}(j))$, gdzie $i < j$. Przykład model Bakisa przedstawiono na rys. 6.17.

Mając ustaloną strukturę modelu HMM oraz dysponując zestawem pomiarów $\mathbf{Y}_K = (\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(K))$ możliwe jest wyznaczenie prawdopodobieństwa otrzymania zestawu obserwacji $P(\mathbf{Y}_K)$ dla sekwencji realizacji stanów $\mathbf{X}_K = (\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K))$. Bezpośrednie wyznaczenie tego prawdopodobień-



Rys. 6.16. Przykład ergodycznego Ukrytego Modelu Markowa



Rys. 6.17. Przykład modelu Bakisa tzw. lewy-prawy

stwa polega na sumowaniu po wszystkich możliwych sekwencjach stanów, tj.

$$P(\mathbf{Y}_K) = \sum_{\text{wszystkie } \mathbf{X}_K} P(\mathbf{X}_K, \mathbf{Y}_K) \quad (6.67)$$

Niestety bezpośrednia implementacja zależności (6.67) jest w praktyce niemożliwa. W praktyce zadanie estymacji $P(\mathbf{Y}_K)$ rozwiązuje się przez zastosowanie metod rekurencyjnych. W jednym z podejść wartość $P(\mathbf{Y}_K)$ wyznaczana jest

6. Poziom podstawowy: fuzja danych

w dwóch etapach. W pierwszym z nich wyznaczane jest prawdopodobieństwo łączne $P(\mathbf{X}_K, \mathbf{Y}_K)$, a w drugim z etapów na podstawie wyznaczonego prawdopodobieństwa łącznego wyznaczane jest prawdopodobieństwo $P(\mathbf{Y}_K)$. Formuła do wyznaczania prawdopodobieństwa łącznego ma w tym przypadku postać:

$$P(\mathbf{Y}_K, \mathbf{x}(k)) = P(\mathbf{y}(k)|\mathbf{x}(k)) \sum_{\mathbf{x}(k-1)=1}^{N_{se}} P_t(\mathbf{x}(k)|\mathbf{x}(k-1))P(\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k-1), \mathbf{x}(k-1)) \quad (6.68)$$

gdzie N_{se} jest liczbą wszystkich sekwencji stanów.

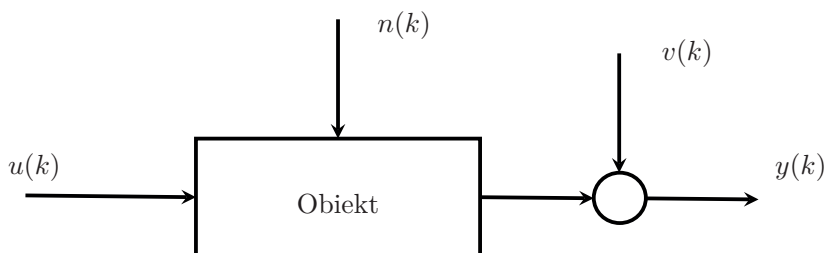
Po wyznaczeniu prawdopodobieństwa $P(\mathbf{Y}_K, \mathbf{x}(k))$ można wyznaczyć prawdopodobieństwo $P(\mathbf{Y}_K)$ z następującej zależności:

$$P(\mathbf{Y}_K) = \sum_{\mathbf{x}(k)=1}^{N_{se}} P(\mathbf{Y}_K, \mathbf{x}(k)) \quad (6.69)$$

Opisana procedura nazywana jest procedurą *forward*. Jest to procedura rekurencyjna, której efektywność jest znacznie lepsza niż w przypadku podejścia bezpośredniego [416].

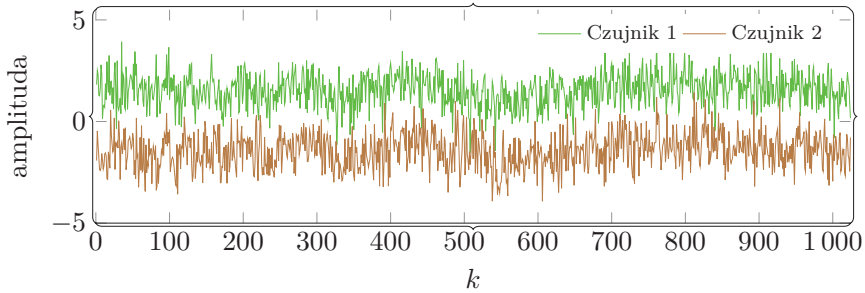
6.5. Fuzja danych z wykorzystaniem filtra Kalmana

W ostatniej części tego rozdziału przedstawimy prosty przykład ilustrujący korzyść z wykorzystania fuzji decyzji na poziomie 1. W omówionym zadaniu zastosowano filtr Kalmana, który użyto do przetwarzania danych z dwóch czujników pomiarowych. Oczywiście przyjmujemy, że dane, którymi dysponujemy, są obciążone błędem (tj. szumem).



Rys. 6.18. Sytuacja pomiarowa dla zadania z podrozdziału 6.5

6.5. Fuzja danych z wykorzystaniem filtra Kalmana



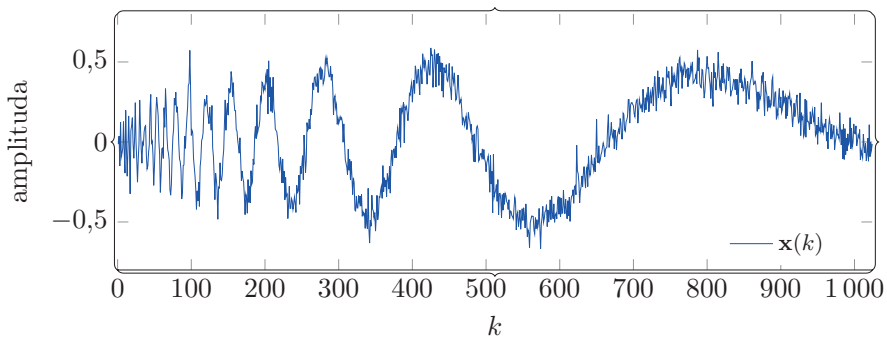
Rys. 6.19. Wizualizacja sygnałów pomiarowych z czujnika 1 i 2

Rozpocznijmy od opisu sytuacji pomiarowej. Wielkością mierzoną jest sygnał *dopplerowski*. Do pomiarów wykorzystujemy dwa czujniki, z których każdy dokonuje pomiaru niezależnie. Dane z czujników zapisujemy w dwóch zmiennych, tj. $y_1(k)$ i $y_2(k)$, $k = 1, 2, \dots, K$. Przebiegi z tych czujników zostały przedstawione na rys. 6.19. Wektor stanu oznaczymy natomiast jako $\mathbf{x}(k)$. Z kolei szum systemowy niech będzie oznaczony jako $\mathbf{n}(k)$, a szum układu pomiarowego jako $\mathbf{v}(k)$. Zarówno szum systemowy, jak i układu pomiarowego będzie białym szumem *gaussowskim*. Jako ostatni element do ustalenia pozostają macierze kowariancji \mathbf{Q} (dla szumu systemowego) oraz macierz \mathbf{R} (dla szumu pomiarowego).

Ustalmy teraz model dla naszego zadania estymacji sygnału na podstawie pomiarów z dwóch czujników pomiarowych:

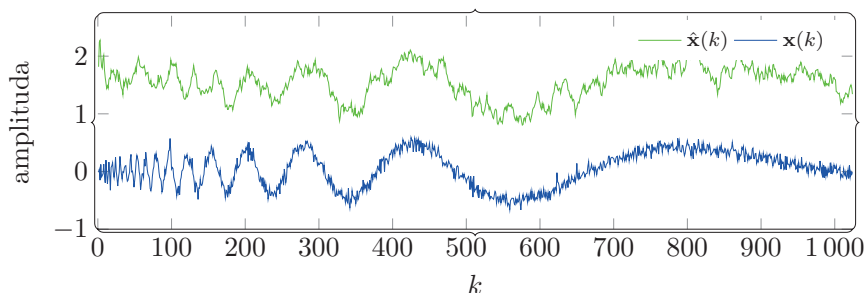
$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{n}(k) \\ \mathbf{y}(k) &= \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k) \end{aligned} \quad (6.70)$$

gdzie \mathbf{A} jest macierzą systemową, z kolei \mathbf{H} reprezentuje macierz pomiarową.

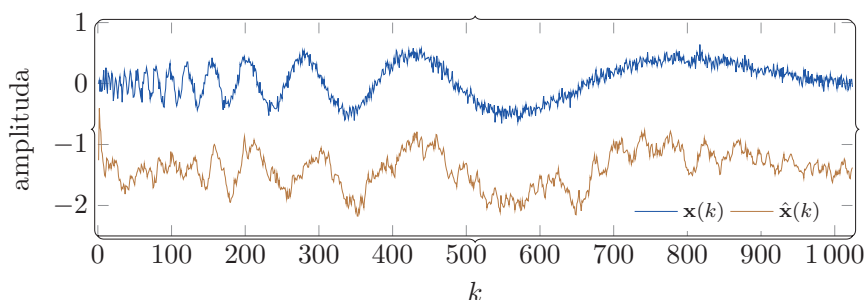


Rys. 6.20. Przebieg zmiennej stanu $\mathbf{x}(k)$

6. Poziom podstawowy: fuzja danych



Rys. 6.21. Wynik estymacji zmiennej stanu $\mathbf{x}(k)$ z wykorzystaniem danych z czujnika 1



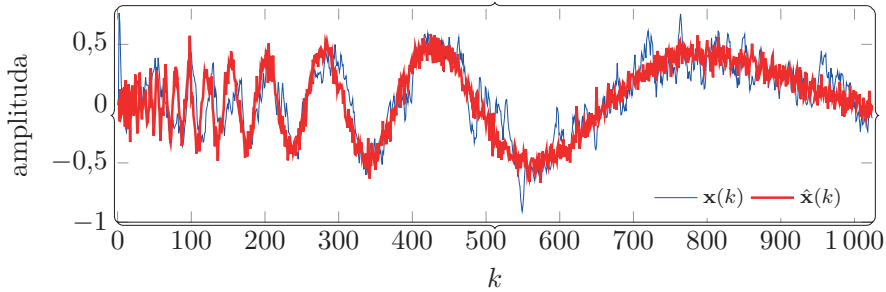
Rys. 6.22. Wynik estymacji zmiennej stanu $\mathbf{x}(k)$ z wykorzystaniem danych z czujnika 2

Przebieg zmian wartości wektora stanu $\mathbf{x}(k)$, który w opisanym zadaniu ma zostać estymowany na podstawie danych z czujników pomiarowych, został przedstawiony na rys. 6.20. W pierwszym podejściu wykorzystamy filtr Kalmana do estymacji wartości wektora stanu z wykorzystaniem danych z jednego czujnika. Dla tego zadania macierz systemowa ma postać $\mathbf{A} = 1$, a macierz pomiarowa $\mathbf{H} = 1$. Wartość kowariancji ustalmy na $Q = 0,005$, a wartość kowariancji dla pomiaru wynosi $R = 0,64$.

Wartość wektora stanu została wyznaczona oddzielnie dla danych z każdego czujnika. Wyniki estymacji tego wektora na podstawie danych z czujnika 1 przedstawiono na rys. 6.21. Natomiast wynik estymacji wektora stanu dla danych z czujnika numer 2 przedstawiono na rys. 6.22. Zarówno w przypadku estymacji wartości wektora stanu z wykorzystaniem danych jedynie z czujnika 1, jak i w przypadku, gdy do estymacji posłużyły jedynie dane z czujnika 2 wartość błędu średniokwadratowego wynosiła $RMSE = 1,51$.

W drugim przypadku wykorzystamy filtr Kalmana i jego zdolność do fuzji danych na poziomie podstawowym. W tym przypadku do estymacji wartości wektora stanu $\mathbf{x}(k)$ wykorzystamy dane z dwóch czujników $y_1(k)$ i $y_2(k)$. Model dla zadania

6.5. Fuzja danych z wykorzystaniem filtra Kalmana



Rys. 6.23. Wynik estymacji zmiennej stanu $\mathbf{x}(k)$ z wykorzystaniem danych z czujników 1 i 2

jest taki sam jak dla poprzedniego przypadku (6.70). Macierz systemowa to:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Macierz pomiarowa przyjmuje postać wektora w formie:

$$\mathbf{H} = [1 \quad 1]$$

Macierze kowariancji dla tego zadania mają postać:

$$\mathbf{Q} = \begin{bmatrix} 0,005 & 0 \\ 0 & 0,005 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 0,64 & 0 \\ 0 & 0,64 \end{bmatrix}$$

Wynik estymacji wartości wektora stanu $\mathbf{x}(k)$ przedstawiono na rys. 6.23. Można zauważyć, że wykorzystanie danych z dwóch czujników do estymacji wektora stanu daje znacznie lepsze wyniki w porównaniu do oddzielnego wykorzystania danych z czujników. Jest to również widoczne w wartości błędu średniokwadratowego, który w tym przypadku wynosi $RMSE = 0,19$.

7. Poziom średni: fuzja właściwości

Żaden powszechnik nie jest substancją, obojętnie, jak byśmy się nań zapatrywali.

William Ockham

7.1. Wstęp

W wielu praktycznych zadaniach jednomodalność danych pomiarowych jest trudna do spełnienia. Może być to spowodowane np. tym, że w rozpatrywanym problemie wymagany jest pomiar wielu różnych zjawisk. W takim przypadku zebrane podczas pomiarów dane mogą nie mieć cech danych jednomodalnych. Dane takie, tj. pochodzące z pomiarów wielu różnych zjawisk, nazywamy danymi wielomodalnymi (patrz podrozdz. 2.1). Do takich danych pomiarowych bezpośrednie zastosowanie algorytmów opisanych w poprzednim rozdziale może okazać się niemożliwe, ponieważ dane takie wymagają algorytmów, które będą przetwarzać je na wyższym poziomie abstrakcji. W takich przypadkach konieczne może być zastosowanie algorytmów fuzji danych na drugim poziomie, tj. na poziomie fuzji właściwości.

W wielu przypadkach przetwarzanie danych na drugim poziomie wiąże się z dodatkowym nakładem obliczeń. Dodatkowe obliczenia związane są m.in. z koniecznością transformacji pomiarów zebranych z różnych źródeł do dziedziny, w której będą one następnie przetwarzane w sposób łączny.

Najważniejsze zadania, które są do wykonania w ramach fuzji danych na tym poziomie to konstrukcja **wektora cech** oraz **konkluzje**, które wyciągane są na jego podstawie. Wektor cech (ang. *feature vector*) jest rodzajem zagregowanego opisu obiektu na wyższym poziomie abstrakcji (tj. wyższym niż ten na poziomie podstawowym). Wektor cech konstruowany jest na podstawie odpowiednio przetworzonych danych pomiarowych. Z kolei konkludowanie wykonywane jest z wykorzystaniem przede wszystkim algorytmów klasyfikacji lub klasteryzacji.

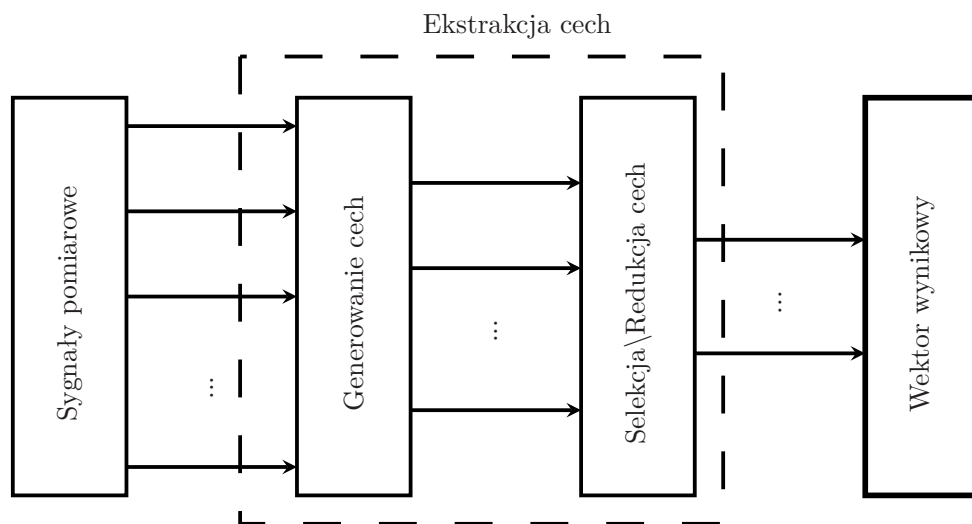
Problem konstrukcji wektora cech składa się z wielu kolejno rozwiązywanych zadań. Najważniejsze z tych zadań to [144]:

7. Poziom średni: fuzja właściwości

- **segmentacja** (ang. *segmentation*), której celem jest podział zbioru danych pomiarowych na mniejsze zbiory,
- **generowanie cech** (ang. *features generation*), które wykonywane jest na podstawie danych pomiarowych i ma na celu wyznaczenie bogatego zestawu cech,
- **selekcja cech** (ang. *features selection*),
- **redukcja cech** (ang. *features reduction*).

Zadania selekcji oraz redukcji cech mają ten sam cel, chociaż sposób jego osiągnięcia jest w każdym przypadku inny. Celem tych dwóch zadań jest zmniejszenie liczności zbioru wygenerowanych cech, przy czym powinno to zostać tak wykonane, by stracić jak najmniej informacji o obiekcie.

Kolejne etapy związane z konstrukcją wektora cech przedstawiono na rys. 7.1. Wszystkie zadania, które służą do przygotowania wektora cech, występują pod wspólną nazwą **ekstrakcji cech** (ang. *feature extraction*).



Rys. 7.1. Etapy konstrukcji wektora cech

7.2. Ekstrakcja cech

Zdefiniujmy na początek zadanie ekstrakcji cech. Celem tego zadania jest transformacja danych z przestrzeni pomiarów do przestrzeni cech. Realizacja zadań wchodzących w skład ekstrakcji cech umożliwi konstrukcję odpowiedniego

wektora cech. Zadanie przekształcenia jednej przestrzeni cech w drugą zapiszmy w następujący sposób:

$$\mathbf{z}(k) = \Upsilon(\mathbf{y}(k)), \quad k = 1, 2, \dots, K \quad (7.1)$$

gdzie $\mathbf{y}(k)$ jest \mathcal{D}_y -wymiarowym wektorem pomiarów, tzn. $\mathbf{y} \in \mathbb{R}^{\mathcal{D}_y}$, $\mathbf{z}(k)$ reprezentuje wynikowy \mathcal{D}_z -wymiarowy wektor cech po ekstrakcji $\mathbf{z} \in \mathbb{R}^{\mathcal{D}_z}$, \mathcal{D}_z określa liczbę współrzędnych wektora $\mathbf{z}(k)$, Υ jest pewnym przekształceniem takim, że $\Upsilon : \mathbb{R}^{\mathcal{D}_y} \rightarrow \mathbb{R}^{\mathcal{D}_z}$, a k to indeks pomiaru.

Postać przekształcenia Υ powinna być dostosowana do zadania i do natury danych, które w tym zadaniu będą podlegać transformacji. Konstrukcja tego przekształcenia jest jednym z głównych zadań, które są do wykonania na średnim poziomie fuzji danych.

Definicja 7.2.1 (Ekstrakcja cech). *Ekstrakcja cech to przekształcenie oryginalnego zbioru danych pomiarowych w zbiór danych bardziej informacyjnych.*

Sformułowanie problemu

Dla zestawu sygnałów pomiarowych:

$$\mathbf{Y}_K = (\mathbf{y}(1) \quad \mathbf{y}(2) \quad \dots \quad \mathbf{y}(K)) \quad (7.2)$$

gdzie K reprezentuje liczbę pomiarów, a $\mathbf{y}(k)$ jest \mathcal{D}_y -wymiarowym wektorem postaci:

$$\mathbf{y}(k) = [y_1 \quad y_2 \quad \dots \quad y_{\mathcal{D}_y}]^T \quad (7.3)$$

wyznacz algorytm ekstrakcji cech:

$$\mathbf{z}(k) = \Upsilon(\mathbf{y}(k)) \quad (7.4)$$

taki, że

$$\mathbf{z}(k) = [z_1 \quad z_2 \quad \dots \quad z_{\mathcal{D}_z}] \quad (7.5)$$

gdzie $\mathcal{D}_z < \mathcal{D}_y$.

Zauważmy, że ekstrakcja cech jest zadaniem wieloetapowym (por. rys. 7.1). W kolejnych podrozdziałach omówimy wszystkie zadania, które rozwiązywane są w ramach zadania ekstrakcji cech, tj. segmentacja, generowanie cech, jak również selekcja oraz redukcja cech.

7.3. Segmentacja

Celem segmentacji jest podział sygnałów pomiarowych na segmenty, w obrębie których mają one podobne właściwości. Rozmiar tych segmentów również zależy od tych właściwości. Do opisu prawdopodobieństwa tych segmentów mogą zostać wykorzystane m.in. cechy statystyczne. Alternatywnym sposobem opisu właściwości tych segmentów jest np. opis w dziedzinie częstotliwości. Metoda oceny podobieństwa podzbiorów powinna wynikać ze specyfiki sygnałów pomiarowych i jest jedną z decyzji projektowych, które należy podjąć podczas konstrukcji algorytmu segmentacji.

Efektom segmentacji sygnału pomiarowego jest ich zestaw o ustalonej długości (tzw. ramki). Segmenty te muszą mieć precyzyjnie ustalony początek i koniec. Do ustalenia początku i końca segmentów wykorzystywane są algorytmy **detekcji zmian** (ang. *change detection*). Zadaniem algorytmów detekcji zmian jest określanie momentów czasowych, w których w sygnale zachodzą istotne zmiany, np. statystyczne.

Miarą jakości algorytmów segmentacji jest liczba niepoprawnych zadziałań (np. wskazanie na zmianę, której nie było lub nie wskazanie na zmianę, która zaistniała) oraz krótki czas zadziałań. Na rysunku 7.2 wskazano kilka typowych błędów, które mogą się pojawić podczas segmentacji sygnałów,

Podsumowując, efektem działania algorytmu segmentacji sygnałów powinien być pewien zestaw N_s segmentów o ustalonej długości:

$$\underbrace{\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k_1)}_{\text{Segment 1}} \quad \underbrace{\mathbf{y}(k_1 + 1), \dots, \mathbf{y}(k_2)}_{\text{Segment 2}} \quad \dots \quad \underbrace{\mathbf{y}(k_{N_s} + 1), \dots, \mathbf{y}(K)}_{\text{Segment } N_s} \quad (7.6)$$

Niech \mathbf{s}_i reprezentuje każdy i -ty segment, wówczas wynikiem segmentacji będzie następująca sekwencja przylegających N_s segmentów:

$$\mathbb{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{N_s}) \quad (7.7)$$

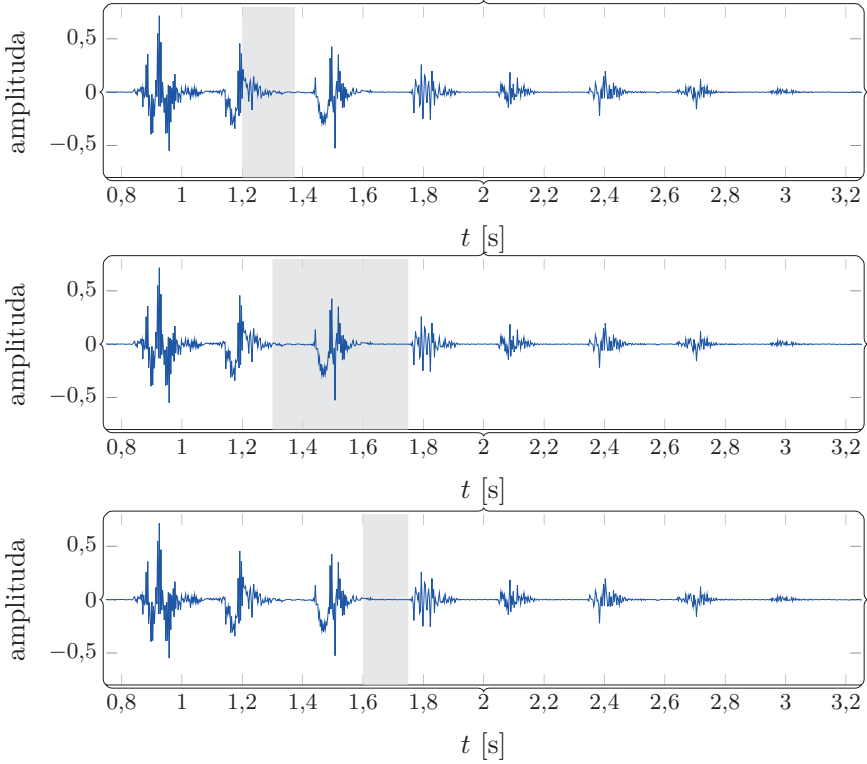
spełniająca następujące właściwości $\mathbf{s}_1 < \mathbf{s}_2 < \dots < \mathbf{s}_{N_s-1} < \mathbf{s}_{N_s}$, $\mathbb{S} = \bigcup_{i=1}^{N_s} \mathbf{s}_i$

oraz $\text{int}(\mathbf{s}_i) \cap \text{int}(\mathbf{s}_j) = \emptyset$ dla wszystkich $\forall i \neq j$. Opisane zadanie segmentacji sygnałów można również przedstawić jako zadanie optymalizacyjne.

Sformułowanie problemu

Dla zestawu sygnałów pomiarowych:

$$\mathbf{Y}_K = (\mathbf{y}(1) \ \mathbf{y}(2) \ \dots \ \mathbf{y}(K)) \quad (7.8)$$



Rys. 7.2. Typowe błędy popełniane podczas segmentacji sygnałów

gdzie K jest liczbą pomiarów, a $\mathbf{y}(k)$ jest \mathcal{D}_y -wymiarowym wektorem:

$$\mathbf{y}(k) = [y_1 \quad y_2 \dots y_{\mathcal{D}_y}]^\top \quad (7.9)$$

oraz zbioru indeksów:

$$\mathcal{I}_K = (1, 2, \dots, k, \dots, K) \quad (7.10)$$

należy znaleźć taką sekwencję N_S indeksów:

$$\mathcal{I}_{N_s} = (k_1, k_2, \dots, k_{N_s}) \quad (7.11)$$

powiązanych ze zmianami w zbiorze (7.8), że

$$(k_1^*, k_2^*, \dots, k_{N_s}^*) = \min_{1 < k < K} Q(\mathbf{Y}_K, \mathcal{I}_K) \quad (7.12)$$

7. Poziom średni: fuzja właściwości

gdzie

$$Q(\mathbf{Y}_K, \mathcal{I}_K) = \sum_{i=1}^{N_S} \sum_{k=k_i}^{k_{i+1}-1} (y(k) - \hat{y}(i))^2 \quad (7.13)$$

a $\hat{y}(i)$ jest średnią wyznaczoną dla N_S -tego segmentu.

W wyniku rozwiązania zadania segmentacji, sygnał pomiarowy zostanie podzielony na segmenty o równej (segmentacja równomierna, ang. *fixed window segmentation*) lub o różnej długości (segmentacja nierównomierna, ang. *variable window segmentation*). Pierwsze z podejść jest relatywnie prostym zadaniem, w którym nie wymaga się dużej dokładności [23]. Drugie z podejść jest znacznie trudniejsze, ponieważ rozmiar segmentu musi być dostosowany do lokalnych właściwości analizowanego sygnału. Stąd też inna nazwa tego podejścia, tj. segmentacja adaptacyjna (ang. *adaptive segmentation*). Trudność w tym przypadku związana jest z tym, że sygnał musi być analizowany na bieżąco i na bieżąco muszą być wychwytywane zmiany w jego właściwościach.

Analiza sygnału przeprowadzana jest zazwyczaj nie na całym dostępnym sygnale, ale na jego części. Próbkę sygnału, które aktualnie są analizowane, wybierane są na podstawie okna czasowego. Okno, w którym odbywa się analiza sygnału, jest sukcesywnie przesuwane podczas analizy sygnału (ang. *window sliding*). Przesuwanie okna może odbywać się w dwóch trybach, tj. z lub bez zakładki okna (ang. *overlapping*). Przesuwanie okna z zakładką umożliwi poprawę jakości analizy sygnału, tj. wpływa korzystnie na jakość detekcji. Na rysunku 7.3 przedstawiono różnicę pomiędzy analizą sygnału z zakładką i bez zakładki okna.

Typowe okno wykorzystywane, nie tylko w segmentacji, ale również w innych metodach analizy sygnałów, ma **kształt prostokątny**. W wielu zadaniach korzystniejsze może okazać się zastosowanie innego niż prostokątne okno czasowe [83]. Inne typy okien czasowych, które mogą być wykorzystane w zadaniu segmentacji to, m.in.:

- okno **Hamminga** (rys. 7.4a)

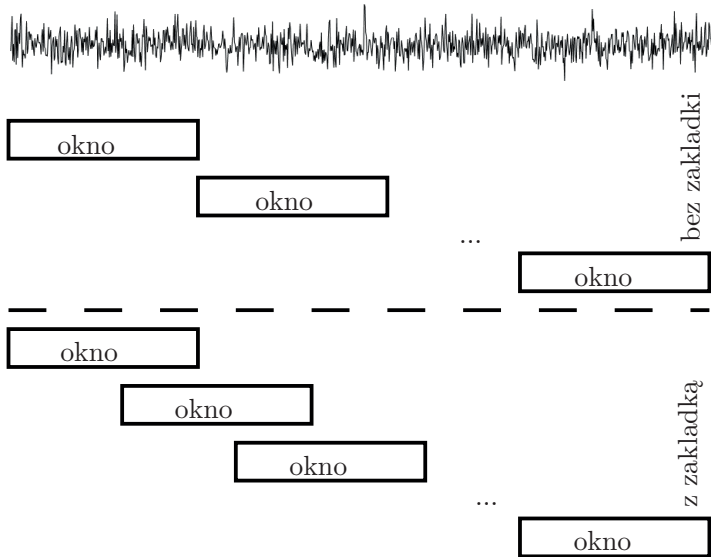
$$w(k) = 0,54 - 0,46 \cos\left(\frac{2\pi k}{K}\right), \quad 0 \leq k \leq K \quad (7.14)$$

- okno **Hanna** (rys. 7.4b)

$$w(k) = 0,5 - 0,5 \cos\left(\frac{2\pi k}{K}\right), \quad 0 \leq k \leq K \quad (7.15)$$

- okno **Blackmana** (rys. 7.4c)

$$w(k) = 0,42 - 0,5 \cos\left(\frac{2\pi k}{K}\right) + 0,08 \cos\left(\frac{4\pi k}{K}\right), \quad 0 \leq k \leq K \quad (7.16)$$



Rys. 7.3. Ilustracja segmentacji z nakładką i bez zakładki okna

- okno **Kaisera** (rys. 7.4d)

$$w(k) = \frac{\mathcal{B}_0\left(\mathfrak{p}\pi\sqrt{1 - \left(\frac{2k}{K-1} - 1\right)^2}\right)}{\mathcal{B}_0(\mathfrak{p}\pi)} \quad (7.17)$$

\mathcal{B}_0 reprezentuje funkcję Bassela zerowego rzędu.

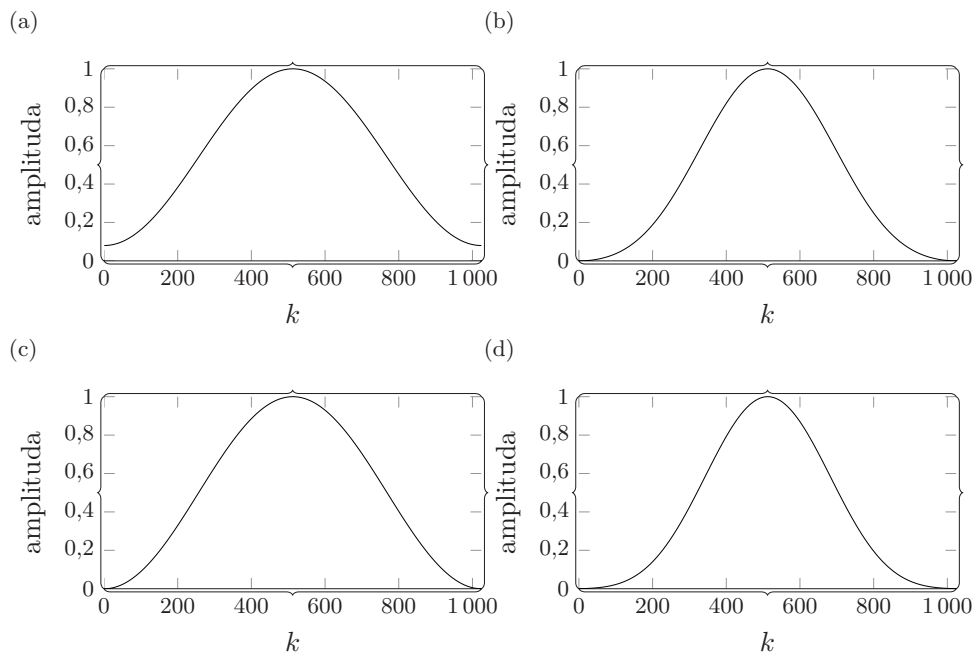
Jak można zauważyć, analizując zależności (7.14)–(7.16) okna Hamminga, Hanninga i Blackmana są ważoną sumą składowych kosinusoidalnych. Z kolei okno Kaisera jest oknem, którego kształt można modyfikować przez dobór odpowiedniej wartości parametru \mathfrak{p} (por. (7.17)).

Oczywiście rodzina okien czasowych jest znacznie liczniejsza, w tym miejscu wybraliśmy i omówiliśmy te z nich, które są najczęściej wykorzystywane. W pracy [77] dokonano analizy, jak kształt okna wpływa na jakość segmentacji sygnałów. Przedstawione w pracy wyniki wskazują, że dobór kształtu okna może wpłynąć na jakość segmentacji sygnału.

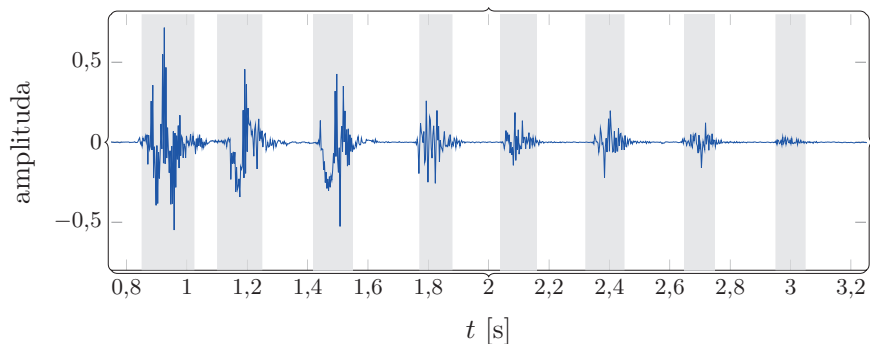
Korzyścią z podziału sygnału na segmenty jest uproszczenie jego analizy oraz zmniejszenie liczby koniecznych do wykonania obliczeń, jednocześnie unikając analizowania tych fragmentów sygnału, w których nie zachodzą istotne zmiany.

Przykład 7.1 (Segmentacja sygnału). Na rysunku 7.5 przedstawiono przykład działania algorytmu segmentacji. W celach ilustracyjnych wybrano sygnał,

7. Poziom średni: fuzja właściwości



Rys. 7.4. Przykłady okien czasowych: a) Hamminga, b) Hann, c) Blackmana oraz d) Kaisera

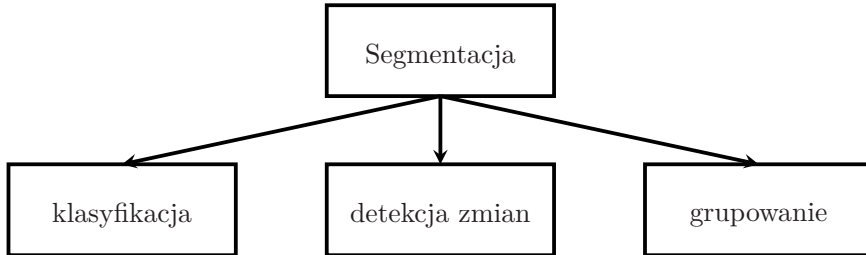


Rys. 7.5. Przykład rozwiązania zadania segmentacji

w którym wyraźnie widać początek i koniec każdego segmentu. Większość spotykanych w praktyce sygnałów ma przebiegi znacznie bardziej złożone i rozróżnienie w nich segmentów może okazać się zadaniem trudnym.

7.3.1. Metody segmentacji

W podrozdziale zostaną omówione podstawowe metody segmentacji sygnałów oraz kluczowe zagadnienia z nią związane. Na rysunku 7.6 przedstawiono systematykę metod segmentacji sygnałów. Wyróżniamy trzy główne podejścia do segmentacji sygnałów. Należą do nich metody segmentacji oparte na detekcji zmian, jak również klasyfikacji i grupowaniu.



Rys. 7.6. Systematyka metod segmentacji

Segmentacja bazująca na detekcji zmian

Jednym z podejść do rozwiązania zadania segmentacji jest wielokrotne rozwiązanie zadania detekcji zmian. Efektem wielokrotnego rozwiązania zadania detekcji zmian są wyznaczone punkty rozpoczęcia i zakończenia każdego segmentu. Istotą metod segmentacji opartych na detekcji zmian jest ustalenie **reguły decyzyjnej**, która służy do podejmowania decyzji pomiędzy dwiema hipotezami na temat sygnału pomiarowego. Najważniejsze cechy, którymi powinna charakteryzować się reguła decyzyjna są:

- jednoznaczność,
- optymalność względem przyjętego kryterium,
- każdy sygnał wejściowy powinien generować decyzję.

Dla zadania segmentacji sygnału hipotezy mogą mieć następującą postać:

$$\begin{aligned}
 \mathcal{H}_0 : y(k) &= s(k) + v(k) & k \in [k_i, k_{i+1}] \\
 \mathcal{H}_1 : y(k) &= v(k) & k \in [k_i, k_{i+1}]
 \end{aligned}
 \tag{7.18}$$

gdzie $y(k)$ jest sygnałem pomiarowym, $v(k)$ jest szumem, a $s(k)$ jest nieznanym sygnałem. W tym przypadku należy zweryfikować hipotezę zerową, tj. sygnał pomiarowy $y(k)$ zawiera, oprócz zakłóceń $v(k)$, również sygnał $s(k)$, wobec hipotezy \mathcal{H}_1 , że analizowany sygnał zawiera jedynie szum.

7. Poziom średni: fuzja właściwości

Jednym z prostszych algorytmów detekcji zmian, który można wykorzystać w zadaniu segmentacji sygnałów, jest **diagram Shewharta** (ang. *Shewhart Control Chart*) [30]. Algorytm ten na bieżąco wykrywa zmiany w wartościach statystycznych sygnału (np. średniej lub wariancji) każdego ustalonego segmentu. Następnie wyznaczana jest reguła decyzyjna, która służy do wyboru pomiędzy dwiema hipotezami (7.18).

Oprócz diagramu Shewharta do detekcji zmian można wykorzystać inne algorytmy, takie jak **diagram geometrycznej średniej ruchomej** (ang. *Geometric Moving Average Control Chart*), **diagram skończonej średniej ruchomej** (ang. *Finite Moving Average Control Charts*) oraz algorytm GLR (ang. *Generalized Likelihood Ratio*) [30].

Popularną metodą detekcji zmian w sygnałach jest **algorytm CUSUM** (ang. *CUmulative SUM*). Algorytm ten charakteryzuje się tym, że umożliwia wykrywanie niewielkich zmian w analizowanym sygnale [30].

Segmentacja bazująca na klasyfikacji

Zadanie segmentacji można również rozwiązać, wykorzystując w tym celu algorytmy klasyfikacji. Jest to podejście, które nazywane jest również podejściem z **uczeniem nadzorowanym** (ang. *supervised learning*). Zaletą tego podejścia jest to, że nie wymaga ono żadnej wiedzy dziedzinowej. Do opracowania stosownych algorytmów potrzebne są natomiast odpowiednio przygotowane zbiory danych uczących. Zbiory te wykorzystywane są na etapie uczenia algorytmów klasyfikacji.

Jedną z wad tego podejścia jest to, że wymagają one bardzo dużych zbiorów danych w celu zapewnienia dobrej jakości segmentacji. W wielu przypadkach problemem może okazać się to, że zbiory, które wykorzystywane są do uczenia klasyfikatorów, są **niefalansowane**. Praktycznie oznacza to, że pewne segmenty w zbiorze uczącym mają wielu reprezentantów, a inne niewielu. Może to prowadzić do problemów podczas uczenia klasyfikatora i mieć wpływ na jakość segmentacji.

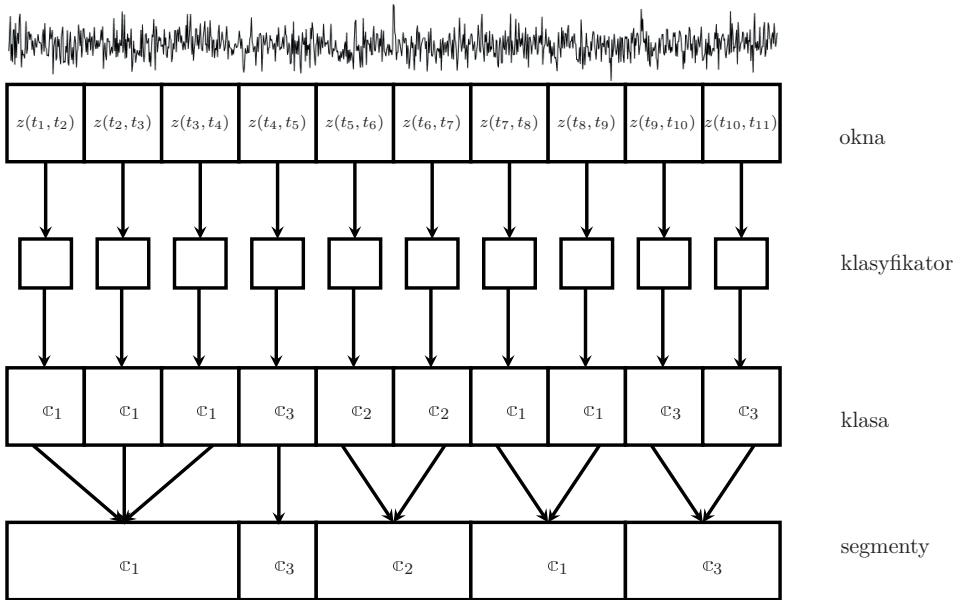
W metodach segmentacji bazujących na klasyfikatorach wyróżniamy dwa główne podejścia, tj. z równomierną i nierównomierną długością segmentu sygnału. Niezależnie od długości segmentu, zadanie klasyfikacji rozwiązywane jest w każdym segmencie oddzielnie, a wyniki następnie łączone w celu uzyskania finalnej segmentacji sygnału. Oznaczmy zbiór sygnałów pomiarowych jako $\mathbf{y}(k)$, $k = 1, 2, \dots, K$, wówczas:

$$c_i(k) = \Upsilon_{sk}(y(k)), \quad k \in [k_i; k_{k+1}] \quad (7.19)$$

gdzie c_i jest i -tą klasą, a $i \in \{1, 2, \dots, N_c\} = \mathcal{N}_c$. \mathcal{N}_c jest zbiorem indeksów klas. Z kolei Υ_{sk} nazywamy algorytmem klasyfikacji. Zadaniem algorytmu klasyfikacji jest przyporządkowanie obiektu do jednej z wcześniej zdefiniowanych klas.

W celu podziału analizowanego sygnału pomiarowego na segmenty, w metodzie bazującej na klasyfikacji wymagany jest drugi krok, którego celem jest ustalenie segmentów na bazie wyników klasyfikacji z poprzedniego kroku. Metoda zastosowana na drugim etapie jest ściśle związana z podejściem zastosowanym na pierwszym etapie.

Jeżeli w każdym z okien klasyfikator (7.19) określa jednoznacznie klasę, to wówczas typowe podejście bazuje na łączeniu wyników klasyfikacji w segmenty. Segmenty te oczywiście mogą mieć różne rozmiary. Metoda ta została zilustrowana na rys. 7.7.



Rys. 7.7. Segmentacja bazująca na klasyfikacji [136]

W przypadku zastosowania algorytmu, którego efektem działania jest wyznaczenie prawdopodobieństw dla każdej z klas, przedstawione wcześniej rozwiązanie można zastąpić podejściem, w którym końcowe przyporządkowanie obiektu do klasy bazuje na wcześniej wyznaczonych klasach. Przykładowym podejściem jest podejście oparte na wygładzaniu (ang. *probability smoothing*) [136].

Pewną niedogodnością metod segmentacji bazujących na algorytmach klasyfikacji jest to, że w zbiorach uczących wszystkie przykłady muszą mieć etykiety. Wymaganie to może okazać się trudne do spełnienia, zwłaszcza w przypadku, gdy zbiór uczący jest duży.

Segmentacja bazująca na grupowaniu

Omówione wcześniej metody segmentacji sygnałów mają pewne wady, które zawężają potencjalne obszary ich zastosowania. Na przykład w algorytmach segmentacji sygnałów bazujących na detekcji zmian konieczne jest m.in. ustalenie wartości progu służącego do ustalenia, która z testowanych hipotez jest prawdziwa. Z kolei metody oparte na algorytmach klasyfikacji wymagają wcześniej przygotowanych zbiorów uczących. Przygotowanie to łączy się z wybraniem etykiety dla każdego elementu ze zbioru uczącego.

Rozwiązaniem może okazać się zastosowanie metod segmentacji bazujących na algorytmach klasteryzacji. Algorytmy te należą do grupy metod opartych na metodach uczenia nienadzorowanego (ang. *unsupervised learning*).

Segmentacja wykorzystująca algorytmy klasteryzacji ma zastosowanie wówczas, gdy w analizowanych sygnałach można wykryć pewne wzorce. Z tym, że poszukiwane wzorce w pomiarach mogą pojawiać się często (np. powtarzające się zdarzenia) lub bardzo rzadko (np. anomalie). Na przykład w sygnałach EKG typowym zadaniem jest wyszukiwanie zespołów QRS. Jest to przykład zadania, w którym wzorec w sygnale pomiarowym pojawia się często. Przykładem dla drugiego typu wzorca są na przykład anomalie w sygnale EEG, które pojawiają się m.in. podczas zaburzeń snu.

Zadanie klasteryzacji sygnałów pomiarowych możemy formalnie przedstawić w formie:

$$\{\mathbb{k}_1, \mathbb{k}_2, \dots, \mathbb{k}_{N_k}\} = \Upsilon_{sg}(y(k)), \quad k \in [k_i; k_{k+1}] \quad (7.20)$$

gdzie \mathbb{k}_i jest i -tym klastrem, $y(k) = \bigcup_{i=1}^{N_k} \mathbb{k}_i$, $\mathbb{k}_i \cap \mathbb{k}_j = \emptyset$ dla $i \neq j$, a Υ_{sg} reprezentuje algorytm grupowania.

Algorytmy grupowania można podzielić na następujące kategorie [6, 170]:

- partycjonowanie (ang. *Partitional Clustering*),
- grupowanie z wykorzystaniem modelu (ang. *Model-based clustering*),
- grupowanie hierarchiczne (ang. *Hierarchical clustering*),
- grupowanie na siatce (ang. *Grid-based clustering*),
- grupowanie gęstościowe (ang. *density-based clustering*).

Problemy, które należy rozwiązać, projektując algorytm grupowania dla zadania segmentacji, związane są z ustaleniem miary odległości pomiędzy punktami pomiarowymi oraz liczbą różnych segmentów, na które ma zostać podzielony analizowany sygnał pomiarowy.

W zależności od zadania, w którym analizowany sygnał ma zostać podzielony na segmenty, analizie może być poddany oryginalny sygnał pomiarowy lub jego

przekształcenie dokonane w dziedzinie czasu, częstotliwości lub czasowo-częstotliwościowej. W następnym rozdziale dokonano przeglądu metod generowania cech w wymienionych zakresach.

7.4. Generowanie cech

Kolejnym zadaniem, które należy rozwiązać w ramach ekstrakcji cech jest zadanie ich generowania (pozyskiwania). Celem jest wydobyć z sygnału pomiarowego takich cech, które ułatwią na kolejnych etapach zbudować wektor cech opisujący analizowany obiekt. Metod, które mogą być wykorzystane do generowania cech jest wiele, przy czym nie ma ogólnych wytycznych co do tego, jakie algorytmy bądź cechy najlepiej nadają się do konkretnego zastosowania. Dowolność doboru cech jest duża, a jakość tego doboru mierzona jest zazwyczaj jakością wektora cech, na podstawie których został on zbudowany.

W ogólności do konstrukcji wektora cech można wykorzystać cechy generowane w dziedzinie czasu, częstotliwości lub czasowo-częstotliwościowej. Warto w tym miejscu również zaznaczyć, że typowa praktyka jest taka, że generuje się nadmiarową liczbę cech. Następnie spośród wygenerowanych cech wybiera się te, które pozwalają na konstrukcję możliwie najlepszego wektora cech. Do wyboru cech, które zostaną wykorzystane do budowy wektora cech, należy dobrać odpowiednie algorytmy selekcji i redukcji cech. Zagadnienia te będą omówione w kolejnych podrozdziałach.

Do generowania cech wykorzystuje się jeden lub kilka algorytmów. Sformułujmy problem konstrukcji tych algorytmów.

Sformułowanie problemu

Dla zestawu sygnałów pomiarowych:

$$\mathbf{Y}_N = (\mathbf{y}(1) \quad \mathbf{y}(2) \quad \dots \quad \mathbf{y}(K)) \quad (7.21)$$

gdzie K jest liczbą pomiarów, $\mathbf{y}(k)$ jest \mathcal{D}_y -wymiarowym wektorem w postaci:

$$\mathbf{y}(k) = [y_1 \quad y_2 \quad \dots \quad y_{\mathcal{D}_y}]^T \quad (7.22)$$

wyznacz

$$\mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z}(k) = \mathcal{G}_{\text{ch}}(\mathbf{y}(k)), \quad k = 1, 2, \dots, K \quad (7.23)$$

7. Poziom średni: fuzja właściwości

gdzie

$$\mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z}(k) = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,\mathcal{D}_z} \\ z_{2,1} & z_{2,2} & \dots & z_{2,\mathcal{D}_z} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_y,1} & z_{\mathcal{D}_y,2} & \dots & z_{\mathcal{D}_y,\mathcal{D}_z} \end{bmatrix} \quad (7.24)$$

a \mathcal{D}_z jest liczbą wygenerowanych cech dla każdej z \mathcal{D}_y składowych wektora $\mathbf{y}(k)$.

Transformacja z wykorzystaniem metody G_{ch} (7.23) może być wykonana pomiędzy następującymi dziedzinami:

- czasu,
- częstotliwości,
- czasowo-częstotliwościowej.

W dalszej części rozdziału przedstawione zostaną metody analizy i generowania cech na podstawie sygnałów pomiarowych. W tym celu przyjmujemy, że dysponujemy pomiarami $y(k)$, $k = 1, 2, \dots, K$, które mogą być – jeżeli jest to konieczne – podzielone na segmenty, tj.

$$\mathbf{Y}_K = \left(\underbrace{\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(k_1)}_{\text{Segment 1}} \quad \underbrace{\mathbf{y}(k_1 + 1), \dots, \mathbf{y}(k_2)}_{\text{Segment 2}} \quad \dots \quad \dots \quad \underbrace{\mathbf{y}(k_{N_s-1} + 1), \dots, \mathbf{y}(K)}_{\text{Segment } N_s} \right)$$

Przyjmijmy również, że czas próbkowania sygnału obserwowanego to Δt .

7.4.1. Metody generowania cech w dziedzinie czasu

Pierwsza grupa cech, które można wykorzystywać na poziomie fuzji właściwości, generowana jest w zakresie czasu. Typowe cechy, które mogą zostać w tym przypadku wyznaczone to m.in. cechy statystyczne. Dalej omówiliśmy wybrane z nich.

Cechy statystyczne

Cechy te należą do takich, które najłatwiej wyznaczyć i stąd można tłumaczyć ich popularność podczas charakteryzowania sygnałów pomiarowych. W tabeli 7.1 zestawiono najważniejsze z nich.

Tabela 7.1. Zestawienie przykładowych cech statystycznych

| Cecha | Zależność |
|------------------------|--|
| Średnia arytmetyczna | $\hat{y} = \frac{1}{K} \sum_{k=1}^K y(k)$ |
| Odchylenie standardowe | $\sigma = \sqrt{\frac{1}{K} \sum_{k=1}^K (y(k) - \hat{y})^2}$ |
| Autokowariancja | $\sigma_{yy} = \frac{1}{K-1} \sum_{k=1}^K (y(k) - \hat{y})(y(k-m) - \hat{y})$ $m = 0, 1, 2, \dots, M$ |
| Kowariancja | $\sigma_{ij} = \frac{1}{K-1} \sum_{k=1}^K (y_i(k) - \hat{y}_i)(y_j(k) - \hat{y}_j)$ |
| Autokorelacja | $R_{yy} = \frac{\sum_{k=1}^{K-m} (y(k) - \hat{y})(y(k+m) - \hat{y})}{\left(\sum_{k=1}^{K-m} (y(k) - \hat{y})^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^K (y(k+m) - \hat{y})^2\right)^{\frac{1}{2}}}$ dla $m = 0, 1, 2, \dots, M_{\max}$ |
| Korelacja wzajemna | $R_{ji} = \frac{\sum_{k=1}^K (y_j(k) - \hat{y}_j)(y_i(k) - \hat{y}_i)}{\left(\sum_{k=1}^K (y_j(k) - \hat{y}_j)^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^K (y_i(k) - \hat{y}_i)^2\right)^{\frac{1}{2}}}$ |
| Kurtoza | $\mathcal{K}_{urt} = \frac{\frac{1}{K} \sum_{k=1}^K (y(k) - \hat{y})^4}{\sigma^4} - 3$ |
| Współczynnik skośności | $S\mathcal{K}\mathcal{E} = \frac{\frac{1}{K} \sum_{k=1}^K (y(k) - \hat{y})^3}{\left(\frac{1}{K-1} \sum_{k=1}^K (y(k) - \hat{y})^2\right)^{3/2}}$ |

7. Poziom średni: fuzja właściwości

Analiza amplitudy, okresu i liczby przejść przez zero

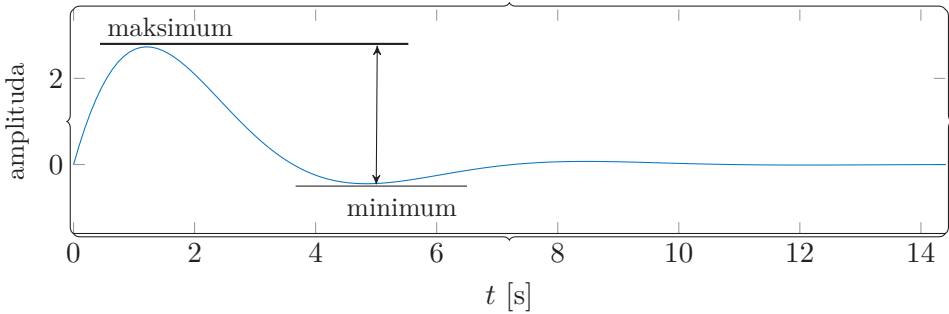
Przejście przez zero to miejsce, w którym sygnał przecina oś odciętych. Jest to prosty sposób na określenie jednej z właściwości sygnału, która związana jest z jego zmianą. W wielu przypadkach dane o przecięciu osi są uzupełniane o dane dotyczące okresu i amplitudy sygnału:

- średni moduł różnicy amplitud (ang. *average magnitude difference function*):

$$amd_M = \frac{1}{K} \sum_{k=1}^K |y(k) - y(k - m)| \quad (7.25)$$

gdzie $m = 1, 2, \dots, M$,

- stosunek maksimum do minimum (ang. *peak-to-valley ratio*) (rys. 7.8),



Rys. 7.8. Ilustracja wyznaczania stosunku maksimum do minimum

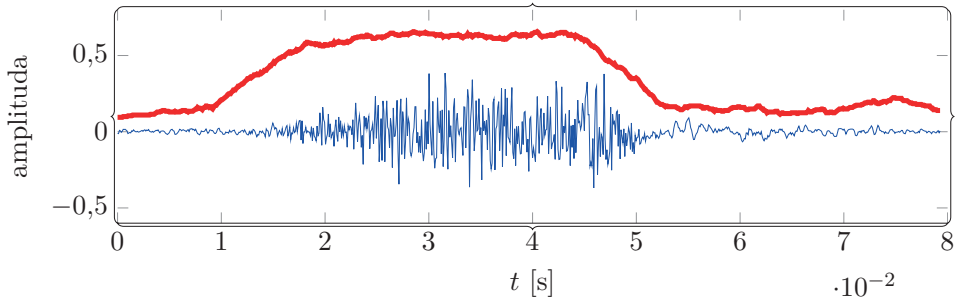
- przejścia przez zero (ang. *zero crossing rate*) (rys. 7.9). Liczbę przejść przez zero wyznaczamy z zależności:

$$zcr = \frac{1}{K} \sum_{k=1}^K |\text{sgn}(y(k)) - \text{sgn}(y(k-1))| \quad (7.26)$$

gdzie

$$\text{sgn}(y(k)) = \begin{cases} 1 & \text{jeżeli } y(k) \geq 0 \\ -1 & \text{jeżeli } y(k) < 0 \end{cases} \quad (7.27)$$

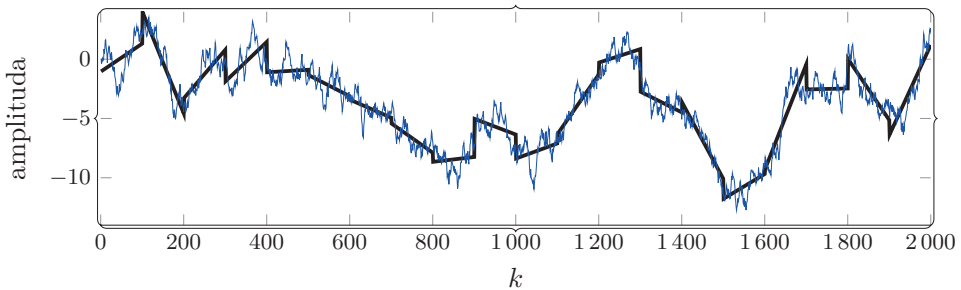
dla $k = 1, 2, \dots, K$.



Rys. 7.9. Ilustracja wyznaczania liczb przejść przez zero

Beztrendowa analiza fluktuacyjna

Beztrendowa analiza fluktuacyjna została opracowana na potrzeby analizy długozasięgowych korelacji w sygnałach [197]. Z czasem znaleziono wiele nowych zastosowań dla tej metody, m.in. w analizie sygnałów biomedycznych. Na przykład wykorzystano algorytm DFA w analizie sygnałów EEG [230, 379]. Na rysunku 7.10 zilustrowano działanie metody DFA. Popularność tego podejścia związana jest z tym, że DFA ułatwia analizę sygnałów niestacjonarnych, zwłaszcza gdy składowe niskoczęstotliwościowe zmieniają się bardzo powoli. Metoda ta znalazła również zastosowanie do generowania cech sygnałów na potrzeby fuzji na poziomie średnim [251].



Rys. 7.10. Beztrendowa analiza fluktuacyjna dla przykładowego sygnału

Operator energetyczny Teagera–Kaisera

Operator Teagera–Kaisera (ang. *Teager–Kaiser Energy Operator*, TKEO) oryginalnie został opracowany na potrzeby pomiarów energii procesów mechanicznych będących źródłem niestacjonarnych sygnałów pomiarowych [193]. Z perspektywy

7. Poziom średni: fuzja właściwości

przetwarzania sygnałów TKEO jest nieliniowym filtrem o wąskim paśmie. W przetwarzaniu sygnałów jest on często wykorzystywany do eliminacji głównej składowej z sygnału [413]. Operator ten może być również wykorzystywany do przekształcenia sygnału pomiarowego na potrzeby ekstrakcji cech [194]. W tym podrozdziale skupimy się na tym zastosowaniu operatora energetycznego Teagera–Kaisera.

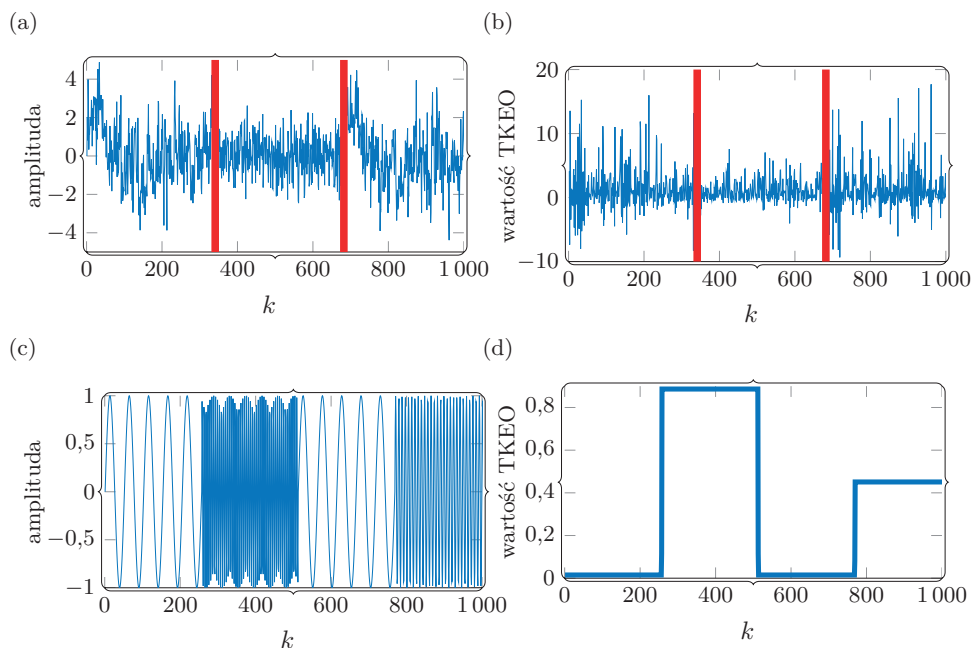
Operator energetyczny Teagera–Kaisera ma postać następującej zależności:

$$tk_{eo}(y(t)) = (\dot{y}(t))^2 - y(t)\ddot{y}(t) \quad (7.28)$$

Jego wersja dyskretna ma z kolei formę:

$$tk_{eo}(y(k)) = y^2(k) - y(k+1)y(k-1) \quad (7.29)$$

Analizując zależność (7.29), zwraca uwagę prostota operatora oraz to, że do jego wyliczenia wystarczą trzy próbki. Na rysunku 7.11 przedstawiono wartość operatora Teagera–Kaisera dla różnych typów sygnałów. Pierwszym z sygnałów, który



Rys. 7.11. Przykład działania operatora Teagera–Kaisera dla różnych typów sygnałów:
 a) szum biały i ułamkowy szum *gaussowski*,
 b) wartość operatora Teagera–Kaisera dla sygnału a, c) sygnał o zmiennej częstotliwości,
 d) wartość operatora Teagera–Kaisera dla sygnału c

poddano analizie z wykorzystaniem TKEO był sygnał będący połączeniem białego szumu z ułamkowym szumem *gaussowskim* (lewa strona rys. 7.11). Analizując wykresy, możemy zauważyć, że zastosowanie operatora Teagera–Kaisera ułatwia wychycenie momentu zmiany typu szumu z białego na ułamkowy szum *gaussowski*.

Podobną analizę przeprowadzono dla sygnału drugiego, który był sygnałem o zmiennej częstotliwości. Na rysunku (prawa strona rys. 7.11) możemy zauważyć, że operator TKEO wykrywa zmiany w częstotliwości analizowanego sygnału. Z analizy wynika, że operator można wykorzystać nie tylko do generowania cech na potrzeby fuzji właściwości, ale również do segmentacji sygnału.

Operator TKEO znalazł zastosowanie w zadaniach związanych z ekstrakcją cech. Na przykład w pracy [342] wykorzystano operator do przekształcenia sygnału pomiarowego do dziedziny TKEO, a następnie rozwiązano zadanie ekstrakcji cech. Wykorzystanie tego operatora może znacząco poprawić jakość analizy sygnałów niestacjonarnych. Podobne zastosowanie TKEO omówiono w pracy [411]. W tym przypadku operator Teagera–Kaisera wykorzystano w analizie sygnału EMG.

Entropia Tsallisa

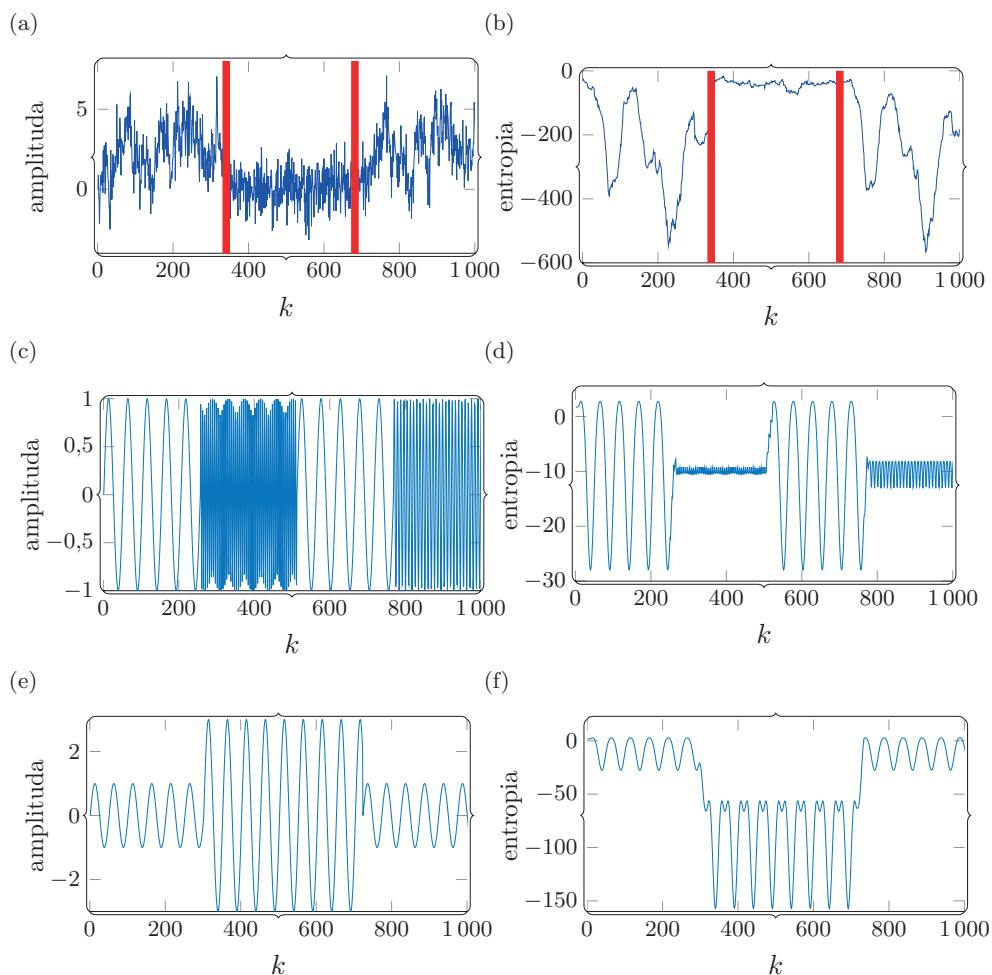
Entropia Tsallisa (ang. *Tsallis Entropy*, TE) jest historycznie pierwszym nielogarytmicznym sformułowaniem entropii. TE wyliczana jest z następującej zależności [99, 69]:

$$\text{ent}_{\text{te}}(\mathbf{q}) = -\mathfrak{d} \frac{1 - \sum_{i=1}^{N_{\text{ms}}} p^{\mathfrak{q}}(i)}{\mathfrak{q} - 1} \quad (7.30)$$

gdzie \mathfrak{d} jest dodatnią stałą, $\mathfrak{q} \in \mathbb{R}^1$ reprezentuje parametr nieekstensywności systemu, p to prawdopodobieństwo wystąpienia pewnego zdarzenia, N_{ms} stanowi liczbę mikrostanów systemu. Warto zaznaczyć, że entropia Tsallisa jest uogólnieniem entropii Boltzmanna–Gibbsa [32].

Na rysunku 7.12 przedstawiono, w jaki sposób entropia Tsallisa może zostać wykorzystana do analizy i generowania cech z sygnałów pomiarowych. Do badania właściwości TE wykorzystano trzy sygnały, tj. sygnał będący połączeniem szumu białego z szumem ułamkowym *gaussowskim*, sygnał o zmiennej częstotliwości oraz sygnał o zmiennej amplitudzie (lewa strona rys. 7.12). Z prawej strony tego rysunku przedstawiono z kolei wartość entropii Tsallisa dla tych sygnałów. Przykłady demonstrują, że entropia Tsallisa umożliwia wychwytywanie zmian zarówno w częstotliwości, jak i w amplitudzie sygnału. Jak wynika z rysunków, TE może być również wykorzystywana do tego, by odróżnić szum biały od szumu ułamkowego *gaussowskiego*.

7. Poziom średni: fuzja właściwości



Rys. 7.12. Przykład działania entropii Tsallisa dla różnych typów sygnałów:
 a) szum biały i ułamkowy szumu *gaussowski* b) entropia Tsallisa dla sygnału a,
 c) sygnał o zmiennej częstotliwości, d) entropia Tsallisa dla sygnału c,
 e) sygnał o zmiennej amplitudzie, f) entropia Tsallisa dla sygnału e

Entropia Tsallisa może być wykorzystana nie tylko, jak to wynika z przeprowadzonej analizy, do detekcji zmian w sygnale, a co jest z tym związane również do ich segmentacji. Jest to narzędzie, które może być wykorzystywane do generowania cech, np. z sygnałów niestacjonarnych takich, jak sygnał EEG [3, 20].

Energia sygnału

Inną wielkością, która może być wyznaczana w dziedzinie czasu jest energia sygnału:

$$\text{eny} = \sum_{k=1}^K |y(k)|^2 \quad (7.31)$$

Przykład zastosowania energii sygnału do generacji cech przedstawiono m.in. w pracy [251]. Energia wyznaczana z sygnału pomiarowego może być również wykorzystana do wstępnego przygotowania danych, co zostało zaproponowane m.in. w pracy [465].

7.4.2. Metody generowania cech w dziedzinie częstotliwości

Do konstrukcji wektora cech można wykorzystywać nie tylko cechy wygenerowane w czasie, ale również w dziedzinie częstotliwości. Metody te bazują głównie na analizie widmowej (spektralnej). Główne narzędzia do wyznaczania cech wykorzystują zarówno metody statystyczne, jak również przekształcenie takie, jak transformacja Fouriera czy Hilberta. Wyróżniamy szczególnie następujące podejścia do analizy spektralnej sygnałów:

- nieparametryczne (klasyczne),
- parametryczne.

W kolejnych podpunktach omówimy wybrane, najbardziej reprezentatywne metody.

Metody nieparametryczne

Metody nieparametryczne opierają się głównie na transformacji Fouriera lub dyskretnej transformacji Fouriera. Algorytmy należące do tej grupy charakteryzują się tym, w przeciwieństwie do metod parametrycznych, że nie bazują one na żadnych modelach matematycznych. Zadaniem tych metod jest estymacja funkcji gęstości widmowej mocy (GWM). Metody te mogą być wykorzystane m.in. do generowania cech sygnałów.

Jedną z podstawowych metod, która jest wykorzystywana do estymacji GWM jest **periodogram**. Periodogram jest kwadratem modułu z transformacji Fouriera sygnału. Wyniki estymacji GWM uzyskane tą metodą charakteryzują się niską jakością. Szczególnie jest to widoczne w przypadku, gdy analizie poddawane są sygnały nieperiodyczne. Najlepsze wyniki estymacji widma z wykorzystaniem periodogramu można otrzymać dla sygnałów okresowych. Wśród estymatorów widma należy również wymienić inne metody, które są poprawionymi wersjami periodogramu. Do tych metod zaliczamy m.in. **zmodyfikowany periodogram**,

7. Poziom średni: fuzja właściwości

uśredniony periodogram. Zmodyfikowany periodogram wyznaczany jest jako kwadrat modułu dyskretnej transformacji Fouriera iloczynu K próbek analizowanego sygnału $s(k)$ z próbkami funkcji okna. Z kolei uśredniony periodogram jest wynikiem uśrednienia N_{pd} periodogramów dla rozłącznych fragmentów analizowanego sygnału. Inne metody, których efekty działania przedstawiono na rys. 7.14 oraz 7.15 to metody **uśrednionego zmodyfikowanego periodogramu** (nazywana metodą **Welcha**), oraz **wygładzonego periodogramu** (metoda **Blackmana–Tukeya**).

Jedną z głównych wad metod nieparametrycznych jest zjawisko przecieku widma (ang. *spectral leakage*). Przecieki widma to zjawisko, które sprawia, że w widmie pojawiają się składowe częstotliwości, których nie ma w analizowanym sygnale. Z tego powodu estymowane widmo jest tylko przybliżeniem widma analizowanego sygnału.

Metody z przedstawionej grupy umożliwiają uzyskanie zadowalających wyników wówczas, gdy liczba próbek sygnału jest dostatecznie duża. Metody te charakteryzują się również niską rozdzielczością oraz dużą wariancją (zmiennością) estymaty widmowej [484],

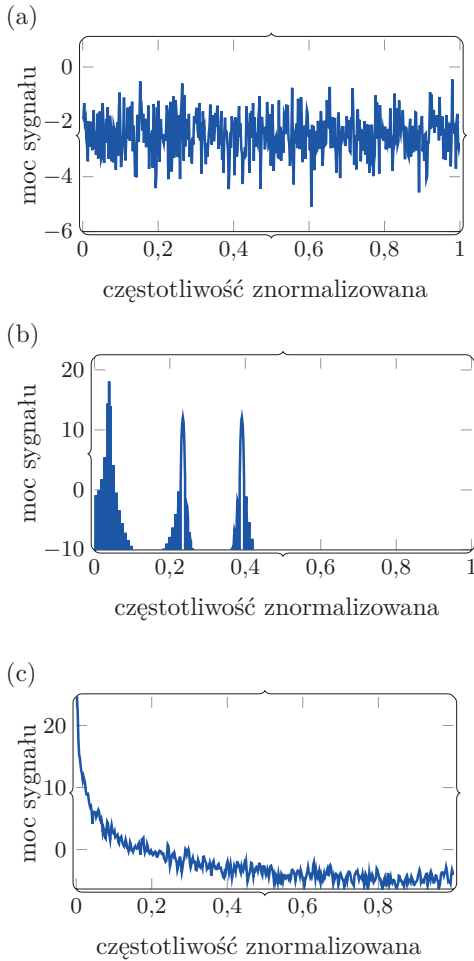
Na koniec dokonamy porównania trzech z wymienionych metod, tj. periodogramu, metody Welcha oraz Blackmana–Tukeya. W badaniach wykorzystano szum biały, szum ułamkowy *gaussowski* oraz sygnał o zmiennej częstotliwości. Na rysunkach 7.13–7.15 przedstawiono przykładowe oszacowania GWM tych sygnałów z wykorzystaniem periodogramu, metody Welcha i metody Blackmana–Tukeya. Zauważmy, że na rys. 7.14 widoczny jest efekt przecieku widma.

Metody parametryczne

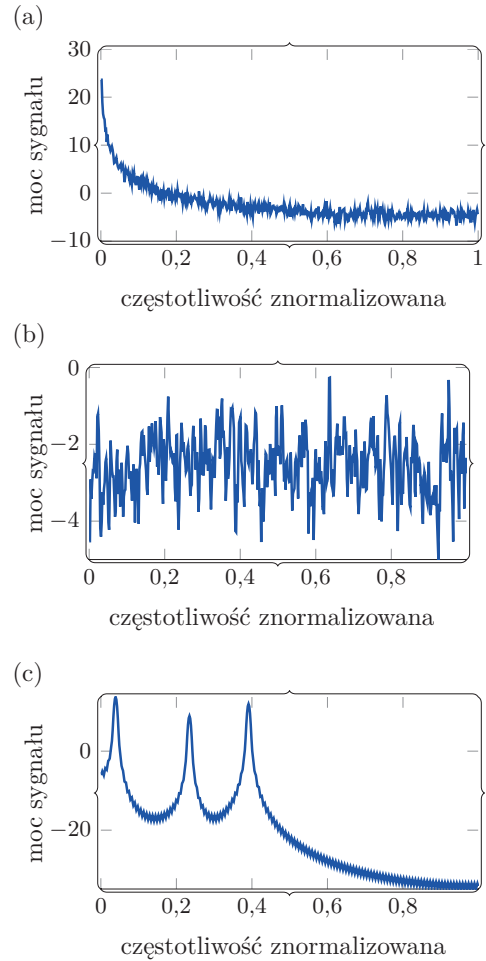
Metody parametryczne, w przeciwieństwie do metod nieparametrycznych, do estymacji gęstości widma mocy sygnału wykorzystują modele matematyczne. Zaletą tych metod jest wyższa rozdzielczość częstotliwościowa estymowanego widma w porównaniu do metod nieparametrycznych. W metodach parametrycznych nie występuje również problem przecieku widma.

Parametryczne metody estymacji widma głównie bazują na modelach typu AR, MA lub ARMA (podrozdz. 5.7.1). Wykorzystanie metod parametrycznych do estymacji GWM sygnału wiąże się z wyborem odpowiedniego modelu, jego rzędu i rozwiązaniem zadania estymacji parametrów modelu. Może się to wiązać z większym nakładem obliczeń oraz wymaga większego doświadczenia od osoby stosującej metody parametryczne.

Metody korzystające z modelu MA dają dobre wyniki estymacji widm ciągłych, np. kolorowego szumu. Z kolei metody oparte na modelu AR wykorzystywane są w przypadku widm prążkowych, tj. takich, które złożone są z sygnałów sinuso-



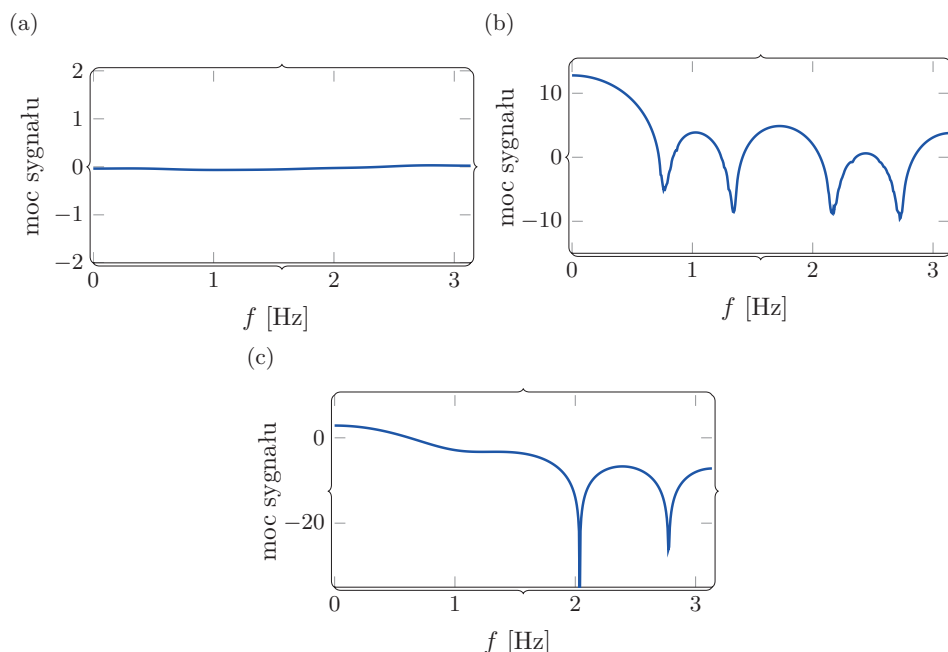
Rys. 7.13. Periodogramy dla wybranych sygnałów: a) szum biały, b) ułamkowy szum *gaussowski*, c) sygnał o zmiennej częstotliwości (patrz rys. 7.11)



Rys. 7.14. Widmo mocy sygnału estymowanego metodą Welch: a) szum biały, b) ułamkowy szum *gaussowski*, c) sygnał o zmiennej częstotliwości (patrz rys. 7.11)

idealnych. Natomiast modele ARMA są efektywne w przypadku widm ciągłych z wybijającymi się prążkami. Widma tego typu są efektem nałożenia na siebie kilku sygnałów sinusoidalnych wraz z kolorowym szumem [484].

7. Poziom średni: fuzja właściwości



Rys. 7.15. Widmo mocy sygnału estymowanego metodą Blackmana–Tukeya:

- a) szum biały, b) ułamkowy szumu *gaussowski*,
- c) sygnał o zmiennej częstotliwości (patrz rys. 7.11)

Jeśli model będzie właściwie dobrany do sygnału, to jego rząd nie będzie wysoki. Podstawowym zadaniem konstrukcyjnym w tej grupie metod jest odpowiedni dobór modelu oraz ustalenie jego rzędu.

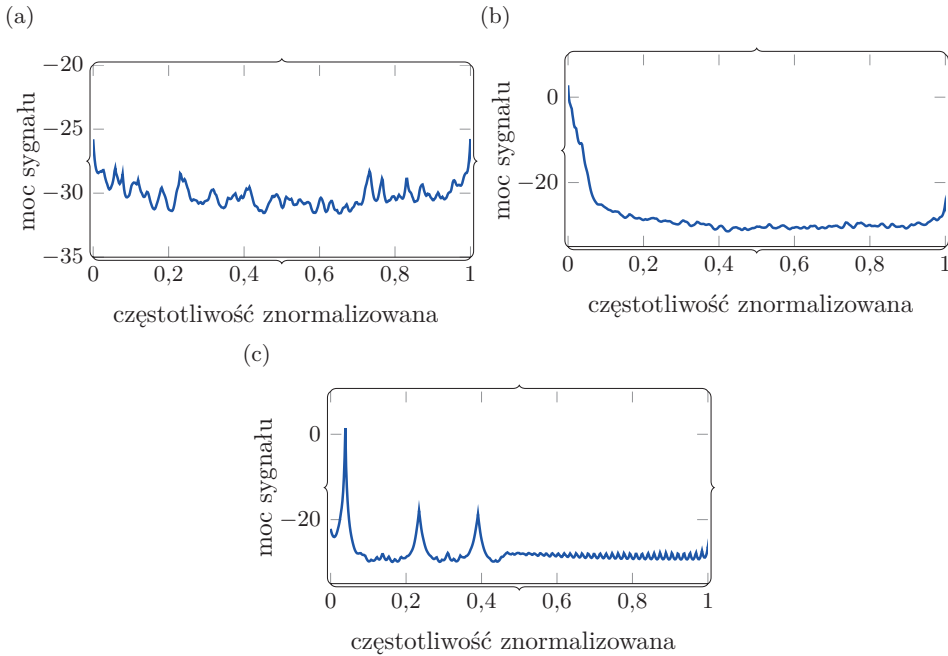
Warto zaznaczyć, że parametry modeli wykorzystanych do analizy spektralnej mogą być wykorzystane jako cechy opisujące analizowany sygnał.

Metody podprzestrzeni

Metody podprzestrzeni (ang. *subspace methods*) są odmianą metod parametrycznych. Metody z tej kategorii to efekt badań nad macierzą autokorelacji. Zauważono, obserwując wektory własne tej macierzy, że gdy sygnał jest sumą kilku składowych sinusoidalnych i szumu białego, to część z tych wektorów zależy głównie od składowych harmonicznych sygnału, a część wyłącznie od szumu [484]. Istnieje grupa metod, która wykorzystuje wnioski z tej obserwacji. Jedną z metod należących do tej kategorii jest algorytm MUSIC (ang. *MUltiple Signal Classification*) [484]. W metodzie tej, co jest charakterystyczne dla metod z tej grupy, wykorzystuje się kilka wektorów własnych z podprzestrzeni szumów w celu wyznaczania widma sygnału.

Metody z tej grupy charakteryzują się dobrą rozdzielczością oraz odpornością na szumy. Kosztem z kolei jest duża złożoność obliczeniowa w porównaniu do metod wykorzystujących transformację Fouriera.

Na rysunku 7.16 przedstawiono przykład działania algorytmu MUSIC dla szumu białego, ułamkowego szumu gaussowskiego oraz sygnał o zmiennej częstotliwości.



Rys. 7.16. Widmo mocy sygnału estymowane metodą MUSIC: a) biały szum, b) ułamkowy szum *gaussowski*, c) sygnał o zmiennej częstotliwości (patrz rys. 7.11)

Analiza spektralna wyższego rzędu

Analiza spektralna wyższego rzędu (ang. *Higher-Order Spectral Analysis*, HOS) to narzędzie przeznaczone do analizy sygnałów niestacjonarnych i z szumem o *nie-gaussowskiej* charakterystyce. Klasyczne metody analizy sygnałów, które zostały omówione w poprzednich rozdziałach, bazują na momentach drugiego rzędu, tj. funkcje korelacji lub widmo sygnału etc. Natomiast HOS jest widmową reprezentacją statystyk wyższego rzędu [80], tj. rzędu trzeciego, czwartego i wyższych. Oprócz tego w analizie HOS wprowadza się nieliniowe kombinacje momentów wyższych rzędów, które nazywane są **kumulantami**. Podejście to charakteryzuje

7. Poziom średni: fuzja właściwości

się właściwościami, które mogą być przydatne w zadaniach przetwarzania i fuzji danych, na przykład na poziomie średnim [124].

Jedną z wad omawianych dotychczas klasycznych metod estymacji widma sygnału jest to, że zawierają one jedynie informacje o jego amplitudzie [235]. Oznacza to, że tracona jest informacja dotycząca przesunięć fazowych pomiędzy poszczególnymi składowymi analizowanego sygnału. W przypadku sygnałów o rozkładach innych niż normalny, informacja która jest tracona podczas przetwarzania sygnałów z wykorzystaniem metod klasycznych może być źródłem błędów i może uniemożliwić precyzyjną analizę sygnału [124].

Rozwiązaniem jest stosowanie metod analizy spektralnej wyższego rzędu, których specyficzne właściwości umożliwiają analizę sygnałów o rozkładach innych niż normalne. Kumulanty, które są nieliniowymi kombinacjami momentów wyższych rzędów charakteryzują się tym, że ich wartość jest całkowicie niezależna od sygnałów o rozkładzie normalnym [234]. Wniosek praktyczny z tej właściwości kumulantów jest taki, że stosując metody wyższych rzędów dla sygnałów o rozkładach innych niż rozkład *gaussowski*, ale z szumem o rozkładzie normalnym, możemy znacznie poprawić jakość takiego sygnału.

Na podstawie wyznaczonych momentów i kumulantów można wyznaczyć widma wyższych rzędów **polispektra**, tj. **bispektrum**, **trisppektrum** etc. Dodajmy, że kumulanty i polispektra zawierają więcej informacji w porównaniu do korelacji i widm, tj. te pierwsze zawierają informację zarówno o amplitudzie, jak i fazie, podczas gdy drugie z wymienionych dają jedynie informację o amplitudzie sygnału.

Przedstawmy teraz praktyczną metodę obliczania momentów wyższych rzędów. Zależności do wyznaczenia momentów n -tego rzędu są następujące [80]:

$$\hat{y}_1 = E[y(k)] \quad (7.32a)$$

$$\hat{y}_2 = E[y(k)y(k+m)] \quad (7.32b)$$

$$\hat{y}_3 = E[y(k)y(k+m_1)y(k+m_2)] \quad (7.32c)$$

$$\hat{y}_4 = E[y(k)y(k+m_1)y(k+m_2)y(k+m_3)] \quad (7.32d)$$

gdzie $\mathbf{Y}_K = (y(1) \ y(2) \ \dots \ y(K))$, m_i jest opóźnieniem (przesunięciem wstecz) sygnału. Z kolei kumulanty wyznaczamy (bazując na formułach (7.32a)–(7.32d)) z następujących zależności [80]:

$$\mathbf{u}_1 = \hat{y}_1 \quad (7.33a)$$

$$\mathbf{u}_2(m) = \hat{y}_2(m) - (\hat{y}_1)^2 \quad (7.33b)$$

$$\mathbf{u}_3(m_1, m_2) = \hat{y}_3(m_1, m_2) - \hat{y}_1[\hat{y}_2(m_1) + \hat{y}_2(m_2) + \hat{y}_2(m_2 - m_1)] + 2(\hat{y}_1)^3 \quad (7.33c)$$

$$\begin{aligned}
u_4(m_1, m_2, m_3) = & \hat{y}_4(m_1, m_2, m_3) - \hat{y}_2(m_1)\hat{y}_2(m_3 - m_2) - \hat{y}_2(m_2)\hat{y}_2(m_3 - m_1) + \\
& - \hat{y}_2(m_3)\hat{y}_2(m_2 - m_1) - \hat{y}_1[\hat{y}_3(m_2 - m_1, m_3 - m_1) + \hat{y}_3(m_2, m_3) + \\
& + \hat{y}_3(m_1, m_2)] + (\hat{y}_1)^2[\hat{y}_2(m_1) + \hat{y}_2(m_2) + \hat{y}_2(m_3) + \\
& + \hat{y}_2(m_3 - m_1) + \hat{y}_2(m_3 - m_2) + \hat{y}_2(m_2 - m_1)] - 6(\hat{y}_1)^4 \quad (7.33d)
\end{aligned}$$

Na podstawie obliczonych momentów i kumulantów możliwe jest wyznaczenie widm wyższego rzędu. Na przykład widmo bazujące na kumulantach trzeciego rzędu nazywamy **bispektrum**. Z kolei **trisppektrum** wykorzystuje kumulanty czwartego rzędu. Bispektrum i trisppektrum są uogólnieniami widma mocy, które wyznaczane jest dla kumulantów drugiego rzędu [80]. Ogólna formuła na wyliczanie widma n -tego rzędu dla sygnału $y(k)$ jest następująca:

$$\begin{aligned}
S_n(f_1, f_2, \dots, f_n) = \\
= \sum_{m_1=-\infty}^{\infty} \dots \sum_{m_{n-1}=-\infty}^{\infty} u(m_1, m_2, \dots, m_{n-1}) \exp\left(-j \sum_{i=1}^{n-1} 2\pi f_i m_i\right) \quad (7.34)
\end{aligned}$$

gdzie f_i reprezentuje i -tą składową częstotliwości sygnału.

Dla przykładu, jeśli przyjmiemy, że $n = 3$, to formuła (7.34) do wyznaczenia widma trzeciego rzędu, przyjmuje postać:

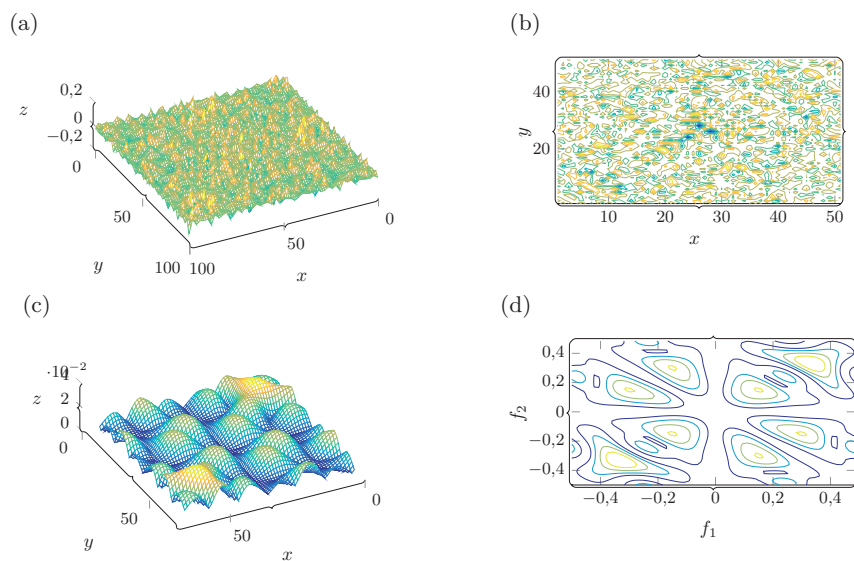
$$S_3(f_1, f_2) = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} u_2(m_1, m_2) \exp(-j(2\pi f_1 m_1 + 2\pi f_2 m_2)) \quad (7.35)$$

Przykłady wykorzystania statystyk wyższego rzędu do konstrukcji wektora cech można znaleźć m.in. w pracach [305, 464]. W pierwszej pracy podejście HOS zostało wykorzystane do ekstrakcji cech z sygnału EEG. W drugiej pracy wykorzystano to narzędzie do ekstrakcji cech z szmerów oddechowych (ang. *respiratory sounds*).

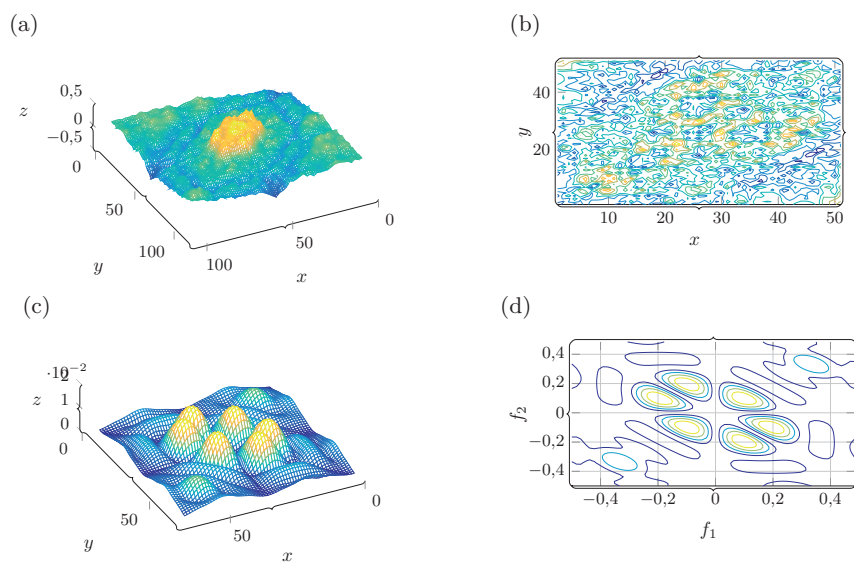
Przykład 7.2 (Przykłady cech wyznaczonych z wykorzystaniem analizy spektralnej wyższego rzędu). *Przykład ilustruje sposób wykorzystania analizy HOS. Zarówno kumulanty jak i bispektrum zostały wyznaczone dla trzech sygnałów, tj. szum biały, ułamkowy szum gaussowski i sygnał o zmiennej częstotliwości. Wyniki przedstawiono kolejno na rys. 7.17–7.19.*

Porównując kumulanty dla szumu białego i szumu ułamkowego gaussowskiego (rys. 7.17 oraz 7.18), widoczne są wyraźne różnice pomiędzy nimi. Z kolei różnica pomiędzy kumulantami dla szumu białego i sygnału o zmiennej częstotliwości nie jest tak wyraźna (rys. 7.17 oraz 7.19). Obserwacje potwierdzają, że metody HOS pozwalają na oddzielenie sygnałów niegaussowskich od szumów o rozkładzie normalnym.

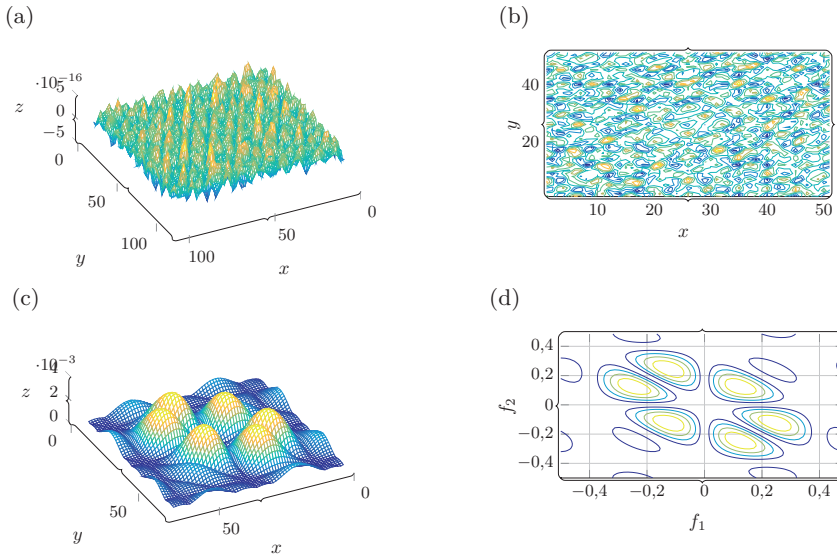
7. Poziom średni: fuzja właściwości



Rys. 7.17. Kumulanty i bispektrum oraz ich przekroje dla szumu białego: a) kumulanty, b) kumulanty (przekrój), c) bispektrum, d) bispektrum (przekrój)



Rys. 7.18. Kumulanty i bispektrum oraz ich przekroje dla szumu ułamkowego *gaussowskiego*: a) kumulanty, b) kumulanty (przekrój), c) bispektrum, d) bispektrum (przekrój)



Rys. 7.19. Kumulanty i bispektrum oraz ich przekroje dla sygnału o zmiennej częstotliwości (patrz rys. 7.11): a) kumulanty, b) kumulanty (przekrój), c) bispektrum, d) bispektrum (przekrój)

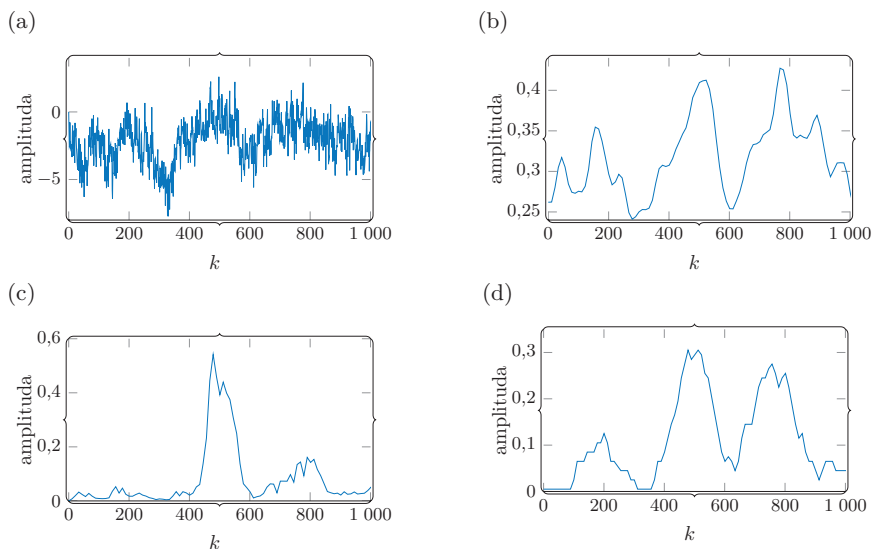
Inne cechy w dziedzinie częstotliwości

Oprócz omówionych do tej pory cech, które generowane są w dziedzinie częstotliwości, wyróżnić można wiele innych, które przydatne są na poziomie fuzji właściwości. Do najważniejszych cech, które mogą być wykorzystane do konstrukcji wektora cech, należą:

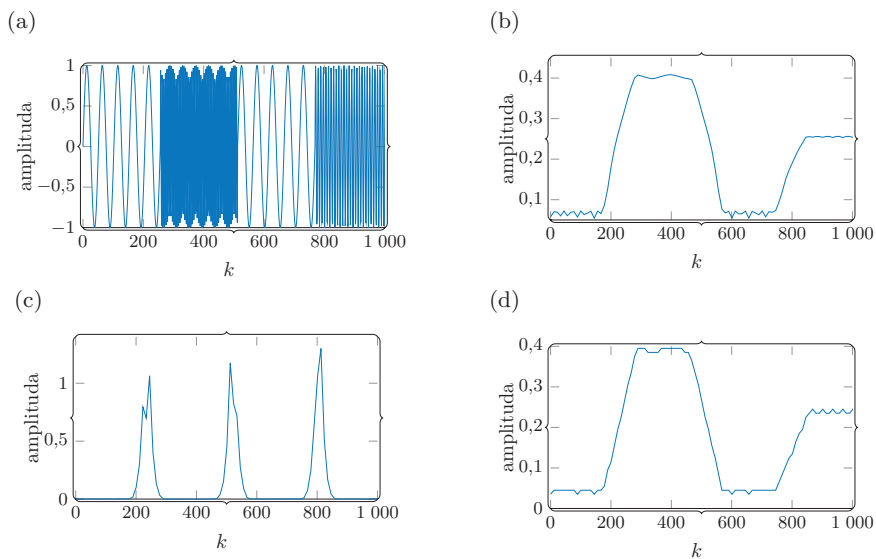
- środek ciężkości widma (ang. *Spectral Centroid*, SC),
- opadanie widma (ang. *Spectral Roll-Off*, SRO),
- płaskość widma (ang. *Spectral Flatness*, SF1),
- przepływ widma (ang. *Spectral Flux*, SF),
- nachylenie widma (ang. *Spectral Slope*, SS),
- przejścia przez zero (ang. *Zero Crossing Rate*, ZCR),
- entropia widma (ang. *Spectral Entropy*, SE),
- rozproszenie widma (ang. *Spectral Spread*, SSP),
- współczynniki *mel-cepstralne* (ang. *Mel-Frequency Cepstral Coefficients*, MFCC).

Zestawienie cech spektralnych wraz z odpowiednimi zależnościami zamieszczono w tab. 7.2. Na rysunkach 7.20 i 7.21 zilustrowano wybrane cechy widmowe dla sygnałów będących ułankowym szumem *gaussowskim* (rys. 7.20a) oraz sygnałem o zmiennej częstotliwości (rys. 7.21a).

7. Poziom średni: fuzja właściwości



Rys. 7.20. Ilustracja wybranych cech widmowych dla ułamkowego szumu *gaussowskiego*: a) sygnał, b) przepływ widma, c) środek ciężkości widma, d) przejście przez zero



Rys. 7.21. Ilustracja wybranych cech widmowych dla sygnału o zmiennej częstotliwości (patrz rys. 7.11): a) sygnał, b) przepływ widma, c) środek ciężkości widma, d) przejście przez zero

Tabela 7.2. Zestawienie najważniejszych cech widmowych

| Cecha | Zależność |
|------------------------|--|
| Środek ciężkości widma | $scw_i = \frac{\sum_{j=0}^{N_{\text{DFT}}} j Y_{\text{DFT}i}(j)}{\sum_{j=0}^{N_{\text{DFT}}} Y_{\text{DFT}i}(j)}$ |
| Płaskość widma | $pw = \frac{\sqrt{\prod_{j=0}^{N_{\text{DFT}}} Y(j)}}{\frac{1}{N_{\text{DFT}}} \sum_{j=0}^{N_{\text{DFT}}} Y(j)}$ |
| Przepływ widma | $prw_{i,i-1} = \sum_{j=1}^{N_{\text{DFT}}} \left(\frac{Y_{\text{DFT}i}(j)}{\sum_{l=1}^{N_{\text{DFT}}} Y_{\text{DFT}i}(l)} - \frac{Y_{\text{DFT}i-1}(j)}{\sum_{l=1}^{N_{\text{DFT}}} Y_{\text{DFT}i-1}(l)} \right)$ |
| Entropia widma | $ew = - \sum_{f=0}^{L-1} \frac{\text{eny}_f}{\sum_{f=0}^{L-1} \text{eny}_f} \cdot \log_2 \left(\frac{\text{eny}_f}{\sum_{f=0}^{L-1} \text{eny}_f} \right)$ |

Z wymienionych cech, które mogą być wykorzystane do opisu sygnału w zakresie częstotliwości omówimy teraz szczegółowo współczynniki *mel-cepstralne*. Współczynniki *mel-cepstralne* są efektem zastosowania **analizy mel-cepstralnej**, która wywodzi się wprost z *fourierowskiej* analizy widmowej. W ostatnich latach wykazano, że analiza *mel-cepstralna* dobrze nadaje się do analizy różnorodnych sygnałów. Przykłady jej zastosowania omówimy w dalszej części książki, teraz natomiast skupimy się na algorytmie wyznaczania cech *mel-cepstralnych*.

Wyznaczanie współczynników MFCC rozpoczynamy od podzielenia analizowanego sygnału $y(k)$ na N_{frm} ramek (ramkowanie), tj. $y_i(k)$ dla $i = 1, 2, \dots, N_{\text{frm}}$. Następnie należy wyznaczyć liczbę próbek w ramce (L_{frm}), która zależy od częstotliwości próbkowania F_s oraz od czasu trwania ramki F_d . Do wyznaczenia długości ramki korzystamy z zależności $L_{\text{frm}} = F_s F_d$ [460].

7. Poziom średni: fuzja właściwości

Kolejnym krokiem jest wyznaczenie dla każdej i -tej ramki Dyskretnej Transformy Fouriera z następującej zależności:

$$Y_{\text{DFT}i}(n) = \sum_{l=1}^{L_{\text{frm}}} y_i(l) w(l) \exp\left(\frac{-j2\pi nl}{L_{\text{frm}}}\right) \quad \text{dla } 1 \leq n \leq N_{\text{DFT}} \quad (7.36)$$

gdzie L_{frm} jest długością ramki, $y(l)$ oznacza sygnał wejściowy, l jest numerem próbki w ramce [422], i jest numerem ramki, N_{DFT} jest liczbą próbek, a $w(l)$ reprezentuje okno czasowe. W metodzie szacowania cech *mel-cepstralnych* najczęściej wykorzystywane jest okno czasowe Hamminga (7.14).

Z zależności (7.36) otrzymujemy widmo mocy sygnału $y(k)$. Następnie wyznaczoną moc sygnału wykorzystujemy do określenia energii w podpasmach, wykorzystując następującą zależność [466]:

$$P_i(q) = \log\left(\sum_{n=1}^{N_{\text{DFT}}} |Y_{\text{DFT}i}(n)|^2 H_q(n)\right) \quad (7.37)$$

gdzie $H_q(n)$ oznacza zespół N_{QF} filtrów trójkątnych. W kolejnym kroku wyznaczone widmo mocy sygnału $y(k)$ poddane zostaje działaniu zestawu filtrów trójkątnych H_q . Liczba filtrów dobierana jest zazwyczaj z przedziału 20–40 [460]. Filtry trójkątne H_q charakteryzują się tym, że są one równomiernie rozmieszczone w danym zakresie częstotliwości w skali mel. Filtry trójkątne H_q wyznaczamy w następujący sposób:

$$H_q(n) = \begin{cases} 0 & \text{dla } n < f(q-1) \\ \frac{2(f(q+1) - n)}{(f(q+1) - f(q-1))(f(q+1) - f(q))} & \text{dla } f(q-1) \leq n < f(q) \\ \frac{2(n - f(q-1))}{(f(q+1) - f(q-1))(f(q) - f(q-1))} & \text{dla } f(q-1) \leq n < f(q) \\ 0 & \text{dla } n \geq f(q+1) \end{cases} \quad (7.38)$$

Częstotliwości $f(q)$ we wzorze (7.38) wyznaczamy z następującej zależności:

$$f(q) = \left(\frac{N_{\text{DFT}}}{F_s}\right) B_{\text{mel}}^{-1}\left(B_{\text{mel}}(f_{\text{min}}) + q \frac{B_{\text{mel}}(f_{\text{max}}) - B_{\text{mel}}(f_{\text{min}})}{N_{\text{QF}} + 1}\right) \quad (7.39)$$

gdzie f_{max} i f_{min} to odpowiednio największa i najmniejsza częstotliwość zespołu filtrów podawana jest w Hz. Natomiast $B_{\text{mel}}(f)$ jest skalą melową, wyznaczaną

z zależności:

$$B_{\text{mel}}(f) = 1125 \log \left(1 + \frac{f \text{ [Hz]}}{700} \right) \quad (7.40)$$

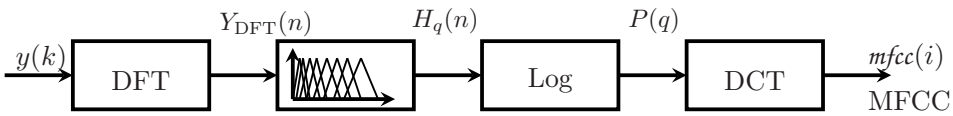
Z kolei $B_{\text{mel}}(f)^{-1}$ to jej odwrotność, wyznaczana w następujący sposób:

$$B_{\text{mel}}(f)^{-1} = 700 \left(\exp \left(\frac{f \text{ [mel]}}{1125} \right) - 1 \right) \quad (7.41)$$

W ostatnim kroku algorytmu MFCC wykorzystujemy dyskretną transformację kosinusową (ang. *Discrete Cosine Transform*, DCT) do logarytmu dziesiętnego z nowej skali wyrażonej w melach. W ten sposób otrzymujemy współczynniki *mel-cepstralne*:

$$mfcc(i) = \sum_{q=0}^{N_{\text{QF}}-1} P_i(q) \cos \left(\frac{\pi i}{2N_{\text{QF}}} (2q+1) \right) \quad \text{dla } 0 \leq i < N_{\text{QF}} \quad (7.42)$$

Algorytm do wyznaczania współczynników *mel-cepstralnych* przedstawiono na rys. 7.22.

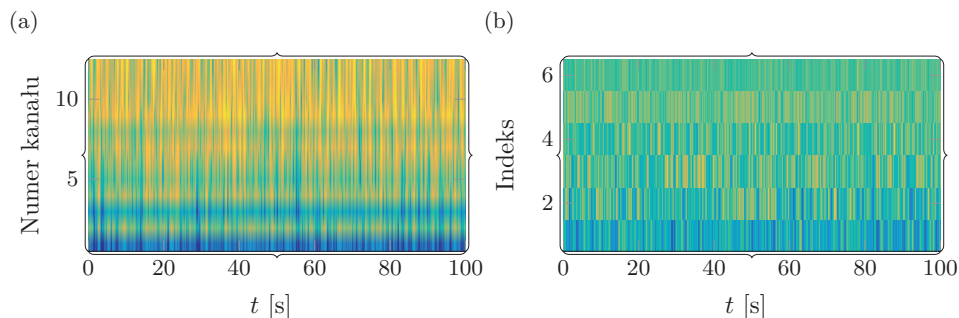


Rys. 7.22. Algorytm wyznaczania współczynników *mel-cepstralnych*

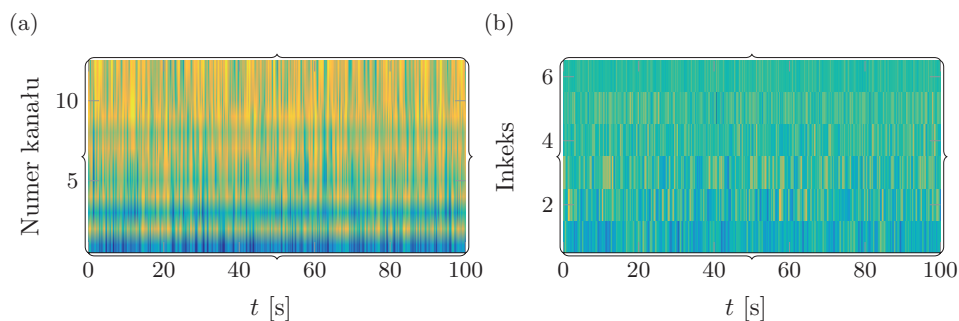
Cechy MFCC są popularnym narzędziem ekstrakcji cech z sygnałów. Na przykład w pracy [52] cechy *mel-cepstralne* zostały wykorzystane do rozpoznawania zagrywek tenisowych. Natomiast w pracy [210] podejście to zostało wykorzystane do ekstrakcji cech z sygnałów EEG na potrzeby rozpoznawania napadów padaczkowych (ang. *seizure event*). W pracy [28] cechy MFCC zostały z kolei wykorzystane do rozpoznawania szmerów oddechowych.

Przykład 7.3 (Wyznaczanie cech MFCC). W przykładzie zilustrowano działanie algorytmu wyznaczającego współczynniki MFCC (rys. 7.23–7.25). W analizie właściwości algorytmu wykorzystano szum biały, ułamkowy szum gaussowski i sygnał o zmiennej częstotliwości. Z analizy wykresów wynika, że, wykorzystując metodę MFCC, trudno jest odróżnić szum biały od ułamkowego szumu gaussowskiego. Natomiast analiza wyników przedstawionych na rys. 7.25 wskazuje, że można ustalić momenty zmian w sygnale o zmiennej częstotliwości.

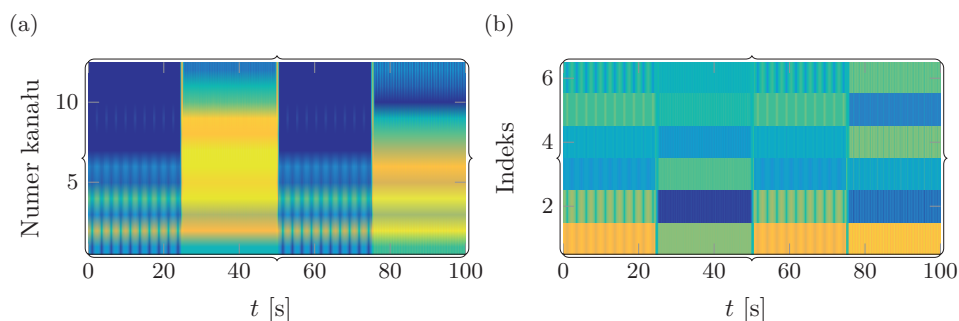
7. Poziom średni: fuzja właściwości



Rys. 7.23. Analiza cech *mel-cepstralnych* dla szumu białego: a) energia współczynników mela w skali logarytmicznej, b) cepstrum częstotliwości mela



Rys. 7.24. Analiza cech *mel-cepstralnych* dla ułamkowego szumu *gaussowskiego*: a) energia współczynników mela w skali logarytmicznej, b) cepstrum częstotliwości mela



Rys. 7.25. Analiza cech *mel-cepstralnych* dla sygnału o zmiennej częstotliwości (patrz rys. 7.11): a) energia współczynników mela w skali logarytmicznej, b) cepstrum częstotliwości mela

Kolejną grupę cech wykorzystywanych do opisu sygnału w dziedzinie częstotliwości można wyznaczyć, stosując transformację Hilberta (ang. *Hilbert Transform*, HT). Transformacja Hilberta podobnie jak transformacja Fouriera jest przekształceniem całkowym. W tym przypadku transformacja ta odwzorowuje sygnał pomiarowy w inny, nowy sygnał. Różnica polega na tym, że w transformacji Hilberta oryginał, jak i transformata są w tej samej dziedzinie, tj. w dziedzinie czasu. Efektem działania transformacji Hilberta jest sygnał nazywany **sygnałem analitycznym** [258] (ze względu na brak składowych widmowych dla ujemnych częstotliwości). Transformację Hilberta wyznacza się z następującej zależności:

$$Y_{\text{HT}}(t) = \mathcal{H}[y(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y(\tau)}{t - \tau} d\tau \quad (7.43)$$

gdzie $y(t)$ jest sygnałem pierwotnym, $Y_{\text{HT}}(t)$ jest sygnałem analitycznym (zespolonym). Dla tego sygnału zachodzi:

$$y_{\text{ht}}(t) = y(t) + j\mathcal{H}[y(t)] = y(t) + jY_{\text{HT}}(t) \quad (7.44)$$

gdzie $y_{\text{ht}}(t)$ jest sygnałem zespolonym. Znając postać sygnału analitycznego $y_{\text{ht}}(t)$, możliwe jest z kolei wyznaczenie:

- **chwilowej amplitudy** lub **obwiedni** (ang. *instantaneous amplitude, envelope*)

$$A(t) = |y_{\text{ht}}(t)| \quad (7.45)$$

- **chwilowej fazy** (ang. *instantaneous phase*)

$$\rho(t) = \arg(y_{\text{ht}}(t)) \quad (7.46)$$

Bazując na powyższych wyliczeniach, można również wyznaczyć:

- **chwilową częstotliwość fazową** (ang. *instantaneous angular frequency*)

$$\omega(t) = \frac{d\rho(t)}{dt} \quad (7.47)$$

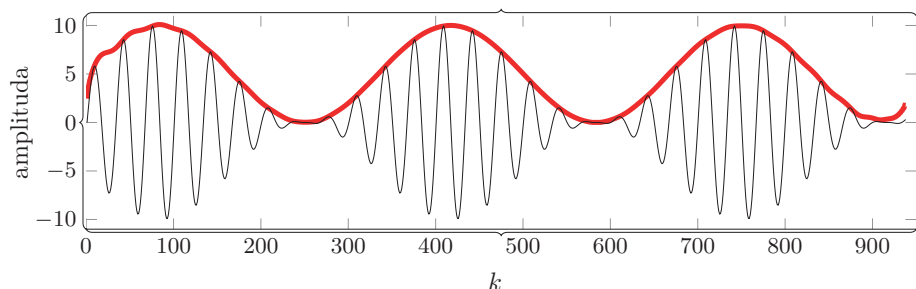
- **chwilową częstotliwość** (ang. *instantaneous frequency*)

$$f(t) = \frac{1}{2\pi}\omega(t) \quad (7.48)$$

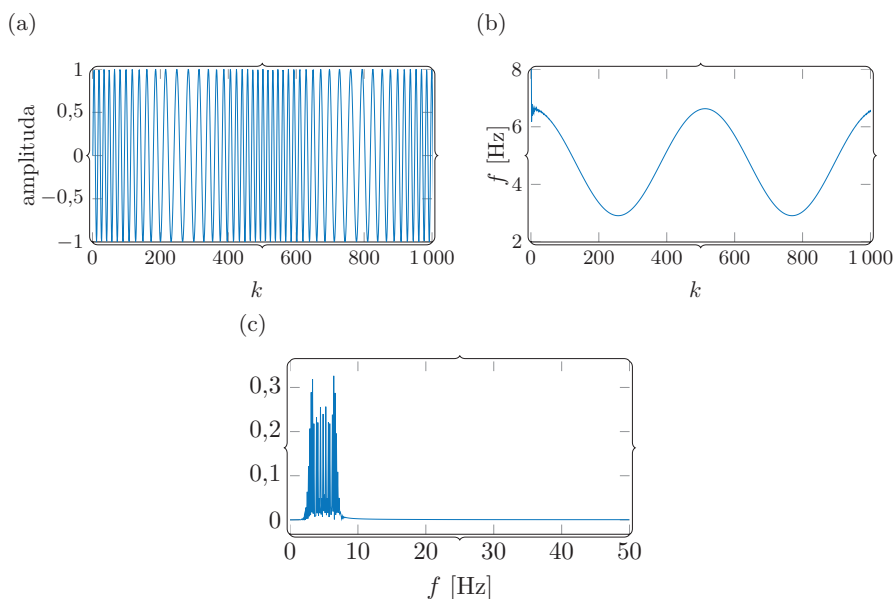
W wielu rzeczywistych zadaniach przetwarzania sygnałów prostsze jest operowanie sygnałem analitycznym $y_{\text{ht}}(t)$ niż sygnałem oryginalnym $y(t)$. Również wielkości, które mogą zostać wyznaczone z wykorzystaniem sygnału analitycznego (np. obwiednia lub chwilowa częstotliwość), są przydatne na średnim poziomie fuzji danych.

Transformacja Hilberta znalazła również wiele zastosowań w zadaniach ekstrakcji cech. W pracach [302] oraz [360] HT została wykorzystana do ekstrakcji cech z sygnału EEG.

7. Poziom średni: fuzja właściwości



Rys. 7.26. Obwiednia wyznaczona z wykorzystaniem transformacji Hilberta



Rys. 7.27. Częstotliwość chwilowa wyznaczona z wykorzystaniem transformacji Hilberta:
a) sygnał o modulowanej częstotliwości, b) estymowana częstotliwość,
c) widmo sygnału o modulowanej częstotliwości

Przykład 7.4 (Zastosowanie transformacji Hilberta). *Przykład wykorzystania transformacji Hilberta do wyznaczenia obwiedni i chwilowej częstotliwości przedstawiono na rys. 7.26 oraz 7.27. Z wykresów wynika, że zastosowanie transformacji Hilberta umożliwia wyznaczenie obwiedni sygnału, jak również odtworzenie częstotliwości chwilowej sygnału.*

7.4.3. Metody generowania cech w dziedzinie czasowo-częstotliwościowej

Dotychczas omówiono metody ekstrakcji cech, które były wyznaczane w dziedzinie czasu lub częstotliwości. W tej części pracy zostaną omówione metody, które umożliwiają wyznaczenie opisu zmienności widma częstotliwościowego sygnału w czasie. Właściwość taką mają metody wykorzystujące czasowo-częstotliwościową reprezentację sygnałów. Jest to szczególnie przydatna w analizie sygnałów niestacjonarnych. Analizę widma sygnału w czasie umożliwiają metody bazujące na metodzie czasowo-częstotliwościowej analizy sygnałów. Wśród metod, które najczęściej wykorzystywane są w tym celu, wyróżniamy:

- krótkoczasową transformację Fouriera (ang. *short-term Fourier Transform*),
- transformację falkową (ang. *wavelet transform*),
- transformację S (ang. *S Transform*),
- transformację Wignera–Villego (ang. *Wigner–Ville distribution*),
- transformację Hilberta–Huanga (ang. *Hilbert–Huang transform*).

Przedstawione we wcześniejszych rozdziałach cechy należały do dwóch obszarów, tj. czasu i częstotliwości. Przydatność tych cech, również w zadaniu fuzji danych, została wielokrotnie potwierdzona. Niestety metody te mają również wady. W wielu przypadkach cechy wyznaczone w dziedzinie czasu czy częstotliwości opisują analizowane sygnały tylko w sposób przybliżony. Jest to szczególnie widoczne w przypadku sygnałów niestacjonarnych. A ponieważ większość rzeczywistych sygnałów jest niestacjonarna, to uzasadnione jest poszukiwanie innych metod reprezentacji sygnałów, które będą umożliwiały analizę takich sygnałów. Alternatywą są wymienione czasowo-częstotliwościowe metody reprezentacji sygnałów.

Transformacja S

Transformacja S jest uogólnieniem szybkiej transformacji Fouriera oraz rozwinięciem koncepcji znanych z ciągłej transformacji falkowej [381]. Związek transformacji S z transformacją Fouriera jest silny również z tego powodu, że odwrotna transformacja S jest transformacją Fouriera [476]. Z tej obserwacji można wyciągnąć ciekawy wniosek praktyczny, tj. możliwa jest transformacja sygnału z zakresu czasu do zakresu czasowo-częstotliwościowego, a następnie do częstotliwości. Wykorzystując odwrotną transformację Fouriera, możemy powrócić do dziedziny czasu.

Cechą charakterystyczną transformacji S jest jej wysoka rozdzielczość czasowo-częstotliwościowa. Cecha ta jest szczególnie istotna w przypadku sygnałów o bogatym widmie częstotliwościowym. Również w przypadku analizy sygnałów niestacjonarnych wysoka rozdzielczość czasowo-częstotliwościowa jest przydatna.

7. Poziom średni: fuzja właściwości

Transformację S wyznaczamy z zależności:

$$S_{ST}(t, f) = \int_{-\infty}^{\infty} y(\tau) |f| e^{-\pi(t-\tau)^2 f^2} e^{-j2\pi f\tau} d\tau \quad (7.49)$$

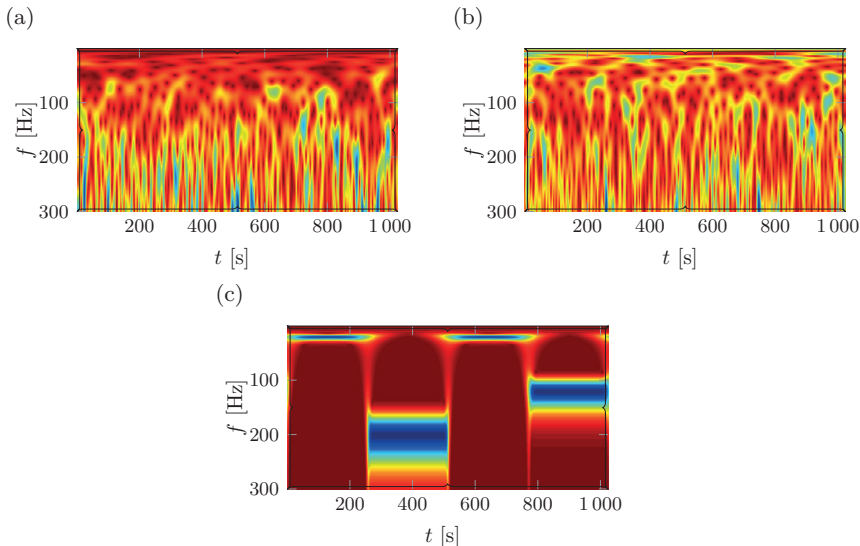
natomiast odwrotną transformację S z zależności:

$$y(\tau) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} S_{ST}(t, f) dt \right] e^{j2\pi f\tau} df \quad (7.50)$$

Warto zauważyć, że w transformacji S wykorzystuje się *gaussowskie* okno przesuwne.

Transformacja S jest często wykorzystywana w analizie sygnałów biomedycznych. Na przykład w pracy [483] transformacja S została wykorzystana do wykrywania w sygnale EKG kompleksów QRS. Natomiast w pracy [295] transformacja ta została zastosowana do segmentacji sygnałów fonokardiograficznych (ang. *Phono-CardioGram*).

Przykład 7.5 (Zastosowanie Transformacji S). W przykładzie tym zilustrujemy działanie transformacji S zastosowanej do szumu białego, ułamkowego szumu gaussowskiego oraz do sygnału o zmiennej częstotliwości. Efekt działania transformacji S na tych sygnałach przedstawiono na rys. 7.28.



Rys. 7.28. Transformacja S dla wybranych sygnałów dla: a) białego szumu, b) ułamkowego szumu *gaussowskiego*, c) sygnału o modulowanej częstotliwości (patrz rys. 7.11)

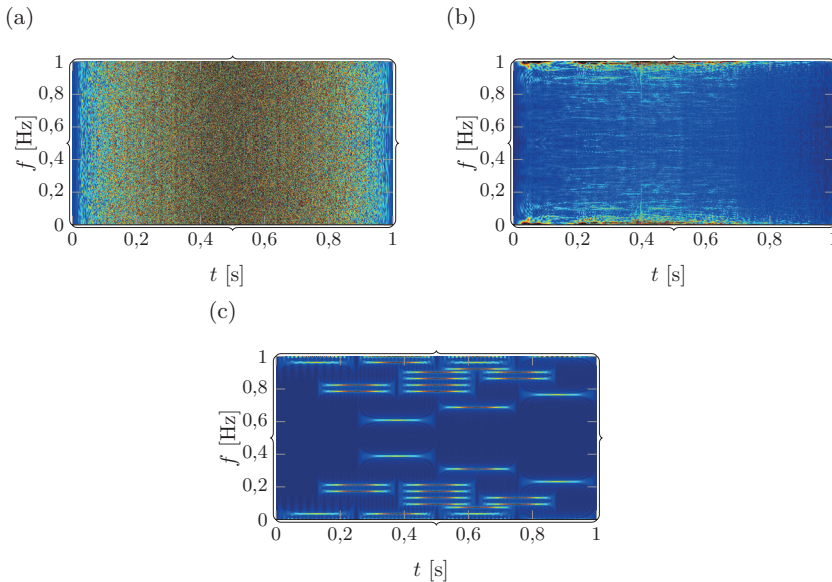
Transformacja Wignera–Villego

Transformacja Wignera–Villego (WV) jest jednym z popularniejszych narzędzi do analizy sygnałów niestacjonarnych. Wyróżnia się tym, że idealnie odwzorowuje w przestrzeni czasowo-częstotliwościowej liniową zmianę częstotliwości [484]. Ciągłą transformatę Wignera–Villego wyznacza się z zależności:

$$S_{WV}(t, f) = \int_{-\infty}^{\infty} y\left(t + \frac{\tau}{2}\right) y^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \quad (7.51)$$

gdzie y^* jest sygnałem zespolonym sprzężonym z sygnałem y .

Transformacja ta charakteryzuje się największą koncentracją energii w przestrzeni czasowo-częstotliwościowej, a co się z tym wiąże, wyróżnia się spośród innych metod wysoką rozdzielczością czasową [391]. Dzięki temu jest przydatnym narzędziem analizy sygnałów niestacjonarnych. Metoda WV charakteryzuje się również tym, że w widmie estymowanym nie występuje zjawisko przecieku widma.



Rys. 7.29. Transformata Wignera–Villego dla wybranych sygnałów dla:

- a) białego szumu, b) ułamkowego szumu *gaussowskiego*,
- c) sygnału o modulowanej częstotliwości (patrz rys. 7.11)

Podstawową wadą tej metody jest występowanie interferencji wzajemnych o charakterze oscylacyjnym pomiędzy różnymi składowymi widma analizowanego

7. Poziom średni: fuzja właściwości

sygnału (ang. *cross-terms*) [484, 487]. Oznacza to, że reprezentacja Wignera–Villego nie jest liniowa, co może doprowadzić do trudności interpretacyjnych.

Transformacja Wignera–Villego znalazła również zastosowanie do ekstrakcji cech rzeczywistych sygnałów. Na przykład w pracy [101] transformacja WV została wykorzystana do analizy sygnałów EKG i wykrywania w nich arytmii. Z kolei w pracy [25] transformacja ta została wykorzystana do ekstrakcji cech z sygnałów EEG na potrzeby analizy faz snu człowieka.

Przykład 7.6 (Zastosowanie Transformaty Wignera–Villego). *Przykład wykorzystania transformacji Wignera–Villego do analizy szumu białego i ułamkowego gaussowskiego oraz sygnału o zmiennej częstotliwości przedstawiono na rys. 7.29.*

7.5. Metody generowania cech nieliniowych

W omawianych dotychczas metodach zakładano, że dane pomiarowe są realizacją pewnego procesu losowego. Jednakże istnieje wiele metod, w których przyjmuje się, że dane pomiarowe generowane są przez pewien obiekt dynamiczny o nieliniowej charakterystyce.

Omówimy teraz wybrane metody analizy sygnałów pomiarowych o takich charakterystykach. Wykażemy również przydatność wygenerowanych tymi metodami wielkości do konstrukcji wektora cech na potrzeby fuzji danych na poziomie średnim.

Wymiar fraktalny

Wymiar fraktalny (ang. *fractal dimension*, FD) jest miarą obiektu geometrycznego charakteryzującego się właściwością samopodobieństwa [309]. Koncepcja wymiaru fraktalnego wywodzi się z geometrii i początkowo służyła do opisu obiektów, które charakteryzowały się skomplikowanymi kształtami. Obecnie wymiar fraktalny oznacza ogólną miarę określającą tempo zwiększania się objętości obiektu fraktalnego w miarę redukcji skali pomiarowej [309]. FD może być również wykorzystywany jako miara złożoności sygnału.

W odróżnieniu od klasycznie rozumianego wymiaru, wymiar fraktalny może przyjmować również wartości niecałkowite. Na przykład prosta ma wymiar fraktalny jeden. Płaszczyzna natomiast ma wymiar fraktalny równy dwa. Obiekt o wymiarze fraktalnym pomiędzy jeden a dwa to linia zamknięta bez samoprzecięć o właściwościach samopodobieństwa, która jest umieszczona na powierzchni. Przykładem takiej krzywej jest krzywa Kocha.

Koncepcję wymiaru fraktalnego można również wykorzystać do analizy sygnałów. Opracowano wiele metod wyznaczania wymiaru fraktalnego dla sygnałów. Jak wykazano m.in. w pracy [82], wyniki wyliczeń wymiaru fraktalnego mogą

być wykorzystane również do konstrukcji wektora cech. Poniżej przedstawiliśmy najważniejsze metody do szacowania wymiaru fraktalnego.

- **wymiar fraktalny Higuchiego** (ang. *Higuchi Fractal Dimension*, HFD) [201, 419]. Na podstawie pomiarów, tj. $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$ wyznaczany jest nowy zbiór szeregów czasowych, który jest pochodny względem oryginalnego zbioru:

$$hg(m_h, n_h) : \left(y(m_h), y(m_h + n_h), y(m_h + 2n_h), \dots, \left(m_h + \left\lfloor \frac{K - m_h}{n_h} \right\rfloor n_h \right) \right) \quad (7.52)$$

dla $m_h = 1, 2, \dots, n_h$, gdzie m_h reprezentuje czas początkowy, n_h to interwał czasowy taki, że $n_h = 1, 2, \dots, N_h$, N_h to parametr, którego wartość należy ustalić, $\lfloor \cdot \rfloor$ to operacja wyznaczania części całkowitej z liczby rzeczywistej. Dysponując zbiorem pochodnych szeregów czasowych, wymiar fraktalny Higuchiego wyznaczamy z zależności:

$$FD_{HG}(m_h, n_h) = \frac{1}{m_h} \left(\sum_{i=1}^{\left\lfloor \frac{K - m_h}{n_h} \right\rfloor} |y(m_h + in_h) - y(m_h + (i - 1)n_h)| \right) \left\lfloor \frac{K - m_h}{n_h} \right\rfloor n_h \quad (7.53)$$

po wyliczeniu wartości $FD_{HG}(m_h, n_h)$ jest ona uśredniania.

- **wymiar fraktalny Katza** (ang. *Katz Fractal Dimension*, KFD) [9, 183]. Do obliczeń w tym przypadku korzystamy z następującej zależności:

$$FD_{KZ} = \frac{\log(\mathbf{ns})}{\log(\mathbf{ns}) + \log\left(\frac{\mathbf{m}\mathbf{d}}{\mathbf{t}\mathbf{l}}\right)} \quad (7.54)$$

gdzie \mathbf{ns} wyznacza się z zależności:

$$\mathbf{ns} = \frac{t\mathbf{l}}{\mathbf{m}\mathbf{n}_{i,i+1}} \quad (7.55)$$

gdzie $\mathbf{m}\mathbf{n}_{i,j}$ jest średnią odległością pomiędzy dwoma punktami, a $\mathbf{t}\mathbf{l}$ wyliczane jest w następujący sposób:

$$\mathbf{t}\mathbf{l} = \sum_{k=0}^K \text{dist} \|y(k+1) - y(k)\|_2 \quad (7.56)$$

7. Poziom średni: fuzja właściwości

Z kolei wielkość $m\delta$ należy wyznaczyć z zależności:

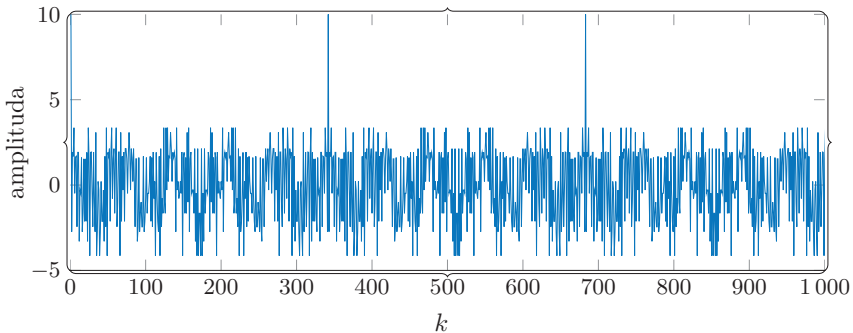
$$m\delta = \max(\text{dist}(y(i), y(j))) \quad i, j = 1, 2, \dots, K \quad (7.57)$$

jest maksymalną odległością pomiędzy dwoma punktami sygnału. Funkcja dist służy do wyznaczania maksymalnej odległości euklidesowej pomiędzy dwoma punktami sygnału pomiarowego, tj. $y(k)$ oraz $y(k+1)$.

Przykład 7.7 (Wymiar fraktalny). *Przykład wykorzystania wymiaru fraktalnego przedstawiono na rys. 7.30–7.35. W analizie wykorzystano następujące sygnały testowe: a) funkcję Weierstrassa-kosinusową (ang. Weierstrass-cosine function), b) ułamkowy szum gaussowski, c) sygnał świergotliwy oraz d) sygnał o zmiennej amplitudzie.*

Funkcja Weierstrassa-kosinusowa to ciągła funkcja rzeczywista, która jest nieróżniczkowalna w całej dziedzinie. Ponieważ funkcja ta spełnia właściwość samopodobieństwa, została ona wykorzystana w badaniach nad właściwościami algorytmów do estymacji wymiaru fraktalnego. Na rysunku 7.30 przedstawiono sekwencję wygenerowaną z wykorzystaniem funkcji Weierstrassa-kosinusowej.

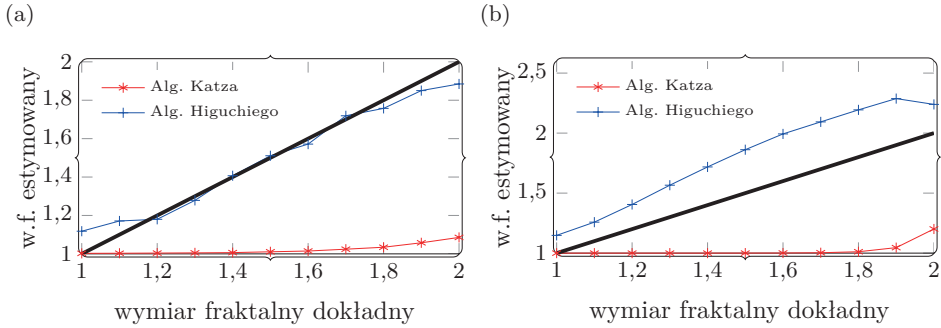
Sygnał w formie ułamkowego szumu gaussowskiego był już prezentowany wcześniej. Przykład sygnału świergotliwego przedstawiono z kolei na rys. 7.34a, a sygnału o zmiennej amplitudzie na rys. 7.35.



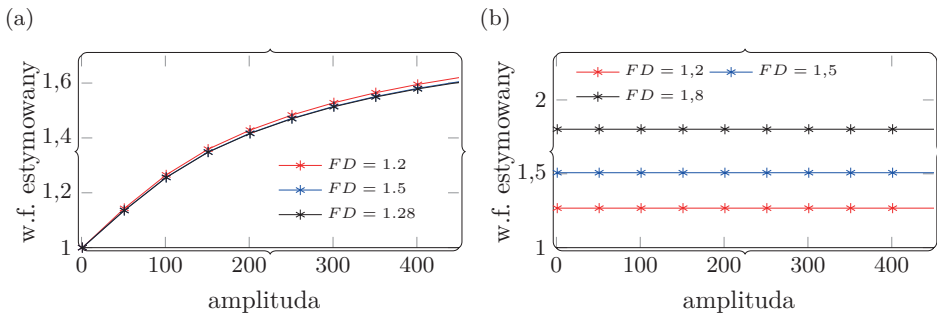
Rys. 7.30. Przykładowa sekwencja wygenerowana dla funkcji Weierstrassa-kosinusowej

Na rysunku 7.31 przedstawiono wyniki badań nad jakością estymacji wymiaru fraktalnego z wykorzystaniem metody Katza i Higuchiego. Z wykresów wynika, że zarówno dla ułamkowego szumu gaussowskiego, jak i dla sekwencji wygenerowanej z wykorzystaniem funkcji Weierstrassa-kosinusowej metoda Katza nie estymuje poprawnie wartości wymiaru fraktalnego. Podobne wnioski można wyciągnąć, obserwując wynik badania wpływu zmian amplitudy sygnału na jakość estymacji

7.5. Metody generowania cech nieliniowych



Rys. 7.31. Porównanie dokładności estymacji wymiaru fraktalnego z wykorzystaniem algorytmu Katza i Higuchiego. Z lewej oszacowany wymiar fraktalny dla ułamkowego szumu *gaussowskiego*. Z prawej oszacowany wymiar fraktalny dla sygnału wygenerowanego dla funkcji *Weierstrassa-kosinusowej*



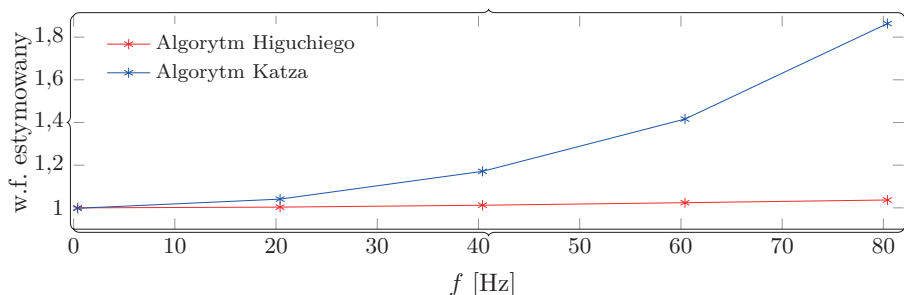
Rys. 7.32. Porównanie dokładności estymacji wymiaru fraktalnego algorytmem Katza (a) i Higuchiego (b) dla zmiennej amplitudy sygnału i przy różnych wartościach wymiaru fraktalnego

wymiaru fraktalnego (rys. 7.32). Można zaobserwować, że w przypadku algorytmu Katza amplituda sygnału ma wpływ na jakość oszacowania. Z kolei w przypadku algorytmu Higuchiego amplituda sygnału nie wpływa na jakość estymacji.

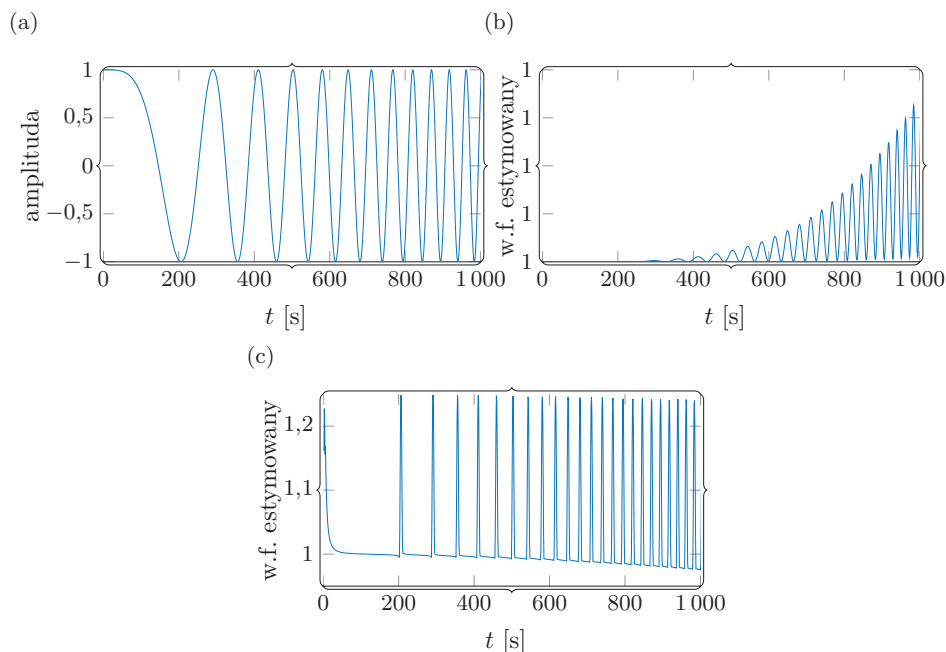
W kolejnym eksperymencie zbadano wpływ zmian w częstotliwości sygnału na jakość estymacji wymiaru fraktalnego. Również i w tym eksperymencie algorytm Katza niedokładnie szacował wymiar fraktalny (rys. 7.33). Na wykresie widać, że jakość oszacowania metodą Katza spada, gdy częstotliwość sygnału wzrasta.

W kolejnym eksperymencie badano jakość estymacji wymiaru fraktalnego dla sygnału świergotliwego (rys. 7.34). Wyniki wskazują, że również i w tym przypadku jakość oszacowania wymiaru fraktalnego wykonanego metodą Katza jest mało

7. Poziom średni: fuzja właściwości



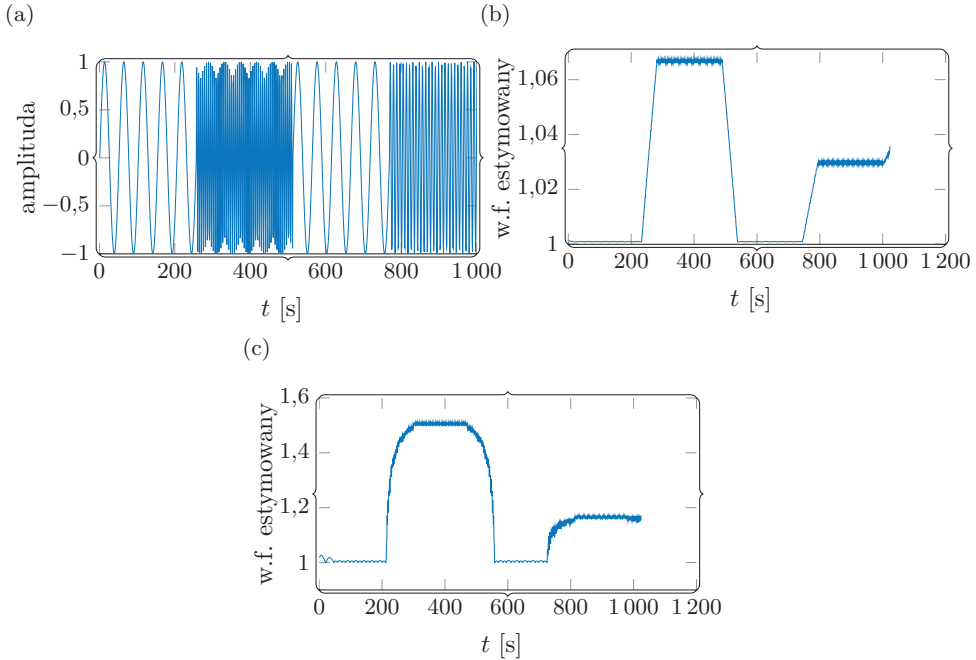
Rys. 7.33. Porównanie dokładności estymacji wymiaru fraktalnego algorytmem Katza i Higuchiego dla zmiennej częstotliwości sygnału. Wymiar fraktalny w badaniu wyniósł 1



Rys. 7.34. Estymacja wymiaru fraktalnego algorytmem Katza i Higuchiego dla sygnału świergotliwego: a) sygnał, b) algorytm Katza, c) algorytm Higuchiego

wiarygodna. Jedyne poprawne oszacowanie wymiaru fraktalnego uzyskano dla sygnału o zmiennej częstotliwości. Zostało to przedstawione na rys. 7.35.

Z przeprowadzonej analizy wynika, że algorytm Katza nie nadaje się do analizy sygnałów niestacjonarnych. Wskazują na to wyniki uzyskane dla ułamkowego szumu *gaussowskiego*, sekwencji wygenerowanej funkcją Weierstrassa-kosinusowej



Rys. 7.35. Estymacja wymiaru fraktalnego algorytmem Katza i Higuchiego dla sygnału o zmiennej częstotliwości: a) sygnał, b) algorytm Katza, c) algorytm Higuchiego

oraz dla sygnału świergotliwego. Niedokładność algorytmu Katza została omówiona m.in. w pracy [62]. Jedyne poprawne wyniki uzyskano dla sygnału o zmiennej częstotliwości. Z przedstawionych badań wynika również, że wymiar fraktalny (zwłaszcza wyznaczany metodą Higuchiego) wykrywa zmiany w częstotliwości analizowanego sygnału.

Wymiar korelacyjny

Wymiar korelacyjny (ang. *correlation dimension*, CD) jest przybliżeniem wymiaru fraktalnego. Metoda ta jest bardzo popularnym narzędziem do opisu systemów, w których wykorzystuje się dane pomiarowe [453]. Jest to szczególnie przydatne narzędzie do opisu systemów dynamicznych nieliniowych. Wymiar korelacyjny wyliczany jest w kilku etapach. Niech sygnał pomiarowy będzie reprezentowany przez $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$. Następnie, wykorzystując metodę opóźnień czasowych (ang. *time-delay embedding*), odtwarzana jest przestrzeń fazowa z następującego zestawu zależności [453]:

$$x_{em}(1) = (y(1), y(1+m), \dots, y(1+(j-1)m)) \quad (7.58a)$$

7. Poziom średni: fuzja właściwości

$$\dots$$

$$x_{\text{em}}(k) = (y(k), y(k+m), \dots, y(k+(z-1)m)) \quad (7.58b)$$

$$\dots$$

$$x_{\text{em}}(K-(z-1)m) = (y(K-(z-1)m), y(K-(z-2)m), \dots, y(K)) \quad (7.58c)$$

gdzie z jest wymiarem zanurzenia, a m jest opóźnieniem czasowym odtwarzanej przestrzeni fazowej. Oznaczmy dodatkowo przez K_{po} liczbę wszystkich wygenerowanych punktów odtwarzanej przestrzeni, tj. $K_{\text{po}} = (K - (z - 1)m)$. Następnie spośród K_{po} wygenerowanych punktów (7.58) wybieramy losowo jeden z nich, który będzie punktem referencyjnym. Oznaczmy ten punkt jako $x_{\text{em}}(i)$. Następnie, z zależności:

$$\delta_{i,j} = \text{dist}(x_{\text{em}}(i), x_{\text{em}}(j)) = \left(\sum_{l=0}^{z-1} (x_{\text{em}}(i+lm) - x_{\text{em}}(j+lm))^2 \right)^{\frac{1}{2}} \quad (7.59)$$

wyznaczamy odległości pomiędzy punktem referencyjnym a wszystkimi pozostałymi $K_{\text{po}} - 1$ punktami w odtwarzanej przestrzeni. Zmienna j jest indeksem wszystkich wygenerowanych punktów, natomiast i jest indeksem losowo wybranego punktu z rekonstruowanej przestrzeni. W przypadku sygnału jednowymiarowego, CD wyznaczane jest bezpośrednio z następującej zależności [453]:

$$\text{CD}(z, \delta_i) = \frac{2}{K_{\text{po}}(K_{\text{po}} - 1)} \sum_{i,j=1}^{K_{\text{po}}} \mathbf{1}(\delta_i - \delta_{i,j}), \quad i \neq j \quad (7.60)$$

gdzie $\mathbf{1}()$ jest funkcją Heaviside, $\delta_{i,j}$ jest odległością pomiędzy dwoma punktami w zrekonstruowanej przestrzeni stanu.

Wymiar korelacyjny wykorzystuje się w analizie sygnałów. Wykorzystywany jest on np. do generowania cech wykorzystywanych do konstrukcji wektora cech. Na przykład w pracy [129] wymiar korelacyjny posłużył do analizy sygnałów EKG. Z kolei w pracy [319] badano wpływ wymiaru korelacyjnego na poprawę jakości algorytmu do detekcji arytmii serca w sygnale EKG.

Entropia

Pierwotnie entropia była wykorzystywana jako statystyczna miara nieuporządkowania w układach termodynamicznych. Następnie pojęcie to zostało uogólnione i aktualnie jest wykorzystywane jako statystyczna miara ilości informacji [441].

Tak rozumiana entropia jest z powodzeniem wykorzystywana w analizie i przetwarzaniu sygnałów. W literaturze zaproponowano wiele metod szacowania entropii w sygnałach. Do najważniejszych z nich zaliczamy:

- **entropię przybliżoną** (ang. *approximate entropy*, ApEn), która umożliwia ocenę złożoności sygnału bez konieczności przyjmowania wstępnych założeń co do charakteru tego sygnału; duża wartość ApEn odpowiada stanom wysokiej złożoności systemu generującemu analizowany sygnał [143]. Entropia przybliżona wyznaczana jest na podstawie pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$. Następnie, bazując na tych pomiarach, odtwarzana jest wielowymiarowa przestrzeń stanów z zależności (7.58) [154, 271, 316]. W kolejnym kroku, po odtworzeniu przestrzeni stanów, wyznaczone są następujące wielkości:

$$C_i(\mathfrak{z}, \mathfrak{p}\mathfrak{d}) = \frac{1}{K - (\mathfrak{z} - 1)m} \{ \text{dist}[x_{\text{em}}(i), x_{\text{em}}(j)] < \mathfrak{p}\mathfrak{d} \} \quad (7.61)$$

gdzie

$$\text{dist}(x_{\text{em}}(i), x_{\text{em}}(j)) = \max_{l=1,2,\dots,\varepsilon} (|x_{\text{em}}(i + l - 1) - x_{\text{em}}(j + l - 1)|) \quad (7.62)$$

a $\mathfrak{p}\mathfrak{d}$ reprezentuje kryterium podobieństwa [143]. Przybliżoną entropię wyznaczamy z kolei z następującej zależności:

$$\text{ApEn}(m, \mathfrak{p}\mathfrak{d}) = \lim_{K \rightarrow \infty} (\aleph_m(\mathfrak{p}\mathfrak{d}) - \aleph_{m+1}(\mathfrak{p}\mathfrak{d})) \quad (7.63)$$

gdzie

$$\aleph_m(\mathfrak{p}\mathfrak{d}) = \frac{1}{K - (\mathfrak{z} - 1)m} \sum_{i=1}^{K - (\mathfrak{z} - 1)m} \ln C_i(\mathfrak{z}, \mathfrak{p}\mathfrak{d}) \quad (7.64)$$

Ponieważ wartość entropii przybliżonej jest wyznaczana ze skończonego zbioru danych pomiarowych, w praktycznych zastosowaniach korzystamy z zależności [143]:

$$\text{ApEn}(m, \mathfrak{p}\mathfrak{d}, K) = \aleph_m(\mathfrak{p}\mathfrak{d}) - \aleph_{m+1}(\mathfrak{p}\mathfrak{d}) \quad (7.65)$$

- **entropia próbkowana** (*sample entropy*, SampEn) jest poprawioną wersją entropii przybliżonej. Wadą tej entropii jest to, że jej wartość silnie zależy od długości analizowanego sygnału. Dodatkowo wielkość ta charakteryzuje się brakiem zgodności, co oznacza, że jest to wielkość czuła na zmianę parametrów. Podobnie jak dla entropii przybliżonej, entropię próbkowaną wyznacza się dla zbioru następujących pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$. Dla tego zbioru pomiarów również konstruujemy wielowymiarową przestrzeń stanów

7. Poziom średni: fuzja właściwości

[71], korzystając z zależności (7.58). Z kolei do wyznaczania maksymalnej odległości pomiędzy dwoma punktami w odtworzonej przestrzeni, tj. $x_{em}(i)$ i $x_{em}(j)$, skorzystamy z zależności (7.62). Do zliczania punktów przestrzeni, które spełniają kryterium podobieństwa $\mathfrak{p}\mathfrak{d}$ korzystamy z zależności (7.61). Następnie wyznaczyć należy:

$$C(\mathfrak{z}, \mathfrak{p}\mathfrak{d}) = \frac{1}{K - \mathfrak{z} + 1} \sum_{i=1}^{K-\mathfrak{z}+1} C_i(\mathfrak{z}, \mathfrak{p}\mathfrak{d}) \quad (7.66)$$

Następnie biorąc pod uwagę (7.66), wartość entropii próbkowanej wyznaczamy z zależności:

$$\text{SampEn}(\mathfrak{z}, \mathfrak{p}\mathfrak{d}, K) = -\ln \left(\frac{C(\mathfrak{z} + 1, \mathfrak{p}\mathfrak{d})}{C(\mathfrak{z}, \mathfrak{p}\mathfrak{d})} \right) \quad (7.67)$$

- **entropia Rényi** (ang. *Rényi entropy*, REn) jest uogólnieniem entropii Shannona [73, 139, 446].

$$\text{REn}(\mathfrak{r}\mathfrak{m}) = -\frac{1}{\mathfrak{r}\mathfrak{m} - 1} \ln \left(\sum_{n=1}^{N_{\text{kf}}} P_n^{\mathfrak{r}\mathfrak{m}} \right) \quad (7.68)$$

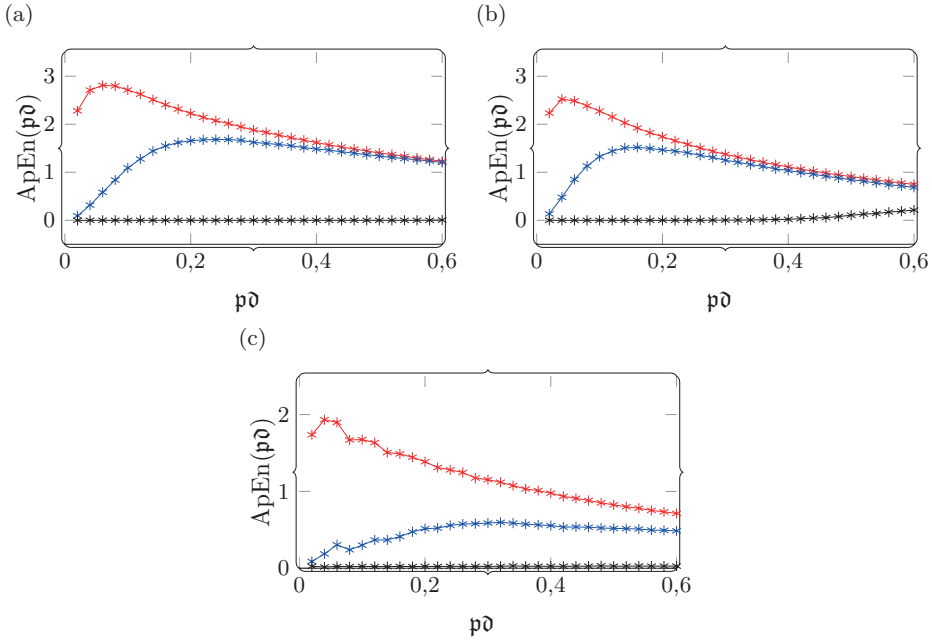
gdzie P_n jest prawdopodobieństwem wystąpienia n -tej kopii fraktala, N_{kf} to liczba kopii fraktala, a $\mathfrak{r}\mathfrak{m}$ to rząd momentu [320].

Przykład 7.8 (Entropia). W przykładzie zbadamy wpływ zmian amplitudy i częstotliwości sygnałów pomiarowych na wartość szacowanej entropii. W przykładzie porównaliśmy entropię przybliżoną z entropią próbkowaną. Wyniki analizy przedstawiono na rys. 7.36–7.39. W pierwszym badaniu oceniono wpływ parametrów \mathfrak{z} oraz $\mathfrak{p}\mathfrak{d}$ na jakość estymacji entropii przybliżonej i próbkowanej. W badaniach wykorzystaliśmy szum biały, fraktalny oraz sygnał o zmiennej częstotliwości. Z badań wynika, że dla sygnałów testowych entropia przybliżona jest czuła na dobór parametrów. Z kolei entropia próbkowana wykazuje mniejszą czułość, jednakże dla sygnału o zmiennej częstotliwości jest ona duża.

W drugim badaniu sprawdziliśmy wpływ zmian amplitudy i częstotliwości sygnału testowego na estymację entropii przybliżonej i próbkowanej. Z rysunków 7.38 oraz 7.39 wynika, że zmiany amplitudy nie wpływają na estymację entropii zarówno przybliżonej i próbkowanej. Z kolei inaczej jest w przypadku zmiany częstotliwości. W tym przypadku widać, że zmiany częstotliwości mają wpływ na estymację wartości entropii obiema metodami.

Wykładnik Lapunowa

Wykładnik Lapunowa (ang. *Lyapunov exponent*, LE) określa średnie tempo rozbieżności orbit dynamicznych wzdłuż osi układu współrzędnych. LE umożliwia



Rys. 7.36. Przybliżona entropia $ApEn(p\delta)$ dla: a) szumu białego, b) ułamkowego szumu *gaussowskiego* oraz dla c) sygnału o zmiennej częstotliwości i dla różnych wartości wymiaru zanurzenia

szacowanie wrażliwości układu na zmiany warunków początkowych [309]. Późniejsze prace z zakresu m.in. przetwarzania pokazały, że wykładnik Lapunowa może być podstawą algorytmów generowania cech.

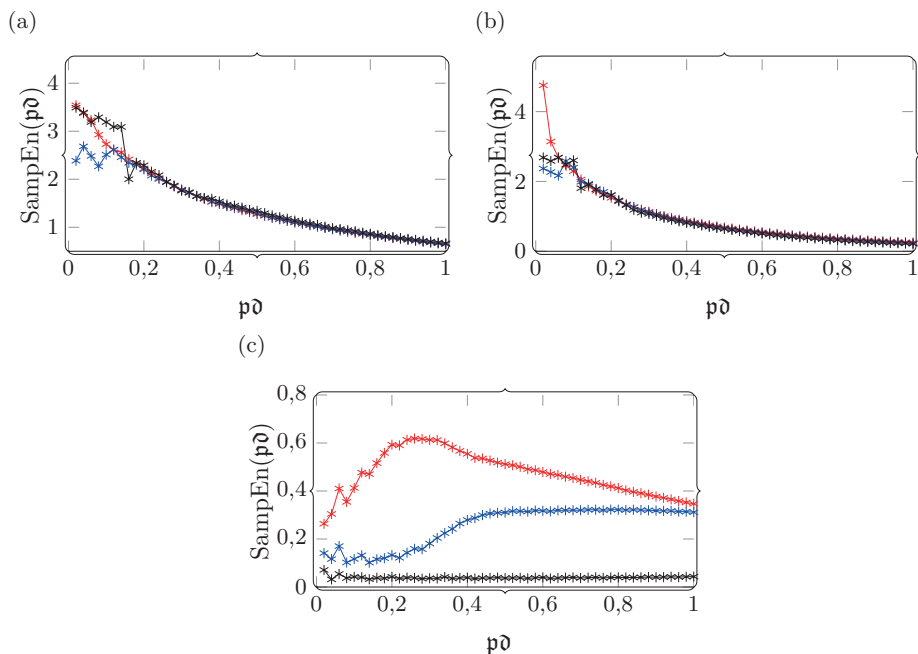
Wartość wykładnika Lapunowa dla systemów, których opis jest znany, można wyznaczyć na drodze obliczeń matematycznych. W przypadkach praktycznych, gdy opis systemu nie jest znany, możliwe jest jedynie szacunkowe wyznaczenie tej wielkości, bazując na danych pomiarowych i stosując metody numeryczne do jego wyznaczania.

Zaproponowano wiele metod służących do wyznaczania wykładnika Lapunowa. Poniżej przedstawiono jedną z metod, która została zaproponowana w pracy [349].

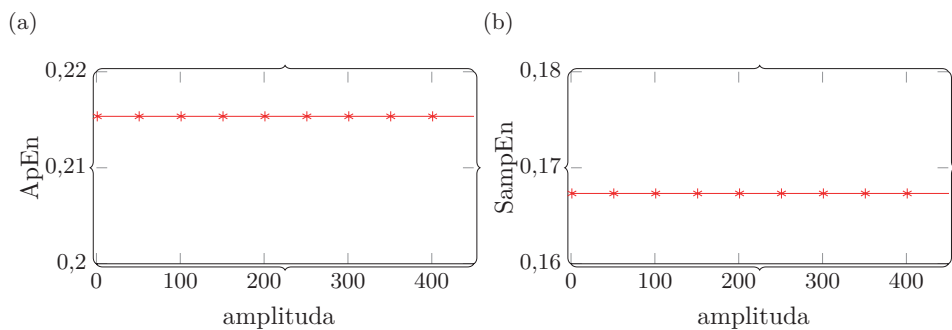
Dla pewnego zbioru pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$ konstruujemy wielowymiarową przestrzeń stanów, korzystając z zależności (7.58). Następnie wyznaczamy najbliższego sąsiada dla każdego punktu w odtworzonej przestrzeni stanów (zależności (7.58a)–(7.58c)), tj.

$$nsp_i = \min_{y^{(i)}} (||x_{em}(i), x_{em}^*(i)||) \quad (7.69)$$

7. Poziom średni: fuzja właściwości

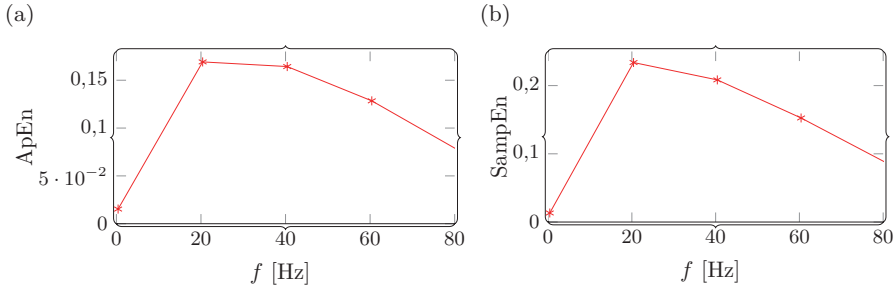


Rys. 7.37. Entropia próbkowana $\text{SampEn}(p\delta)$ dla: a) szumu białego, b) ułamkowego szumu *gaussowskiego* c) sygnału o zmiennej częstotliwości i dla różnych wartości wymiaru zanurzenia



Rys. 7.38. Wpływ zmiany wartości amplitudy sygnału na wartość wyznaczonej entropii dla entropii: a) przybliżonej ApEn , b) próbkowanej SampEn

gdzie $x_{em}^*(i)$ jest najbliższym sąsiadem punktu $x_{em}(i)$, a $(\|\cdot\|)$ reprezentuje normę euklidesową. W następnym kroku można już wyznaczyć wykładnik Lapunowa



Rys. 7.39. Wpływ zmiany wartości częstotliwości sygnału na wartość wyznaczonej entropii dla entropii: a) przybliżonej ApEn, b) próbkowanej SampEn

z następującej zależności [56, 349]:

$$\text{LapExp}(i) = \frac{1}{jF_s} \frac{1}{(K_{\text{po}} - j)} \sum_{i=1}^{K_{\text{po}}-j} \ln \frac{\text{nsp}_i(j)}{\text{nsp}_i(0)} \quad (7.70)$$

gdzie F_s to czas próbkowania, $K_{\text{po}} = (K - (j - 1)m)$.

Przykład 7.9 (Wykładniki Lapunowa). Przykład wykorzystania wykładników Lapunowa w analizie sygnałów przedstawiono na rys. 7.40 oraz 7.41.

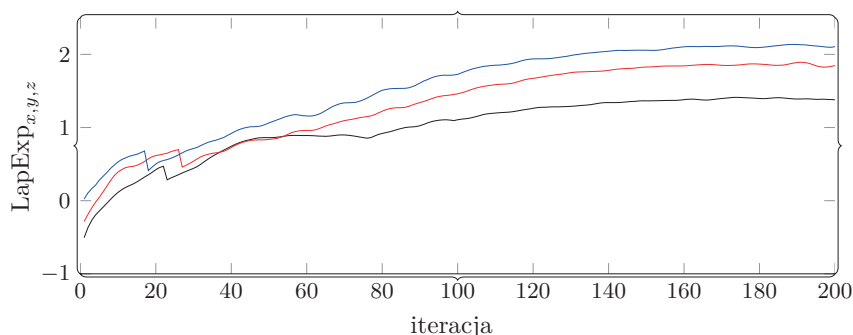
Wykrywanie chaosu w sekwencji danych przetestowano dla układu Lorenza. Rozważono dwa przypadki, tj. taki zestaw parametrów, w którym dobrane ich wartości sprawiają, że układ przejawia zachowanie chaotyczne oraz taki ich zestaw, dla którego nie występuje zachowanie chaotyczne w układzie.

Dla układów chaotycznych wartość wykładnika Lapunowa jest dodatnia, co widoczne jest na rys. 7.40. Z kolei w układach niechaotycznych wartość wykładnika Lapunowa jest ujemna. Na rysunku 7.41 widać, że wartość wykładnika Lapunowa dla każdej zmiennej jest ujemna. W tym przypadku parametry układu Lorenza zostały tak dobrane, że w układzie nie ma chaosu (por. 7.41a).

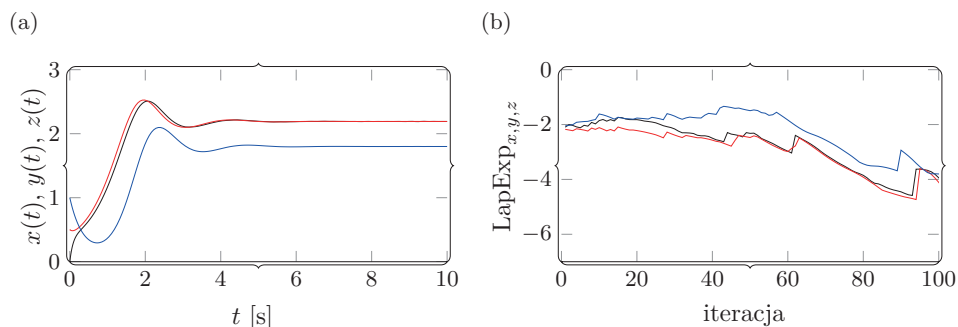
Wykładnik Hursta

Wykładnik Hursta (ang. *Hurst exponent*, HE) jest miarą długookresowej pamięci w sygnałach. Jest on również wykorzystywany jako wskaźnik stopnia chaotyczności obiektu, z którym powiązany jest sygnał pomiarowy [117, 177, 309]. Istnieje kilka podejść do estymacji wykładnika Hursta. Jednym z nich jest zastosowanie analizy przeskalowanego zakresu R/S [142] (ang. *rescaled range statistic*). Na pierwszym etapie zestaw pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$ dzielony jest na N_{si} nienakładających się podsekwencji ($\mathcal{A}_{Li}, i = 1, 2, \dots, N_{si}$), gdzie ich długości wyznaczone są z zależności $L_{si} = \left\lfloor \frac{K}{N_{si}} \right\rfloor$.

7. Poziom średni: fuzja właściwości



Rys. 7.40. Estymowana wartość wykładnika Lapunowa dla chaotycznego układu Lorenza metodą Rosensteina. Wartości estymowanych wykładników to $\text{LapExp}_x = 21,33$ (zmienna $x(t)$), $\text{LapExp}_y = 21,47$ (zmienna $y(t)$), $\text{LapExp}_z = 30,06$ (zmienna $z(t)$)



Rys. 7.41. Estymowana wartość wykładnika Lapunowa dla układu Lorenza. Z lewej rozwiązanie układu Lorenza dla układu bez chaosu. Z prawej wartości wykładnika Lapunowa (estymowanych metodą Rosensteina) to $\text{LapExp}_x = -49,96$ (zmienna $x(t)$), $\text{LapExp}_y = -37,59$ (zmienna $y(t)$), $\text{LapExp}_z = -33,67$ (zmienna $z(t)$)

W kolejnym kroku dla każdego segmentu wyznaczana jest wartość średnia:

$$\widehat{\mathcal{V}}_{Li} = \frac{1}{L_{si}} \sum_{j=1}^{L_{si}} \mathcal{V}_{Li}(j), \quad i = 1, 2, \dots, N_{si} \quad (7.71)$$

Bazując na wyznaczonych wartościach średnich dla każdej z podsekwencji, w kolejnym kroku wyliczane jest:

$$\mathcal{T}_i(k) = \sum_{j=1}^k (\mathcal{V}_{Li}(j) - \widehat{\mathcal{V}}_{Li}), \quad i = 1, 2, \dots, N_{si} \quad (7.72)$$

Następnie wyznaczamy kolejne dwie wielkości:

$$R_N = \max_{1 \leq j \leq L_{si}} \mathcal{V}_i(j) - \min_{1 \leq j \leq L_{si}} \mathcal{V}_i(j), \quad i = 1, 2, \dots, N_{si} \quad (7.73)$$

oraz

$$S_N = \sqrt{\frac{1}{M} \sum_{j=1}^{L_{si}} (\mathcal{V}_{Li}(j) - \widehat{\mathcal{V}}_{Li})^2}, \quad i = 1, 2, \dots, N_{si} \quad (7.74)$$

Uwzględniając (7.73) oraz (7.74) wykładnik Hursta wyznaczany jest z uśrednionego przeskalowanego zakresu R/S :

$$\left(\frac{R_N}{S_N} \right) = \frac{1}{N_{si}} \sum_{n=1}^{N_{si}} \left(\frac{R_n}{S_n} \right) \quad (7.75)$$

Zależność (7.75) jest proporcjonalna do [292]:

$$E \left[\frac{R_N}{S_N} \right] = \mathfrak{g} N^{\mathfrak{H}} \quad (7.76)$$

gdzie \mathfrak{g} jest pewną stałą (jej typowa wartość to $\mathfrak{g} = 0,5$) [425], a \mathfrak{H} jest wykładnikiem Hursta.

Wykładnik Hursta przyjmuje wartości z zakresu $\mathfrak{H} = \langle 0, 1 \rangle$. Wartość tego wykładnika wyznaczona dla analizowanego sygnału umożliwia jego scharakteryzowanie:

- jeżeli $\mathfrak{H} = \langle 0, 0,5 \rangle$, to wówczas sygnał jest **antypersystentny**, tj. mają tendencję do częstych zwrotów kierunków przemieszczania [425],
- jeżeli $\mathfrak{H} = 0,5$, to wówczas sygnał jest sygnałem losowym,
- jeżeli $\mathfrak{H} = \langle 0,5, 1,0 \rangle$, to wówczas sygnał jest **persystentny**, tj. w sygnałach pojawia się efekt długiej pamięci oraz mają tendencję do wzmacniania trendów [425].

Do wyznaczenia wykładnika Hursta można posłużyć się również innymi metodami, takimi jak metoda bazująca na omówionym już algorytmie DFA (7.4.1) [255, 292]. Dla zbioru pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$ w pierwszym kroku wyznaczamy wartość średnią:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K y(k) \quad (7.77)$$

Następnie należy wyliczyć skumulowaną sumę z zależności:

$$u(k) = \sum_{i=1}^k (y(i) - \hat{y}), \quad k = 1, 2, \dots, K \quad (7.78)$$

7. Poziom średni: fuzja właściwości

Wynikowa suma skumulowana $\mathcal{U}(k)$ dzielona jest następnie na równych N_{si} przedziałów o długości L_{si} , gdzie $N_{si} = \lfloor K/L_{si} \rfloor$. Następnie w każdym z przedziałów wyznacza się linię trendu (\mathcal{U}_{tr}). W typowym przypadku będzie to prosta. Zauważmy, że do estymacji trendu można wykorzystać np. wielomiany wyższego rzędu [243]. W kolejnym kroku wyznaczamy szereg czasowy, z którego usunięto linię trendu:

$$F_{tr} = \sqrt{\frac{1}{K} \sum_{k=1}^K (\mathcal{U}(k) - \mathcal{U}_{tr}(k))^2} \quad (7.79)$$

Wyznaczona zależność (7.79) jest proporcjonalna do [243]:

$$F_{tr} = L_{si}^{\mathfrak{H}} \quad (7.80)$$

gdzie \mathfrak{H} jest wykładnikiem Hursta.

Inny przykład metody szacowania wykładnika Hursta przedstawiono w [2, 420]. W pracy zaproponowano metodę wyznaczania wykładnika Hursta bazującą na transformacji falkowej. W pierwszym kroku dla zbioru pomiarów $\mathbf{Y}_K = (y(1), y(2), \dots, y(K))$ wyznaczane są współczynniki transformacji falkowej (patrz podrozdz. 3:

$$Y_{WVT}(\mathbf{a}, \mathbf{b}) = \frac{1}{\sqrt{\mathbf{a}}} \int_{-\infty}^{\infty} y(t) \Psi\left(\frac{t - \mathbf{b}}{\mathbf{a}}\right) dt \quad (7.81)$$

Następnie zapisujemy warunek samopodobieństwa w postaci [65, 309, 371]:

$$Y_{WVT}(S\mathbf{a}, S\mathbf{b}) \approx S^{(1/2)+\mathfrak{H}} Y_{WVT}(\mathbf{a}, \mathbf{b}) \quad (7.82)$$

gdzie $S \rightarrow 0$. Jeśli dokonamy uśrednienia współczynników względem \mathbf{b} , to wówczas zależność (7.82) możemy zapisać jako:

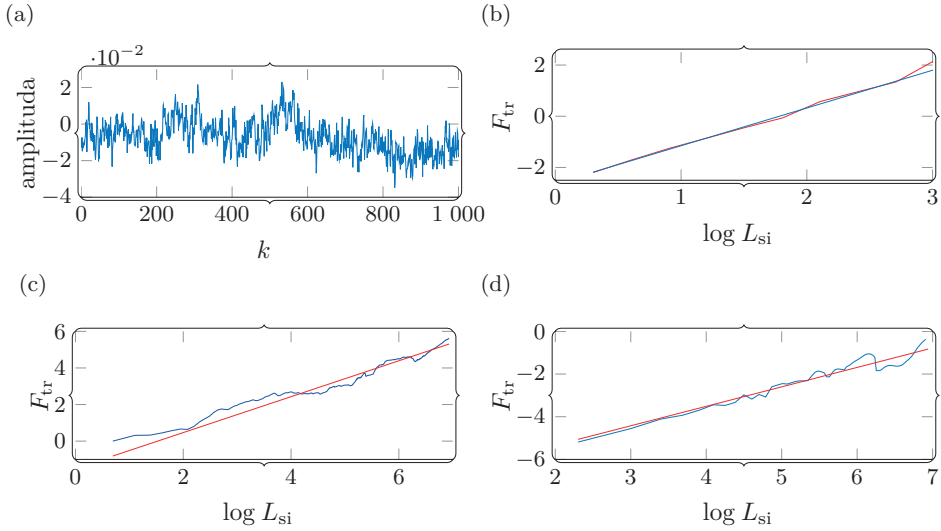
$$Y_{WVT}(S\mathbf{a}) \approx S^{(1/2)+\mathfrak{H}} Y_{WVT}(\mathbf{a}) \quad (7.83)$$

Wykładnik Hursta \mathfrak{H} można estymować, wyznaczając współczynnik nachylenia prostej w :

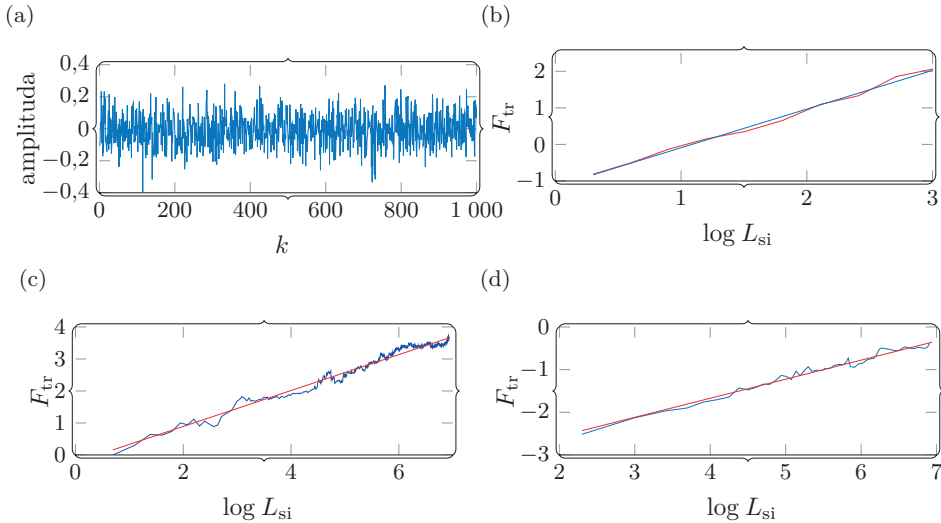
$$\mathfrak{H} = w - \frac{1}{2} \quad (7.84)$$

przez analizę zależności $Y_{WVT}(\mathbf{a})$ i w w skali logarytmicznej.

Przykład 7.10 (Wykładnik Hursta). *Przykład ilustruje działanie metod wykorzystywanych do estymacji wykładnika Hursta dla ułamkowego szumu gaussowskiego wygenerowanego dla wykładnika Hursta $\mathfrak{H} = 0,95$, $\mathfrak{H} = 0,5$ oraz $\mathfrak{H} = 0,10$. Wyniki przedstawiono kolejno na rys. 7.42–7.44.*

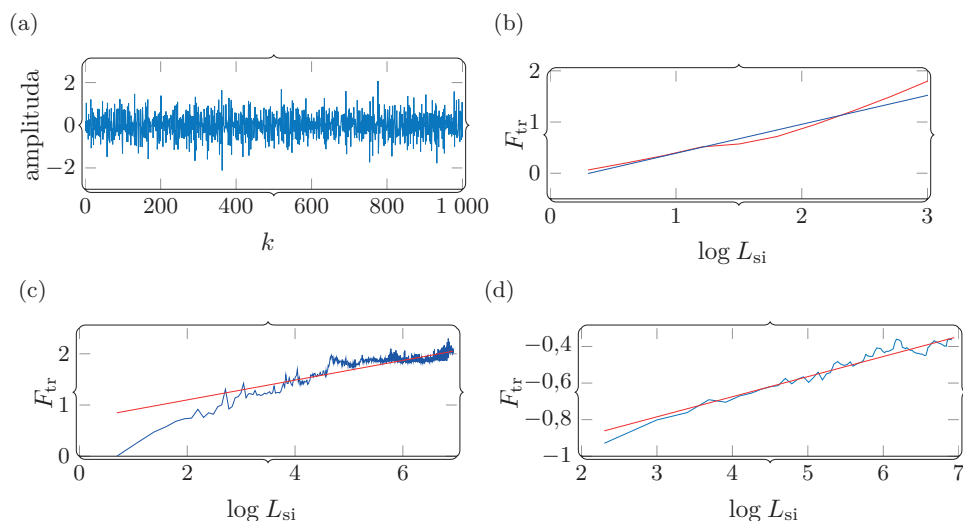


Rys. 7.42. Estymowana wartość wykładnika Hursta dla:
 a) ułamkowego szumu *gaussowskiego* ($\mathfrak{H} = 0,95$),
 b) algorytm bazujący na transformacie falkowej ($\mathfrak{H} = 0,97$),
 c) algorytm bazujący na analizie *R/S* ($\mathfrak{H} = 0,98$), d) algorytm DFA ($H = 0,91$)



Rys. 7.43. Estymowana wartość wykładnika Hursta dla: a) ułamkowego szumu *gaussowskiego* ($\mathfrak{H} = 0,50$), b) algorytm bazujący na transformacie falkowej ($\mathfrak{H} = 0,55$),
 c) algorytm bazujący na analizie *R/S* ($H = 0,56$), d) algorytm DFA ($\mathfrak{H} = 0,44$)

7. Poziom średni: fuzja właściwości



Rys. 7.44. Estymowana wartość wykładnika Hursta dla:

- a) ułamkowego szumu *gaussowskiego* ($\mathfrak{H} = 0,10$),
- b) algorytm bazujący na transformacie falkowej ($\mathfrak{H} = 0,07$),
- c) algorytm bazujący na analizie *R/S* ($\mathfrak{H} = 0,19$), d) algorytm DFA ($\mathfrak{H} = 0,10$)

Również wykładnik Hursta znalazł zastosowanie w projektowaniu algorytmów do konstrukcji wektora cech. Na przykład w pracy [82] wykładnik ten został wykorzystany do przetwarzania sygnałów EMG. Z kolei w pracy [241] został on wykorzystany do ekstrakcji cech z sygnałów EEG.

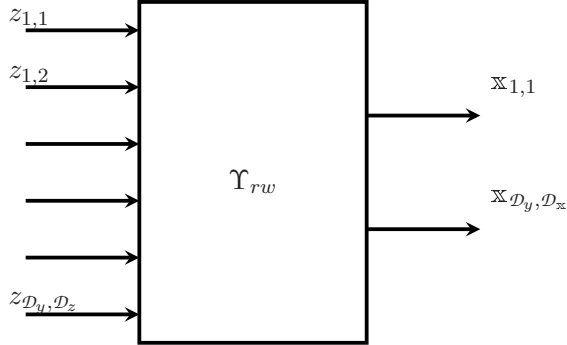
7.6. Redukcja wymiaru

Jednym z efektów wykorzystania metod generowania cech z sygnałów pomiarowych jest to, że liczba wyznaczonych cech może być bardzo duża. Niestety, wraz ze wzrostem liczby cech zwiększa się ryzyko wzajemnych zależności pomiędzy wyznaczonymi cechami. Oznacza to, że informacje wnoszone przez niektóre z tych zmiennych mogą być redundantne.

Oznacza to, że część z cech nie wnosi nic nowego do opisu obiektu, a jedynie powiela informacje, które dostarcza inna z cech. Istnienie w jednym zbiorze cech, które są ze sobą powiązane, może wpływać niekorzystnie na złożoność obliczeniową algorytmów fuzji właściwości. Ponadto cechy najlepiej charakteryzujące obiekt nie są znane *a priori*, a ich liczba zazwyczaj przekracza liczbę cech, które niezbędne są do opisanego analizowanego obiektu. Zagadnienia związane z analizą cech i ich przydatnością w konstrukcji wymiaru nazywamy **redukcją wymiarowości** [41].

Definicja 7.6.1. Redukcja wymiaru to poszukiwanie najmniejszego zbioru cech, na podstawie którego można dokonać opisu obiektu z akceptowalnym błędem.

Na rysunku 7.45 przedstawiono ilustrację zadania redukcji wymiaru, natomiast teraz przejdźmy do sformułowania tego zadania.



Rys. 7.45. Ilustracja zadania redukcji wymiaru

7.6.1. Sformułowanie problemu

Dla zestawu cech (por. (7.24)):

$$\mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z}(k) = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1, \mathcal{D}_z} \\ z_{2,1} & z_{2,2} & \dots & z_{2, \mathcal{D}_z} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_y, 1} & z_{\mathcal{D}_y, 2} & \dots & z_{\mathcal{D}_y, \mathcal{D}_z} \end{bmatrix} \quad (7.85)$$

wyznacz

$$\mathbb{X}_{\mathcal{D}_y, \mathcal{D}_x}(k) = \begin{bmatrix} \mathbb{X}_{1,1} & \mathbb{X}_{1,2} & \dots & \mathbb{X}_{1, \mathcal{D}_x} \\ \mathbb{X}_{2,1} & \mathbb{X}_{2,2} & \dots & \mathbb{X}_{2, \mathcal{D}_x} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{X}_{\mathcal{D}_y, 1} & \mathbb{X}_{\mathcal{D}_y, 2} & \dots & \mathbb{X}_{\mathcal{D}_y, \mathcal{D}_x} \end{bmatrix} \quad (7.86)$$

gdzie $\mathcal{D}_x \ll \mathcal{D}_z$. Co oznacza, że po rozwiązaniu zadania redukcji wymiaru liczba cech powinna zostać zmniejszona, przy jednoznacznej maksymalizacji informacyjności cech, które pozostały. Zadanie redukcji cech możemy przedstawić jako

7. Poziom średni: fuzja właściwości

zadanie poszukiwania takiego odwzorowania, które:

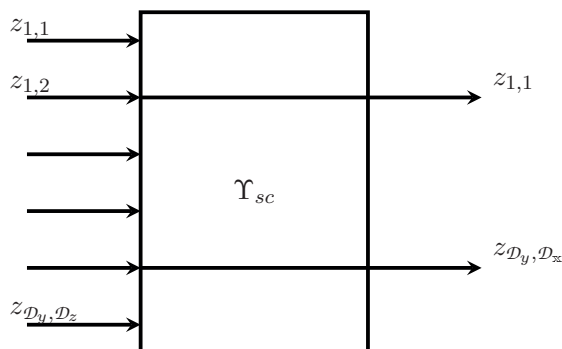
$$\Upsilon_{rw} : \mathbf{Z} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_z} \rightarrow \mathbb{X} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_z} \quad (7.87)$$

Wyznaczony nowy wektor cech może być rozpięty na tej samej przestrzeni co wektor oryginalny. Alternatywnie konstruowany wektor cech może być rozpięty na innej przestrzeni niż wektor oryginalny. W pierwszym przypadku, gdy szukamy podzbioru ze zbioru oryginalnego, zadanie takie nazywamy zadaniem **selekcji cech**. W drugim z wymienionych przypadków zadanie takie nazywa się zadaniem **redukcji cech**. Zadania te zostaną omówione w kolejnych podrozdziałach.

7.6.2. Selekcja cech

Omówimy teraz zadanie selekcji cech oraz odpowiednie algorytmy służące do jego rozwiązania. Celem metod selekcji cech jest wybranie ze zbioru wszystkich wygenerowanych cech takiego ich podzbioru, który dla możliwie najmniejszej ich liczby najlepiej opisuje analizowany obiekt. Selekcja zbioru cech ma wiele zalet, wśród których można wymienić m.in. zmniejszony nakład obliczeń na dalszych etapach fuzji danych oraz redukcję kosztów związanych z transmisją danych.

Ilustracja tego zadania została przedstawiona na rys. 7.46. Przedstawimy teraz definicję oraz sformułowanie problemu selekcji cech [434, 435].



Rys. 7.46. Ilustracja zadania selekcji cech

Definicja 7.6.2 (Selekcja cech). *Selekcja cech to wybór zbioru cech charakteryzujących obiekt ze zbioru oryginalnych cech pochodzących z pomiaru obiektu.*

Sformułowanie problemu

Dla zestawu cech:

$$\mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z}(k) = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,\mathcal{D}_z} \\ z_{2,1} & z_{2,2} & \dots & z_{2,\mathcal{D}_z} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_y,1} & z_{\mathcal{D}_y,2} & \dots & z_{\mathcal{D}_y,\mathcal{D}_z} \end{bmatrix} \quad (7.88)$$

należy wyznaczyć:

$$\mathbf{X}_{\mathcal{D}_y, \mathcal{D}_x}(k) = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,\mathcal{D}_x} \\ z_{2,1} & z_{2,2} & \dots & z_{2,\mathcal{D}_x} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_y,1} & z_{\mathcal{D}_y,2} & \dots & z_{\mathcal{D}_y,\mathcal{D}_x} \end{bmatrix} \quad (7.89)$$

gdzie $\mathcal{D}_x \ll \mathcal{D}_z$. Co oznacza, że po rozwiązaniu zadania selekcji cech otrzymujemy ich zredukowaną liczbę, z zachowaniem informacyjności cech, które pozostały. Zadanie selekcji cech polega na poszukiwaniu takiego algorytmu wyboru cech z oryginalnego ich zbioru, który:

$$\Upsilon_{sc} : \mathbf{Z} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_z} \rightarrow \mathbf{X} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_x} \quad (7.90)$$

Jeden ze sposobów rozwiązania sformułowanego problemu bazuje na przeglądzie zupełnym. Niestety wraz ze wzrostem liczby cech, z których ma zostać zbudowany wektor cech, liczba koniecznych operacji rośnie zgodnie z zależnością:

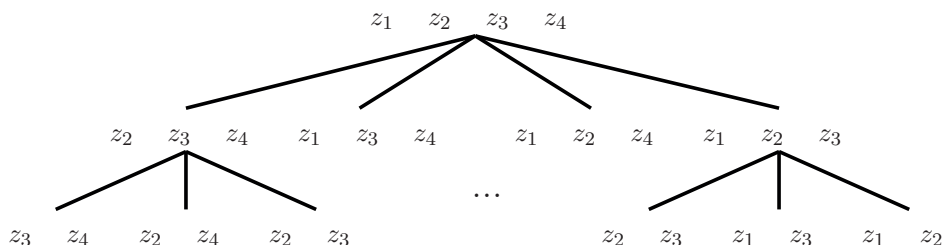
$$\binom{\mathcal{D}_z}{\mathcal{D}_x} = \frac{\mathcal{D}_z!}{\mathcal{D}_x!(\mathcal{D}_z - \mathcal{D}_x)!} \quad (7.91)$$

Zależność ta pokazuje, że szukanie podzbioru cech już dla zadań średniego rozmiaru może być złożone obliczeniowo. Dlatego tak istotne jest projektowanie algorytmów selekcji cech, które umożliwią rozwiązanie tego zadania w akceptowalnym czasie. Warto w tym miejscu dodać, że większość opracowanych metod nie gwarantuje znalezienie optymalnego zbioru cech. Natomiast metod, które umożliwiają rozwiązanie optymalne, jest niewiele. Jedną z nich, która to ułatwia, jest metoda selekcji cech bazująca na metodzie podziału i ograniczeń.

Metoda podziału i ograniczeń

Podstawą podejścia wykorzystującego schemat podziału i ograniczeń jest odpowiednio przygotowane drzewo, które na kolejnych etapach jest systematycznie przeszukiwane. Przykład takiego drzewa przedstawiono na rys. 7.47.

7. Poziom średni: fuzja właściwości



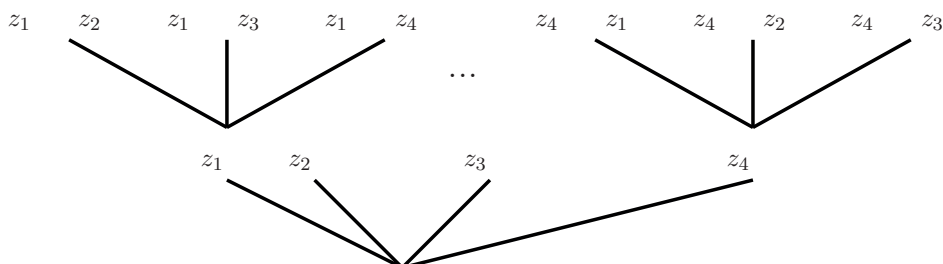
Rys. 7.47. Drzewo dla zadania selekcji cech (strategia *top-down*)

Proces przeszukiwania drzewa rozpoczyna się od najwyższego poziomu. Następnie przechodzi się do kolejnych, niższych poziomów. Ponieważ przeszukiwanie całego drzewa byłoby zadaniem bardzo czasochłonnym, dlatego algorytm ten na każdym etapie odznacza obiecujące rozwiązania i w ten sposób zawęża obszar przeszukiwań.

Metody suboptymalne

Wadą podejścia wykorzystującego metodę podziału i ograniczeń jest to, że w szczególnie niekorzystnym przypadku selekcja zbioru wygenerowanych cech może sprowadzić się do przeglądu zupełnego. Inną wadą tego podejścia jest to, że konstrukcja rozwiązania rozpoczyna się od całego zbioru dostępnych cech i dopiero w kolejnych krokach zbiór ten jest zmniejszany (ang. *top-down strategy*). Dlatego w praktycznych zadaniach preferowane są metody, w których rozwiązanie konstruuje się od najmniejszej liczby cech, a w następnych krokach dodaje się do zbioru wynikowego kolejne elementy (ang. *bottom-up strategy*). Przykład takiej strategii podano na rys. 7.48.

Wśród metod, które opierają się na tym podejściu, wyróżniamy m.in sekwencyjną selekcję postępującą (ang. *Sequential Forward Selection, SFS*) oraz uogólnioną



Rys. 7.48. Drzewo dla zadania selekcji cech (strategia *bottom-up*)

sekwencyjną selekcję postępującą (ang. *Generalized Sequential Forward Selection*, GSFS).

7.6.3. Redukcja cech

Drugie z podejść do redukcji wymiaru nazywane jest zadaniem redukcji cech. W tym podejściu dokonuje się pewnej transformacji na zbiorze wygenerowanych cech. Przy czym liczność zbioru po transformacji jest mniejsza w porównaniu do zbioru źródłowego. Zauważmy, że cel redukcji cech jest taki sam jak w przypadku zadania selekcji cech, tj. modyfikacja oryginalnego zbioru wygenerowanych cech do zbioru o mniejszej liczności z zachowaniem informacyjności zbioru wynikowego. Oznacza to, że do konstrukcji wektora cech można wykorzystać algorytm redukcji cech. Podamy teraz definicję zadania redukcji cech [433] oraz sformułujemy to zadanie.

Definicja 7.6.3. *Redukcja cech to problem transformacji wygenerowanych cech do przestrzeni cech o niższym wymiarze.*

Efektom zastosowania redukcji cech na pewnym zbiorze wejściowym jest nowy zbiór cech, których nie było w zbiorze wejściowym.

Sformułowanie problemu

Dla zestawu $\mathcal{D}_y \times \mathcal{D}_z$ cech wejściowych:

$$\mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z}(k) = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,\mathcal{D}_z} \\ z_{2,1} & z_{2,2} & \dots & z_{2,\mathcal{D}_z} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_y,1} & z_{\mathcal{D}_y,2} & \dots & z_{\mathcal{D}_y,\mathcal{D}_z} \end{bmatrix} \quad (7.92)$$

wyznacz $\mathcal{D}_y \times \mathcal{D}_x$ zestaw cech wynikowych:

$$\mathbf{X}_{\mathcal{D}_y, \mathcal{D}_x}(k) = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,\mathcal{D}_x} \\ x_{2,1} & x_{2,2} & \dots & x_{2,\mathcal{D}_x} \\ \vdots & \vdots & \ddots & \vdots \\ x_{\mathcal{D}_y,1} & x_{\mathcal{D}_y,2} & \dots & x_{\mathcal{D}_y,\mathcal{D}_x} \end{bmatrix} \quad (7.93)$$

gdzie $\mathcal{D}_x \ll \mathcal{D}_z$. Oznacza to, że algorytm redukcji cech Υ_{rc} dokonuje przekształcenia:

$$\Upsilon_{rc} : \mathbf{Z} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_z} \rightarrow \mathbf{X} \in \mathbb{R}^{\mathcal{D}_y \times \mathcal{D}_x} \quad (7.94)$$

Tak jak to było w przypadku zadania selekcji cech, również i dla zadania redukcji cech opracowano wiele metod. W dalszej części omówiono wybrane z nich.

Analiza głównych składowych

Jedną z najczęściej wykorzystywanych metod redukcji cech jest algorytm analizy głównych składowych (ang. *Principal Component Analysis*, PCA). Celem metody PCA jest przekształcenie oryginalnego wektora cech do nowego rozpiętego na nowej przestrzeni. Obliczenia rozpoczyna się od wyznaczenia macierzy transformacji \mathbf{W}_{PCA} :

$$\mathbf{x} = \mathbf{W}_{\text{PCA}}\mathbf{z} \quad (7.95)$$

gdzie \mathbf{W}_{PCA} to macierz o wymiarach $\mathcal{D}_x \times \mathcal{D}_z$, z kolei wektor cech \mathbf{z} jest \mathcal{D}_z -wymiarowym wektorem w postaci:

$$\mathbf{z} = \text{row}_i \mathbf{Z}_{\mathcal{D}_y, \mathcal{D}_z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\mathcal{D}_z} \end{bmatrix}^T \quad (7.96)$$

gdzie $i = 1, 2, \dots, \mathcal{D}_y$, a \mathbf{x} jest wektorem o \mathcal{D}_x składowych:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{\mathcal{D}_x} \end{bmatrix}^T \quad (7.97)$$

Składowe nowego wektora \mathbf{x} są liniową kombinacją oryginalnego wektora cech \mathbf{z} oraz wag macierzy \mathbf{W}_{PCA} :

$$x_i = \sum_{j=1}^{\mathcal{D}_z} w_{pi,j} z_j = w_{pi,1} z_1 + w_{pi,2} z_2 + \dots + w_{pi,\mathcal{D}_z} z_{\mathcal{D}_z}, \quad i = 1, 2, \dots, \mathcal{D}_x \quad (7.98)$$

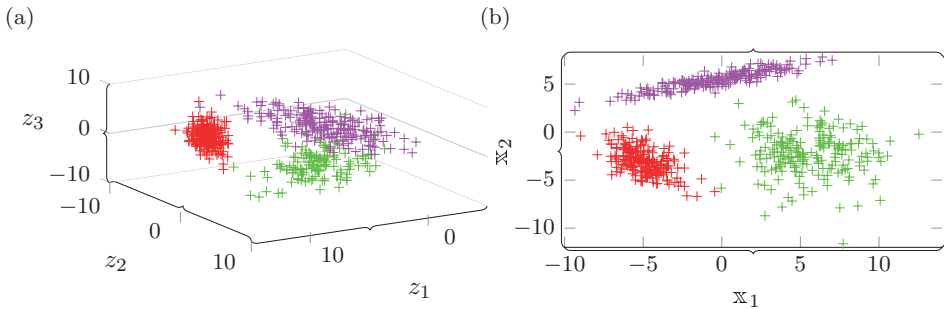
gdzie $w_{pi,j}$ są elementami macierzy \mathbf{W}_{PCA} . Dodajmy, że geometrycznie macierz \mathbf{W}_{PCA} jest macierzą rotacji (patrz podrozdz. 9.2).

Alternatywne podejście do redukcji cech wykorzystujące metodę PCA bazuje na metodzie SVD (Rozkład według wartości osobliwych, ang. *Singular Value Decomposition*). Metoda SVD różni się tym od metody PCA, że ta ostatnia jest przeznaczona do danych, w których występują liniowe zależności.

Algorytm PCA w zadaniu redukcji przedstawiono np. w pracy [343]. Autorzy tego opracowania wykorzystali metodę analizy głównych składowych zespołów QRS w sygnale EKG. Z kolei w pracy [61] zaprezentowano podejście bazujące na algorytmie PCA również do analizy sygnału EKG w celu analizy odcinków

ST–T. Wyniki analizy sygnału EKG wspomagają diagnostykę takich chorób, jak niedokrwienie mięśnia sercowego.

Przykład 7.11 (Analiza głównych składowych). W przykładzie zilustrowano działanie metody PCA dla zbioru danych pomiarowych przedstawionego na rys. 7.49a. Wynik działania algorytmu dla przykładowego zbioru danych przedstawiono na rys. 7.49b. Rysunek ten przedstawia, w jaki sposób, wykorzystując algorytm PCA, dokonano transformacji oryginalnego zbiorów wektorów cech w zbiór wynikowy \mathbb{X}_{D_y, D_x} .



Rys. 7.49. Ilustracja działania algorytmu PCA: a) dane pomiarowe, b) składowe główne

Analiza składowych niezależnych

Kolejną metodą, którą można wykorzystać w zadaniu redukcji cech, jest metoda analizy składowych niezależnych (ang. *Independent Component Analysis*, ICA). Oryginalnie metoda ta została opracowana na potrzeby zadania ślepej separacji źródeł (ang. *Blind Source Separation*, BSS). Istotą metody ICA jest dekompozycja sygnału wejściowego na zestaw addytywnych sygnałów składowych. Przy czym dekompozycja wykonywana jest bez żadnych informacji dotyczących źródła sygnału, jak również bez żadnych danych o warunkach, w jakich wykonano pomiary. W metodzie zakłada się natomiast to, że składowe są od siebie statystycznie niezależne. Celem metody ICA jest znalezienie następującej liniowej kombinacji sygnału pomiarowego $\mathbf{z}(k)$ [257, 374]:

$$\hat{\mathbf{s}}(k) \approx \mathbf{W}_{\text{ICA}} \mathbf{z}(k) \quad (7.99)$$

gdzie \mathbf{W}_{ICA} nazywa się macierzą separującą (ang. *separation matrix*), a $\mathbf{z}(k)$ wyznaczone jest z zależności:

$$\mathbf{z}(k) = \mathbf{A}_{\text{mix}} \mathbf{s}(k) \quad (7.100)$$

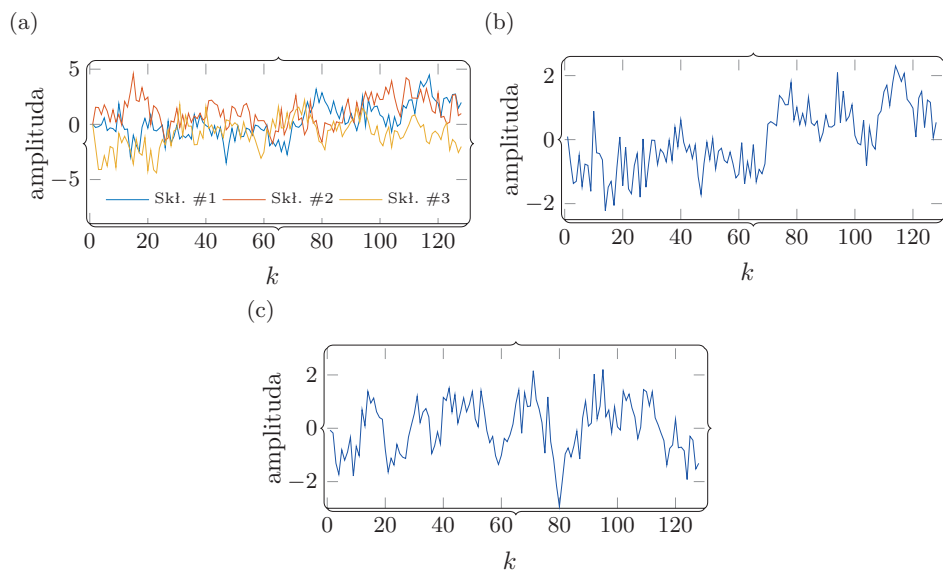
7. Poziom średni: fuzja właściwości

gdzie z kolei $\mathbf{s}(k)$ jest nieznanym sygnałem, a \mathbf{A}_{mix} jest macierzą współczynników nazywaną macierzą mieszającą (ang. *mixing matrix*).

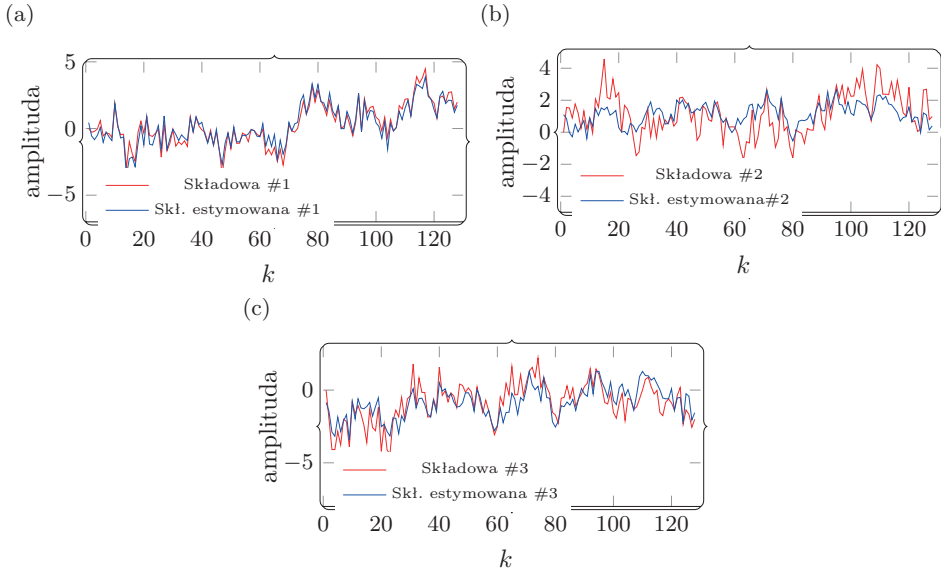
Warto zaznaczyć, że nie jest możliwe znalezienie dokładnego rozwiązania zadania (7.100), ponieważ dokładne postaci macierzy mieszającej \mathbf{A}_{mix} oraz $\mathbf{s}(k)$ nie są znane, dlatego w metodzie bazuje się na ich oszacowaniach. Podstawowym narzędziem do rozwiązania zadania (7.100) jest wykorzystanie statystyk wyższego rzędu [456].

Metoda ICA znalazła szerokie zastosowanie w zadaniach przetwarzania sygnałów. Na przykład w pracy [261] metodę tę wykorzystano do separacji sygnału EEG. Podobne zastosowanie przedstawiono w artykułach [166] oraz [268].

Przykład 7.12 (Analiza składowych niezależnych). *Przykładowy efekt działania metody ICA przedstawiono na rys. 7.50 i 7.51. Zadanie wykonano dla dwóch przypadków. W pierwszym z nich dokonano dekompozycji na dwie składowe, w drugim na trzy składowe. Dane pomiarowe przedstawiono na rys. 7.50a. Wynik dekompozycji na dwie składowe na rys. 7.50b i 7.50c. Z kolei wynik dekompozycji na trzy składowe na rys. 7.51.*



Rys. 7.50. Wynik działania algorytmu ICA: a) sygnały pomiarowe, b) składowa niezależna #1, c) składowa niezależna #2



Rys. 7.51. Wynik działania algorytmu ICA (ciąg dalszy).
Oryginalne pomiary składowych oraz ich wartości estymowane

Liniowa Analiza Dyskryminacyjna

Liniowa Analiza Dyskryminacyjna (ang. *Linear Discriminant Analysis*, LDA) oryginalnie była wykorzystywana w zadaniach, w których występowały dwie klasy. Później opracowano również wersję metody dla wielu klas. Cechą charakterystyczną podejścia jest to, że w metodzie tej dokonuje się wstępnego podziału danych wejściowych ze względu na przynależność do jednej z klas. Zauważmy, że w metodzie LDA wykorzystuje się dwie macierze, które służą do opisu wewnątrzgrupowej i międzygrupowej zmienności zbioru wektorów cech.

Metoda ta, może być również stosowana w zadaniach redukcji cech [185]. Niech wejściowy wektor cech ma postać:

$$\mathbf{z}_{\mathcal{D}_z}(k) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\mathcal{D}_z} \end{bmatrix} \quad (7.101)$$

7. Poziom średni: fuzja właściwości

Wektor ten wykorzystujemy do budowy następującej macierzy:

$$\mathbf{Z}_{K, \mathcal{D}_z} = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,K} \\ z_{2,1} & z_{2,2} & \dots & z_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_z,1} & z_{\mathcal{D}_z,2} & \dots & z_{\mathcal{D}_z,K} \end{bmatrix} \quad (7.102)$$

gdzie $\mathbf{Z} \in \mathbb{R}^{K \times \mathcal{D}_z}$, K jest liczbą pomiarów, a \mathcal{D}_z wymiarem oryginalnego wektora cech.

W metodzie należy przyjąć, że znana jest przynależność każdego z wektorów do jednej z N_c klas. Bazując na tym założeniu, dokonujemy podziału macierzy cech $\mathbf{Z}_{K, \mathcal{D}_z}$ (7.102) na N_c macierzy. Każda z tych macierzy powiązana jest z jedną z N_c klas, tj. $\mathbf{Z}_{K, \mathcal{D}_z} = [\mathbf{Z}_{K, \mathcal{D}_z, 1}, \mathbf{Z}_{K, \mathcal{D}_z, 2}, \dots, \mathbf{Z}_{K, \mathcal{D}_z, N_c}]$. Do każdej klasy przynależy K_{c_i} punktów pomiarowych takich, że $K = \sum_{i=1}^{N_c} K_{c_i}$.

Celem metody LDA jest wyznaczenie macierzy rzutowania \mathbf{W}_{LDA} (ang. *projection matrix*), której kolumny będą stanowiły \mathcal{D}_x -wymiarową przestrzeń liniową [165, 430, 475, 478]:

$$\mathbf{x} = \mathbf{W}_{\text{LDA}} \mathbf{z} \quad (7.103)$$

Optymalną macierz rzutowania można wyznaczyć, rozwiązując zadanie optymalizacji, w którym wskaźnik jakości ma postać:

$$Q(\mathbf{W}_{\text{LDA}}) = \frac{\text{tr}(\mathbf{S}_b)}{\text{tr}(\mathbf{S}_w)} \quad (7.104)$$

gdzie \mathbf{S}_b jest macierzą rozproszenia między klasami, z kolei \mathbf{S}_w jest macierzą rozproszenia wewnątrz klas. Macierze te wyznaczane są z następujących zależności:

$$\mathbf{S}_w = \frac{1}{K} \sum_{i=1}^{N_c} \sum_{j \in c_i} (\mathbf{x}_j - \hat{\mathbf{x}}_i)(\mathbf{x}_j - \hat{\mathbf{x}}_i)^\top \quad (7.105)$$

$$\mathbf{S}_b = \frac{1}{K} \sum_{i=1}^{N_c} K_{c_i} (\hat{\mathbf{x}}_i - \hat{\mathbf{x}})(\hat{\mathbf{x}}_i - \hat{\mathbf{x}})^\top \quad (7.106)$$

gdzie $\hat{\mathbf{x}}$ jest globalną średnią dla całego wektora cech, $\hat{\mathbf{x}}_i$ jest średnią wektora cech związanego z i -tą klasą. Podstawiając $\mathbf{x} = \mathbf{W}_{\text{LDA}} \mathbf{z}$, otrzymujemy:

$$\mathbf{S}_w = \frac{1}{K} \sum_{i=1}^{N_c} \sum_{j \in c_i} \mathbf{W}_{\text{LDA}}^\top (\mathbf{z}_j - \hat{\mathbf{z}}_i)(\mathbf{z}_j - \hat{\mathbf{z}}_i)^\top \mathbf{W}_{\text{LDA}} \quad (7.107)$$

$$\mathbf{S}_b = \frac{1}{K} \sum_{i=1}^{N_c} K_{ci} \mathbf{W}_{\text{LDA}}^T (\hat{\mathbf{z}}_i - \hat{\mathbf{z}}) (\hat{\mathbf{z}}_i - \hat{\mathbf{z}})^T \mathbf{W}_{\text{LDA}} \quad (7.108)$$

Bazując na zależnościach (7.107) oraz (7.108), zapiszmy wskaźnik jakości w następującej formie:

$$Q(\mathbf{W}_{\text{LDA}}) = \frac{\text{tr}(\mathbf{W}_{\text{LDA}}^T \mathbf{G}_b \mathbf{W}_{\text{LDA}})}{\text{tr}(\mathbf{W}_{\text{LDA}}^T \mathbf{G}_w \mathbf{W}_{\text{LDA}})} \quad (7.109)$$

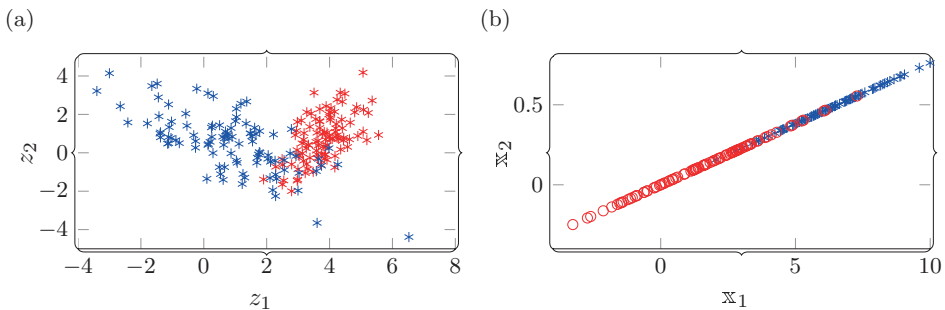
gdzie

$$\mathbf{G}_w = \frac{1}{K} \sum_{i=1}^{N_c} \sum_{j \in c_i} (\mathbf{z}_j - \hat{\mathbf{z}}_i) (\mathbf{z}_j - \hat{\mathbf{z}}_i)^T \quad (7.110)$$

$$\mathbf{G}_b = \frac{1}{K} \sum_{i=1}^{N_c} K_{ci} (\hat{\mathbf{z}}_i - \hat{\mathbf{z}}) (\hat{\mathbf{z}}_i - \hat{\mathbf{z}})^T \quad (7.111)$$

W literaturze opisano również wiele wariantów metody LDA, m.in. zaproponowano metodę LDA dostosowaną do przetwarzania sygnałów. W pracy [321] zaproponowano przyrostową metodę LDA. Również w pracy [79] zaproponowano podejście przyrostowe dla metody LDA.

Przykład 7.13 (Liniowa analiza dyskryminacyjna). *Przykład działania metody LDA przedstawiono na rys. 7.52. Na rysunku 7.52a przedstawiono zbiór danych wejściowych (z_1 oraz z_2), z kolei na rys. 7.52b wynik przekształcenia zbiorów wejściowych.*



Rys. 7.52. Ilustracja działania Liniowej Analizy Dyskryminacyjnej:
 a) wizualizacja zbioru danych pomiarowych,
 b) dekompozycja z wykorzystaniem metody LDA

Korelacja kanoniczna

Metoda analizy korelacji kanonicznych (ang. *Canonical Correlation Analysis*, CCA) bazuje na poszukiwaniu zależności pomiędzy dwoma wektorami. Zmienne, które służą do opisu tych zależności nazywa się kanonicznymi. Wyznaczane są one jako sumy ważone elementów pierwszego i drugiego zbioru. Wagi w tych sumach ważonych nazywa się z kolei wagami kanonicznymi. Celem metody jest wyznaczenie takich wag, aby uzyskać możliwie największą korelację pomiędzy wektorami. Wartości wag mają prostą interpretację, którą można wykorzystać w konstrukcji wektora cech, tj. im większa jest bezwzględna wartość wagi kanonicznej, tym większy wkład danego elementu wektora do zmiennej kanonicznej. W celu łatwiejszego porównania zmiennych są one standaryzowane. Zauważmy, że liczba zmiennych kanonicznych jest równa liczbie zmiennych tego wektora, w którym jest ich najmniej.

Prezentację metody rozpoczniemy od skonstruowania dwóch macierzy na podstawie K pomiarów. Pierwszy z wektorów ma postać:

$$\mathbf{z}_1(k) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{d_{z_1}} \end{bmatrix} \quad (7.112)$$

a drugi:

$$\mathbf{z}_2(k) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{d_{z_2}} \end{bmatrix} \quad (7.113)$$

gdzie $k = 1, 2, \dots, K$.

Na podstawie wektorów z pomiarami (7.112) oraz (7.113) konstruujemy dwie macierze, odpowiednio dla pierwszego wektora

$$\mathbf{Z}_1 = \begin{bmatrix} z_{1,1} & z_{1,2} & \dots & z_{1,K} \\ z_{2,1} & z_{2,2} & \dots & z_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ z_{d_{z_1},1} & z_{d_{z_1},2} & \dots & z_{d_{z_1},K} \end{bmatrix} \quad (7.114)$$

oraz drugiego:

$$\mathbf{Z}_2 = \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,K} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathcal{D}_2,1} & z_{\mathcal{D}_2,2} & \cdots & z_{\mathcal{D}_2,K} \end{bmatrix} \quad (7.115)$$

gdzie $\mathbf{Z}_1 \in \mathbb{R}^{K \times \mathcal{D}_{z1}}$, \mathcal{D}_{z1} jest wymiarem pierwszego wektora cech, a \mathcal{D}_{z2} reprezentuje wymiar drugiego wektora cech, z kolei $\mathbf{Z}_2 \in \mathbb{R}^{K \times \mathcal{D}_{z2}}$. K jest liczbą wykonanych pomiarów.

W metodzie CCA należy oszacować wartości dwóch wektorów, tj. \mathbf{u}_{CCA} oraz \mathbf{v}_{CCA} , które będą maksymalizowały następujący wskaźnik jakości [75, 106]:

$$Q(\mathbf{u}_{CCA}, \mathbf{v}_{CCA}) = \frac{\mathbf{u}_{CCA}^T \mathbf{S}_{z_1, z_2} \mathbf{v}_{CCA}}{\sqrt{(\mathbf{u}_{CCA}^T \mathbf{S}_{z_1, z_1} \mathbf{u}_{CCA}) (\mathbf{v}_{CCA}^T \mathbf{S}_{z_2, z_2} \mathbf{v}_{CCA})}} \quad (7.116)$$

Macierze w zależności (7.116), tj. \mathbf{S}_{z_1, z_2} , \mathbf{S}_{z_1, z_1} oraz \mathbf{S}_{z_2, z_2} są macierzami kowariancji, wyznaczanymi z zależności:

$$\mathbf{S}_{z_1, z_2} = \frac{1}{K} \mathbf{Z}_1^T \mathbf{Z}_2 \quad \mathbf{S}_{z_1, z_1} = \frac{1}{K} \mathbf{Z}_1^T \mathbf{Z}_1 \quad \mathbf{S}_{z_2, z_2} = \frac{1}{K} \mathbf{Z}_2^T \mathbf{Z}_2 \quad (7.117)$$

W celu maksymalizacji wskaźnika (7.116) formułujemy zadanie optymalizacji do rozwiązania:

$$\begin{aligned} (\mathbf{u}_{CCA}^*, \mathbf{v}_{CCA}^*) = \max_x & \quad \mathbf{u}_{CCA}^T \mathbf{S}_{z_1, z_2} \mathbf{v}_{CCA} \\ \text{p.o.} & \quad \mathbf{u}_{CCA}^T \mathbf{S}_{z_1, z_1} \mathbf{u}_{CCA} = 1 \\ & \quad \mathbf{v}_{CCA}^T \mathbf{S}_{z_2, z_2} \mathbf{v}_{CCA} = 1 \end{aligned} \quad (7.118)$$

Metoda widma osobliwego

Metoda widma osobliwego (ang. *Singular Spectrum Analysis*, SSA) jest nieparametrycznym podejściem do analizy spektralnej sygnałów pomiarowych [58, 260, 432, 454]. W metodzie tej dokonuje się dekompozycji sygnału pomiarowego na trzy składowe. Jedna z tych składowych reprezentuje trend, druga składowa jest składową *quasi*-periodyczną, a ostatnia z nich to szum. Zaletą metody SSA jest to, że w wyniku dekompozycji sygnału uzyskuje się składowe, które mają swoją interpretację fizyczną. Niech

$$\mathbf{z}_{\mathcal{D}_z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\mathcal{D}_{z1}} \end{bmatrix} \quad (7.119)$$

7. Poziom średni: fuzja właściwości

będzie \mathcal{D}_z -wymiarowym wektorem cech. W kolejnym kroku należy dokonać mapowania (ang. *mapping*) tego wektora na macierz \mathbf{H}_{SSA} [222, 232]:

$$\mathbf{H}_{\text{SSA}} = \begin{bmatrix} z(1) & z(2) & \dots & z(N_{\text{sg}}) \\ z(2) & z(3) & \dots & z(N_{\text{sg}} + 1) \\ \vdots & \vdots & \vdots & \vdots \\ z(L_{\text{frm}}) & z(L_{\text{frm}} + 1) & \dots & z(N_{\text{sg}}) \end{bmatrix} \quad (7.120)$$

gdzie $N_{\text{sg}} = \mathcal{D}_z - L_{\text{frm}} + 1$, a L_{frm} jest zmienną reprezentującą długość okna, dla której zachodzi $L_{\text{frm}} \leq \mathcal{D}_z/2$.

Macierz \mathbf{H}_{SSA} to macierz Hankela (trajektorii) [18, 232]. W kolejnym kroku dekomponuje się macierz \mathbf{H}_{SSA} , korzystając z metody SVD:

$$\mathbf{H}_{\text{SSA}} = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{V}_s^T = \sum_{i=1}^Q \sigma_{si} \mathbf{u}_{si} \mathbf{v}_{si}^T \quad (7.121)$$

gdzie $Q = \text{rank}(\mathbf{H}_{\text{SSA}}) \leq L_{\text{frm}}$, macierze \mathbf{U}_s i \mathbf{V}_s są macierzami ortonormalnymi. Z kolei elementy \mathbf{u}_{si} (wektory własne) oraz \mathbf{v}_{si} (wartości własne) są składowymi macierzy odpowiednio \mathbf{U}_s i \mathbf{V}_s . Natomiast macierz $\mathbf{\Sigma}_s$ jest macierzą diagonalną, której elementy leżące na diagonalii σ_{si} są wartościami osobliwymi macierzy \mathbf{H}_{SSA} . Składową główną w metodzie SSA wyznaczamy z zależności:

$$\mathbf{w}_{si} = \mathbf{H}_{\text{SSA}}^T \mathbf{u}_{si} \quad (7.122)$$

dla $i = 1, 2, \dots, Q$, a \mathbf{w}_{si} jest wektorem o wymiarze $N_{\text{sg}} \times 1$. Poszczególne elementy wektora \mathbf{w}_{si} wyznaczane są z następującej zależności:

$$\mathbf{w}_{si,j} = \sum_{l=1}^{L_{\text{frm}}-1} u_{si,j} z_{l+j}, \quad j = 0, 1, \dots, N_{\text{sg}} \quad (7.123)$$

W następnym kroku przeprowadzana jest operacja wydzielenia struktur, której celem jest analiza widma wartości własnych. Efektem tej analizy jest macierz trajektorii wyrażona w formie:

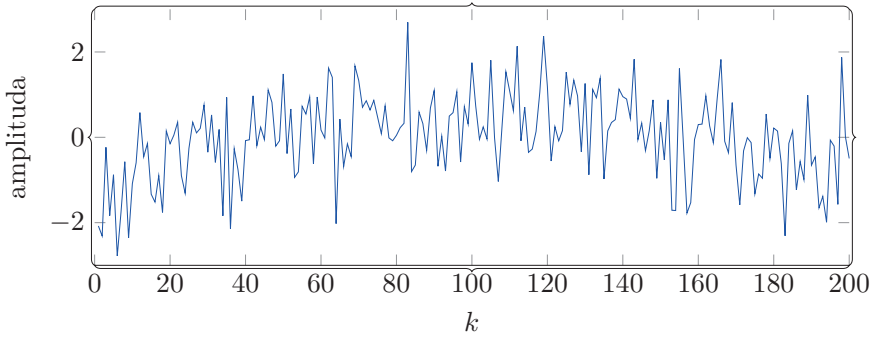
$$\mathbf{H}_{\text{SSA}} = \sum_{j=1}^{N_{\text{ds}}} \mathbf{H}_{\text{SSA}} \mathcal{I}_j \quad (7.124)$$

gdzie \mathcal{I}_j reprezentuje zbiór indeksów, N_{ds} to liczba podzbiorów, na który podzielono zbiór indeksów $1, 2, \dots, Q$.

Ostatnim etapem metody widma osobliwego jest rekonstrukcja sygnału z wykorzystaniem następujących formuł:

$$\tilde{\mathbf{z}}_{si}(k) = \begin{cases} \frac{1}{k} \sum_{j=1}^k u_{si,j} w_{si,k-i} & \text{dla, } 1 \leq k < L_{frm} \\ \frac{1}{L_{frm}} \sum_{j=1}^{L_{frm}-1} u_{si,j} w_{si,k-i} & \text{dla, } L_{frm} \leq k \leq N_{sg} \\ \frac{1}{K-k} \sum_{i=k-K+1}^{N_{sg}-N} u_{si,j} w_{si,k-i} & \text{dla, } N_{sg} + 1 \leq n \leq K \end{cases} \quad (7.125)$$

Przykład 7.14 (Metoda widma osobliwego). *Przykład działania metody widma osobliwego przedstawiono na rys. 7.53–7.56. Na rysunku 7.53 przedstawiono sygnał pomiarowy wykorzystany w przykładzie, a na rys. 7.54 i 7.55 odpowiednio komponenty, na które został rozłożony sygnał (rys. 7.53) w trakcie dekompozycji oraz komponenty po ich odtworzeniu (estymacji) (rys. 7.55). Na rysunku 7.56 przedstawiono wynik dekompozycji sygnału na podstawie wszystkich składowych (górna część rys. 7.56) oraz dla pierwszych dwóch składowych (dolna część rys. 7.56).*

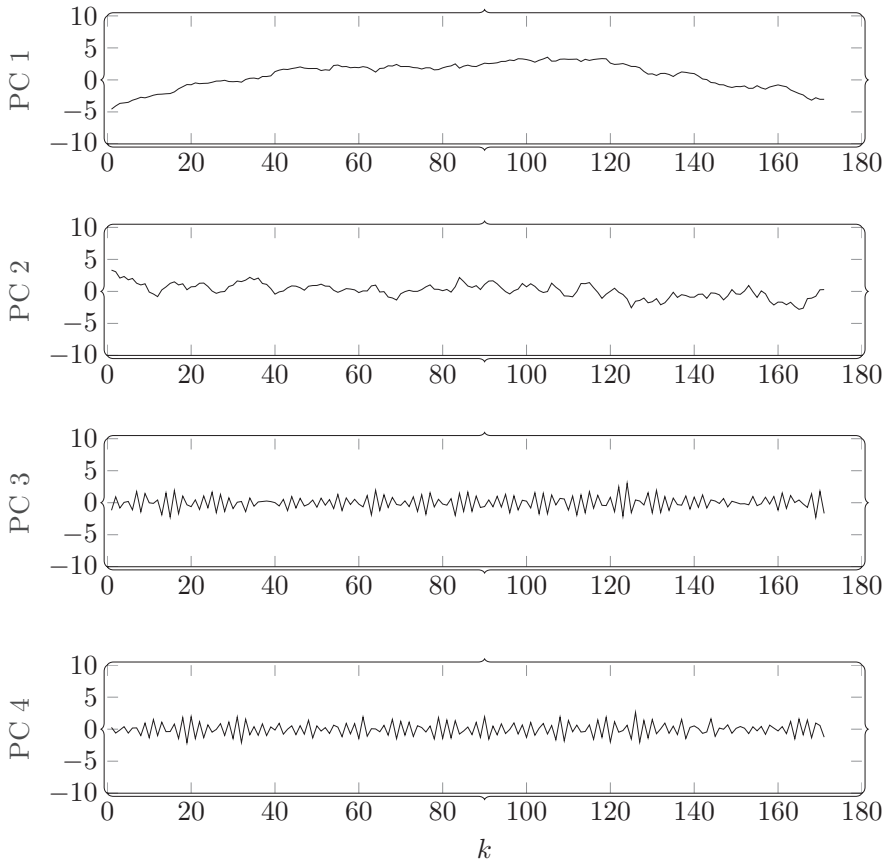


Rys. 7.53. Przykładowy sygnał pomiarowy

Autoenkoder

Autoenkoder (ang. *Autoencoder*, AE) jest siecią neuronową, w której wyróżnia się warstwy neuronów wejściowych, wyjściowych oraz więcej niż jedną warstwę ukrytą. Warstwy ukryte zawierają neurony, których liczba jest ustalana na etapie uczenia. Z reguły liczba neuronów w warstwach ukrytych powinna być mniejsza niż w warstwie wejściowej, ale na tyle duża, by umożliwić odtworzenie danych wejściowych na wyjściu autoenkodera. Zaznaczmy, że cechą charakterystyczną autoenkoderów jest to, że do ich uczenia wykorzystuje się algorytmy uczenia nienadzorowanego [16, 196]. Opisanie właściwości autoenkodera znalazły zastosowanie m.in. w redukcji cech.

7. Poziom średni: fuzja właściwości

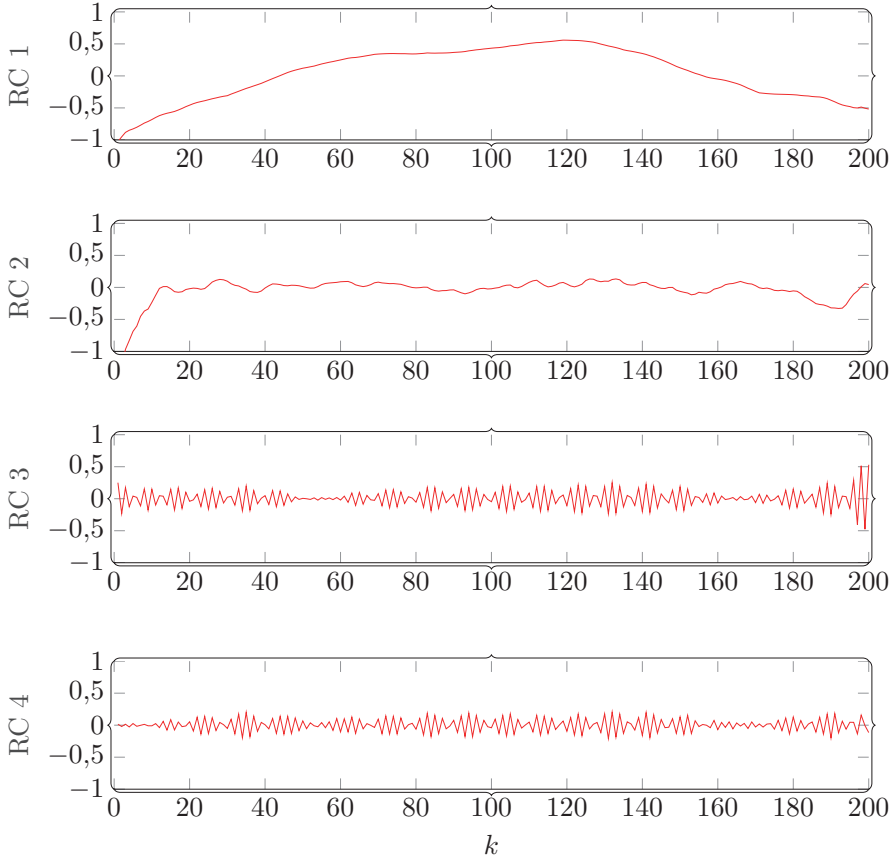


Rys. 7.54. Składowe zdekomponowane algorytmem SSA

Rozpocznijmy od opisu danych wejściowych. Niech

$$\mathbf{z}_{\mathcal{D}_z}(k) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\mathcal{D}_z} \end{bmatrix} \quad (7.126)$$

będzie \mathcal{D}_z -wymiarowym wektorem wejściowym, którego wymiar ma zostać zredukowany. Wykonując K pomiarów, zadaniem (tak jak poprzednio) jest redukcja wymiaru wektora $\mathbf{z} \in \mathbb{R}^{\mathcal{D}_z}$ do wymiaru wektora $\mathbf{x} \in \mathbb{R}^{\mathcal{D}_x}$, gdzie $\mathcal{D}_x < \mathcal{D}_z$. Przedstawimy teraz koncepcję redukcji wymiaru z wykorzystaniem *autoenkodera*.



Rys. 7.55. Składowe zrekonstruowane (algorytm SSA)

W ogólności autoenkoder składa się z dwóch części, tj. kodującej i dekodującej, co zapisujemy w następującej postaci:

$$\mathbf{x} = \mathcal{G}_{\text{kod}}(\mathbf{W}_{\text{aeh}}\mathbf{z} + \mathbf{b}_h) \quad (7.127a)$$

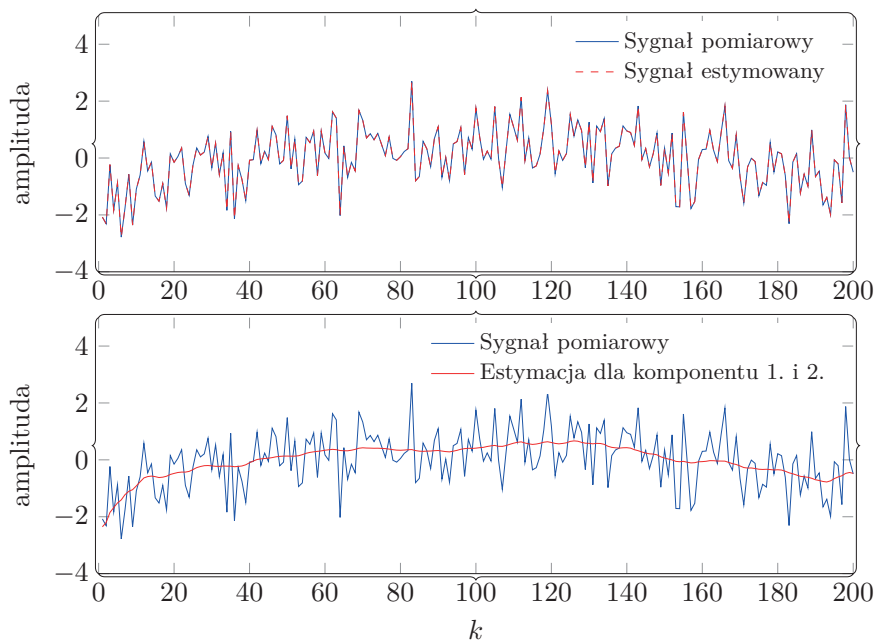
$$\tilde{\mathbf{z}} = \mathcal{G}_{\text{dek}}(\mathbf{W}_{\text{aeo}}\mathbf{x} + \mathbf{b}_o) \quad (7.127b)$$

gdzie \mathbf{W}_{aeh} jest macierzą wag dla warstw ukrytych, \mathbf{W}_{aeo} jest macierzą wag dla warstwy wyjściowej. Z kolei \mathcal{G}_{kod} to funkcja kodująca, a \mathcal{G}_{dek} reprezentuje funkcję dekodującą. Zależności (7.127a) oraz (7.127b), możemy zapisać również jako:

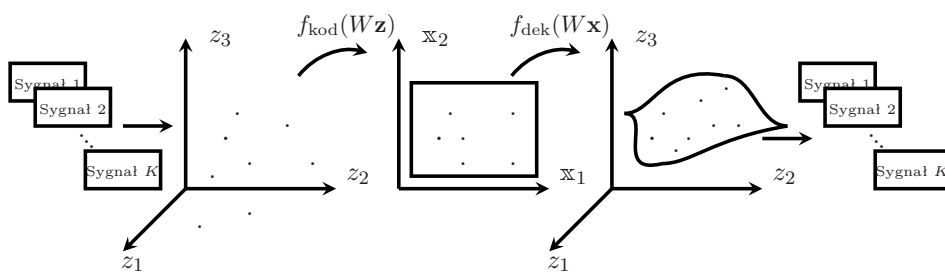
$$\tilde{\mathbf{z}} = \mathcal{G}_{\text{dek}}(\mathbf{W}_{\text{aeo}}\mathcal{G}_{\text{kod}}(\mathbf{W}_{\text{aeh}}\mathbf{z} + \mathbf{b}_h) + \mathbf{b}_o) \quad (7.128)$$

Działanie autoenkodera zilustrowano na rys. 7.57.

7. Poziom średni: fuzja właściwości



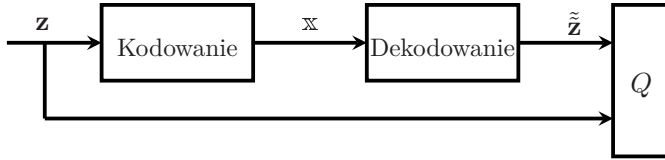
Rys. 7.56. Sygnał zrekonstruowany (algorytm SSA)



Rys. 7.57. Ilustracja redukcji wymiaru z wykorzystaniem autoenkodera (na podstawie [137])

Parametry autoenkodera (7.128) można estymować rozwiązując zadanie optymalizacji. Przyjmijmy dla tego zadania następujący wskaźnik jakości:

$$Q = \frac{1}{2} \sum_{k=1}^K \|\mathbf{z}(k) - \tilde{\mathbf{z}}(k)\|^2 = \frac{1}{2} \sum_{k=1}^K \|\mathbf{z}(k) - \mathcal{G}_{\text{dek}}(\mathbf{W}_{\text{aéo}} \mathcal{G}_{\text{kod}}(\mathbf{W}_{\text{aéh}} \mathbf{z} + \mathbf{b}_h) + \mathbf{b}_o)\|^2 \quad (7.129)$$

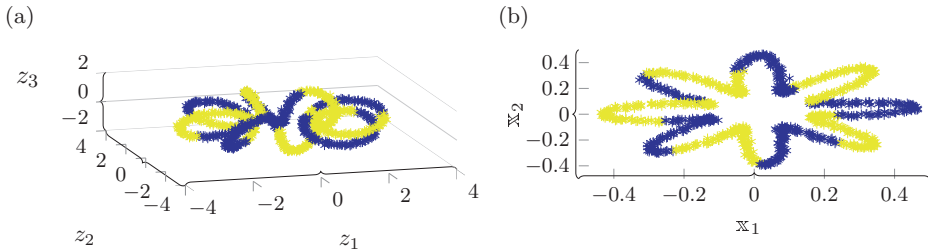


Rys. 7.58. Wykorzystanie autoenkodera w zadaniu redukcji wymiaru

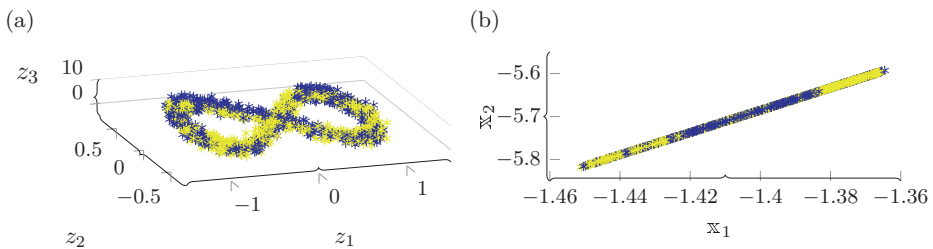
Koncepcja redukcji wymiaru została przedstawiona na rys. 7.58.

W pracy [315] omówiono metody wykorzystania autoenkoderów do fuzji właściwości w zadaniu śledzenia aktywności fizycznej człowieka. Podobne zastosowanie przedstawiono w pracach [67] oraz [259].

Przykład 7.15 (Autoenkoder w zadaniu redukcji cech). *Przykład działania autoenkodera przedstawiono na rys. 7.59 oraz 7.60. Wykorzystane w badaniach zbiory przedstawiono na rys. 7.59a i 7.60a, z kolei przekształcenia tych zbiorów z wykorzystaniem autoenkodera na rys. 7.59b i 7.60b.*



Rys. 7.59. Ilustracja działania autoenkodera dla pierwszego zbioru danych (a) i wynik transformacji (b)



Rys. 7.60. Ilustracja działania autoenkodera dla drugiego zbioru danych (a) i wynik transformacji (b)

7.7. Klasyfikacja

Kolejnym krokiem, po przygotowaniu wektora cech opisującego analizowany obiekt, jest fuzja tych cech. Typowymi narzędziami fuzji stosowanymi na tym poziomie są algorytmy klasyfikacji. Celem stosowania algorytmów klasyfikacji na drugim poziomie fuzji danych jest wyciąganie wniosków (konkludowanie) dotyczących badanego obiektu na podstawie wyselekcjonowanych cech, które opisują ten obiekt (w formie wektora cech). Rozróżniamy dwa podejścia do rozwiązania zadania fuzji właściwości. Jedno z nich to podejście *periodyczne*, a drugie to podejście *nieperiodyczne*. Klasyfikacja *periodyczna* nazywana jest **klasyfikacją na żądanie** (ang. *on demand*) i wykonywana jest jedynie wówczas, gdy wystąpiły warunki do ponownego rozwiązania zadania klasyfikacji. Może być to związane z pojawieniem się nowych danych, które wymuszają powtórne wykonanie klasyfikacji. Innym powodem może być pojawienie się istotnych zmian w przetwarzanych zbiorach danych. Cechą charakterystyczną klasyfikacji *nieperiodycznej* jest brak ustalonej częstotliwości wykonywania tego zadania.

Alternatywą jest **klasyfikacja periodyczna**, która wykonywana jest w ustalonych momentach czasu. Zaletą tego podejścia jest to, że możliwe jest z góry ustalenie momentu kolejnych obliczeń związanych z rozwiązaniem zadania klasyfikacji. Jest to istotne, zwłaszcza w przypadku systemów o ograniczonych zasobach, np. obliczeniowych. Wadą tego rozwiązania jest natomiast to, że zasoby projektowanego systemu będą wykorzystywane nawet wówczas, gdy czynności związane z klasyfikacją nie są uzasadnione (np. brak nowych danych w systemie lub ich istotna zmiana).

Konieczność stosowania klasyfikacji *nieperiodycznej* (na żądanie) może wiązać się z cechami rzeczywistych sygnałów pomiarowych albo ze zjawiskami, które mogą w nich wystąpić. Pierwszy przypadek dotyczy sygnałów *niestacjonarnych*, a drugi związany jest ze zjawiskiem **dryftu modelu** (ang. *concept drift*).

Zagadnienie *niestacjonarności* w danych pomiarowych zostało omówione w rozdziale 4. Natomiast zjawisko *dryftu modelu* związane jest ze zmianami w czasie właściwości klasy, które były brane pod uwagę podczas budowy klasyfikatora. Zjawisko to wpływa niekorzystnie na jakość klasyfikacji, której jakość w długim horyzoncie czasowym może ulegać pogorszeniu. Zarówno *niestacjonarność*, jak i *dryft modelu*, może być wykrywana i uwzględniana zarówno w algorytmie klasyfikacji, jak i w samym systemie fuzji danych.

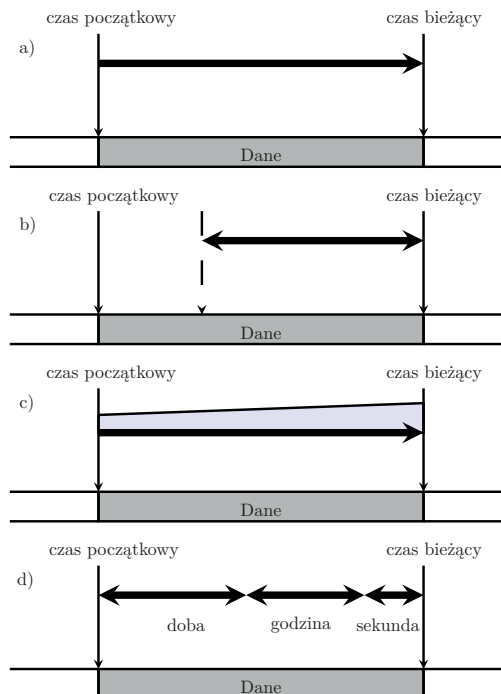
Decyzja o tym, czy w projektowanym systemie zadanie klasyfikacji będzie rozwiązywane *periodycznie* czy *nieperiodycznie*, zarówno w systemach o architekturze centralnej, jak i rozproszonej, może wiązać się z tym, czy system ma pracować w *reżimie czasu rzeczywistego*. Wiele zagadnień praktycznych wymaga przetwarzania danych w czasie rzeczywistym, np. system ciągłego monitorowania pacjentów. W takich przypadkach należy liczyć się z:

- **ograniczonymi zasobami pamięci**; jest to szczególnie istotne zagadnienie, gdy wykorzystywanych czujników pomiarowych w systemie jest wiele, a częstotliwość ich prac oraz transferu danych jest duża,
- **ograniczonymi zasobami obliczeniowymi**; problem może być szczególnie istotny, gdy dostarczone do jednostki obliczeniowej dane wymagają dużego nakładu obliczeń.

7.7.1. Okno czasowe

Częstotliwość klasyfikacji, która związana jest ze specyfiką rozwiązywanego problemu lub przetwarzanych danych, ma również wpływ na metodę przetwarzania gromadzonych danych oraz typ okna czasowego. Na potrzeby przetwarzania danych wykorzystywane są następujące typy okien czasowych:

- okno z punktem początkowym (ang. *landmark window*) rys. 7.61a,
- okno przesuwne (ang. *sliding window*) rys. 7.61b,
- okno ważone (ang. *fading window*) rys. 7.61c,



Rys. 7.61. Typy okien: a) okno z punktem początkowym, b) okno przesuwne, c) okno ważne i d) okno o zmiennej skali czasowej

7. Poziom średni: fuzja właściwości

- okno o zmiennej skali czasowej (ang. *tilted time window*) rys. 7.61d.

Okno z punktem początkowym (rys. 7.61a) charakteryzuje się tym, że dane są gromadzone raz i wykorzystywane np. podczas konstrukcji klasyfikatora. Wszystkie dane, które zostały zgromadzone od punktu początkowego do bieżącego, są tak samo ważne i nie są rozróżniane podczas ich przetwarzania na potrzeby klasyfikacji. Rozwiązanie to jest niepraktyczne, zwłaszcza w przypadku gdy w danych pomiarowych występuje zjawisko dryftu modelu lub gdy dane są niestacjonarne.

W przypadku danych, których charakterystyka może zmieniać się w czasie, można wykorzystać okno przesuwne (rys. 7.61b). W tym przypadku przetwarzaniu podlega zbiór ostatnio zebranych danych, liczność tego zbioru określona jest przez szerokość okna, na którym wykonywane są obliczenia. Rozróżniamy dwa typy okien przesuwnych, tj. okno o stałej oraz zmiennej szerokości.

Kolejnym typem jest okno ważone (rys. 7.61c). Podobnie jak okno z punktem początkowym, w tym przypadku gromadzone są wszystkie dane. Różnica polega na tym, że przetwarzanym danym nadawana jest waga. Wartość wagi uzależniona jest od czasu, jaki upłynął od momentu pozyskania danych. Dane, które pojawiły się dawno, mają niższą wagę od danych, które pojawiły się później. Informacja o nadanych wagach wykorzystywana jest np. w procesie budowy klasyfikatora.

W przetwarzaniu z oknem o zmiennej skali czasowej dane są grupowane w zbiory o różnych skalach czasowych. Okno podzielone jest na ramki o różnych rozdzielczościach czasowych. W ramach o wysokiej rozdzielczości czasowej zbierane są dane bieżące, w pozostałych o niższej rozdzielczości – gromadzone są starsze dane. Na rys. 7.61d podano przykład okna czasowego o zmiennej skali czasowej, gdzie w oknie gromadzone są dane z ostatniej sekundy, z ostatniej godziny i z ostatniej doby.

7.7.2. Metody przetwarzania danych

Oprócz ustalenia typu okna czasowego, w systemach fuzji decyzji wyróżniamy dwa główne podejścia do przetwarzania danych:

- przyrostowe (ang. *incremental learning*),
- blokowe (ang. *block base learning*).

Przetwarzanie przyrostowe

Przetwarzanie przyrostowe charakteryzuje się adaptacyjną modyfikacją modelu klasyfikatora, która wykonywana jest sukcesywnie przy pojawieniu się nowych danych. Przetwarzanie przyrostowe zilustrowano na rys. 7.62.

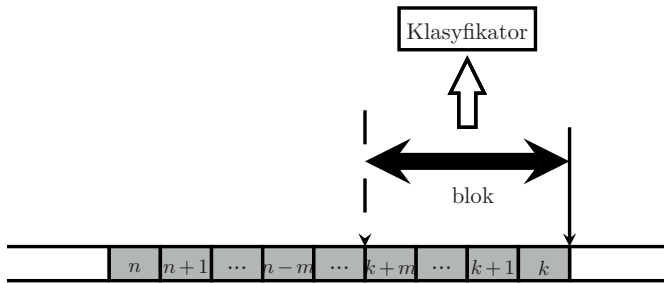
Przetwarzanie przyrostowe jest wymagające obliczeniowo i dlatego rekomendowane jest dla przypadku, gdy nowe dane pojawiają się wraz z informacją o typie klasy, do której należą. Zaletą podejścia jest to, że wynik klasyfikacji dostępny jest natychmiast.



Rys. 7.62. Ilustracja przetwarzania przyrostowego

Przetwarzanie blokowe

Alternatywą dla przetwarzania przyrostowego jest przetwarzanie blokowe. W tym przypadku dane przetwarzane są w blokach danych. Rozmiar bloku danych definiowany jest przez okna czasowe. Na rysunku 7.63 zilustrowano, w jaki sposób wykonywane jest przetwarzanie blokowe.



Rys. 7.63. Ilustracja przetwarzania blokowego

Ponieważ dane w podejściu blokowym nie są przetwarzane na bieżąco, lecz z pewnym opóźnieniem, nie jest to rozwiązanie tak wymagające obliczeniowo, jak omówione poprzednio przetwarzanie przyrostowe. Wadą podejścia jest to, że wynik klasyfikacji nie jest dostępny natychmiast, lecz z pewnym opóźnieniem.

Pewnym wariantem przetwarzania blokowego jest podejście, w którym wykorzystywana jest detekcja zmian. Wówczas rozmiar bloku zależy od wykrytych momentów istotnych zmian w przetwarzanych danych.

Podejście blokowe rekomendowane jest w przypadku, gdy dane nie mają etykiety z klasą. Wówczas podejście blokowe umożliwia wykonanie tej operacji w czasie, gdy dane oczekują w kolejce na ich przetworzenie.

7.7.3. Zbiory danych

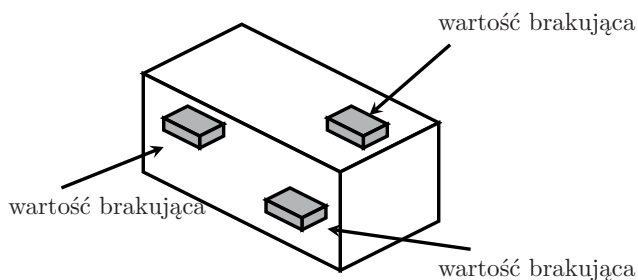
Wybór klasyfikatora dla zadania fuzji właściwości należy powiązać z analizą zawartości zbiorów danych, które mają zostać wykorzystane w zadaniu konstrukcji

7. Poziom średni: fuzja właściwości

modelu na potrzeby klasyfikacji. Dwa główne zagadnienia, które są rozpatrywane w tym kontekście to niekompletność danych (ang. *incomplete data*, *missing data*) oraz brak ich zbalansowania (ang. *imbalanced data*).

Dane niekompletne

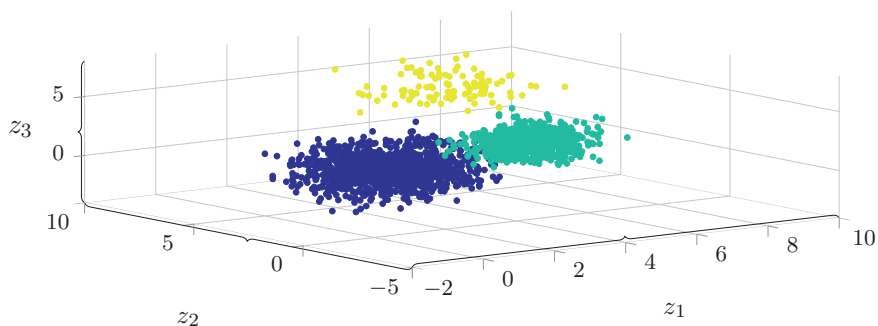
Problem danych niekompletnych (rys. 7.64), w przypadku systemów fuzji danych z czujników pomiarowych, może być związany z awarią czujnika lub zespołu czujników pomiarowych. Inną przyczyną mogą być problemy związane z transmisją danych, np. przekłamania podczas transmisji danych.



Rys. 7.64. Zbiór danych niekompletnych

Dane niezbalansowane

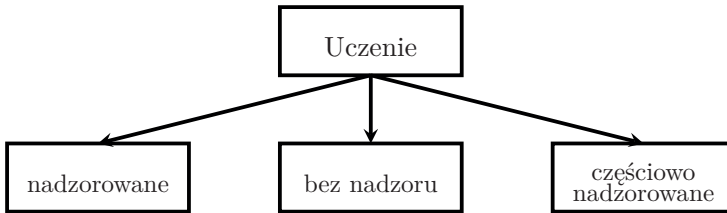
Zbiory uczące zawierające dane niezbalansowane charakteryzują się tym, że część klas obiektów w nich zawarta jest nadreprezentowana w zbiorze. Nadreprezentacja danych z pewnej klasy może wynikać ze specyfiki problemu, którego zbiór dotyczy. Przykład zbioru niezbalansowanego przedstawiono na rys. 7.65.



Rys. 7.65. Przykład zbioru danych niezbalansowanych

7.7.4. Metody uczenia

Konstrukcja klasyfikatora składa się z dwóch etapów, tj. **uczenia** i **testowania**. Na etapie uczenia budowany jest odpowiedni model dla algorytmu klasyfikacji z wykorzystaniem danych uczących. Celem drugiego z wymienionych etapów jest weryfikacja algorytmu klasyfikacji. Wyróżniamy trzy główne podejścia do uczenia modeli, tj. **uczenie nadzorowane** (ang. *supervised learning*), **uczenie bez nadzoru** (ang. *unsupervised learning*) i **uczenie częściowo nadzorowane** (ang. *semi-supervised learning*) (rys. 7.66).



Rys. 7.66. Klasyfikacja metod uczenia

Uczenie nadzorowane

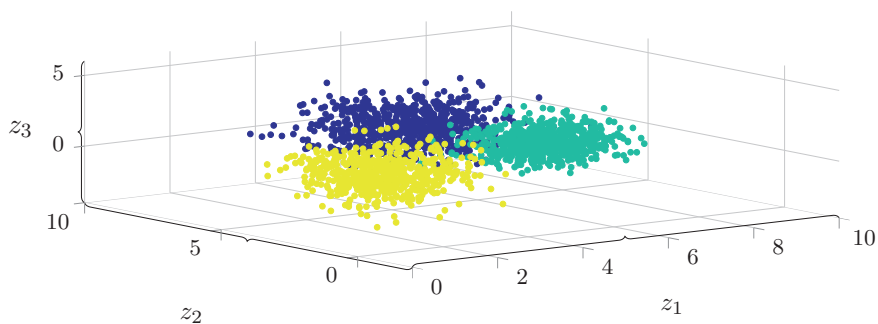
W przypadku uczenia nadzorowanego przyjmujemy, że dysponujemy zbiorem danych, w którym każdy element ma przyporządkowaną klasę. Typową reprezentacją elementu w zbiorze uczącym jest wektor cech $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_{d_x}]^T$. Natomiast do opisu elementów zbioru wykorzystuje się klasy z następującego N_c -elementowego zbioru, tj. $c \in \{1, 2, \dots, N_c\} = \mathcal{N}_c$. Zdefiniujemy jeszcze zbiór uczący jako $TS_{N_{le}} = \{(\mathbf{x}_i, c_i)\}$ dla $i = 1, 2, \dots, N_{le}$ (rys. 7.67), gdzie przez N_{le} oznaczamy liczbę przykładów uczących w zbiorze.

Podejście nadzorowane do uczenia klasyfikatorów, w przypadku fuzji danych strumieniowych, rekomendowane jest wówczas, gdy możliwe jest nadanie etykiety każdemu nowemu pomiarowi.

Uczenie bez nadzoru

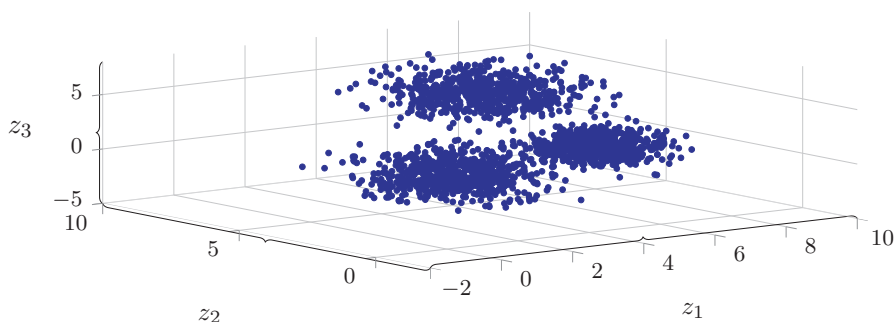
Cechą charakterystyczną uczenia bez nadzoru jest to, że nie dysponujemy typowym zbiorem uczącym, tak jak to jest w przypadku uczenia z nadzorem. Różnica w budowie zbioru uczącego polega na tym, że elementy tego zbioru nie mają przyporządkowanych klas, tj. $TN_{N_{le}} = \{(\mathbf{x}_i)\}$ dla $i = 1, 2, \dots, N_{le}$ (rys. 7.68). Jedną z technik wykorzystywanych w uczeniu nienadzorowanym jest grupowanie. Grupowanie to technika uczenia, która pozwala na grupowanie podobnych

7. Poziom średni: fuzja właściwości



Rys. 7.67. Przykład zbioru z elementami przyporządkowanymi do klas

elementów ze zbioru uczącego w grupy (tzw. klastry). Każdy klastrow powinien charakteryzować się tym, że podobieństwo elementów w nim zgromadzonych jest możliwie największe, a podobieństwo pomiędzy elementami w innych klastrach jest możliwie najmniejsze.



Rys. 7.68. Przykład zbioru z elementami bez przyporządkowania do klas

W zadaniu klasyfikacji uczenie bez nadzoru jest wykorzystywane wówczas, gdy nie ma możliwości nadawania etykiet nowym elementom w zbiorze uczącym. W przypadku, gdy nadanie etykiet możliwe jest jedynie dla części z elementów w zbiorze uczącym, wówczas zalecanym podejściem do uczenia jest uczenie **częściowo nadzorowane**.

Uczenie częściowo nadzorowane

Uczenie częściowo nadzorowane w zagadnieniu klasyfikacji danych w strumieniach wykorzystywane jest wówczas, gdy w zbiorze uczącym znajdują się dane, z których część ma przyporządkowane klasy, a część z nich nie ma takie-

go przyporządkowania. W tym przypadku zbiór uczący ma postać: $TH_{N_{le}} = \{(\mathbf{x}_1, \mathbf{c}_1), (\mathbf{x}_2, \mathbf{c}_2), \dots, (\mathbf{x}_{N_{le1}}, \mathbf{c}_{N_{le1}}), \mathbf{x}_{N_{le1}+1}, \mathbf{x}_{N_{le1}+2}, \dots, \mathbf{x}_{N_{le1}+N_{le2}}\}$, gdzie N_{le1} reprezentuje liczbę próbek z przyporządkowanymi klasami, a N_{le2} to liczba próbek bez przyporządkowania do klasy.

7.8. Algorytmy klasyfikacji

Omówimy teraz podstawowe zagadnienia związane z klasyfikacją danych strumieniowych, tj. sformułujemy problem klasyfikacji, omówimy istotne zagadnienia związane z tym zadaniem oraz dokonamy przeglądu reprezentatywnych algorytmów służących do jego rozwiązania.

Budowa klasyfikatora jest zwykle zadaniem dwuetapowym. Na pierwszym etapie, nazywanym uczeniem, buduje się klasyfikator, korzystając z wcześniej przygotowanego zbioru uczącego. Drugi etap polega na testowaniu klasyfikatora z wykorzystaniem wcześniej przygotowanego zbioru testowego.

Drugie z wymienionych zadań, tj. testowanie klasyfikatorów, w przypadku strumieni danych, nie jest zadaniem trudnym. Nie wymaga również specyficznych metod, które różniłyby się od metod stosowanych w przypadku danych niebędących strumieniami. Jest to podyktowane tym, że testy wykonywane są sekwencyjnie dla każdego zbioru walidacyjnego.

Specyfikę strumieni danych należy natomiast uwzględnić na etapie uczenia klasyfikatora. Jest to związane z metodą jego przeprowadzania, która polega na wykorzystaniu wielu przykładów uczących, które pochodzą z różnych chwil czasowych. Ze względu na kolejność pojawiania się danych w różnych chwilach czasu, ich właściwości (np. statystyczne) mogą się od siebie różnić. Zjawiskiem tym jest wspomniany już dryft modelu. Warto zaznaczyć, że zjawisko to ma niekorzystny wpływ na jakość klasyfikatora i może wpłynąć na proces jego uczenia. Często jest to również związane ze zwiększonym zapotrzebowaniem na moc obliczeniową.

W pracy [4] zwrócono uwagę na kilka innych zagadnień, które należy rozważyć podczas uczenia i/lub testowania algorytmów klasyfikacji. Jednym z nich jest problem zarządzania zasobami wykorzystywanymi do uczenia i testowania algorytmów klasyfikacji. Jest to spowodowane tym, że konieczność wykonania jednego z dwóch podstawowych zadań w procesie konstrukcji klasyfikatora może pojawić się w różnych momentach. Warto zaznaczyć, że w większości przypadków brak jest możliwości wpływania na częstotliwość występowania żądań przeprowadzenia aktualizacji lub budowy modelu.

Innym problemem jest mała częstość występowania klas w zbiorze uczącym. Może to prowadzić do obniżenia jakości klasyfikacji. Dodatkowo klasy, które występują rzadko w przykładach uczących, są znacznie bardziej istotne niż te, które

7. Poziom średni: fuzja właściwości

występują w nich często. Mogą one dotyczyć istotnych zdarzeń, takich jak np. awarie, wypadki itp.

Kolejny problem, który należy uwzględnić w projektowaniu klasyfikatora, związany jest z dostępnością przykładów uczących. Dostępność może być związana np. z miejscem. Prowadzi to do kolejnego problemu, jakim jest udostępnianie danych z jednej lokalizacji do drugiej. W tym przypadku może okazać się konieczne zaprojektowanie rozproszonego systemu konstrukcji algorytmów klasyfikacji.

7.8.1. Algorytm k -Najbliższych Sąsiadów

Metoda k -Najbliższych Sąsiadów (ang. *k-Nearest Neighbors*, k -NN) jest popularną nieparametryczną metodą klasyfikacji [87, 392, 168]. Idea metody polega na przyporządkowaniu nowego obiektu do tej klasy, do której należy k najbliższych jej sąsiadów. W tym celu należy ustalić miarę, która będzie służyła do oceny stopnia podobieństwa nowego obiektu do innych obiektów, które już są w zbiorze uczącym. W przypadku metody k -NN sprowadza się to do wyboru miary odległości w przestrzeni obserwacji. Najpopularniejszą miarą odległości stosowaną dla metody k -Najbliższych Sąsiadów jest odległość euklidesowa:

$$\text{dist}(\mathbf{z}, \mathbf{z}(k)) = \sqrt{\sum_{i=1}^{D_z} (\mathbf{z}^{(i)} - \mathbf{z}^{(i)}(k))^2} \quad (7.130)$$

gdzie $\mathbf{z}(k)$ reprezentuje zestaw danych ze zbioru $\mathbf{Z}_K = (\mathbf{z}(1), \mathbf{z}(2), \dots, \mathbf{z}(K))$.

W kolejnym kroku konstruowany jest ranking sąsiadów przez nadanie każdemu z nich odpowiedniej wagi. Ranking w metodzie k -NN budujemy w następujący sposób:

$$\text{score}(\mathbf{z}_{\text{nowy}}, \mathfrak{c}_i) = \sum_{\mathbf{z}(k) \in \text{KNN}(\mathbf{z}_{\text{nowy}})} \text{dist}(\mathbf{z}_{\text{nowy}}, \mathbf{z}(k)) \delta_K(\mathbf{z}(k), \mathfrak{c}_i) \quad (7.131)$$

gdzie \mathfrak{c}_i jest i -tą klasą ze zbioru wszystkich dostępnych N_c klas, $\text{KNN}(\mathbf{z})$ jest zbiorem najbliższych sąsiadów nowego elementu \mathbf{z}_{nowy} , a $\delta_K(\mathbf{z}(k), \mathfrak{c}_i)$ jest funkcją postaci:

$$\delta(\mathbf{z}(k), \mathfrak{c}_i) = \begin{cases} 1 & \text{jeżeli } \mathbf{z}(k) \in \mathfrak{c}_i \\ 0 & \text{jeżeli } \mathbf{z}(k) \notin \mathfrak{c}_i \end{cases} \quad (7.132)$$

Biorąc pod uwagę zależności od (7.130) do (7.132), reguła decyzyjna w algorytmie k -NN jest następująca:

$$\mathbb{c} = \arg \min_{\mathbb{c} \in \mathcal{N}_c} \text{score}(\mathbf{z}_{\text{nowy}}, \mathbb{c}_i) = \arg \min_{\mathbb{c} \in \mathcal{N}_c} \left(\sum_{\mathbf{z}(k) \in \text{KNN}(\mathbf{z}_{\text{nowy}})} \text{dist}(\mathbf{z}_{\text{nowy}}, \mathbf{z}(k)) \delta_K(\mathbf{z}(k), \mathbb{c}_i) \right) \quad (7.133)$$

W przypadku strumieni danych odległość euklidesowa może prowadzić do niezadowolających wyników. Rozwiązaniem jest zastosowanie algorytmu o nazwie Dynamiczna Transformacja Czasowa (ang. *Dynamic Time Warping*, DTW) [105, 293, 397]. Metoda DTW umożliwia porównywanie strumieni danych, które na przykład są przesunięte względem siebie w czasie.

Metoda DTW bazuje na lokalnych pomiarach odległości pomiędzy porównywanymi strumieniami danych, przy czym wyznaczone odległości gromadzone są w odpowiedniej tabeli. Liczba wierszy i kolumn w konstruowanej tabeli zależy od długości porównywanych strumieni. Do lokalnych pomiarów odległości najczęściej wykorzystywana jest miara euklidesowa. Na kolejnym etapie opracowana tabela jest przeszukiwana i wyznaczana jest ścieżka dopasowania (ang. *warping curve*) łącząca pierwszy i ostatni element w tabeli. Ścieżka jest wyznaczana przez minimalizację kosztu przejścia pomiędzy elementem początkowym i docelowym w analizowanej tabeli. Zaznaczmy, że wyznaczana ścieżka musi spełniać warunki monotoniczności i ciągłości.

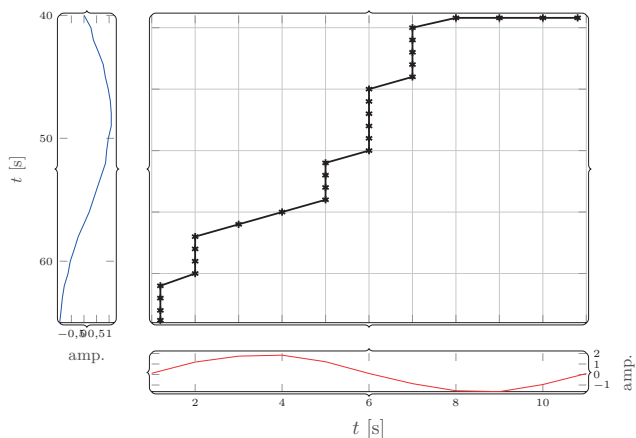
Niech $\mathbf{Z}_{K_1} = (z_1(1) \ z_1(2) \ \dots \ z_1(K_1))$ oraz $\mathbf{Z}_{K_2} = (z_2(1) \ z_2(2) \ \dots \ z_2(K_2))$ będą dwoma sygnałami o różnych długościach. Zadaniem naszym jest oszacowanie podobieństwa tych dwóch sygnałów. Do rozwiązania zadania wykorzystujemy macierz odległości (ang. *distance matrix*), którą wyznaczamy z następującej zależności:

$$\mathbf{C}_{DM}(i, j) = \text{dist}(z_1(i), z_2(j)) + \min \begin{bmatrix} \mathbf{C}_{DM}(i, j-1) \\ \mathbf{C}_{DM}(i-1, j) \\ \mathbf{C}_{DM}(i-1, j-1) \end{bmatrix} \quad (7.134)$$

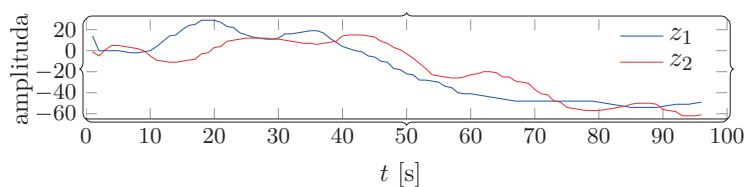
gdzie $\text{dist}(z_1(i), z_2(j))$ jest odległością pomiędzy pomiarami z dwóch sekwencji, tj. $z_1(i)$, $i = 1, 2, \dots, K_1$ oraz $z_2(j)$, $j = 1, 2, \dots, K_2$. W tym przypadku można wykorzystać różne metryki, np. euklidesową, p -normę lub kosinusową. Macierz \mathbf{C}_{DM} nazywamy macierzą odległości lokalnych (ang. *Cumulative Distance Map*). Macierz ta reprezentuje odległości pomiędzy wszystkimi pomiarami sygnałów \mathbf{Z}_{K_1} oraz \mathbf{Z}_{K_2} . Przykład konstrukcji macierzy odległości przedstawiono na rys. 7.69.

Przykład 7.16 (Algorytm DTW). Na rysunkach 7.69–7.72 przedstawiono działanie algorytmu DTW.

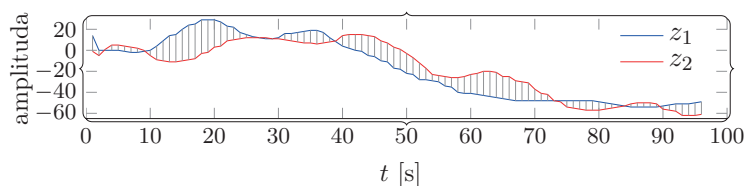
7. Poziom średni: fuzja właściwości



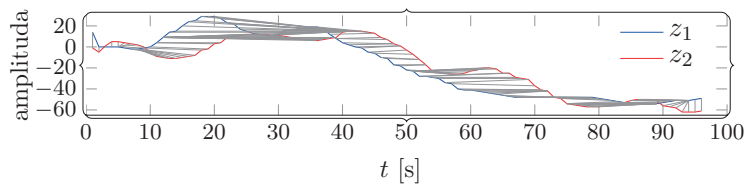
Rys. 7.69. Przykład konstrukcji macierzy odległości w algorytmie DTW



Rys. 7.70. Przykład dwóch strumieni danych



Rys. 7.71. Porównanie strumieni danych z wykorzystaniem odległości euklidesowej



Rys. 7.72. Porównanie strumieni danych z wykorzystaniem metody DTW

7.8.2. Regresja logistyczna

Regresja Logistyczna (ang. *Logistic Regression*, RL) jest algorytmem klasyfikacji będącym uogólnieniem Regresji Liniowej [42, 298]. Zależność dla Regresji Logistycznej ma postać:

$$P(c|\mathbf{z}) = \Phi_{rl}(\mathbf{h}^T \mathbf{z}) \quad (7.135)$$

gdzie \mathbf{h} reprezentują wagi, a funkcja Φ_{rl} jest funkcją aktywacji (np. *sigmoidea*).

Przedstawiony model ma zestaw parametrów, które należy oszacować. W celu wyznaczenia \mathbf{h} rozwiązujemy zadanie optymalizacyjne z następującym wskaźnikiem jakości:

$$Q(\mathbf{h}) = - \sum_{i=1}^{N_{le}} \left(c_i \ln \Phi_{rl}(\mathbf{h}^T \mathbf{z}_i) + (1 - c_i) \ln (1 - \Phi_{rl}(\mathbf{h}^T \mathbf{z}_i)) \right) \quad (7.136)$$

gdzie N_{le} liczbą przykładów w zbiorze uczącym, \mathbf{z} jest wektorem wejściowym, a c_i przyporządkowaną jej klasą ze zbioru uczącego. Do wyznaczenia parametrów Regresji Logistycznej wykorzystać można jedną z metod optymalizacji numerycznej.

7.8.3. Maszyna wektorów nośnych

Maszyny wektorów nośnych (ang. *Support Vector Machines*, SVM) są bardzo popularną metodą analizy danych do rozwiązywania zadań klasyfikacji lub regresji [4, 107, 394, 431, 433]. Algorytm SVM należy do grupy algorytmów, które wymagają uczenia ze zbiorem uczącym.

Rozważmy klasyfikator binarny bazujący na metodzie SVM. Niech $TS_{le} = \{(\mathbf{x}_i, c_i)\}$, $i = 1, 2, \dots, N_{le}$ będzie zbiorem uczącym, zawierającym wektor cech oraz przyporządkowany numer klasy ze zbioru $\mathcal{N}_c = \{-1, 1\}$. Zdefiniujemy teraz hiperpłaszczyznę decyzyjną (ang. *decision hyperplane*):

$$\mathbf{h}^T \mathbf{z} + h_0 = 0 \quad (7.137)$$

gdzie wielkości \mathbf{h} oraz h_0 są parametrami dla hiperpłaszczyzny decyzyjnej. Funkcja decyzyjna dla algorytmu SVM ma ogólną formę:

$$\mathcal{D}_{SVM}(\mathbf{z}) = \text{sgn}(\mathbf{h}^T \mathbf{z} + h_0) \quad (7.138)$$

Dla zadania estymacji parametrów algorytmu SVM zdefiniujemy jeszcze zadanie optymalizacyjne:

7. Poziom średni: fuzja właściwości

$$\{\mathbf{h}^*, b_0^*, q^*\} = \min_{\mathbf{h}, b_0, q} \left(\frac{1}{2} \mathbf{h}^\top \mathbf{h} + p \sum_{i=1}^{N_{sw}} q_i \right)$$

p.o. (7.139)

$$q_i \geq 0, \quad i = 1, \dots, N_{sw}$$

$$p \geq 0$$

gdzie \mathbf{h} jest wektorem normalnym do hiperpłaszczyzny, h_0 określa przesunięcie hiperpłaszczyzny od środka układu współrzędnych, p reprezentuje parametr funkcji kary, natomiast q_i to zmienne luźne (ang. *slack variables*).

Wykorzystanie algorytmu SVM w klasyfikacji strumieni danych jest trudne. Klasyfikator ten wymaga uczenia przyrostowego, co pociąga za sobą konieczność aktualizacji na bieżąco modelu SVM. Dodatkową trudność może sprawiać liczba przykładów uczących, która zwiększa się wraz z upływem czasu. Rozwiązywanie tych zadań w przypadku algorytmu SVM wiąże się z dużymi nakładami obliczeń, które mogą sprawić, że niemożliwe będzie wykonanie aktualizacji modelu SVM z wymaganą częstotliwością. W pracy [221] podjęto próbę opracowania algorytmu bazującego na algorytmie maszyn wektorów nośnych, dostosowanego do warunków, które wymusza przetwarzanie strumieni danych. Ogólna koncepcja zmodyfikowanego algorytmu SVN jest taka, że z każdą zmienną luźną q_i powiązany jest dodatkowy parametr \mathbf{h}_{qi} . Zaproponowana modyfikacja prowadzi do następującego zadania optymalizacji [221]:

$$\{\mathbf{h}^*, h_0^*, q^*\} = \min_{\mathbf{h}, b_0, q} \left(\frac{1}{2} \mathbf{h}^\top \mathbf{h} + p \sum_{i=1}^{N_{sw}} \mathbf{h}_{qi} q_i \right) \quad \text{p.o.} \quad p \geq 0, \quad \mathbf{h}_{qi} \geq 0 \quad (7.140)$$

7.8.4. Głębokie sieci neuronowe

Głęboka sieć neuronowa (ang. *Deep Neural Networks*, DNN) to nowe podejście, które wykorzystywane jest m.in. w problemach klasyfikacji. DNN tym różni się m.in. od sieci neuronowej, że sieć ta konstruowana jest z wykorzystaniem większej liczby warstw ukrytych [249]. Jeżeli sieć neuronowa ma więcej niż trzy warstwy (wliczając w to warstwę wejściową i wyjściową), to taką sieć nazywamy siecią głęboką. Głębokie sieci neuronalne odróżniają się od sieci klasycznych również podejściem do ich uczenia. W przypadku sieci głębokich uczenie (tj. wyznaczanie wag sieci) odbywa się stopniowo, tzn. do wyznaczania wag w danej warstwie wykorzystywane jest wyjście warstwy poprzedzającej. Skutkuje to tym, że każda kolejna warstwa umożliwia konstrukcję reprezentacji bardziej złożonych cech [164]. Jest to tak zwana hierarchia cech (ang. *feature hierarchy*).

Wstępne przygotowanie cech, które posłużą następnie do wyznaczania ostatecznych ich wartości w procesie uczenia, jest istotną cechą głębokich sieci neuronalnych.

Warunkowa ograniczona maszyna Boltzmann

W przetwarzaniu danych strumieniowych znalazł zastosowanie m.in. jeden z typów sieci głębokich, tj. warunkowa ograniczona maszyna Boltzmann (ang. *Conditional Restricted Boltzmann Machine*, CRBM), która jest wariantem ograniczonej maszyny Boltzmann [171, 226, 469]. Ograniczona maszyna Boltzmann jest nieskierowanym, dwudzielnym grafem.

Niech

$$E(\mathbf{z}, \mathbf{h}) = \mathbf{z}^\top \mathbf{W}_{\text{crbm}} \mathbf{h} + \mathbf{b}_0^\top \mathbf{z} + \mathbf{b}_1^\top \mathbf{h} \quad (7.141)$$

będzie funkcją energii, a:

$$P(\mathbf{z}, \mathbf{h}) = \exp(E(\mathbf{z}, \mathbf{h})) \quad (7.142)$$

rozkładem łącznym dla zmiennych z warstw ukrytych (\mathbf{h}) oraz wejściowej (\mathbf{z}). W metodzie tej odpowiednie rozkłady warunkowe wyznaczane są z następujących zależności:

$$P(\mathbf{h}_j | \mathbf{z}) = \Phi_{\text{crbm}} \left(\mathbf{b}_{0j} + \sum_i \mathbf{W}_{\text{crbm}i,j} \mathbf{z}_i + \sum_m \sum_i \mathbf{B}_{i,j,m} \mathbf{z}_i(k-m) \right) \quad (7.143a)$$

$$P(\mathbf{z}_i | \mathbf{h}) = \Phi_{\text{crbm}} \left(\mathbf{b}_{1i} + \sum_j \mathbf{W}_{\text{crbm}i,j} \mathbf{h}_j + \sum_m \sum_i \mathbf{A}_{i,j,m} \mathbf{z}_i(k-m) \right) \quad (7.143b)$$

gdzie \mathbf{A} oraz \mathbf{B} to macierze opisujące związki pomiędzy elementami z warstwy wejściowej z elementami z warstw ukrytych [226].

7.9. Metody oceny jakości algorytmów fuzji właściwości

Ważnym zagadnieniem związanym z fuzją na poziomie właściwości jest ocena wykorzystanych algorytmów. Do oceny jakości tych algorytmów, oprócz odpowiednio przygotowanego zestawu danych testowych, potrzebne jest zgromadzenie oraz odpowiednia analiza uzyskanych wyników. Jednym z narzędzi do analizy wyników fuzji właściwości jest **tablica pomyłek** (ang. *confusion matrix*). Tablicę pomyłek przedstawiono w tab. 7.3.

Korzystając z tabeli prawdy, można wyliczyć kilka wskaźników umożliwiających ocenę wykorzystanych algorytmów fuzji właściwości. Dalej przedstawiliśmy najważniejsze z nich:

7. Poziom średni: fuzja właściwości

Tabela 7.3. Tablica pomyłek

| | | Klasa rzeczywista | | |
|----------------------|-----------|---------------------------------------|---------------------------------------|----------------|
| | | Pozytywna | Negatywna | |
| Klasa predykowana | Pozytywna | prawdziwie pozytywna (TP) | falszywie pozytywna (FP) | TP + FP |
| | Negatywna | falszywie negatywna (FN) | prawdziwie negatywna (TN) | FN + TN |
| | | TP + FN | FP + TN | |

- **czułość** (ang. *sensitivity*, SE) określa stosunek prawidłowo wykrytych (TP) przypadków do sumy prawidłowo wykrytych (TP) i niewykrytych (FN) przypadków:

$$SE = \frac{TP}{TP + FN} \quad (7.144)$$

- **specyficzność** (ang. *specificity*, SP) określa stosunek liczby prawidłowo niewykrytych (TN) przypadków do sumy prawidłowo niewykrytych (TN) i błędnie wskazanych (FP) przypadków (tj. błąd pierwszego rodzaju):

$$SP = \frac{TN}{TN + FP} \quad (7.145)$$

gdzie TN to prawdziwie negatywna.

- **dodatnia wartość prognostyczna** (ang. *Positive Predictive Value*, PPV) określa stosunek prawidłowo wykrytych (TP) przypadków do sumy prawidłowo wykrytych (TP) i błędnie wskazanych (FP) przypadków:

$$PPV = \frac{TP}{TP + FP} \quad (7.146)$$

- **ujemna wartość prognostyczna** (ang. *Negative Predictive Value*, NPV) określa stosunek prawidłowo niewykrytych (TN) do sumy prawidłowo niewykrytych oraz błędnie niewykrytych (FN):

$$NNV = \frac{TN}{TN + FN} \quad (7.147)$$

- **dokładność** (ang. *Classification Accuracy*, ACC) określa sumę prawidłowo wykrytych (TP) przypadków i prawidłowo niewykrytych (TN) przypadków do

7.9. Metody oceny jakości algorytmów fuzji właściwości

sumy wszystkich możliwych przypadków:

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7.148)$$

- **prawdopodobieństwo fałszywego alarmu** (ang. *False Alarm*, FA) określa liczbę błędnie wskazanych (FP) przypadków do liczby wszystkich prawidłowo niewykrytych (TN) przypadków i błędnie wskazanych:

$$\text{FA} = \frac{\text{FP}}{\text{TN} + \text{FP}} \quad (7.149)$$

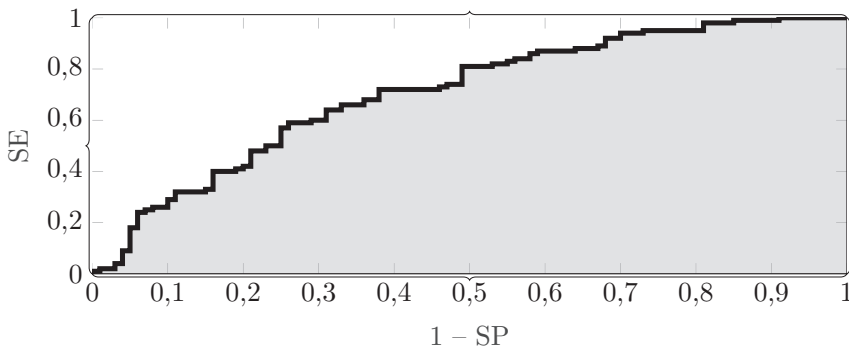
- **błąd całkowity** (ang. *Total Error*, TE) określa zdolność dyskryminacyjną algorytmu fuzji właściwości i jest to stosunek liczby błędnie wskazanych (FP) do sumy wszystkich możliwych przypadków:

$$\text{TE} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7.150)$$

- **miara F_1** (ang. *F1 score*, F1) wyznaczana jest z następującej zależności:

$$F1 = 2 \frac{\text{PPV} \cdot \text{SE}}{\text{PPV} + \text{SE}} \quad (7.151)$$

Do badania wpływu cech klasyfikatora można również wykorzystać krzywą ROC (ang. *Receiver Operating Characteristic*). Postać krzywej ROC przedstawiono na rys. 7.73.



Rys. 7.73. Przykład krzywej ROC

7.10. Rozpoznawanie aktywności ruchowej człowieka – fuzja właściwości

Przedstawimy teraz przykład wykorzystania omówionych w tym rozdziale algorytmów fuzji właściwości. Przykład dotyczy rozpoznawania aktywności ruchowej osoby podczas treningu. W przedstawionym przykładzie osobą jest sportowiec (tenisista), a aktywnością ruchową są ćwiczenia podczas treningu technicznego. Zadaniem jest dokonanie analizy danych z czujnika przyśpieszenia oraz ustalenie na podstawie tych pomiarów, jaka zagrywka została wykonana. Wyniki rozpoznawania na podstawie danych z czujnika przyśpieszenia mogą być wykorzystane np. do wykonania statystyk dla pojedynczego treningu czy treningów wykonanych w dłuższym przedziale czasu. Przedstawione w tym podrozdziale wyniki opracowano na podstawie prac [52, 388].

Dane niezbędne do rozwiązania zadania zostały zebrane z wykorzystaniem czujnika *The Pebble Watch*. Czujnik ten został umieszczony na nadgarstku osoby grającej. Dane z czujnika wysyłane były bezprzewodowo do aplikacji gromadzącej te dane. Czujnik *The Pebble Watch* umożliwia pomiar przyśpieszenia w trzech osiach do ± 4 G z częstotliwością próbkowania 10 Hz, 25 Hz, 50 Hz lub 100 Hz.

Badania zostały wykonane na krytym korcie. Przed dokonaniem pomiarów, każdy z tenisistów podczas rozgrzewki wykonał kilka serwisów, *forehandów* i *backhandów*.

Tenisisci zostali podzieleni na dwie grupy, tj. grupę amatorską i zawodową. W pierwszej grupie znalazło się 20 mężczyzn i 2 kobiety. W drugiej grupie natomiast było 7 mężczyzn i 5 kobiet. Zadaniem każdej badanej osoby było wielokrotne wykonanie następujących zagrywek, tj. serwisu, *forehandu* i *backhandu*. Podsumowanie zamieszczono w tab. 7.4. Z kolei liczba wszystkich zgromadzonych danych została przedstawiona w tab. 7.5.

Tabela 7.4. Liczba prób pomiarowych dla każdej z zagrywek w podziale na amatorów i zawodowców

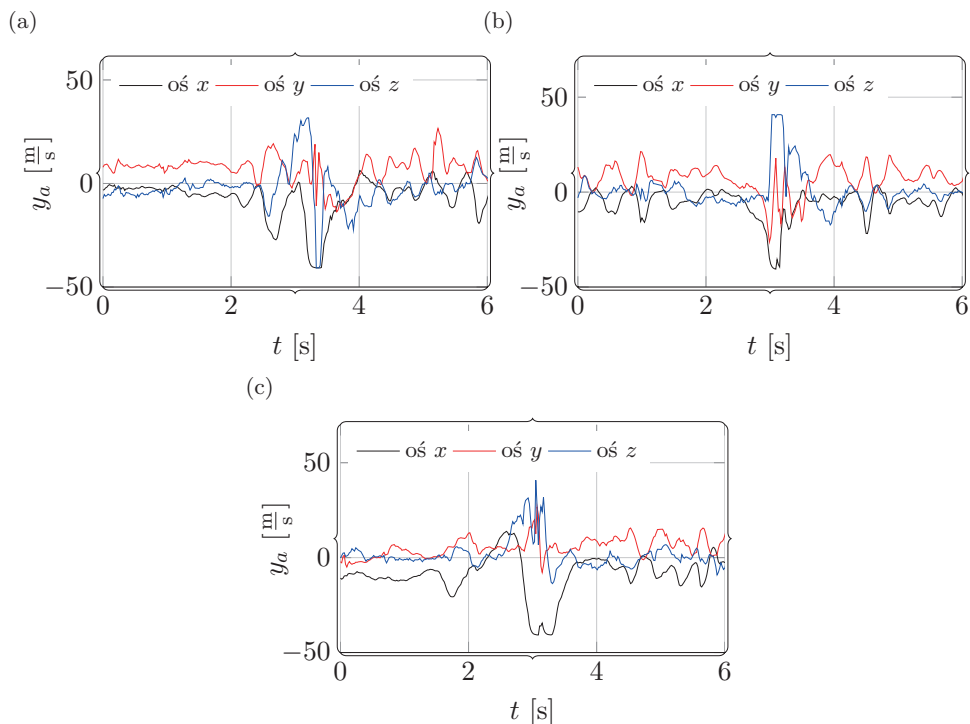
| Zagrywka | Amatorzy | Zawodowcy |
|-----------------|---------------|-----------|
| | Liczba próbek | |
| <i>Forehand</i> | 961 | 271 |
| <i>Backhand</i> | 491 | 257 |
| Serwis | 342 | 93 |

Przykładowe próby pomiarowe przyśpieszenia dla *forehandu*, *backhandu* i serwisu przedstawiono na rys. 7.74.

7.10. Rozpoznawanie aktywności ruchowej człowieka – fuzja właściwości

Tabela 7.5. Liczba prób pomiarowych w podziale na amatorów i zawodowców

| Poziom | Liczba badanych | Liczba prób |
|------------|-----------------|-------------|
| Amator | 22 | 1794 |
| Zawodowiec | 12 | 621 |

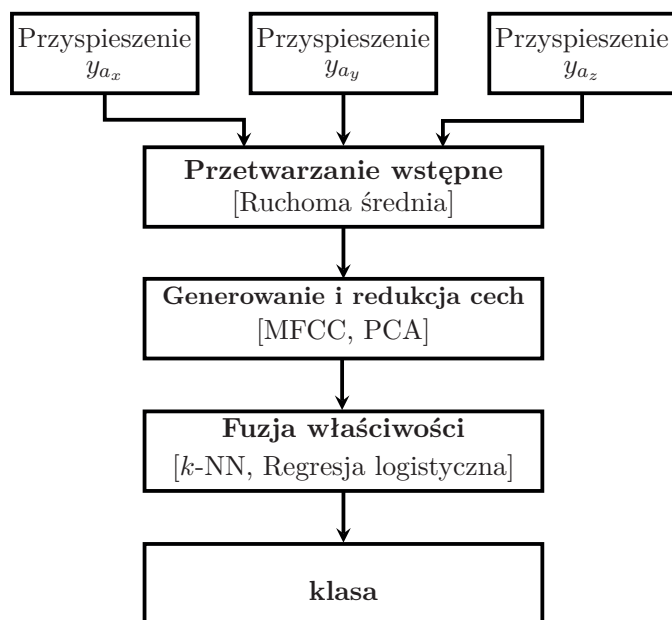


Rys. 7.74. Przykładowe dane pomiarowe przyspieszenia dla:
a) *forehandu*, b) *backhandu* i c) *serwisu*

7.10.1. Algorytm fuzji właściwości

Zaproponowany algorytm fuzji właściwości przedstawiono na rys. 7.75. Główne kroki algorytmu, zgodnie z przyjętą metodyką fuzji danych, zawierają przetwarzanie wstępne, generowanie i ekstrakcję cech oraz fuzję właściwości. Wyniki pomiarów z czujnika przyspieszenia oznaczyliśmy $\mathbf{Y}_a = [\mathbf{y}_a(1), \mathbf{y}_a(2), \dots, \mathbf{y}_a(K)]$, $k = 1, 2, \dots, K$. K reprezentuje liczbę wszystkich pomiarów a $\mathbf{y}_a = [y_{ax} \ y_{ay} \ y_{az}]^\top$.

7. Poziom średni: fuzja właściwości



Rys. 7.75. Zaproponowany algorytm fuzji właściwości dla zadania klasyfikacji zagrywek tenisowych

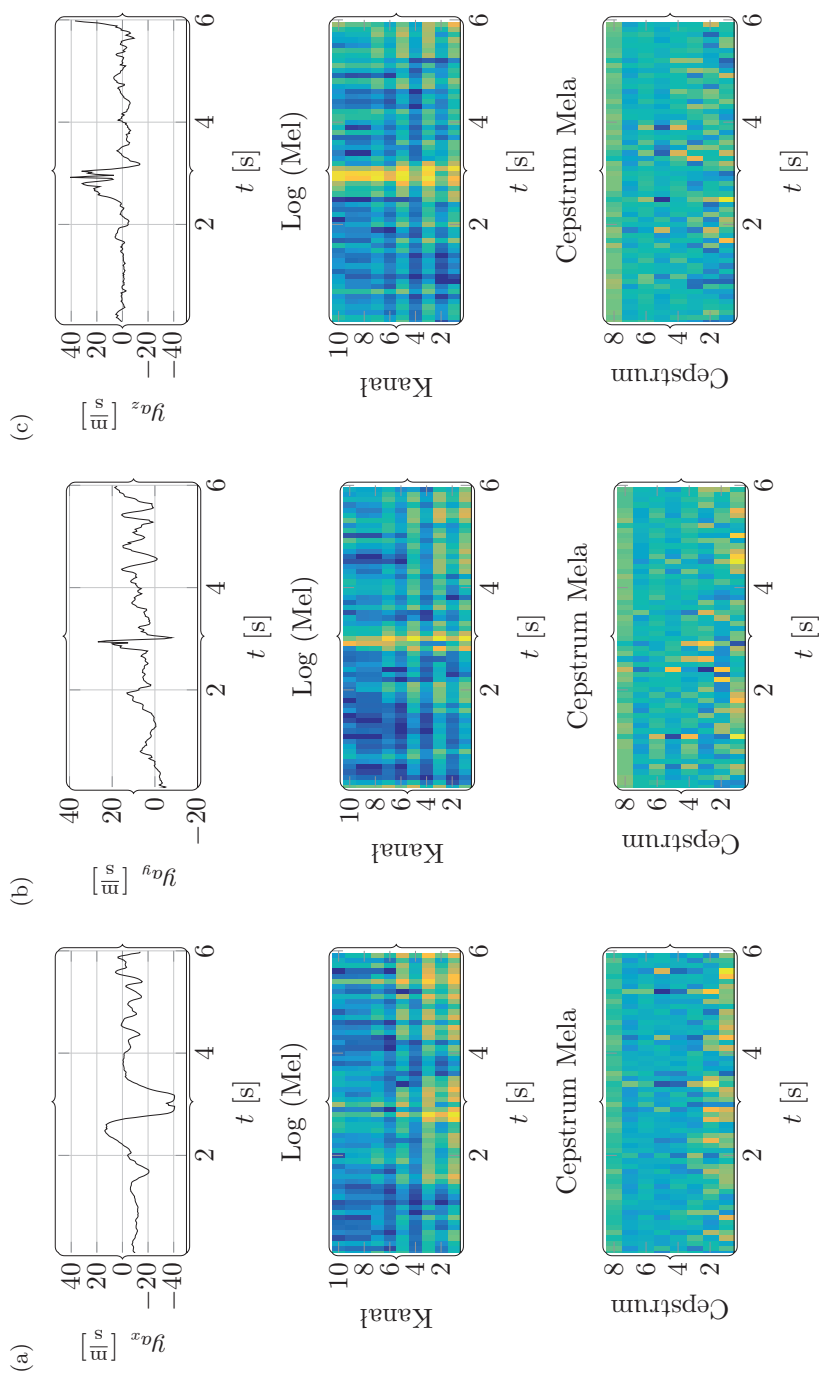
Przetwarzanie wstępne

Zadaniem przetwarzania wstępnego jest przygotowanie pomiarów z czujnika przyspieszenia do kolejnego etapu, tj. generowania i ekstrakcji cech. W prezentowanym przykładzie do przetwarzania wstępnego wykorzystano metodę uśredniania w celu usunięcia niepożądanych składowych z sygnału pomiarowego. Algorytm jest prosty, ale na potrzeby tego zadania uzyskane wyniki okazały się wystarczająco dobre.

Generowanie cech

Kolejny krok w zaproponowanym rozwiązaniu związany jest z generowaniem cech na podstawie danych z czujnika przyspieszenia. W tym przypadku wykorzystaliśmy współczynniki MFCC (por. podrozdz. 7.4.2). Zaletą współczynników MFCC, która była istotna w rozwiązywanym zadaniu, jest ich niewielka złożoność obliczeniowa oraz możliwość reprezentowania sygnału w zwartej formie [94]. Przykład wyznaczonych współczynników MFCC dla danych z zadania (pomiar przyspieszenia) przedstawiono na rys. 7.76.

7.10. Rozpoznawanie aktywności ruchowej człowieka – fuzja właściwości



Rys. 7.76. Współczynniki MFCC pomiarów z czujnika przyspieszenia dla serwisu dla: a) osi x , b) osi y oraz c) osi z

Redukcja cech

Kolejnym krokiem w zaproponowanym rozwiązaniu (rys. 7.75) jest redukcja cech. W prezentowanym podejściu wykorzystano metodę PCA (patrz podrozdz. 7.6.3), która umożliwia zredukowanie wymiaru wektora wejściowego przez jego rzutowanie na przestrzeń o mniejszej wymiarowości.

Fuzja właściwości

Do rozwiązania zadania fuzji właściwości zaproponowano dwa algorytmy klasyfikacji, których przydatność oceniono w badaniach eksperymentalnych. Pierwszym badanym algorytmem był algorytm k -Najbliższych Sąsiadów, drugim z kolei Regresja Logistyczna. Metody te zostały omówione w podrozdziałach odpowiednio 7.8.1 oraz 7.8.2.

Wyniki fuzji właściwości

Do oceny jakości zaproponowanego rozwiązania wykorzystano dwie metody, tj. 10-krotną walidację krzyżową oraz *leave-one-out*. Metoda *leave-one-out* jest wariantem metody K -krotnej walidacji krzyżowej, w której podzielono N elementową próbę na N podzbiorów. Do oceny wyników fuzji właściwości w zadaniu do oceny jakości fuzji właściwości wykorzystano wskaźnik ACC (dokładność) (por. podrozdz. 7.9).

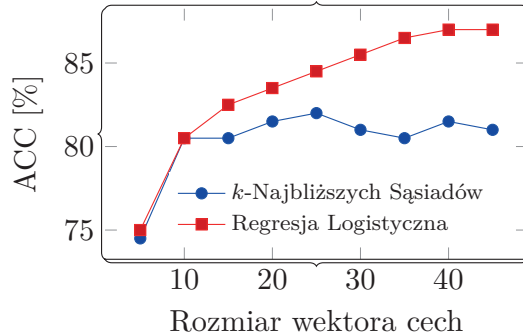
W przeprowadzonych badaniach ocenie poddano wpływ rozmiaru wektora cech, szerokość okna przetwarzania oraz liczbę współczynników MFCC na jakość klasyfikacji.

Wpływ rozmiaru wektora właściwości

Liczba wygenerowanych, na podstawie danych (w tym przypadku z czujnika przyspieszenia), współczynników MFCC może być duża. Duża ich liczba może wpłynąć niekorzystnie na liczbę obliczeń niezbędnych do rozwiązania zadania fuzji właściwości, bazując na wektorze cech zbudowanym ze współczynników MFCC. Dlatego zasadne jest wybranie spośród wygenerowanych cech tylko tych z nich, które są najbardziej informacyjne. Dlatego kolejne zadanie, które należy rozwiązać to zadanie redukcji cech. W przedstawionym przykładzie wykorzystano metodę PCA. Na rysunku 7.77 przedstawiliśmy jakość fuzji właściwości dla różnej liczby współczynników MFCC. Badania przeprowadzono dla algorytmu k -Najbliższych Sąsiadów oraz Regresji Logistycznej. Z wykresu wynika, że dla przygotowanego zbioru danych algorytm k -Najbliższych Sąsiadów uzyskał najlepsze wyniki dla wektora składającego się z 25 cech wyznaczonych na podstawie współczynników MFCC.

7.10. Rozpoznawanie aktywności ruchowej człowieka – fuzja właściwości

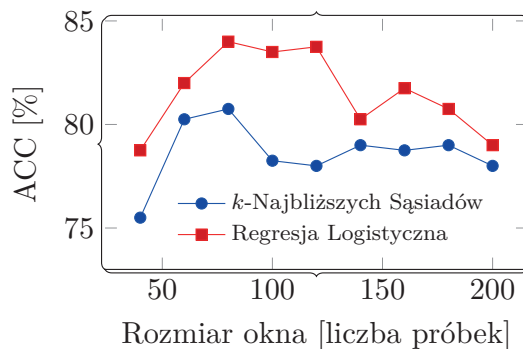
Natomiast znacznie lepszą jakość klasyfikacji uzyskano dla Regresji Logistycznej i to przy znacznie większym rozmiarze wektora właściwości, tj. 50 cech.



Rys. 7.77. Jakość klasyfikacji dla różnych rozmiarów wektora właściwości

Wpływ rozmiaru okna czasowego

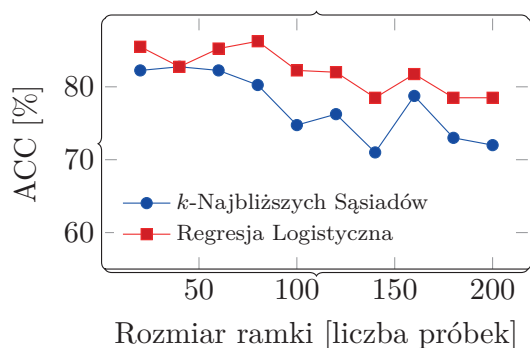
Z przeprowadzonych badań wynika, że w opisywanym zadaniu można ustalić optymalny rozmiar okna czasowego. Zarówno zbyt krótkie, jak i za długie okno czasowe wpływa na pogorszenie jakości fuzji właściwości zarówno dla metody k -Najbliższych Sąsiadów, jak i Regresji Logistycznej. Rozmiar okna czasowego dla rozpatrywanych danych i metod jest podobny i powinien wynosić ok. 80 próbek (rys. 7.78).



Rys. 7.78. Jakość klasyfikacji dla różnych rozmiarów okna czasowego

Wpływ parametrów algorytmu do wyznaczania MFCC

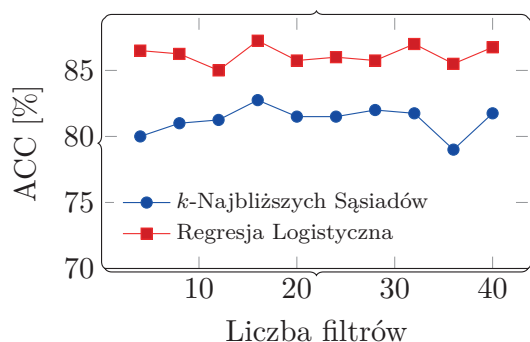
Zbadajmy teraz jak na jakość fuzji właściwości wpływa dobór parametrów algorytmów do wyznaczania współczynników MFCC. Na rysunku 7.79 pokazano wpływ rozmiaru ramki na jakość fuzji właściwości (parametr N_w) wykorzystanej do wyznaczania Dyskretnej Transformaty Fouriera. W przypadku algorytmu



Rys. 7.79. Jakość klasyfikacji zależna od rozmiaru ramki w algorytmie MFCC

k -Najbliższych Sąsiadów rozmiar ten wynosi 40 próbek. Z kolei dla Regresji Logistycznej optymalny rozmiar ramki to 80 próbek. Dalsze zwiększanie rozmiaru ramki znacząco wpływa na pogorszenie jakości fuzji właściwości.

Na rysunku 7.80 przedstawiono wpływ liczby filtrów (algorytmu do wyznaczania współczynników MFCC) na jakość fuzji na poziomie średnim. Z badań i analizy wykresu wynika, że liczba filtrów wykorzystanych w algorytmie do wyznaczania cech MFCC wynosi 16. Liczba ta jest najlepsza zarówno dla algorytmu k -Najbliższych Sąsiadów, jak i dla Regresji Logistycznej.



Rys. 7.80. Jakość klasyfikacji zależna od liczby filtrów w algorytmie MFCC

Analiza z wykorzystaniem walidacji krzyżowej

Oceńmy teraz jakość zaproponowanego podejścia w 10-krotnej walidacji krzyżowej. Procedura polega na dziesięciokrotnym losowaniu zbioru uczącego i testowego oraz uśrednianiu uzyskanych wyników. Badania przeprowadzono oddzielnie dla zbioru danych zebranych od tenisistów amatorów i zawodowców. Wykonano również badania dla zbioru, w którym znalazły się pomiary od amatorów i zawodowców. Wyniki zarówno dla metody k -Najbliższych Sąsiadów, jak i Regresji Logistycznej zamieszczono w tab. 7.6. Z badań wynika, że algorytm k -Najbliższych Sąsiadów uzyskał wyższą jakość klasyfikacji dla zbiorów zawodowców, natomiast algorytm Regresji Logistycznej lepsze wyniki uzyskał dla zbioru amatorów. Regresja Logistyczna uzyskała lepszą jakość klasyfikacji w przypadku zbioru danych, w którym znalazły się pomiary zarówno od amatorów, jak i zawodowców.

Tabela 7.6. Jakość klasyfikacji dla algorytmów k -Najbliższych Sąsiadów i Regresji Logistycznej w 10-krotnej walidacji krzyżowej

| Algorytm | Amatorzy | Zawodowcy | Amatorzy i Zawodowcy |
|----------------------------|---------------|---------------|----------------------|
| | ACC [%] | | |
| k -Najbliższych Sąsiadów | 79,65 ± 11,36 | 91,33 ± 10,56 | 82,22 ± 11,30 |
| Regresja Logistyczna | 85,55 ± 8,24 | 88,44 ± 11,12 | 87,99 ± 8,48 |

Analiza z wykorzystaniem *leave-one-out*

Procedura *leave-one-out* [212] umożliwiła ocenę wrażliwości algorytmu na dane, które nie były wykorzystane w procesie uczenia klasyfikatora. W rzeczywistym przypadku można to odnieść do sytuacji, gdy algorytm ma dokonać klasyfikacji zagrywek dla użytkownika, którego dane nie były wykorzystane do konstrukcji zbioru uczącego.

Tabela 7.7. Jakość klasyfikacji dla algorytmów k -Najbliższych Sąsiadów i Regresji Logistycznej w teście *leave-one-out*.

| Algorytm | Amatorzy | Zawodowcy | Amatorzy i Zawodowcy |
|----------------------------|---------------|---------------|----------------------|
| | ACC [%] | | |
| k -Najbliższych Sąsiadów | 81,96 ± 10,54 | 92,22 ± 12,22 | 82,16 ± 9,79 |
| Regresja Logistyczna | 85,12 ± 5,16 | 88,66 ± 12,37 | 87,16 ± 6,86 |

7. Poziom średni: fuzja właściwości

W tabeli 7.7 prezentujemy wyniki badania algorytmów fuzji właściwości dla zbiorów utworzonych z danych pochodzących od amatorów i zawodowców. Również i w tym przypadku algorytm k -Najbliższych Sąsiadów uzyskał lepszy wynik w zbiorze pomiarów od zawodowców, a metoda Regresji Logistycznej w zbiorze danych od amatorów. Z kolei we wspólnym zbiorze danych Regresja Logistyczna uzyskała lepsze wyniki klasyfikacji.

8. Poziom wysoki: fuzja decyzji

*Dobra decyzja bazuje na wiedzy,
a nie na liczbach.*

Platon

8.1. Wstęp

W rozdziale omówione zostały zagadnienia związane z fuzją decyzji. Zgodnie z przyjętym w rozdziale 2 podziałem metod fuzji danych (ze względu na poziom abstrakcji) fuzja decyzji jest na jego najwyższym poziomie. Celem fuzji na tym poziomie jest wyznaczenie decyzji na bazie zgromadzonych na niższych poziomach danych, wyekstrahowanych cech oraz konkluzji (z poziomu drugiego).

Rozdział ten został podzielony na dwie części. W pierwszej z nich rozwinięto zagadnienia, które zostały wprowadzone w rozdziale 2. Szczególnie skupiono się na zagadnieniach związanych z typami architektur dla algorytmów fuzji decyzji.

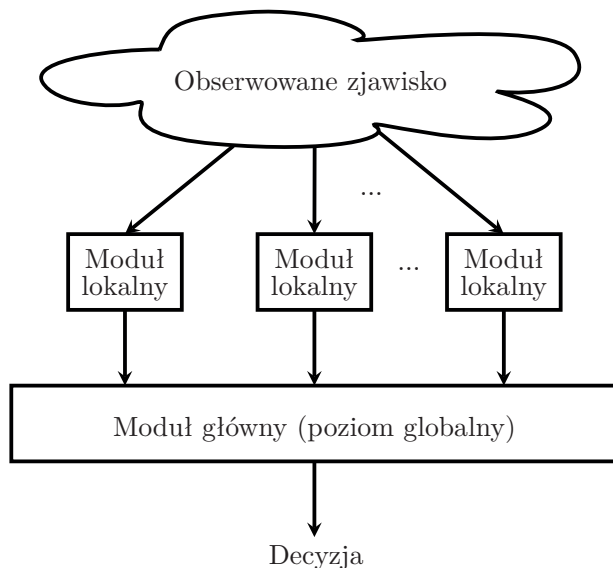
Druga część została poświęcona algorytmom fuzji decyzji. W części tej omówiono różne metody fuzji decyzji m.in. z podziałem na metody bazujące na wartościach dyskretnych i ciągłych oraz rankingach klas.

8.2. Podstawy fuzji decyzji

Na rysunku 8.1 przedstawiono przykład przetwarzania danych na poziomach **lokalnym** i **globalnym**. Wynikiem obserwacji pewnego zjawiska jest zbiór danych pomiarowych. Dane te na pierwszym etapie przetwarzane są w modułach lokalnych. Natomiast w celu budowy całościowego obrazu analizowanego zjawiska, przetworzone w lokalnych modułach dane pomiarowe, właściwości i konkluzje muszą zostać przesyłane do modułu głównego (poziom globalny).

Opisany problem przetwarzania danych można przedstawić np. w formie schematu, który został zaproponowany na rys. 8.2. Do obserwacji obiektu można wykorzystać np. czujniki pomiarowe. Dane zebrane z czujników przesyłane są z kolei do modułów lokalnych (w modułach tych implementowane są głównie algorytmy fuzji na poziomie podstawowym oraz średnim) w celu ich przetworzenia

8. Poziom wysoki: fuzja decyzji



Rys. 8.1. Przykład przetwarzania na poziomie lokalnym i globalnym

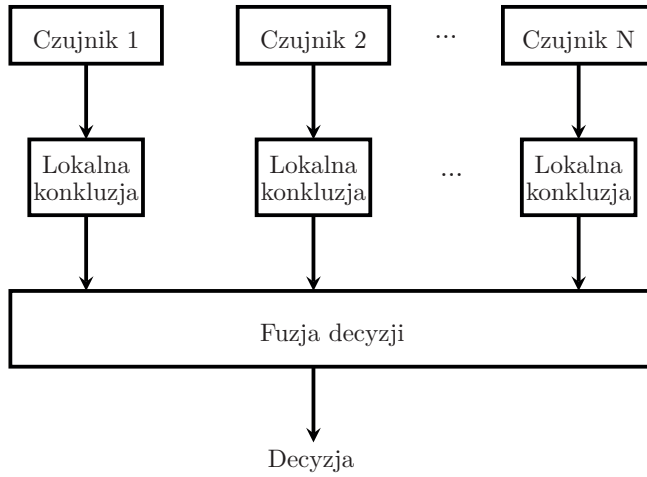
i wyciągnięcia na ich podstawie wniosków (konkluzji). W większości przypadków wnioski wyciągane w modułach lokalnych dotyczą jedynie wybranych fragmentów obserwowanego zjawiska, tj. tych, z którymi bezpośrednio związane jest dane urządzenie pomiarowe. Konkluzje na poziomie lokalnym formułowane są głównie na podstawie algorytmów fuzji właściwości, które z kolei bazują w większości przypadków na metodach klasyfikacji.

W kolejnym kroku konkluzje z poziomu lokalnego przesyłane są do modułu głównego (poziom globalny), gdzie podejmowana jest na ich podstawie decyzja. Jest to poziom fuzji decyzji.

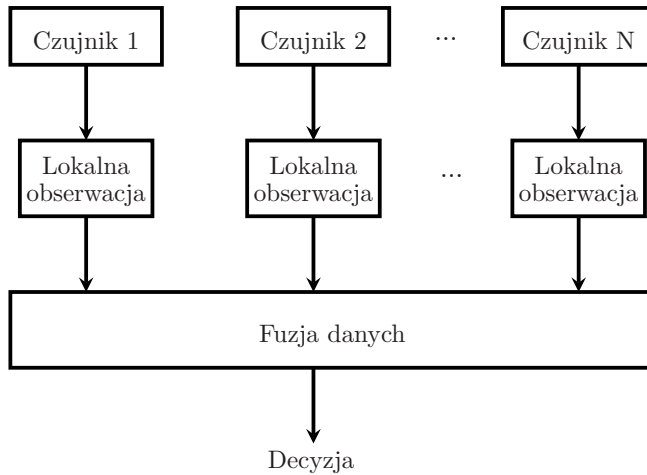
Przedstawiona architektura nazywana jest **architekturą rozproszoną**. Warto podkreślić, że zarówno algorytmy przetwarzania, jak i wnioskowania w każdym z modułów mogą być inne, co musi zostać uwzględnione podczas projektowania algorytmów dla modułu głównego.

Alternatywą dla architektury rozproszonej jest **architektura centralna**. W tym przypadku moduły lokalne wykorzystywane są głównie do gromadzenia obserwacji z urządzeń pomiarowych i przesłania ich do modułu głównego. Proces podejmowania decyzji odbywa się jedynie w module głównym, na podstawie obserwacji dostarczonych przez moduły lokalne. Architektura centralna została przedstawiona na rys. 8.3.

Ustalając architekturę dla projektowanego systemu fuzji danych, należy również przeprowadzić analizę typów źródeł danych w systemie. W ogólności źródła te



Rys. 8.2. Przykład wykorzystania fuzji decyzji dla sytuacji z rys. 8.1 w architekturze rozproszonej



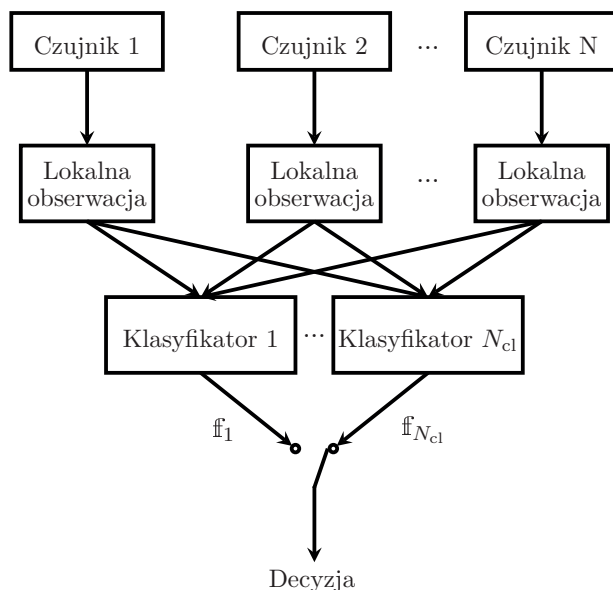
Rys. 8.3. Przykład wykorzystania fuzji decyzji dla sytuacji z rys. 8.1 w architekturze centralnej

mogą być ilościowo **jednorodne** tzn. jednomodalne (np. zestaw czujników pomiarowych do pomiaru przyspieszeń i prędkości kątowych) lub **niejednorodne** tzn. wielomodalne (np. zestaw pomiarowy złożony z kamer wideo, odbiorników sygnału GPS i czujników ciśnienia atmosferycznego). W przypadku źródeł niejednorodnych może okazać się niezbędne zaprojektowanie oddzielnych algorytmów dla każdego

8. Poziom wysoki: fuzja decyzji

ze źródeł. Wymaganie to może mieć wpływ na wybór architektury projektowanego systemu fuzji decyzji.

Rozwiązując zadanie dla przypadku niejednorodnych (wielomodalnych) danych źródłowych, w którym wykorzystano wiele algorytmów fuzji właściwości, możemy rozważyć kilka wariantów architektury projektowanego systemu. To który z wariantów zostanie wykorzystany w danym zadaniu zależy m.in. od tego, jakiego typu wyniki dostępne są na wyjściu każdego z algorytmów fuzji właściwości (poziom drugi). W ogólności wyjścia te możemy podzielić na wyjścia **liczbowe dyskretne** (tj. klasy) oraz **liczbowe ciągłe**. W zależności od typu wyjścia możemy wyróżnić dwa warianty architektury rozproszonych systemów fuzji decyzji (rys. 8.2). Pierwszy z tych wariantów został przedstawiony na rys. 8.4.

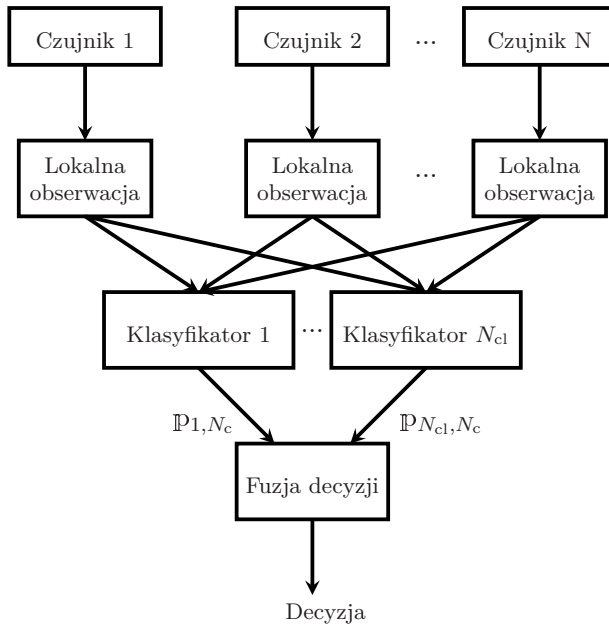


Rys. 8.4. Fuzja decyzji z wykorzystaniem klasyfikatorów z wyjściem liczbowym dyskretnym

Przyjmijmy, że na poziomie lokalnym dysponujemy zestawem N_{cl} klasyfikatorów. Zadanie na tym poziomie polega na przypisaniu obserwowanego zjawiska do jednej z klas c_i (gdzie $i \in \mathcal{N}_c = \{1, 2, \dots, N_c\}$) na podstawie wcześniej przygotowanego wektora cech \mathbf{x} . Pamiętajmy, że wartości, będące składowymi wektora \mathbf{x} , związane są z obserwowanym zjawiskiem i zostały wygenerowane przez algorytm fuzji właściwości na poziomie lokalnym. Efektem działania N_{cl} klasyfikatorów jest N_{cl} -

-elementowy wektor z wynikami dostarczonymi przez każdy z klasyfikatorów $\mathbf{f} = [f_1, f_2, \dots, f_{N_{cl}}]^T$. Składowe tego wektora przyjmują tylko wartości ze zbioru \mathcal{N}_c .

Drugi z przypadków dotyczy systemów złożonych z klasyfikatorów, których wyjście jest liczbowe ciągle (rys. 8.5). W tym przypadku każdy z N_{cl} klasyfikatorów generuje na swoim wyjściu N_c wymiarowy wektor $\mathbb{P}_{j,N_c} = [\mathbb{P}_{j,1} \ \mathbb{P}_{j,2} \ \dots \ \mathbb{P}_{j,N_c}]^T$. Wektor taki interpretujemy jako ilościowe oszacowanie stopnia przynależności obserwowanego obiektu, opisanego wektorem \mathbf{x} , do każdej z klas, przez każdy j -ty klasyfikator. Bez straty ogólności dalszych rozważań możemy przyjąć, że klasyfikatory przyjmują wartości ze zbioru liczb rzeczywistych w przedziale $[0, 1]$. Zauważmy, że dla N_{cl} klasyfikatorów i N_c klas otrzymamy macierz określającą stopień przynależności obiektu do każdej z N_c klas.

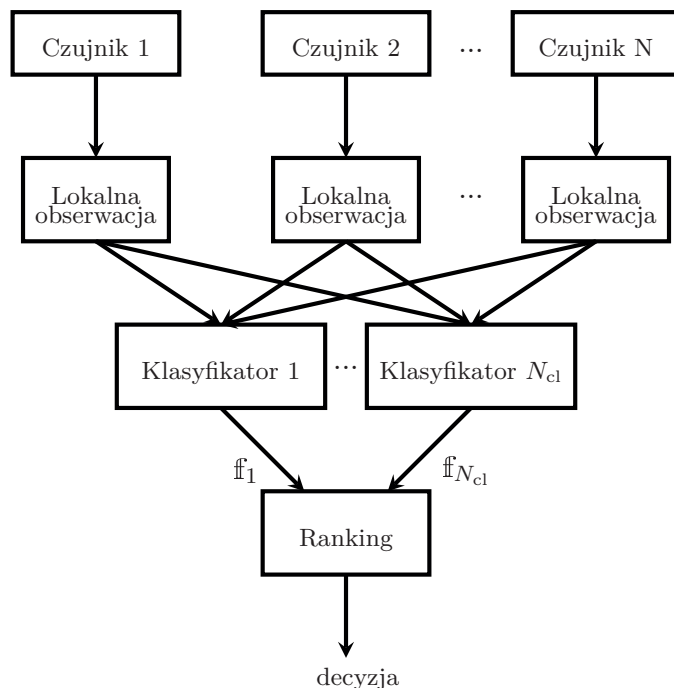


Rys. 8.5. Fuzja decyzji z wykorzystaniem klasyfikatorów z wyjściem liczbowym ciągłym

Rozważmy jeszcze jeden przypadek, który może wystąpić w omawianym zadaniu. Przypadek ten dotyczy algorytmów fuzji właściwości, których wynikiem działania są wyjścia liczbowe dyskretne. Jednym z możliwych podejść jest w tym przypadku utworzenie rankingów klas i na podstawie wyników przygotowanego rankingów rozwiązanie zadania fuzji decyzji. Ranking tworzony jest na podstawie wyników uzyskanych od każdego z klasyfikatorów. W tym przypadku wyjściami

8. Poziom wysoki: fuzja decyzji

klasyfikatorów mogą być liczby dyskretne lub rzeczywiste. Podejście to zostało zilustrowane na rys. 8.6.

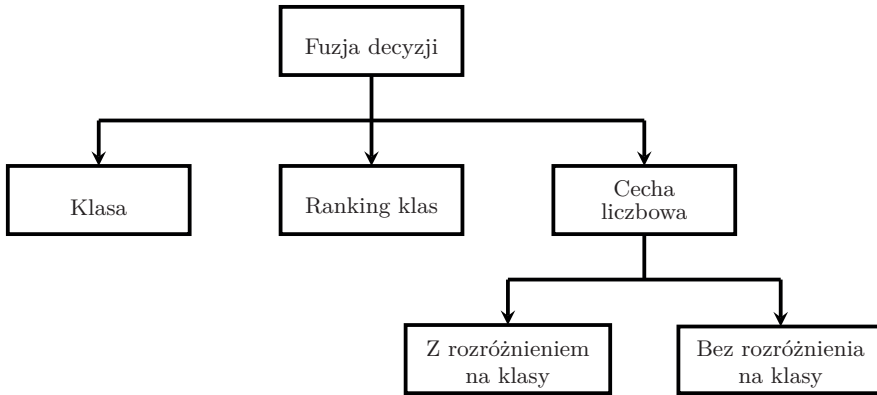


Rys. 8.6. Fuzja decyzji z wykorzystaniem klasyfikatorów z wyjściami liczbowymi dyskretnymi, które uporządkowano według rankingu

8.3. Metody fuzji decyzji

Z dotychczasowych rozważań wynika, że fuzja decyzji jest procesem wysokopoziomowym służącym do przetwarzania danych pozyskanych na niższych poziomach, tj. na poziomie podstawowym i średnim. W ogólności możemy stwierdzić, że istotą fuzji decyzji jest wnioskowanie lub predykcja bazująca zarówno na danych pomiarowych (głównie przetworzonych), ich reprezentacji (tj. cechach) lub konkluzjach z niższych poziomów fuzji danych (głównie z poziomu drugiego).

Bazując na opisie z poprzedniego rozdziału proponujemy następującą hierarchię metod fuzji decyzji (rys. 8.7). Na schemacie wyróżniliśmy fuzje na podstawie



Rys. 8.7. Systematyka metod fuzji decyzji

klas (cech dyskretnych) oraz cech liczbowych ciągłych. Opis został uzupełniony o fuzję decyzji bazującą na rankingach klas. W przypadku cech liczbowych ciągłych wyróżniliśmy dwa podejścia, tj. z rozróżnianiem oraz bez rozróżniania klas (odpowiednio ang. *class-conscious* i *class indifferent*).

8.3.1. Fuzja decyzji na podstawie wielkości dyskretnych

W kolejnych podrozdziałach omówione zostaną metody fuzji decyzji dla przypadku, gdy na poziomie trzecim dostępne są jedynie wielkości dyskretnie wygenerowane na niższych poziomach fuzji danych (tj. poziomie pierwszym i drugim).

Podejście z wykorzystaniem bayesowskiej teorii decyzji

Zadanie fuzji cech dyskretnych można rozwiązać, bazując m.in. na *bayesowskiej* teorii decyzji. Istotą metod bazujących na tym modelu jest estymacja prawdopodobieństwa *a posteriori*, tj. $P(d_i|c_j)$. W tym przypadku $d \in \mathcal{N}_d = \{1, 2, \dots, N_d\}$. \mathcal{N}_d jest zbiorem możliwych N_d decyzji. Następnie, wykorzystując oszacowane prawdopodobieństwo *a posteriori*, dokonujemy wyboru jednej z decyzji d_i ze zbioru \mathcal{N}_d . W tym celu korzystamy z zestawu wcześniej przygotowanych N_{cl} klasyfikatorów.

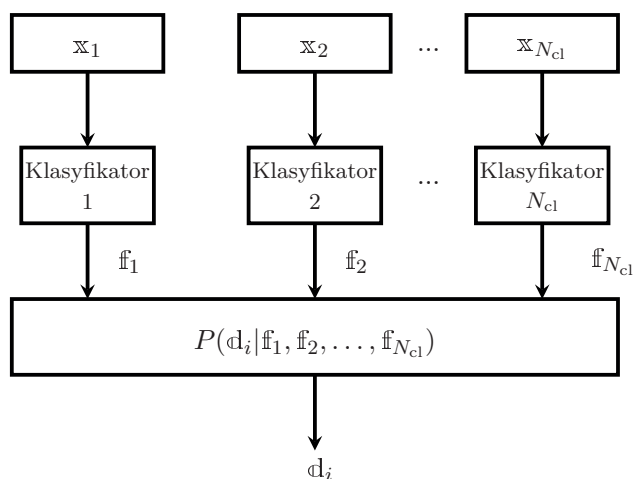
Zgodnie z regułą Bayesa, prawdopodobieństwo *a posteriori* wyznaczamy z następującej zależności [399]:

$$P(d_i|f_1, f_2, \dots, f_{N_{cl}}) = \frac{P(d_i) \prod_{j=1}^{N_{cl}} P(f_j|d_i)}{P(f_1, f_2, \dots, f_{N_{cl}})} \quad (8.1)$$

8. Poziom wysoki: fuzja decyzji

gdzie N_{cl} reprezentuje liczbę klasyfikatorów, a i jest indeksem decyzji, tj. $i = 1, 2, \dots, N_d$.

Równość (8.1) zachodzi po założeniu statystycznej niezależności wszystkich N_{cl} klasyfikatorów. Na rysunku 8.8 zilustrowano opisany schemat fuzji decyzji w modelu probabilistycznym.



Rys. 8.8. Schemat fuzji decyzji w modelu probabilistycznym

Wadą przedstawionego rozwiązania jest m.in. wymaganie o statystycznej niezależności wszystkich N_{cl} klasyfikatorów. W praktycznych zastosowaniach jest to warunek trudny do spełnienia. Kolejną trudnością w praktycznym stosowaniu formuły (8.1) jest trudność w oszacowaniu prawdopodobieństwa *a priori* $P(f_j | d_i)$. Dlatego w zadaniach praktycznych, w których pojawia się potrzeba rozwiązania zadania fuzji decyzji, zwykle stosuje się inne metody.

Podejście bazujące na głosowaniu

W poprzednim podrozdziale przedstawiliśmy metodę fuzji decyzji, która bazuje na bayesowskiej teorii decyzji. Wskazaliśmy na jej wady i problem z praktycznym jej zastosowaniem. Alternatywą dla tej metody jest podejście bazujące na metodach głosowania. W podejściu tym konkluzje z niższych poziomów wykorzystywane są przez algorytmy fuzji decyzji na poziomie trzecim. Zazwyczaj w tym przypadku wykorzystuje się kilka algorytmów fuzji decyzji. Istotą tego podejścia jest to, że zamiast poszukiwać najlepszego algorytmu fuzji decyzji, ostateczny wynik fuzji bazuje na wynikach dostarczonych przez każdy z N_{cl} algorytmów fuzji decyzji.

Głosowanie większościowe

W przypadku głosowania większościowego przyporządkowanie obserwowanego zjawiska do danej klasy następuje wówczas, gdy:

- wszystkie klasyfikatory wskazały tę samą klasę; głosowanie takie nazywamy jednomyślnym (ang. *unanimous voting*),
- na tę samą klasę wskazała połowa plus przynajmniej jeszcze jeden z klasyfikatorów; jest to zwykle głosowanie większościowe (ang. *simple majority*),
- wybrana zostaje ta klasa, na którą wskazało najwięcej klasyfikatorów (ang. *plurality voting*).

Omówimy teraz każdy z typów głosowania. Reguła dla pierwszego z wymienionych, tj. dla głosowania większościowego, ma następującą postać:

$$d_{i^*} \rightarrow \sum_{j=1}^{N_{cl}} \mathbb{P}_{j,i^*} = \max_{i=1,2,\dots,N_c} \sum_{j=1}^{N_{cl}} \mathbb{P}_{j,i} \quad (8.2)$$

gdzie $\mathbb{P}_{j,i}$ reprezentuje wektor binarny, a d_i jest decyzją ze zbioru decyzji \mathcal{N}_d .

Dalej scharakteryzowano każdy z typów głosowania. Reguła dla pierwszego z wymienionych, tj. dla głosowania większościowego, jest optymalną regułą fuzji decyzji przy spełnieniu następujących warunków:

- liczba N_{cl} klasyfikatorów dla problemu z dwoma klasami jest nieparzysta,
- wyjścia klasyfikatorów są od siebie niezależne.

Warto również zauważyć, że w tych warunkach głosowanie większościowe i proste są identyczne.

Podejście oparte na głosowaniu większościowym, po spełnieniu pewnych warunków, umożliwia uzyskanie optymalnego rozwiązania (dowód zamieszczono m.in. w pracy [224]).

Głosowanie większościowe ważone

W przypadku, gdy jakość klasyfikatorów w systemie na poziomie niższym jest zróżnicowana, możliwe jest wykorzystanie alternatywy w postaci głosowania większościowego z wagami. W tym przypadku każdy z N_{cl} klasyfikatorów ma nadaną wagę w_{clj} , gdzie $j = 1, 2, \dots, N_{cl}$. Wartość wagi nadana każdemu j -temu klasyfikatorowi związana jest z jakością tego klasyfikatora. Biorąc pod uwagę to, że każdemu klasyfikatorowi przyporządkowano wagę, formuła (8.2) przyjmuje następującą zmodyfikowaną postać:

$$d_{i^*} \rightarrow \sum_{j=1}^{N_{cl}} b_j \mathbb{P}_{j,i^*} = \max_{i=1,2,\dots,N_c} \sum_{j=1}^{N_{cl}} w_{clj} \mathbb{P}_{j,i} \quad (8.3)$$

8. Poziom wysoki: fuzja decyzji

gdzie w_{clj} jest wagą nadaną każdemu j -temu klasyfikatorowi. W celu łatwiejszej interpretacji uzyskanych wyników możliwa jest normalizacja nadanych klasyfikatorom wag.

W przypadku fuzji decyzji bazującej na ważonej regule głosowania większościowego zasadne jest zadanie pytania o metodę nadawania wag poszczególnym algorytmom fuzji decyzji. Oczywiście jest, że algorytm, który dla konkretnego zadania będzie dokładniejszy, powinien otrzymać wyższe wagi od tych, których rozwiązania nie są tak dokładne. W pracy [224] przedstawiono dowód twierdzenia o optymalności reguły ważonego głosowania większościowego.

Algorytm BKS

Inne podejście do algorytmów fuzji decyzji zaproponowano w pracy [172], jest to algorytm BKS (ang. *Behavior Knowledge Space*). Istotą metody jest estymacja prawdopodobieństwa *a posteriori* dla wszystkich klas i wszystkich kombinacji głosów. Zadanie to rozwiązywane jest przez konstrukcję odpowiedniej tabeli (tzw. tabela BKS). Do tabeli tej wstawiana jest liczba powtórzeń wszystkich możliwych kombinacji decyzji, przy czym należy pamiętać, że do konstrukcji tabeli BKS wykorzystywany jest zbiór uczący. W kolejnym kroku na podstawie przygotowanej tabeli szacowane jest prawdopodobieństwo *a posteriori*. Szacunki te wykonywane są przez wyszukanie tych komórek, w których zapamiętano największą liczbę powtórzeń. Następnie dokonuje się ich zliczenia. Przykład działania metody został przedstawiony na rys. 8.9 (na podst. [331]).

| K_1 | K_2 | K_3 | c_1 | c_2 | c_3 |
|-------|----------|-------|-------|-------|-------|
| c_1 | c_1 | c_1 | 35 | 8 | 21 |
| c_1 | c_1 | c_2 | 13 | 30 | 4 |
| c_1 | c_1 | c_3 | 19 | 11 | 18 |
| c_1 | c_2 | c_1 | 5 | 15 | 28 |
| c_1 | c_2 | c_2 | 11 | 16 | 9 |
| | \vdots | | | | |
| c_3 | c_3 | c_3 | 20 | 21 | 2 |

Rys. 8.9. Ilustracja działania algorytmu BKS

Fuzja decyzji na podstawie rankingów

W metodzie bazującej na rankingach każdy z N_{cl} klasyfikatorów generuje zbiór klas na bazie zbioru dostępnych klas N_{cl} . Wygenerowany przez każdy j -ty klasyfikator zbiór klas jest podzbiorem wszystkich możliwych klas ze zbioru dostępnych klas. Następnie z tych zbiorów klas wygenerowanych przez każdy j -ty klasyfikator tworzony jest ranking. W rankingu tym wyniki posortowane są w pewnym porządku według stopnia przynależności tych klas do danego obiektu. Wyróżniamy dwie metody fuzji decyzji na podstawie tak przygotowanych rankingów [162]:

- redukcji klas (ang. *Class Set Reduction*),
- zmiana uporządkowania klas (ang. *Class Set Reordering*).

W pierwszej z metod, tj. w metodzie redukcji klas, dokonuje się sukcesywnej redukcji rozpatrywanych klas. Istotną kwestią w tym podejściu jest sposób redukcji klas. Redukcja powinna być wykonywana w ten sposób, by z jednej strony liczność zbioru klas była możliwie mała, z drugiej natomiast by podczas redukcji maksymalizować szanse pojawienia się poprawnego rozwiązania w końcowym rankingu. W metodzie tej wyróżniamy dwa podejścia:

- IoN (ang. *Intersection of Neighborhoods*); w podejściu tym budowane są rankingi klasyfikatorów i po ustaleniu progu te z nich, które nie przekroczyły go są z niego usuwane,
- UoN (ang. *Union of Neighborhoods*), buduje końcowy ranking przez sukcesywne włączanie kolejnych klasyfikatorów.

Zaznaczmy, że wymienione podejście UoN jest szczególnie przydatne wówczas, gdy algorytmy fuzji decyzji mają wyjścia różnego typu [270].

Druga metoda fuzji decyzji na podstawie rankingów bazuje na zmianie uporządkowania klas. Metoda ta umożliwia uzyskanie lepszych wyników w przypadku, gdy prawdziwa (poszukiwana) klasa znacznie odróżnia się od pozostałych [270]. Metoda ta charakteryzuje się tym, że modyfikuje wejściowe uporządkowanie klas tak, by prawdziwa cecha była jak najwyżej w rankingu. Wyróżniamy następujące metody do modyfikacji rankingu cech:

- preferencyjna (ang. *highest rank method*),
- Borda (ang. *Borda count method*),
- regresja logistyczna.

Pierwszym z podejść jest metoda preferencyjna (ang. *Highest Rank Method*). Jest to metoda, która nie wymaga etapu uczenia. Przy czym, charakteryzuje się również niską złożonością obliczeniową. Metoda składa się z kilku etapów. Pierwszy z nich związany jest z przygotowaniem wstępnego rankingu dla każdego z algorytmów fuzji decyzji. W kolejnym kroku wyznaczana jest wartość minimalna dla tych rankingów, a następnie jest ona przypisywana do każdego z tych rankingów. W następnym kroku dokonujemy kolejnej konstrukcji rankingu, tym razem uwzględniając wyliczoną w poprzednim kroku wartość. Ostateczna decyzja

8. Poziom wysoki: fuzja decyzji

podejmowana jest na podstawie rankingu zmodyfikowanego w opisany sposób. Wartość będąca na jego szczycie jest traktowana jako decyzja końcowa.

Cechą tej metody jest to, że wykorzystuje ona zalety każdego z algorytmów fuzji decyzji. Metoda sprawdza się głównie w przypadku systemów, w których wykorzystywana jest niewielka liczba algorytmów fuzji decyzji [270].

Kolejną metodą, w której edytuje się początkowo ustalony ranking jest metoda Borda (ang. *Borda count*). Istotą metody jest przekształcenie rankingów każdego z N_{cl} klasyfikatorów na jeden ranking, na podstawie którego wyznaczana jest końcowa decyzja. Cechą metody jest to, że nie usuwa się z analizy tych wyników, które źle rokują.

Przy czym, w metodzie przyjmujemy, że algorytmy fuzji decyzji są od siebie niezależne. Warunek ten w praktyce nie zawsze jest spełniony. Kolejną wadą metody jest to, że wyniki uzyskane dla wszystkich algorytmów traktowane są jednakowo i nie uwzględniają jakości każdego z algorytmów osobno.

Do rozwiązania zadania fuzji decyzji bazującej na podejściu, w którym dokonujemy edycji początkowo przygotowanego rankingu jest metoda bazująca na regresji logistycznej. Metoda wprowadza pewne ulepszenia do uprzednio omówionej metody Borda. Wadą metody Borda jest to, że brakuje w niej możliwości uwzględnienia jakości każdej z metod. Z kolei w metodzie bazującej na regresji logistycznej jest to możliwe przez dodanie do każdego z klasyfikatorów wag. Ze szczegółami metody można zapoznać się m.in. w pracy [270].

8.3.2. Fuzja decyzji na podstawie wielkości ciągłych

Na wyjściu klasyfikatora mogą pojawić się nie tylko wartości liczbowe dyskretne, ale również liczbowe ciągłe. W metodach, w których fuzja decyzji bazuje na danych typu ciągłego, wykorzystuje się dedykowaną tablicę, tj. **profil decyzyjny** (ang. *Decision Profile*, DP) [224]. Poszczególne elementy wzoru (8.4) określają stopień przynależności obiektu do klasy.

Profil decyzyjny można utworzyć z wyjść każdego z algorytmów fuzji właściwości i dla zadanego wejściowego wektora cech. Profil decyzyjny można reprezentować następującą tablicą:

$$\begin{array}{c}
 \text{Ocena pochodząca od } N_{cl} \\
 \text{klasyfikatorów dla } i\text{-tej klasy} \\
 \downarrow \\
 DP(\mathbf{x}) = \begin{bmatrix} \mathbb{P}_{1,1} & \mathbb{P}_{1,1} & \cdots & \mathbb{P}_{1,N_c} \\ \mathbb{P}_{2,1} & \mathbb{P}_{2,2} & \cdots & \mathbb{P}_{2,N_c} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{P}_{N_{cl},1} & \mathbb{P}_{N_{cl},i} & \cdots & \mathbb{P}_{N_{cl},N_c} \end{bmatrix} \leftarrow \begin{array}{l} \text{Ocena pochodząca} \\ \text{od } j\text{-tego klasyfikatora} \\ \text{dla } N_c \text{ klas} \end{array} \quad (8.4)
 \end{array}$$

W tym przypadku elementy w DP (8.4) mogą przyjmować wartości rzeczywiste z przedziału $\langle 0, 1 \rangle$.

Wśród metod bazujących na profilu decyzyjnym wyróżniamy dwie metody. Pierwsza z metod wykorzystuje informację zawartą w kolumnach tablicy decyzyjnej, która dotyczy tego, w jakim stopniu każda z klas jest wskazywana przez każdy z algorytmów.

Podjęcie to charakteryzuje się tym, że możliwe jest w nim rozróżnienie na klasy (ang. *Class-conscious Methods*). W metodach tych dokonuje się pewnych operacji na tablicy decyzji w celu określenia, w jakim stopniu konkretna decyzja jest wspierana przez każdy z algorytmów fuzji decyzji.

Fuzja decyzji z rozróżnieniem na klasy

W podrozdziale przedstawiono najczęściej wykorzystywane metody analizy profilu decyzyjnego DP. Jedną z nich jest:

$$rule_i = \sum_{j=1}^{N_{cl}} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.5)$$

kolejnymi natomiast:

$$rule_i = \prod_{j=1}^{N_{cl}} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.6)$$

oraz:

$$rule_i = \frac{1}{N_{cl}} \sum_{j=1}^{N_{cl}} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.7)$$

Jeżeli chcemy uwzględnić klasyfikatory w fuzji decyzji w różnym stopniu, to wówczas zamiast średniej, można wykorzystać średnią ważoną:

$$rule_i = \sum_{j=1}^{N_{cl}} w_{clj} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.8)$$

Do wyliczenia wartości wskaźnika *rule* na podstawie profilu decyzyjnego można wykorzystać również:

$$rule_i = \min_{j=1,2,\dots,N_{cl}} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.9)$$

$$rule_i = \max_{j=1,2,\dots,N_{cl}} \mathbb{P}_{j,i} \quad i = 1, 2, \dots, N_c \quad (8.10)$$

8. Poziom wysoki: fuzja decyzji

$$rule_i = \text{mediana}(\mathbb{P}_{j,i}) \quad i = 1, 2, \dots, N_c, \quad j = 1, 2, \dots, N_{cl} \quad (8.11)$$

Fuzja decyzji bez rozróżnienia na klasy

W drugim z podejść fuzja decyzji bazuje na analizie całej tablicy decyzyjnej (8.4), tj. bez rozróżniania na poszczególne klasy. W metodzie tej nie uwzględnia się informacji o tym, w jakim stopniu wykorzystane algorytmy klasyfikacji wskazują na konkretną decyzję (ang. *Class-Indifferent Methods*).

Podstawowe metody używane w tym podejściu, które należy wyróżnić, to **szablony decyzyjne** (ang. *Decision Templates*, DT) oraz metoda bazująca na teorii **Dempstera–Schafera** (nazywana też teorią funkcji przekonania lub matematyczną teorią ewidencji).

Pierwsza z metod do rozwiązywania zadania fuzji decyzji wykorzystuje specjalnie przygotowane szablony, które tworzone są z wykorzystaniem zbioru uczącego. Szablon decyzyjny konstruowany jest na podstawie tablicy decyzyjnej z wykorzystaniem następującej zależności [224]:

$$DT_i = \frac{1}{N_{lei}} \sum_{j=1}^{N_{le}} DP(TS_{N_{lej}}). \quad (8.12)$$

gdzie $i = 1, 2, \dots, N_c$, N_{lei} reprezentuje liczbę elementów w zbiorze uczącym $TS_{N_{lej}}$, które wskazują na i -tą klasę c . Efektem zastosowania zależności (8.12) do profilu decyzyjnego jest szablon decyzyjny w formie:

$$DT = \begin{bmatrix} dt_{1,1} & dt_{1,2} & \cdots & dt_{1,N_c} \\ dt_{2,1} & dt_{2,2} & \cdots & dt_{2,N_c} \\ \vdots & \vdots & \ddots & \vdots \\ dt_{N_{cl},1} & dt_{N_{cl},2} & \cdots & dt_{N_{cl},N_c} \end{bmatrix} \quad (8.13)$$

Istotą metody jest wyznaczenie dla każdej z klas najbardziej reprezentatywnego szablonu decyzyjnego DT, który następnie porównywany jest z wyznaczonym wcześniej profilem decyzyjnym. W tym celu należy określić miary podobieństwa. W metodzie bazującej na szablonach decyzyjnych, najbardziej typowe miary to:

$$rule_i = 1 - \frac{1}{N_{cl}} N_c \sum_{j=1}^{N_{cl}} \sum_{i=1}^{N_c} (dt_{j,i} - \mathbb{P}_{j,i})^2 \quad (8.14)$$

oraz jej modyfikacja w postaci:

$$rule_i = \sum_{j=1}^{N_{cl}} \sum_{i=1}^{N_c} (dt_{j,i} - \mathbb{P}_{j,i})^2 \quad (8.15)$$

Kolejną miarą jest symetryczna różnica (ang. *symmetric difference*):

$$rule_i = 1 - \frac{1}{N_{cl}N_c} \sum_{j=1}^{N_{cl}} \sum_{i=1}^{N_c} \max \left(\min \left(dt_{j,i}(1 - \mathbb{P}_{j,i}) \right), \min \left((1 - \mathbb{P}_{j,i}), \mathbb{P}_{j,i} \right) \right) \quad (8.16)$$

W każdym z przypadków, tj. (8.14)–(8.16) wybierana jest ta decyzja, dla której wartość podanych miar jest największa.

Podstawą do fuzji decyzji na bazie teorii Dempstera–Schafera jest wyznaczenie funkcji przekonania (lub ewidencji) z następującej zależności:

$$\text{belief}(\mathbb{P}_{j,i}) = \frac{\mathcal{X}_{j,i} \prod_{j \neq i} (1 - \mathcal{X}_{j,i})}{1 - \mathcal{X}_{j,i} \left(1 - \prod_{j \neq i} (1 - \mathcal{X}_{j,i}) \right)}, \quad (8.17)$$

gdzie $\mathcal{X}_{j,i}$ jest wyznaczane z kolei z zależności:

$$\mathcal{X}_{j,i} = \frac{1 + \left(\|\text{DT}_{i,j} - \mathbb{P}_{j,i}^2\| \right)^{-1}}{\sum_{i=1}^{N_c} \left(1 + \|\text{DT}_{j,i} - \mathbb{P}_{j,i}\| \right)^{-1}} \quad (8.18)$$

W zależności (8.18) różnice w liczniku i mianowniku przekształcamy do postaci podobieństw. Wyrażenie w mianowniku jest czynnikiem normalizującym.

Dysponując wyznaczonym przekonaniem oraz wykorzystując regułę kombinacji Dempstera (ang. *Dempster's rule of combination*), możliwe jest wyznaczenie ostatecznej decyzji:

$$rule_i = \prod_{j=1}^{N_{cl}} \text{belief}_i(\mathbb{P}_{j,i}) \quad i = 1, 2, \dots, N_c \quad (8.19)$$

Część III

Modelowanie układu ruchu człowieka

9. Ruch ciała sztywnego

*Zmiana może być powodowana
tylko przez ruch.*

Aldous Huxley

9.1. Wstęp

Rozdział ten rozpoczyna drugą część książki poświęconą modelowaniu i analizie ruchu człowieka. W tej części omówione zostaną m.in. zagadnienia związane z metodami reprezentacji orientacji obiektu w przestrzeni, a także z algorytmami estymacji tej orientacji w przestrzeni na bazie danych pomiarowych. Również w tej części omówione zostaną zagadnienia związane z metodami estymacji położenia obiektu w przestrzeni oraz algorytmy śledzenia jego zmian.

Kolejna grupa zagadnień, która zostanie omówiona w tej części pracy dotyczy problemów związanych z analizą biomechaniczną ludzkiego ciała oraz modelowania jego kinematyki. W części tej przedstawione zostaną modele oraz algorytmy wykorzystywane w tych zadaniach.

Ostatnia grupa zagadnień, które zostaną omówione w tej części pracy, dotyczyć będzie zagadnień estymacji ruchu człowieka na bazie danych pomiarowych.

9.2. Ciało sztywne w przestrzeni trójwymiarowej

Rozważania w tej części pracy rozpoczniemy od omówienia podstawowych zagadnień dotyczących **ciała sztywnego** (ang. *rigid body*) oraz **ruchu ciała sztywnego** w przestrzeni \mathbb{R}^3 .

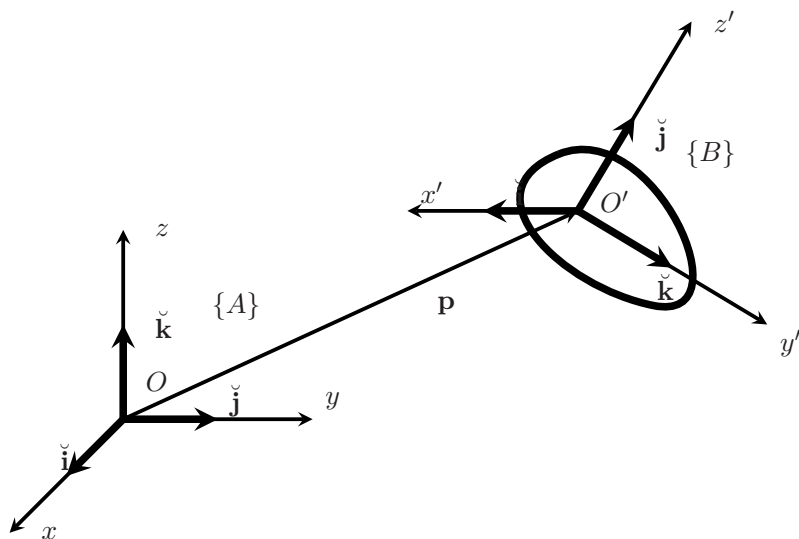
Każdy obiekt fizyczny istnieje w **czasoprzestrzeni** złożonej z przestrzeni i czasu [398]. Przez przestrzeń (przestrzeń fizyczną) będziemy rozumieli zbiór punktów, których położenie względem zadanego prawoskrętnego układu współrzędnych kartezjańskich można określić, wykorzystując trójkę liczb rzeczywistych. Przestrzeń tę będziemy utożsamiać z przestrzenią euklidesową \mathbb{R}^3 , dla której możliwe są operacje mnożenia skalarnego i wektorowego. Drugim składnikiem czasoprzestrzeni jest czas (czas fizyczny), który jest zbiorem chwil powiązanych ze zbiorem liczb rzeczywistych \mathbb{R} .

9. Ruch ciała sztywnego

Umieścimy teraz w czasoprzestrzeni układ fizyczny. Układem fizycznym jest obiekt fizyczny lub zbiór obiektów fizycznych. Przykładem układu fizycznego jest punkt materialny. Dla układu fizycznego **przestrzenią konfiguracyjną** jest przestrzeń \mathbb{R}^3 . Przestrzenią konfiguracyjną nazywamy zbiór wszystkich możliwych stanów układu fizycznego w przestrzeni. W przestrzeni \mathbb{R}^3 stan układu fizycznego opisany jest jego **położeniem** i **orientacją**. Położeniem jest umiejscowienie układu fizycznego w przestrzeni względem wybranego układu współrzędnych. Orientacja określa położenie pomiędzy dwoma układami współrzędnych, tj. układem odniesienia i **układem obiektu**, którym może być np. punkt materialny. **Układem odniesienia** jest układ, względem którego określa się położenie lub zmianę położenia oraz orientację danego ciała sztywnego.

Położenie i orientacja ciała sztywnego w \mathbb{R}^3

Do pełnego opisu ciała sztywnego (układu fizycznego) w układzie współrzędnych $\{B\}$ w przestrzeni \mathbb{R}^3 względem pewnego układu odniesienia $\{A\}$ należy ustalić jego orientację (ang. *orientation*) oraz położenie (ang. *position*) (rys. 9.1). Ustaloną orientację i położenie obiektów w przestrzeni nazywamy **pozycją** (ang. *pose*).



Rys. 9.1. Położenie i orientacja ciała sztywnego w przestrzeni

Przykład układu odniesienia przedstawiono na rys. 9.1. Początek tego układu oznaczono jako O . Wektory $\mathbf{i}, \mathbf{j}, \mathbf{k}$ to jednostkowe wektory ortogonalne osi x, y, z .

9.2. Ciało sztywne w przestrzeni trójwymiarowej

Ciału sztywnemu przyporządkujemy natomiast układ współrzędnych o początku w punkcie O' i wektorach jednostkowych $\check{\mathbf{i}}', \check{\mathbf{j}}', \check{\mathbf{k}}'$. Układ ten jest nieruchomy względem ciała sztywnego.

Do opisu położenia i orientacji ciała sztywnego względem układu odniesienia wykorzystujemy dwa przekształcenia:

- translacja,
- obrót (rotacja).

Przekształcenia te nazywamy przekształceniami sztywnymi. Translację wiążemy z wektorem położenia (wektor przesunięcia) i oznaczamy przez \mathbf{p} . Wykorzystując wektor położenia \mathbf{p} , położenie punktu O' względem układu odniesienia zapisujemy jako:

$$\mathbf{p} = p_x \check{\mathbf{i}} + p_y \check{\mathbf{j}} + p_z \check{\mathbf{k}} \quad (9.1)$$

Ponieważ wektor położenia nie opisuje pełnego usytuowania ciała sztywnego w przestrzeni, wprowadzamy drugą operację, tj. obrót. Z obrotem (rotacją) powiązana jest **macierz rotacji (obrotu)**, którą oznaczymy jako \mathbf{R} . Do budowy macierzy rotacji wykorzystujemy ortogonalne wektory jednostkowe związane z układem ciała [219]:

$$\mathbf{R} = [\check{\mathbf{i}}', \check{\mathbf{j}}', \check{\mathbf{k}}'] = \begin{bmatrix} i'_x & j'_x & k'_x \\ i'_y & j'_y & k'_y \\ i'_z & j'_z & k'_z \end{bmatrix} \quad (9.2)$$

Macierz rotacji \mathbf{R} , która określa wzajemną orientację dwóch obiektów w przestrzeni \mathbb{R}^3 , spełnia następujące warunki:

- \mathbf{R} jest ortogonalna tj. $\mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{I}$,
- $\det(\mathbf{R}) = +1$.

Macierze spełniające podane warunki tworzą specjalną grupę ortogonalną rzędu 3.

Definicja 9.2.1 (Specjalna grupa ortogonalna $\mathbf{SO}(3)$). *Specjalna grupa ortogonalna $\mathbf{SO}(3)$ jest grupą macierzy ortogonalnych stopnia 3 o wyznaczniku +1:*

$$\mathbf{SO}(3) = \{\mathbf{R} \in \mathbb{R}^{3 \times 3} : \mathbf{R}^T \mathbf{R} = \mathbf{I}, \det(\mathbf{R}) = +1\} \quad (9.3)$$

Na specjalnej grupie ortogonalnej określona jest algebra Liego $\mathfrak{so}(3)$. Algebra Liego $\mathfrak{so}(3)$ rozpinana jest przez generatory grupy $\mathbf{SO}(3)$ [218].

Definicja 9.2.2 (Algebra Liego dla specjalnej grupy ortogonalnej $\mathbf{SO}(3)$). *Algebra Liego dla $\mathbf{SO}(3)$ jest zbiorem macierzy takich, że*

$$\mathfrak{so}(3) = \{\mathbf{S} \in \mathbb{R}^{3 \times 3} : \mathbf{S}^T = -\mathbf{S}\} \quad (9.4)$$

gdzie \mathbf{S} jest macierzą antysymetryczną (macierzą skośnie symetryczną).

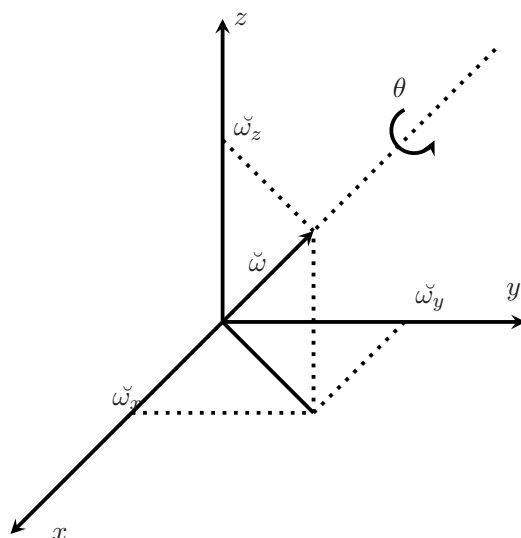
9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3

W poprzednim podpunkcie wprowadziliśmy macierz rotacji jako narzędzie opisu orientacji obiektu w przestrzeni. Jedną z wad tego opisu jest to, że jest on nadmiarowy [219]. W przypadku macierzy rotacji elementów, które wykorzystywane są do opisu obiektu w przestrzeni, jest dziewięć. Przy czym elementy tej macierzy są ze sobą powiązane. Dlatego, oprócz reprezentacji z wykorzystaniem macierzy obrotu, opracowano inne metody reprezentacji rotacji. Metody te wyróżniają się m.in. tym, że do opisu wykorzystuje się mniejszą liczbę zmiennych. Zostaną one omówione w kolejnych podrozdziałach.

9.3.1. Reprezentacja oś–kąt

Z twierdzenia Eulera o obrotach [89] wynika, że możliwe jest przejście z jednego położenia kąowego do drugiego za pomocą jednego obrotu wokół pewnej osi. Wnioski z tego twierdzenia wykorzystano do zaproponowania reprezentacji typu oś–kąt.

W reprezentacji oś–kąt, oś rotacji reprezentowana jest przez wektor jednostkowy osi obrotu $\check{\omega}$, który jest opisany w układzie odniesienia x, y, z i kąta rotacji θ [219, 300]. Wektor osi rotacji oraz kąt rotacji przedstawiono na rys. 9.2.



Rys. 9.2. Wektor jednostkowy osi rotacji oraz kąt rotacji [377]

9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3

Macierz rotacji \mathbf{R} oraz reprezentacja typu ós-kąt są ze sobą powiązane. Zapiszmy macierz rotacji \mathbf{R} w następującej postaci:

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (9.5)$$

Zależności opisujące związek pomiędzy macierzą \mathbf{R} a wektorem jednostkowym osi rotacji $\check{\omega}$ i kątem rotacji θ są następujące [219]:

$$\theta = \arccos \left(\frac{r_{11} + r_{22} + r_{33} - 1}{2} \right) = \arccos \left(\frac{\text{trace}(\mathbf{R}) - 1}{2} \right) \quad (9.6)$$

$$\check{\omega} = \frac{1}{2 \sin \theta} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (9.7)$$

dla $\sin \theta \neq 0$. Dla składowych wektora $\check{\omega}$ zachodzi:

$$\check{\omega}_x^2 + \check{\omega}_y^2 + \check{\omega}_z^2 = 1 \quad (9.8)$$

Reprezentacja typu ós-kąt nie jest jednoznaczna, ponieważ obrót zarówno o kąt θ , jak i $-\theta$ wokół osi $\check{\omega}$ daje ten sam wynik. Jest tak dlatego, że dla $\sin \theta = 0$ macierz \mathbf{R} jest macierzą jednostkową. Stąd wynika nieokreśloność jednostkowego wektora osi rotacji $\check{\omega}$ [377].

9.3.2. Kąty Eulera

Kąty Eulera są kolejną metodą reprezentacji rotacji ciała sztywnego w przestrzeni. Kąty te odpowiadają złożeniu trzech elementarnych rotacji wokół osi x, y, z . Warto zaznaczyć, że opis taki nie jest jednoznaczny, tzn. istnieje dwanaście różnych zbiorów kątów Eulera (tabela 9.1). Wynika to z liczby możliwych sekwencji elementarnych rotacji. Jednym z najczęściej stosowanych złożań są kąty Eulera ZYZ.

Kąty Eulera ZYZ typowo oznaczane są jako α, β, γ (odpowiednio są to przechylenie, pochylenie, odchylenie) [322]. Wypadkowa rotacja, która im odpowiada, jest wynikiem następującej sekwencji rotacji elementarnych:

1. obrót układu o kąt α wokół osi z , tj.

$$\mathbf{R}_z(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

9. Ruch ciała sztywnego

2. obrót układu o kąt β wokół osi y , tj.

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix}$$

3. obrót układu o kąt γ wokół osi z , tj.

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Tabela 9.1. Zestawienie zbioru kątów Eulera

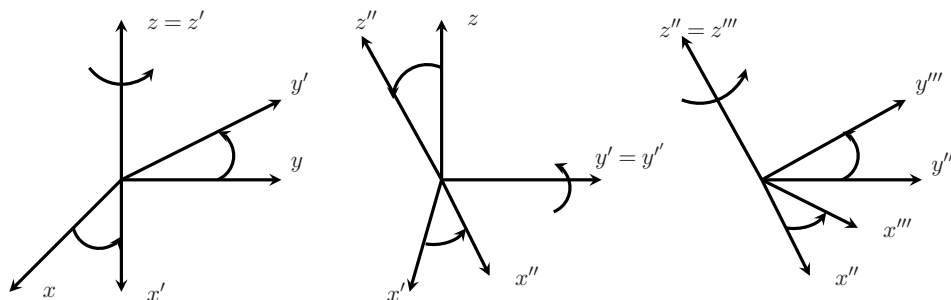
| Lp. | Pierwsza rotacja | Druga rotacja | Trzecia rotacja | Sekwencja rotacji |
|-----|------------------|---------------|-----------------|-------------------|
| 1 | | | Z | X-Y-Z |
| 2 | X | Y | X | X-Y-X |
| 3 | | | X | X-Z-X |
| 4 | | Z | Y | X-Z-Y |
| 5 | | | X | Y |
| 6 | Y | | Z | Y-X-Z |
| 7 | | Z | X | Y-Z-X |
| 8 | | | Y | Y-Z-Y |
| 9 | | | X | Y |
| 10 | Z | | Z | Z-X-Z |
| 11 | | Y | Z | Z-Y-Z |
| 12 | | | X | Z-Y-X |

Składając powyższe macierze rotacji, otrzymujemy następującą macierz rotacji:

$$\begin{aligned} \mathbf{R}_{zyz}(\alpha, \beta, \gamma) &= \mathbf{R}_z(\alpha) \mathbf{R}_y(\beta) \mathbf{R}_z(\gamma) = \\ &= \begin{bmatrix} \cos \alpha \cos \beta \cos \gamma - \sin \alpha \sin \gamma & -\cos \alpha \cos \beta \sin \gamma - \sin \alpha \cos \gamma & \cos \alpha \sin \beta \\ \sin \alpha \cos \beta \cos \gamma + \cos \alpha \sin \gamma & -\sin \alpha \cos \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \sin \beta \\ -\sin \beta \cos \gamma & \sin \beta \sin \gamma & \cos \beta \end{bmatrix} \end{aligned} \quad (9.9)$$

Zauważmy, że ponieważ macierz (9.9) jest złożeniem elementarnych rotacji, to jest ona również macierzą ortogonalną. Na rysunku 9.3 zilustrowano sekwencję rotacji elementarnych dla kątów Eulera ZYZ.

9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3



Rys. 9.3. Sekwencja rotacji elementarnych dla kątów Eulera ZYZ

Możliwe jest również działanie odwrotne, tj. wyznaczenie kątów α, β, γ na podstawie macierzy (9.9). Zapiszmy macierz (9.9) w postaci:

$$\mathbf{R}_{zyz}(\alpha, \beta, \gamma) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (9.10)$$

Zakładając, że $\sin \beta \neq 0$, otrzymamy:

$$\beta = \text{atan2} \left(\sqrt{r_{31}^2 + r_{32}^2}, r_{33} \right) \quad (9.11)$$

$$\alpha = \text{atan2} \left(\frac{r_{23}}{\sin(\beta)}, \frac{r_{13}}{\sin(\beta)} \right) \quad (9.12)$$

$$\gamma = \text{atan2} \left(\frac{r_{32}}{\sin(\beta)}, -\frac{r_{31}}{\sin(\beta)} \right), \quad (9.13)$$

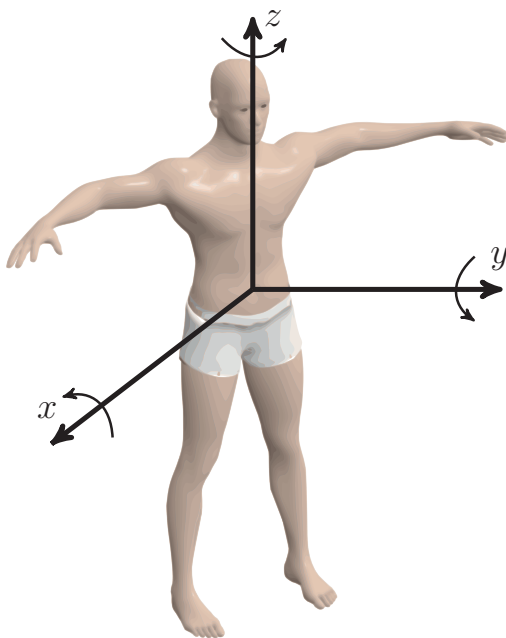
gdzie

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{dla } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{dla } y \geq 0, x < 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{dla } y < 0, x < 0 \\ +\frac{\pi}{2} & \text{dla } y > 0, x = 0 \\ -\frac{\pi}{2} & \text{dla } y < 0, x = 0 \\ \text{nieokreślone} & \text{dla } y = 0, x = 0 \end{cases} \quad (9.14)$$

9. Ruch ciała sztywnego

Rozwiązania równań (9.12) oraz (9.13) stają się niejednoznaczne dla $\sin \beta = 0$. Niejednoznaczność w tych zależnościach wynika z tego, że gdy kąt β przyjmuje wartości $\beta = 0$ lub $\beta = \pi$, to obroty o kąty α i γ następują wokół równoległych osi [219]. W literaturze problem ten znany jest pod nazwą **blokadę przegubów** (ang. *Gimbal Lock*).

Inną popularną sekwencją elementarnych rotacji z wykorzystaniem kątów Eulera jest sekwencja RPY (ang. *Roll-Pitch-Yaw*). Kąty wchodzące w skład sekwencji RPY są oznaczane jako φ , ϑ , ψ . W polskiej nomenklaturze kąty RPY nazywane są kołysaniem, kiwaniem, myszkowaniem [398]. Na rysunku 9.4 przedstawiono ilustrację kątów Eulera RPY.



Rys. 9.4. Ilustracja kątów Eulera RPY

Kąty Eulera RPY otrzymuje się przez złożenie następujących rotacji wykonanych względem układu odniesienia [219]:

1. obrót układu o kąt φ wokół osi x , tj.

$$\mathbf{R}_x(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix}$$

9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3

2. obrót układu o kąt ϑ wokół osi y , tj.

$$\mathbf{R}_y(\vartheta) = \begin{bmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{bmatrix}$$

3. obrót układu o kąt ψ wokół osi z , tj.

$$\mathbf{R}_z(\psi) = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Efekt złożenia podanych macierzy jest następująca macierz:

$$\begin{aligned} \mathbf{R}_{\text{RPY}}(\varphi, \vartheta, \psi) &= \mathbf{R}_z(\varphi) \mathbf{R}_y(\vartheta) \mathbf{R}_x(\psi) = \\ &= \begin{bmatrix} \cos \varphi \cos \vartheta & \cos \varphi \sin \vartheta \sin \psi - \sin \varphi \cos \psi & \cos \varphi \sin \vartheta \cos \psi + \sin \varphi \sin \psi \\ \sin \varphi \cos \vartheta & \sin \varphi \sin \vartheta \sin \psi + \cos \varphi \cos \psi & \sin \varphi \sin \vartheta \cos \psi - \cos \varphi \sin \psi \\ -\sin \vartheta & \cos \vartheta \sin \psi & \cos \vartheta \cos \psi \end{bmatrix} \end{aligned} \quad (9.15)$$

Podobnie jak dla kątów Eulera ZYZ, również dla kątów RPY możliwe jest wyznaczenie kątów φ, ϑ, ψ na podstawie otrzymanej macierzy rotacji (9.15). Zapiszmy macierz (9.15) w postaci:

$$\mathbf{R}_{\text{RPY}}(\varphi, \vartheta, \psi) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (9.16)$$

Wartości kątów $(\varphi, \vartheta, \psi)$ wyznaczamy z następujących zależności [219]:

$$\varphi = \text{atan2}(r_{21}, r_{11}) \quad (9.17)$$

$$\vartheta = \text{atan2}\left(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2}\right) \quad (9.18)$$

$$\psi = \text{atan2}(r_{32}, r_{33}) \quad (9.19)$$

dla $\vartheta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Jeżeli kąt $\vartheta \in \left(\frac{\pi}{2}, 3\frac{\pi}{2}\right)$, to wówczas:

$$\varphi = \text{atan2}(-r_{21}, -r_{11}) \quad (9.20)$$

$$\vartheta = \text{atan2}\left(-r_{31}, -\sqrt{r_{32}^2 + r_{33}^2}\right) \quad (9.21)$$

$$\psi = \text{atan2}(-r_{32}, -r_{33}) \quad (9.22)$$

9. Ruch ciała sztywnego

Kąty Eulera ZYZ oraz RPY są przykładami lokalnych trójelementowych reprezentacji grupy $\mathbf{SO}(3)$. Również w przypadku tej reprezentacji występują osobliwości (możliwość wystąpienia blokady przegubu), których nie można usunąć [219].

9.3.3. Kwanternion

W przedstawionych dotychczas reprezentacjach występują pewne osobliwości, które utrudniają ich wykorzystanie do opisu orientacji obiektu w przestrzeni. Alternatywnym opisem do tych już omówionych jest opis w formie **kwaternionu**. Formalnie kwaternion to czwórka liczb rzeczywistych lub, równoważnie, para złożona ze składowych i wektora o trzech składowych. Kwanternion jest uogólnieniem liczb zespolonych [219]. Tak jak liczby zespolone można wykorzystać do opisu obrotów na płaszczyźnie, tak kwaterniony stosujemy do opisu rotacji w przestrzeni \mathbb{R}^3 . W odróżnieniu od kątów Eulera, będących trójelementową reprezentacją grupy $\mathbf{SO}(3)$, w przypadku kwaternionów nie występuje problem osobliwości. Zaznaczmy, że kwaterniony umożliwiają również globalną parametryzację grupy $\mathbf{SO}(3)$. Warto jednak mieć na uwadze to, że do opisu rotacji z wykorzystaniem kwaternionów wykorzystywane są cztery zmienne [219].

Omówimy teraz związek pomiędzy reprezentacją w postaci kwaternionu z reprezentacją oś-kąt. Formalnie kwaternion zawiera te same informacje, które zawiera reprezentacja oś-kąt, jednak są one zapisane w innej, wygodniejszej do wykorzystania postaci [322]. Kwanternion jest na przykład reprezentacją wygodniejszą do wykonywania operacji składania obrotów. Zapiszmy kwaternion w formie wektora o czterech składowych:

$$\mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}, \quad q_i \in \mathbb{R}, \quad i = 0, \dots, 3 \quad (9.23)$$

gdzie q_0 jest skalar, a $[q_1, q_2, q_3]^T$ to wektor. Na rysunku 9.5 przedstawiono graficzną ilustrację kwaternionu.

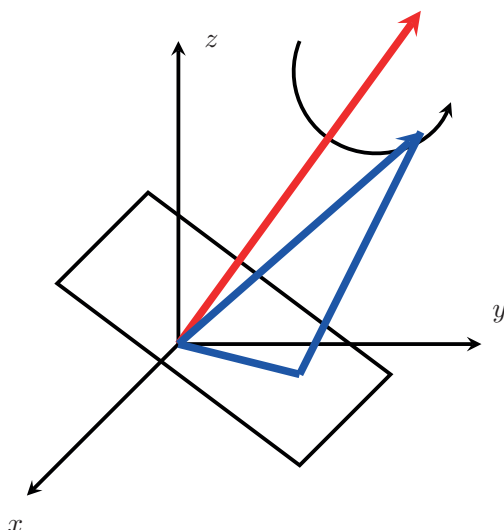
Kwanternion jednostkowy jest szczególnym przypadkiem kwaternionu, który spełnia zależność:

$$\|\mathbf{q}\| = \sqrt{q_0 + q_1 + q_2 + q_3} = 1 \quad (9.24)$$

Warto zaznaczyć, że dla $\check{\omega} \in \mathbb{R}^3$ oraz $\theta \in \mathbb{R}$ zachodzi następująca zależność [377]:

$$\mathbf{q} = \begin{bmatrix} \cos \frac{\theta}{2} \\ \check{\omega} \cos \frac{\theta}{2} \end{bmatrix} \in \mathbb{R}^4 \quad (9.25)$$

9.3. Inne reprezentacje orientacji ciała sztywnego w \mathbb{R}^3



Rys. 9.5. Ilustracja obrotu z wykorzystaniem parametrów kwaternionu

Interpretacja wielkości $\check{\omega}$ oraz θ została przedstawiona w podrozdziale 9.3.1. Oczywiście, jest możliwe również wyznaczenie wielkości $\check{\omega} \in \mathbb{R}^3$ oraz $\theta \in \mathbb{R}$ na podstawie kwaternionu \mathbf{q} [377]:

$$\theta = 2 \cos^{-1} q_0 \quad (9.26)$$

$$\check{\omega} = \begin{cases} \frac{[q_1, q_2, q_3]^T}{\sin\left(\frac{\theta}{2}\right)} & \text{jeżeli } \theta \neq 0 \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (9.27)$$

Istnieje również związek pomiędzy kwaternionem \mathbf{q} a macierzą rotacji \mathbf{R} (patrz rozdz. 9) [377]:

$$\mathbf{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_0q_2 + q_1q_3) \\ 2(q_0q_3 + q_1q_2) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_0q_1 + q_2q_3) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (9.28)$$

Oczywiście mamy również zależność odwrotną, tj. pomiędzy macierzą rotacji \mathbf{R} (rozdz. 9) a kwaternionem \mathbf{q} :

$$q_0 = \frac{1}{2} \sqrt{1 + r_{11} + r_{22} + r_{33}} \quad (9.29)$$

9. Ruch ciała sztywnego

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \frac{1}{4q_0} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (9.30)$$

W przeciwieństwie do przedstawionych wcześniej reprezentacji, tj. w postaci kątów Eulera i oś-kąt w reprezentacji rotacji w postaci kwaternionu, nie występuje problem punktów osobliwych.

Podstawowe operacje, jakie mogą być wykonane na kwaternionach, tj. dodawanie, odejmowanie, mnożenie oraz operacja sprzężenia, zdefiniowano zależnościami (9.31)–(9.35).

Algebra kwaternionów

Rozważmy dwa kwaterniony:

$$\mathbf{q}^a = \begin{bmatrix} q_1^a \\ q_2^a \\ q_3^a \\ q_4^a \end{bmatrix}, \quad \mathbf{q}^b = \begin{bmatrix} q_1^b \\ q_2^b \\ q_3^b \\ q_4^b \end{bmatrix} \quad (9.31)$$

Operacje dodawania i odejmowania kwaternionów wyznaczamy w następujący sposób [78]:

$$q^a \pm q^b = \begin{bmatrix} q_1^a \\ q_2^a \\ q_3^a \\ q_4^a \end{bmatrix} \pm \begin{bmatrix} q_1^b \\ q_2^b \\ q_3^b \\ q_4^b \end{bmatrix} = \begin{bmatrix} q_1^a \pm q_1^b \\ q_2^a \pm q_2^b \\ q_3^a \pm q_3^b \\ q_4^a \pm q_4^b \end{bmatrix} \quad (9.32)$$

Jak można zauważyć, dodawanie kwaternionów jest zwykłym dodawaniem czwórek liczb. Mnożenie kwaternionów jest określone wzorem [409]:

$$q^a \otimes q^b = \begin{bmatrix} q_1^a \\ q_2^a \\ q_3^a \\ q_4^a \end{bmatrix} \otimes \begin{bmatrix} q_1^b \\ q_2^b \\ q_3^b \\ q_4^b \end{bmatrix} = \begin{bmatrix} q_1^a q_1^b - q_2^a q_2^b - q_3^a q_3^b - q_4^a q_4^b \\ q_1^a q_2^b + q_2^a q_1^b - q_3^a q_4^b + q_4^a q_3^b \\ q_1^a q_3^b + q_2^a q_4^b + q_3^a q_1^b - q_4^a q_2^b \\ q_1^a q_4^b - q_2^a q_3^b + q_3^a q_2^b + q_4^a q_1^b \end{bmatrix} \quad (9.33)$$

Mnożenie kwaternionów jest działaniem łącznym, ale jest nieprzemienne, tj. $q^a \otimes q^b \neq q^b \otimes q^a$. Operacja sprzężenia jest definiowana następująco [409]:

$$\mathbf{q}^* = \begin{bmatrix} q_1 \\ -q_2 \\ -q_3 \\ -q_4 \end{bmatrix} \quad (9.34)$$

dla kwaternionu postaci:

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \quad (9.35)$$

9.3.4. Współrzędne wykładnicze

Do opisu rotacji w przestrzeni trójwymiarowej można wykorzystać również współrzędne wykładnicze (ang. *the exponential coordinates*).

Pewien punkt \mathbf{P} w przestrzeni \mathbb{R}^3 , opisany współrzędnymi wektora \mathbf{p} , porusza się wokół osi rotacji $\check{\omega}$. Równanie opisujące prędkość poruszającego się punktu \mathbf{P} wokół osi określonej przez oś obrotu $\check{\omega}$ o ustalony kąt $\theta \in \mathbb{R}$ opisuje równanie [300]:

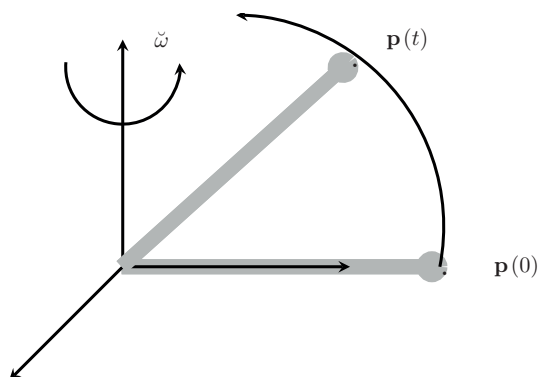
$$\dot{\mathbf{p}}(t) = \check{\omega} \times \mathbf{p}(t) = [\check{\omega}]\mathbf{p}(t) \quad (9.36)$$

gdzie $[\check{\omega}]$ jest macierzą antysymetryczną (skośnie symetryczną) postaci [300]:

$$[\check{\omega}] = \begin{bmatrix} 0 & -\check{\omega}_3 & \check{\omega}_2 \\ \check{\omega}_3 & 0 & -\check{\omega}_1 \\ -\check{\omega}_2 & \check{\omega}_1 & 0 \end{bmatrix} \quad (9.37)$$

Opisana sytuacja została zilustrowana na rys. 9.6. Równanie (9.36) jest równaniem różniczkowym, którego rozwiązanie ma postać:

$$\mathbf{p}(t) = e^{[\check{\omega}]t} \mathbf{p}(0) \quad (9.38)$$



Rys. 9.6. Obrót punktu \mathbf{p} wokół osi rotacji $\check{\omega}$

9. Ruch ciała sztywnego

gdzie $\mathbf{p}(0)$ są współrzędnymi punktu początkowego, $e^{[\dot{\omega}]}$ jest eksponentą macierzy (ang. *matrix exponential*). Eksponentę macierzy rozwijamy w szereg Taylora [180, 300]:

$$e^{[\dot{\omega}]} = \mathbf{I} + [\dot{\omega}] + \frac{[\dot{\omega}]^2}{2!} + \frac{[\dot{\omega}]^3}{3!} + \dots \quad (9.39)$$

Przyjmując, że punkt \mathbf{P} opisany przez współrzędne wektora \mathbf{p} obraca się ze stałą prędkością 1 [rad/s], argument t można zastąpić przez skalar θ będący reprezentacją kąta rotacji. Wówczas równanie (9.38) przyjmuje postać:

$$\mathbf{p}(\theta) = e^{[\dot{\omega}]\theta} \mathbf{p}(0) \quad (9.40)$$

Zależność ta oznacza to, że po t sekundach wektor o współrzędnych \mathbf{p} obróci się o t radianów. Biorąc pod uwagę równość (9.40), zależność (9.39) zapisujemy w postaci [300]:

$$e^{[\dot{\omega}]\theta} = \mathbf{I} + [\dot{\omega}]\theta + \frac{([\dot{\omega}]\theta)^2}{2!} + \frac{([\dot{\omega}]\theta)^3}{3!} + \dots \quad (9.41)$$

Wiedząc, że

$$\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \dots \quad (9.42)$$

$$\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \dots \quad (9.43)$$

podstawmy teraz (9.42) i (9.43) do (9.41). Otrzymujemy wówczas następujące wyrażenie:

$$\begin{aligned} e^{[\dot{\omega}]\theta} &= \mathbf{I} + \theta[\dot{\omega}] + \frac{\theta^2}{2!}[\dot{\omega}]^2 + \frac{\theta^3}{3!}[\dot{\omega}]^3 + \dots = \\ &= \mathbf{I} + \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots \right) [\dot{\omega}] + \left(\frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \frac{\theta^6}{6!} - \dots \right) [\dot{\omega}]^2 \end{aligned} \quad (9.44)$$

Ostatecznie, zależność (9.44) możemy przedstawić w postaci **wzoru Rodrigueza**:

$$e^{[\dot{\omega}]\theta} = \mathbf{I} + \sin \theta [\dot{\omega}] + (1 - \cos \theta) [\dot{\omega}]^2 \quad (9.45)$$

gdzie $\dot{\omega} \in \mathbb{R}^3$, dla której $\|\dot{\omega}\| = 1$ oraz $\theta \in \mathbb{R}$. W przypadku, gdy $\|\dot{\omega}\| \neq 1$ zależność (9.45) przyjmuje postać [300]:

$$e^{[\dot{\omega}]\theta} = \mathbf{I} + \frac{[\dot{\omega}]}{\|\dot{\omega}\|} \sin(\|\dot{\omega}\| \theta) + \frac{[\dot{\omega}]^2}{\|\dot{\omega}\|^2} (1 - \cos(\|\dot{\omega}\| \theta)) \quad (9.46)$$

9.4. Ruch ciała sztywnego w przestrzeni trójwymiarowej

W pracy [300] udowodniono, że eksponenta macierzy (patrz zależn. (9.45) oraz (9.46)) jest macierzą rotacji:

$$\mathbf{R} = e^{[\check{\omega}]^{\theta}} \in \mathbf{SO}(\mathbf{3}) \quad (9.47)$$

W omówionej reprezentacji wykładniczej rotacji informacja o osi i kącie obrotu zawarta jest w postaci pojedynczego wektora. Jest to cecha wspólna z reprezentacją w formie kwaternionów (por. 9.3.3). W przypadku tego opisu wektor osi rotacji określany jest przez kierunek wektora, a długość wektora reprezentuje kąt obrotu. Reprezentacje wykładniczą wyróżnia to, że jest ona ciągła w początku układu współrzędnych. Cecha ta wynika z tego, że kątowi zerowemu przyporządkowano wektor zerowy.

Opisana metoda charakteryzuje się również tym, że istnieją dla niej dobrze określone pochodne. Cecha ta ma znaczenie w przypadku potrzeby przetwarzania np. prędkości kątowych [322]. Wadą tej reprezentacji jest natomiast to, że podobnie jak reprezentacja w postaci kątów Eulera ma ona punkty osobliwe.

9.4. Ruch ciała sztywnego w przestrzeni trójwymiarowej

W tej części omówione zostaną metody wykorzystywane do modelowania ruchu ciała sztywnego w przestrzeni. Ruch w przestrzeni jest ruchem złożonym, który składa się z ruchu postępowego oraz obrotowego (rys. 9.7).

Wróćmy do rysunku 9.1; na rysunku tym przedstawiono orientację i położenie układu współrzędnych ciała sztywnego w relacji do układu współrzędnych odniesienia. Oznaczmy układ odniesienia przez $\{A\}$, a układ współrzędnych ciała sztywnego przez $\{B\}$. Oznaczmy następnie wektor położenia układu współrzędnych ciała sztywnego $\{B\}$ względem układu współrzędnych odniesienia $\{A\}$ jako $\mathbf{p}_{a,b}$, gdzie $\mathbf{p}_{a,b} \in \mathbb{R}^3$. Indeks dolny a oznacza powiązanie z układem $\{A\}$, natomiast b z układem $\{B\}$. Do opisu orientacji jednego układu względem drugiego wykorzystamy macierz rotacji $\mathbf{R}_{a,b}$, gdzie $\mathbf{R}_{a,b} \in \mathbf{SO}(\mathbf{3})$. Operacja przekształcenia układu współrzędnych związanych z układem $\{A\}$ i układu współrzędnych związanych z układem $\{B\}$ opisana jest przez parę $(\mathbf{p}_{a,b}, \mathbf{R}_{a,b})$. Parę tę nazywamy przestrzenią konfiguracyjną. Tak zdefiniowana przestrzeń konfiguracyjna jest specjalną grupą euklidesową i oznaczamy jako $\mathbf{SE}(\mathbf{3})$ [300].

Definicja 9.4.1 (Specjalna grupa euklidesowa $\mathbf{SE}(\mathbf{3})$). *Specjalna grupa euklidesowa $\mathbf{SE}(\mathbf{3})$ spełnia następujące warunki:*

$$\mathbf{SE}(\mathbf{3}) = \{\mathbf{g} = (\mathbf{p}, \mathbf{R}) : \mathbf{p} \in \mathbb{R}^3, \mathbf{R} \in \mathbf{SO}(\mathbf{3})\} = \mathbb{R}^3 \times \mathbf{SO}(\mathbf{3}) \quad (9.48)$$

9. Ruch ciała sztywnego

Oczywiście, specjalną grupę euklidesową można uogólnić na wiele wymiarów:

$$\mathbf{SE}(\mathcal{D}) = \mathbb{R}^{\mathcal{D}} \times \mathbf{SO}(\mathcal{D}) \quad (9.49)$$

gdzie \mathcal{D} to liczba wymiarów. Grupa $\mathbf{SE}(\mathbf{3})$ jest przykładem gładkiej rozmaiłości [398].

Przyjmijmy, że składowe wektora $\mathbf{p}_a \in \mathbb{R}^3$ będą współrzędnymi punktu w układzie $\{A\}$, a $\mathbf{p}_b \in \mathbb{R}^3$ w układzie $\{B\}$. Znając współrzędne \mathbf{p}_b oraz mając zdefiniowany wektor położenia $\mathbf{p}_{a,b}$ i macierz rotacji $\mathbf{R}_{a,b}$, możemy wyznaczyć współrzędne punktu \mathbf{p}_a w układzie $\{A\}$ z następującej zależności:

$$\mathbf{p}_a = \mathbf{p}_{a,b} + \mathbf{R}_{a,b}\mathbf{p}_b \quad (9.50)$$

gdzie parę $(\mathbf{p}_{a,b}, \mathbf{R}_{a,b})$ oznaczmy jako:

$$\mathbf{g}_{a,b} = (\mathbf{p}_{a,b}, \mathbf{R}_{a,b}) \quad (9.51)$$

Zależność (9.50) nazywamy przemieszczeniem ciała sztywnego w przestrzeni euklidesowej [398]. Natomiast zależność (9.51) traktujemy jako konfigurację układu współrzędnych $\{B\}$ w relacji do układu A (rys. 9.1). Parę $\mathbf{g}_{a,b}$ można zapisać w zwartej formie, wprowadzając **współrzędne jednorodne**.

9.4.1. Przekształcenia jednorodne

Zapis w formie (9.50) jest złożonym zapisem w porównaniu do zapisu, w którym operacje rotacji i przesunięć rozpatrywane są samodzielnie. Przekształcenia stają się jeszcze bardziej złożone w przypadku ich składania. Tracą one wówczas jednorodność. W celu zachowania jednorodności opisu dla dowolnych przypadków wzajemnych położenia układów współrzędnych wprowadza się współrzędne jednorodne [184].

Współrzędne jednorodne konstruujemy, uzupełniając współrzędne wektora \mathbf{p} o czwartą współrzędną równą 1. Otrzymujemy wówczas wektor o czterech składowych:

$$\mathbf{p} = \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} \in \mathbb{R}^4 \quad (9.52)$$

gdzie \mathbf{p} nazywamy współrzędnymi jednorodnymi punktu \mathbf{P} . Uwzględniając (9.52), przekształcenie $\mathbf{p}_a = \mathbf{g}_{a,b}(\mathbf{p}_b)$ (9.50) można zapisać w postaci:

$$\mathbf{p}_a = \begin{bmatrix} \mathbf{p}_a \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{a,b} & \mathbf{p}_{a,b} \\ \mathbf{0}_{3 \times 1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p}_b \\ 1 \end{bmatrix} = \mathbf{T}_{a,b}\mathbf{p}_b \quad (9.53)$$

9.4. Ruch ciała sztywnego w przestrzeni trójwymiarowej

Macierz $\underline{\mathbf{T}}_{a,b}$ jest nazywana **macierzą transformacji jednorodnej**, która jest reprezentacją jednorodną pary $\mathbf{g}_{a,b} \in \mathbf{SE}(\mathbf{3})$. W ogólności, jeżeli $\mathbf{g} = (\mathbf{p}, \mathbf{R}) \in \mathbf{SE}(\mathbf{3})$, to:

$$\underline{\mathbf{T}} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ 0 & 1 \end{bmatrix} \quad (9.54)$$

Przedstawiona macierz transformacji jednorodnej opisuje w zwartej formie przekształcenie pomiędzy dwoma układami współrzędnych. Kosztem stosowania wprowadzonego opisu jest zwiększenie liczby obliczeń podczas dokonywania przekształceń, ponieważ przekształceniom podlegają obiekty, które są czterowymiarowe, a nie trójwymiarowe.

Macierz transformacji jednorodnej można wykorzystać również do składania przekształceń. Przyjmijmy, że $\mathbf{g}_{b,c} \in \mathbf{SE}(\mathbf{3})$ jest konfiguracją układu $\{C\}$ w relacji do układu $\{B\}$ oraz $\mathbf{g}_{a,b} \in \mathbf{SE}(\mathbf{3})$ analogicznie jest konfiguracją układu $\{B\}$ w relacji do układu $\{A\}$. Biorąc pod uwagę równanie (9.54), nową konfigurację zapisujemy w postaci:

$$\underline{\mathbf{T}}_{a,c} = \underline{\mathbf{T}}_{a,b} \underline{\mathbf{T}}_{b,c} = \begin{bmatrix} \mathbf{R}_{a,b} \mathbf{R}_{b,c} & \mathbf{R}_{a,b} \mathbf{p}_{b,c} + \mathbf{p}_{a,b} \\ 0 & 1 \end{bmatrix} \quad (9.55)$$

Równanie (9.55) wskazuje, że możliwe jest składanie przekształceń przez wykonanie mnożenia. Podstawowe właściwości grupy $\mathbf{SE}(\mathbf{3})$ są następujące [219]:

- jest zamknięta, tj. jeżeli $\mathbf{g}_1, \mathbf{g}_2 \in \mathbf{SE}(\mathbf{3})$, to również $\mathbf{g}_1 \cdot \mathbf{g}_2 \in \mathbf{SE}(\mathbf{3})$,
- istnieje element jednostkowy: $\mathbf{I} \in \mathbf{SE}(\mathbf{3})$,
- mnożenie macierzy transformacji jednorodnych jest łączne,
- elementem odwrotnym do $\underline{\mathbf{T}} \in \mathbf{SE}(\mathbf{3})$ jest macierz $\underline{\mathbf{T}}^{-1} \in \mathbf{SE}(\mathbf{3})$:

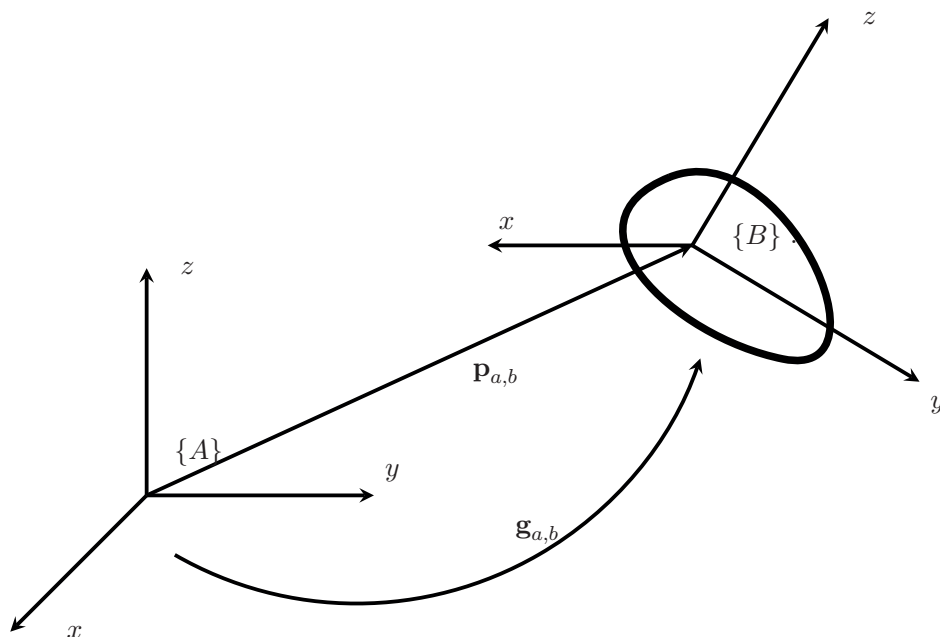
$$\underline{\mathbf{T}}^{-1} = \begin{bmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathbf{p} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}$$

9.4.2. Reprezentacja wykładnicza w ruchu ciała sztywnego

Rozważmy teraz przypadek ogólniejszy. Ruch ciała sztywnego w przestrzeni może być połączeniem ruchu postępowego i obrotowego. Przypadek ten zilustrowano na rys. 9.7.

Niech pewien punkt w przestrzeni \mathbf{P} opisany współrzędnymi wektora $\mathbf{p} \in \mathbb{R}^3$ obraca się wokół pewnej osi $\tilde{\omega}$ i jednocześnie porusza się ruchem postępowym z pewną prędkością liniową \mathbf{v} . Niech $\tilde{\omega} \in \mathbb{R}^3$, gdzie $\|\tilde{\omega}\| = 1$, oraz $\mathbf{v} \in \mathbb{R}^3$. Przyjmując, że obiekt obraca się z prędkością jednostkową, prędkość poruszającego

9. Ruch ciała sztywnego



Rys. 9.7. Ilustracja ruchu ciała sztywnego w przestrzeni

się wokół osi obrotu punktu $\mathbf{p} \in \mathbb{R}^3$ wyrażamy zależnością:

$$\dot{\mathbf{p}}(t) = \check{\boldsymbol{\omega}} \times \mathbf{p}(t) + \mathbf{v} \quad (9.56)$$

gdzie \mathbf{v} jest wektorem prędkości liniowej. Uwzględniając zależność (9.37), wyrażenie (9.56) przyjmuje postać:

$$\dot{\mathbf{p}}(t) = [\check{\boldsymbol{\omega}}]\mathbf{p}(t) + \mathbf{v} \quad (9.57)$$

W formie macierzowej zależność (9.57) zapisujemy w następującej postaci:

$$\begin{bmatrix} \dot{\mathbf{p}}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} [\check{\boldsymbol{\omega}}] & \mathbf{v} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}(t) \\ 1 \end{bmatrix} \quad (9.58)$$

Dokonując z kolei w zależności (9.58) następujących podstawień $[\boldsymbol{\xi}] = \begin{bmatrix} [\check{\boldsymbol{\omega}}] & \mathbf{v} \\ 0 & 0 \end{bmatrix}$, $\dot{\mathbf{p}} = \begin{bmatrix} \dot{\mathbf{p}} \\ 0 \end{bmatrix}$ oraz $\mathbf{p} = \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix}$, równanie opisane zależnością (9.56) zapiszemy w po-

9.4. Ruch ciała sztywnego w przestrzeni trójwymiarowej

staci:

$$\dot{\mathbf{p}}(t) = [\boldsymbol{\xi}]\mathbf{p}(t) \quad (9.59)$$

gdzie $[\boldsymbol{\xi}]$ nazywamy **skrętnikiem** (ang. *twist*) [398]. Rozwiązaniem równania (9.59) jest:

$$\mathbf{p}(t) = e^{[\boldsymbol{\xi}]t}\mathbf{p}(0) \quad (9.60)$$

Zależność (9.60) interpretujemy jako przesunięcie punktu po obrocie równym t radianów od punktu początkowego. Przez analogię do równania (9.40), zapiszmy zależność (9.60) w postaci:

$$\mathbf{p}(\theta) = e^{[\boldsymbol{\xi}]\theta}\mathbf{p}(0) \quad (9.61)$$

Wprowadźmy jeszcze za [300] symbol \vee do oznaczenia następującej operacji:

$$[\boldsymbol{\xi}]^\vee = \begin{bmatrix} [\check{\boldsymbol{\omega}}] & \mathbf{v} \\ 0 & 0 \end{bmatrix}^\vee = \begin{bmatrix} \mathbf{v} \\ \check{\boldsymbol{\omega}} \end{bmatrix} \quad (9.62)$$

Wykonując operację (9.62), otrzymamy wektor [219]:

$$\boldsymbol{\xi} = \begin{bmatrix} \mathbf{v} \\ \check{\boldsymbol{\omega}} \end{bmatrix} \quad (9.63)$$

Z kolei, niech \wedge opisuje operację odwrotną do (9.62), tj.

$$\begin{bmatrix} \mathbf{v} \\ \check{\boldsymbol{\omega}} \end{bmatrix}^\wedge = \begin{bmatrix} [\check{\boldsymbol{\omega}}] & \mathbf{v} \\ 0 & 0 \end{bmatrix} \quad (9.64)$$

Wektor $\boldsymbol{\xi} \in \mathbb{R}^6$ jest reprezentacją macierzy $[\boldsymbol{\xi}] \in \mathfrak{se}(\mathbf{3})$. Gdzie $\mathfrak{se}(\mathbf{3})$ jest algebrą Liego dla specjalnej grupy euklidesowej:

$$\mathfrak{se}(\mathbf{3}) = \{(\mathbf{v}, [\check{\boldsymbol{\omega}}]) : \mathbf{v} \in \mathbb{R}^3, [\check{\boldsymbol{\omega}}] \in \mathfrak{so}(\mathbf{3})\} \quad (9.65)$$

natomiast $\mathfrak{so}(\mathbf{3})$ jest algebrą Liego dla specjalnej grupy obrotów:

$$\mathfrak{so}(\mathbf{3}) = \{S \in \mathbb{R}^{3 \times 3} : S^\top = -S\} \quad (9.66)$$

Z następujących rozważań możemy stwierdzić, że:

$$e^{[\boldsymbol{\xi}]\theta} = \mathbf{SE}(\mathbf{3}) \quad (9.67)$$

gdzie $[\boldsymbol{\xi}] \in \mathfrak{se}(\mathbf{3})$ oraz $\theta \in \mathbb{R}$. Dowód równości (9.67) przedstawiono w pracy [300].

9. Ruch ciała sztywnego

Macierz opisana zależnością (9.67) jest macierzą przekształcenia jednorodnego, które będziemy oznaczać jako:

$$\mathbf{g} = e^{[\xi]\theta} \quad (9.68)$$

9.5. Prędkość ruchu ciała sztywnego

W poprzednim rozdziale wprowadziliśmy aparat służący do opisu związków między współrzędnymi wybranego punktu przestrzeni w różnych układach. Pod uwagę braliśmy dotychczas tylko przypadek statyczny, tj. gdy wzajemne relacje pomiędzy układami nie zmieniały się w czasie. W tym rozdziale rozpatrzemy problem ogólniejszy, w którym analizie będzie podlegała trajektoria poruszających się punktów przestrzeni. W tej części pracy rozważymy zagadnienia związane z prędkościami punktów w różnych układach współrzędnych.

Prędkość kątowna

Oznaczmy przez $\{B\}$ lokalny układ współrzędnych związany z pewnym ciałem sztywnym, a przez $\{A\}$ układ odniesienia, który jest nieruchomy względem ciała sztywnego. Układ współrzędnych związany z układem odniesienia nazywamy układem współrzędnych w układzie przestrzeni, natomiast układ współrzędnych $\{B\}$ jest związany z układem ciała sztywnego. Oznaczmy macierz rotacji, która wiąże układ $\{A\}$ i $\{B\}$ (patrz 9.3) jako $\mathbf{R}_{a,b}$. Macierze rotacji $\mathbf{R}_{a,b}$ tworzą specjalną grupę ortogonalną (9.3), gdzie $\mathbf{R}_{a,b} \in \mathbf{SO}(\mathbf{3})$. Z podrozdziału (9.3.4) wiemy, że:

$$\mathbf{R}(t) = e^{[\check{\omega}]\theta(t)} \quad (9.69)$$

gdzie przyjęto, że $\mathbf{R} = \mathbf{R}(t)$. Pochodna względem czasu $\mathbf{R}(t)$ macierzy \mathbf{R} ma postać:

$$\dot{\mathbf{R}}(t) = [\check{\omega}]\dot{\theta}(t)e^{[\check{\omega}]\theta(t)} \quad (9.70)$$

Zależność może zostać zapisana w formie:

$$\dot{\mathbf{R}}(t) = [\check{\omega}]\dot{\theta}(t)\mathbf{R}(t) \quad (9.71)$$

Podstawiając $[\check{\omega}]_a(\theta) = [\check{\omega}]\dot{\theta}$, otrzymamy następującą zależność:

$$\dot{\mathbf{R}}(t) = [\check{\omega}]_a(\theta)\mathbf{R}(t) \quad (9.72)$$

9.5. Prędkość ruchu ciała sztywnego

gdzie $[\check{\omega}]_a(\theta)$ jest macierzą skośnie symetryczną. Warto zauważyć, że dla pochodnej macierzy \mathbf{R} mamy $\dot{\mathbf{R}} \in \mathbb{R}^{3 \times 3}$. Macierz $[\check{\omega}]_a(\theta)$ może zostać zapisana w formie:

$$[\check{\omega}]_a(\theta) = \dot{\mathbf{R}}(t) \mathbf{R}^{-1}(t) \quad (9.73)$$

Ponieważ macierz obrotu $\mathbf{R}(t) \in \mathbf{SO}(\mathbf{3})$, poprzednie wyrażenie zatem można zapisać jako:

$$[\check{\omega}]_a(\theta) = \dot{\mathbf{R}}(t) \mathbf{R}^T(t) \quad (9.74)$$

Zależność (9.74) wskazuje, że macierz prędkości kątowej ruchu obrotowego może być funkcją czasu. Przeprowadzając analogiczne obliczenia, możemy wykazać, że:

$$[\check{\omega}]_b(\theta) = \mathbf{R}^T(t) \dot{\mathbf{R}}(t) \quad (9.75)$$

10. Kinematyka systemu ciał sztywnych

Czasami lepiej być na łańcuchu niż być wolnym.

Franz Kafka

10.1. Wstęp

W rozdziale tym wprowadzono aparat matematyczny do opisu kinematyki systemu ciał sztywnych (ang. *system of rigid bodies*). Podstawowym narzędziem służącym do modelowania połączonych ciał sztywnych jest łańcuch kinematyczny. W pierwszej części pracy opisane zostaną poszczególne elementy mogące wchodzić w skład łańcucha kinematycznego. Następnie omówiono dwie podstawowe metody reprezentacji łańcucha kinematycznego, tj. reprezentacje Denavita–Hartenberga oraz wykładniczą. Pierwsza z metod reprezentacji jest to klasyczna metoda znana z modelowania manipulatorów i robotów. Druga z metod jest alternatywą dla metody klasycznej. Podejście to bazuje na reprezentacji wykładniczej (patrz podrozdz. 9.3.4). W rozdziale tym omówiono każde z tych podejść oraz porównano je ze sobą. W ostatniej części tego rozdziału sformułowano zadanie kinematyki prostej. Zadanie to jest podstawowym zadaniem rozwiązywanym na potrzeby analizy ruchu człowieka.

10.2. Łańcuch kinematyczny dla systemu ciał sztywnych

Kinematyka systemu ciał sztywnych dotyczy opisu ruchu poszczególnych jego elementów bez uwzględniania przyczyn, które ten ruch wywołują [184], [219]. Systemem ciał sztywnych będziemy nazywali układ ciał sztywnych (**ogniw**), połączonych ze sobą za pomocą **złączy (przegubów)**. Złącza i ogniwa są **elementami systemu ciał sztywnych**. Każdy z przegubów łączący dwa ogniwa reprezentuje jeden stopień swobody. Typowo, przeguby mają charakter:

- **osi obrotów** (przegub obrotowy),
- **osi przesuwu** (przegub przesuwny lub pryzmatyczny).

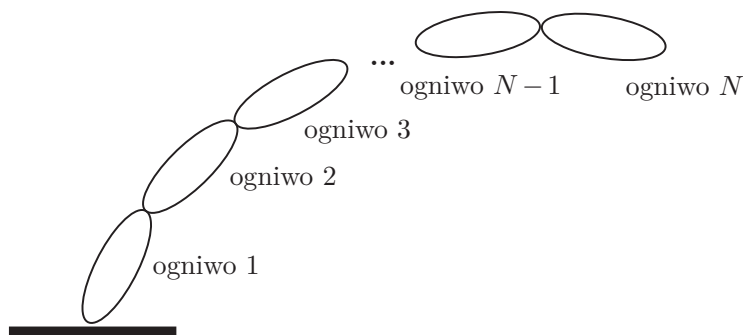
10. Kinematyka systemu ciał sztywnych

Przeguby obrotowe zapewniają ruch obrotowy sąsiednich ogniwi. Natomiast przeguby przesuwne zapewniają wzajemny ruch liniowy pomiędzy sąsiadującymi ogniwami. Każdy z wymienionych typów przegubów zapewnia jeden stopień swobody, który określany jest jedną wielkością skalarną nazywaną **zmienną złączową** [184]. Zmienna złączowa jest **kątem obrotu** (dla przegubu obrotowego) lub **przesunięciem liniowym** (dla przegubu przesuwnego).

Para połączonych ze sobą złączem ciał sztywnych (ogniwi) tworzy **parę kinematyczną**, natomiast połączenie wielu par kinematycznych tworzy **łańcuch kinematyczny**. Rozróżniamy następujące typy łańcuchów kinematycznych:

- **proste otwarte**,
- **proste zamknięte**,
- **złożone**.

W pracy, ze względu na jej tematykę, rozpatrywany będzie jedynie łańcuch kinematyczny prosty otwarty. Przykład systemu ciał sztywnych tworzących łańcuch kinematyczny przedstawiono na rys. 10.1.



Rys. 10.1. Schemat łańcucha kinematycznego otwartego prostego dla systemu ciał sztywnych

Kolejne elementy (ogniwa) systemu ciał sztywnych, które stanowią łańcuch kinematyczny prosty otwarty, typowo numerowane są począwszy od podstawy (ogniwo o numerze 0), aż do ostatniego elementu (ogniwo numer N_{10}). Oczywiście możliwe jest również numerowanie odwrotne, tj. podstawa jest ogniwiem numer N_{10} , a ostatni element w łańcuchu ma numer 0 [219].

Ostatnim etapem opisu łańcucha kinematycznego, wraz z opisem wzajemnego położenia elementów wchodzących w jego skład, jest ich jednoznaczny opis oraz przyporządkowanie każdemu z nich odpowiedniego układu współrzędnych [219].

Dalszą część tego rozdziału poświęcono budowie struktur reprezentujących łańcuchy kinematyczne. Problematyką tą zajmuje się **kinematyka**. Ruch poszcze-

10.3. Reprezentacja łańcucha kinematycznego dla systemu ciał sztywnych

gólnych elementów wchodzących w skład łańcucha kinematycznego określany jest za pomocą hierarchii ruchów. Dwa podejścia do ustalania położenia obiektów tworzących łańcuch kinematyczny to zadanie **kinematyki prostej**, w której położenie wyznaczane jest na podstawie podanych wartości parametrów. Dodajmy, że drugie z podejść to **kinematyka odwrotna**. W tym przypadku określa się pożądanego położenie obiektów wchodzących w skład łańcucha kinematycznego, a zadanie polega na wyznaczeniu odpowiednich wartości parametrów. W dalszej części tej książki skupimy się na zadaniu prostym kinematyki.

10.3. Reprezentacja łańcucha kinematycznego dla systemu ciał sztywnych

Po ponumerowaniu elementów (ogniwi i złączy) systemu ciał sztywnych, tak jak na rys. 10.1, kolejnym krokiem jest przygotowanie opisu łańcucha kinematycznego. W tym celu dokonuje się przyporządkowania każdemu z ogniwi układu współrzędnych. Układ współrzędnych związany z ogniwiem łańcucha kinematycznego jest kartezjańskim układem współrzędnych. Najpopularniejszą metodą opisu (reprezentacji) łańcucha kinematycznego jest opis bazujący na **parametrach Denavita–Hartenberga** [104]. W literaturze rozpatruje się również **zmodyfikowaną reprezentację Denavita–Hartenberga** [89].

Alternatywą dla podanych metod opisu łańcucha kinematycznego jest wykorzystanie **reprezentacji wykładniczej** [300], [369].

10.3.1. Reprezentacja Denavita–Hartenberga

Do opisu łańcucha kinematycznego zgodnie z reprezentacją Denavita–Hartenberga wykorzystuje się zestaw następujących parametrów:

- ϖ_i długość i -tego ogniwa; jest mierzona jako odległość między i oraz $i + 1$ osiami przegubów,
- α_i kąt skręcania i -tego ogniwa prawoskrętnie wokół ϖ_i ; jest mierzony jako kąt między i oraz $i + 1$ osiami przegubów,
- ς_i odległość mierzona wzdłuż osi i -tego przegubu między ϖ_{i-1} i ϖ_i ,
- ϕ_i kąt między ϖ_{i-1} i ϖ_i ; jest określany prawoskrętnie wokół osi i -tego przegubu.

Kolejnym etapem jest przyporządkowanie poszczególnym ogniwom systemu ciał sztywnych układu współrzędnych. Dla reprezentacji Denavita–Hartenberga sposób przyporządkowania osi jest następujący:

- oś z_i pokrywa się z osią przegubu $i + 1$;
- oś x_i jest prostopadła do osi z_{i-1} oraz z_i i jest skierowana od przegubu i do przegubu $i + 1$;
- oś y_i uzupełnia prawoskrętny układ współrzędnych.

10. Kinematyka systemu ciał sztywnych

Dla tak ustalonego układu współrzędnych ortogonalna macierz rotacji (macierz obrotu) przekształcająca i -ty układ współrzędnych w układ współrzędnych $i - 1$ ma postać:

$$\mathbf{R}_{i-1,i} = \begin{bmatrix} \cos \phi_i & -\cos \alpha_i \sin \phi_i & \sin \alpha_i \sin \phi_i \\ \sin \phi_i & \cos \alpha_i \cos \phi_i & -\sin \alpha_i \cos \phi_i \\ 0 & \sin \alpha_i & \cos \alpha_i \end{bmatrix} \quad (10.1)$$

Natomiast wektor położenia początku i -tego układu współrzędnych względem układu $i - 1$:

$$\mathbf{p}_{i-1,i} = \begin{bmatrix} \varpi_i \cos \phi_i \\ \varpi_i \sin \phi_i \\ \varsigma_i \end{bmatrix} \quad (10.2)$$

Macierz rotacji \mathbf{R} (10.1) oraz wektor położenia \mathbf{p} (10.2) można zapisać w postaci zwartej jako macierz transformacji jednorodnej (patrz rozdz. 9.4):

$$\mathbf{T}_{i-1,i} = \begin{bmatrix} \cos \phi_i & -\cos \alpha_i \sin \phi_i & \sin \alpha_i \sin \phi_i & \varpi_i \cos \phi_i \\ \sin \phi_i & \cos \alpha_i \cos \phi_i & -\sin \alpha_i \cos \phi_i & \varpi_i \sin \phi_i \\ 0 & \sin \alpha_i & \cos \alpha_i & \varsigma_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10.3)$$

W przypadku odwrotnej reprezentacji Denavita–Hartenberga macierz rotacji \mathbf{R} oraz wektor położenia \mathbf{p} (odpowiednio) mają postać:

$$\mathbf{R}_{i-1,i} = \begin{bmatrix} \cos \phi_i & -\sin \phi_i & 0 \\ \sin \phi_i \sin \alpha_{i-1} & \cos \phi_i \cos \alpha_{i-1} & -\sin \alpha_i \\ \sin \phi_i \sin \alpha_{i-1} & \cos \phi_i \sin \alpha_{i-1} & \cos \alpha_{i-1} \end{bmatrix} \quad (10.4)$$

$$\mathbf{p}_{i-1,i} = \begin{bmatrix} \varpi_{i-1} \\ -\varsigma_i \sin \alpha_{i-1} \\ \varsigma_i \cos \alpha_{i-1} \end{bmatrix} \quad (10.5)$$

10.3.2. Reprezentacja wykładnicza

Alternatywne podejście do opisu kinematyki systemu ciał sztywnych bazuje na reprezentacji wykładniczej (patrz 9.3.4). Wyczerpujący opis wykorzystania reprezentacji wykładniczej do modelowania kinematyki przedstawiono w pracy [300]. W tym miejscu przedstawimy tylko te jego elementy, które będą niezbędne w dalszej części książki. Formalizm, wykorzystany w tym rozdziale, jak i pozostałych odnoszących się do opisu kinematyki systemu ciał sztywnych bazuje na tym zaprezentowanym w podanym opracowaniu.

10.3. Reprezentacja łańcucha kinematycznego dla systemu ciał sztywnych

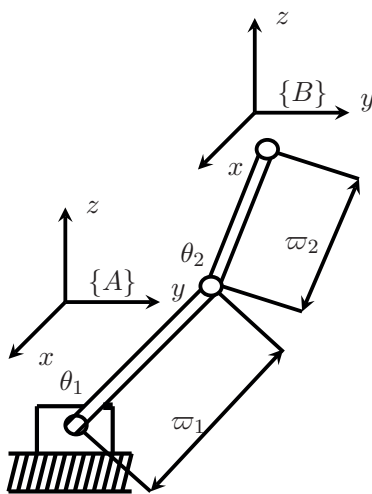
W rozdziałach poprzednich pokazaliśmy, że ruch możemy rozpatrywać jako złożenie ruchu obrotowego wokół osi oraz translacji w kierunku wektora równoległego do tej osi. Ruch taki nazywamy **ruchem śrubowym**.

W rozdziale 9 pokazano, że reprezentacja wykładnicza może zostać wykorzystana do opisu kinematyki systemu ciał sztywnych. Na przykład chcąc opisać transformację z układu $\{A\}$ do układu $\{B\}$, tj. wyznaczyć końcową konfigurację ciała sztywnego $\mathbf{g}_{a,b}(\theta)$ po zakończeniu ruchu, przy zadanej konfiguracji początkowej $\mathbf{g}_{a,b}(0)$, korzystamy z zależności:

$$\mathbf{g}_{a,b}(\theta) = e^{[\check{\xi}]^{\theta}} \mathbf{g}_{a,b}(0) \quad (10.6)$$

Jeżeli macierz $[\check{\xi}]$ opisuje przegub przesuwny, wówczas zmienna złączowa $\theta \in \mathbb{R}$ odnosi się do przesunięcia. Natomiast jeżeli macierz dotyczy $[\check{\xi}]$ przegubu obrotowego $\theta \in \mathbb{S}^1$, to wówczas zmienna złączowa związana jest z kątem obrotu.

Formułę (10.6) można uogólnić na łańcuch kinematyczny o N_{l_0} ogniwach. Na rysunku 10.2 przedstawiono przykładowy system ciał sztywnych składający się z dwóch przegubów i dwóch ogniw.



Rys. 10.2. System ciał sztywnych złożony z dwóch elementów

Po założeniu, że pierwsze ogniwo jest unieruchomione, konfiguracja systemu jest zależna jedynie od zmiennej złączowej (θ_2). Wówczas opis dla rozpatrywanego układu jest następujący:

$$\mathbf{g}_{a,b}(\theta_2) = e^{[\check{\xi}_2]^{\theta_2}} \mathbf{g}_{a,b}(0) \quad (10.7)$$

10. Kinematyka systemu ciał sztywnych

gdzie a to pierwsze ogniwo, a b to ostatnie ogniwo w łańcuchu kinematycznym. Natomiast ξ_2 jest wektorem przestrzennych prędkości dla drugiego przegubu. Teraz unieruchomimy drugi przegub z jednoczesnym wykonywaniem obrotów pierwszego z nich. Wówczas opis konfiguracji systemu ciał sztywnych, po uwzględnieniu zależności (10.7), ma postać:

$$\mathbf{g}_{a,b}(\theta_1, \theta_2) = e^{[\check{\xi}_1]\theta_1} \mathbf{g}_{a,b}(\theta_2) = e^{[\check{\xi}_1]\theta_1} e^{[\check{\xi}_2]\theta_2} \mathbf{g}_{a,b}(0) \quad (10.8)$$

Konfiguracja nie zależy od kolejności ruchów wykonywanych w każdym przegubie, tj. zadanie możemy rozwiązać, poruszając przegubem 1, a następnie 2:

$$\mathbf{g}_{a,b}(\theta_1) = e^{[\check{\xi}_1]\theta_1} \mathbf{g}_{a,b}(0) \quad (10.9)$$

Z przeprowadzonej analizy wynika (zależności (10.6)–(10.9)), że formuła wyznaczająca konfigurację systemu ciał sztywnych składającego się z N_{lo} elementów ma postać [300]:

$$\mathbf{g}_{a,b}(\theta) = e^{[\check{\xi}_1]\theta_1} e^{[\check{\xi}_2]\theta_2} \dots e^{[\check{\xi}_N]\theta_N} \mathbf{g}_{a,b}(0) \quad (10.10)$$

W łańcuchu kinematycznym ogniwa systemu ciał sztywnych połączone są złączkami, przez które na siebie oddziałują. Ruch i -tego elementu w systemie, który zależy od ruchu elementu $i - 1$, można reprezentować przez ruch śrubowy. Orientacja elementu i -tego względem elementu $i - 1$ w systemie ciał sztywnych opisana jest przez macierz transformacji jednorodnej $\mathbf{T}_{i-1,i}$. W celu wyznaczenia przejścia z układu współrzędnych dla elementu 0 do układu współrzędnych dla elementu N należy wykonać sekwencję mnożeń macierzy transformacji jednorodnej \mathbf{T} (10.10).

Przykładowe wektory prędkości przestrzennych dla typowych przegubów mają postać:

- Dla przegubu obrotowego:

$$\xi_i = \begin{bmatrix} -\boldsymbol{\omega}_i \times \mathbf{p}_i \\ \boldsymbol{\omega}_i \end{bmatrix} \quad (10.11)$$

gdzie $\boldsymbol{\omega}_i \in \mathbb{R}^3$, $\mathbf{p}_i \in \mathbb{R}^3$.

- Dla przegubu przesuwnego:

$$\xi_i = \begin{bmatrix} \mathbf{v}_i \\ 0 \end{bmatrix} \quad (10.12)$$

gdzie $\mathbf{v}_i \in \mathbb{R}^3$ to prędkość liniowa związana z ruchem postępowym.

10.4. Metody reprezentacji łańcucha kinematycznego: podsumowanie

Poniżej przedstawiono zestawienie porównujące dwa z omówionych podejść do reprezentacji łańcucha kinematycznego systemu ciał sztywnych (tab. 10.1) [341].

Tabela 10.1. Porównanie reprezentacji Denavita–Hartenberga (D–H) z reprezentacją wykładniczą

| | Reprezentacja D–H | Reprezentacja wykładnicza |
|------------------------------|---|---|
| Opis łańcucha kinematycznego | <ul style="list-style-type: none"> – Zbiór parametrów jest minimalny – Z każdym ogniwem związany jest jeden układ współrzędnych – Wybór układów współrzędnych jest algorytmiczny – Nie zależy od układu odniesienia | <ul style="list-style-type: none"> – Zbiór parametrów nie jest minimalny – Dwa układy współrzędnych dla całego łańcucha kinematycznego – Wybór układów współrzędnych jest arbitralny – Zależy od układu odniesienia |
| Zadanie proste kinematyki | <ul style="list-style-type: none"> – Zmienne złączowe reprezentują przesunięcia względne – Rozwiązywane na podstawie macierzy transformacji jednorodnej | <ul style="list-style-type: none"> – Zmienne złączowe reprezentują rzeczywiste przesunięcia – Rozwiązywanie na podstawie macierzy transformacji jednorodnej |

Jedną z głównych zalet reprezentacji wykładniczej jest dowolność w wyborze układów współrzędnych dla każdego elementu łańcucha kinematycznego. Co więcej, dobierając w odpowiedni sposób konfigurację odniesienia, uzyskujemy zmienną złączową, która związana jest z bieżącym przesunięciem elementu systemu ciał sztywnych.

W pracy do opisu łańcucha kinematycznego zastosowano reprezentację wykładniczą. Opis ten został wykorzystany do rozwiązania zadania prostego kinematyki.

10.5. Zadanie proste kinematyki

Kinematyka to opis ruchu elementów wchodzących w skład systemu ciał sztywnych bez uwzględniania wpływu sił, które ten ruch wywołują. Dwa podstawowe problemy kinematyki to wymienione już zadania proste i odwrotne.

Zadanie proste kinematyki polega na wyznaczeniu położenia i orientacji układu współrzędnych ostatniego elementu w danym łańcuchu kinematycznym względem układu odniesienia. Z kolei zadanie odwrotne kinematyki polega na wyznaczeniu

10. Kinematyka systemu ciał sztywnych

wartości kolejnych zmiennych złączowych z ustalonym położeniem i orientacją systemu ciał sztywnych.

Analizując układ ruchu człowieka, również można brać pod uwagę aspekty związane z kinematyką. Dotyczy to zarówno zadania prostego, jak i odwrotnego. Ze względu na postawione w pracy cele, rozważania dotyczące kinematycznej analizy układu ruchu człowieka ograniczone zostały do zadania prostego. Sformułujmy teraz problem kinematyki prostej dla łańcucha kinematycznego.

Problem 10.1. *Dla danego zbioru $\theta_N \in \Theta$, gdzie N jest liczbą ogniw w łańcuchu należy wyznaczyć: konfigurację elementu końcowego b w relacji do elementu odniesienia a .*

W celu rozwiązania problemu kinematyki prostej należy znaleźć następujące odwzorowanie: $\mathbf{g}_{a,b} : \Theta \rightarrow \mathbf{SE}(\mathbf{3})$.

Rozwiązanie zadania prostego kinematyki

Wykorzystując opis przedstawiony w podrozdziale 10.3.2 ogólna postać rozwiązania prostego zadania kinematyki (10.7) jest następująca:

$$\mathbf{g}_{a,b}(\theta) = e^{[\check{\xi}_1]\theta_1} e^{[\check{\xi}_2]\theta_2} \dots e^{[\check{\xi}_N]\theta_N} \mathbf{g}_{a,b}(0) \quad (10.13)$$

gdzie a to element pierwszy w łańcuchu kinematycznym, b jest elementem ostatnim tego łańcucha, natomiast N_{l_0} jest liczbą ogniw w łańcuchu.

11. Kinematyka układu ruchu człowieka

*Aktywny umysł nie może
funkcjonować w nieaktywnym ciele.*

George S. Patton Jr.

11.1. Wstęp

W rozdziale omówiono zagadnienia związane z zastosowaniem opisu kinematycznego w analizie układu ruchu człowieka. Na początku tego rozdziału omówiono podstawowe zagadnienia dotyczące biomechaniki oraz jej zastosowanie do analizy układu ruchu człowieka. W dalszej części tego rozdziału omówiono modele kinematyki układu ruchu człowieka skonstruowane na bazie omówionej już reprezentacji wykładniczej. Zaproponowano i omówiono modele kończyn dolnych i górnych.

Ponieważ podczas modelowania ruchu człowieka istotna jest również znajomość podstawowych wymiarów ludzkiego ciała, rozdział ten został zakończony krótkim wprowadzeniem do antropometrii.

11.2. Mechanika układu ruchu człowieka

Biomechanika jest interdyscyplinarną nauką zajmującą się działaniem wewnętrznych i zewnętrznych sił na organizmy biologiczne, tj. człowieka czy zwierzęcia [46]. Wewnętrznymi siłami nazywamy te siły, które są generowane przez mięśnie. Przykładem siły zewnętrznej, która może oddziaływać na organizmy biologiczne, jest grawitacja.

Efektom działania sił na organizmy biologiczne są zmiany w ich ruchu. W przypadku człowieka wśród typowych ruchów wyróżniamy m.in. chód, bieg i skok [46]. Do tych typowych ludzkich aktywności ruchowych można dodać kolejne, związane np. z różnymi dyscyplinami sportowymi. Ludzką aktywność ruchową można rozpatrywać z kilku punktów widzenia. Do najważniejszych należy analiza struktury kinematycznej i dynamicznej. Poznanie tych struktur jest niezbędne do oceny potencjału ruchowego człowieka [46]. W dalszej części pracy omawiane pro-

11. Kinematyka układu ruchu człowieka

blemy i propozycje ich rozwiązania dotyczyć będą analizy struktury kinematycznej układu ruchu człowieka.

Kolejnym etapem, następującym po analizie struktury kinematycznej ruchów ciała ludzkiego, jest ich optymalizacja. Zagadnienie optymalizacji ruchów ciała ludzkiego nie jest przedmiotem tego opracowania.

W biomechanice możemy wyróżnić kilka działów, wśród których najważniejsze to:

- biomechanika ogólna,
- biomechanika medyczna,
- biomechanika inżynierska,
- biomechanika pracy,
- biomechanika sportu.

Ponieważ problematyka książki związana jest z biomechaniką sportu, to w kolejnych rozdziałach omówione zostaną podstawowe zagadnienia z zakresu kinematyki (podrozdz. 11.3) oraz z zakresu antropometrii (podrozdz. 11.5).

11.3. Kinematyczna analiza układu ruchu człowieka

W ogólności system ciał sztywnych składa się z dwóch typów elementów: ogniw (elementy sztywne) oraz złączy, które odgrywają rolę elementów ruchowych w systemie. W odniesieniu do ciała ludzkiego elementami sztywnymi (ogniwami) są **kości**, a elementami ruchowymi łączącymi elementy sztywne są **stawy**.

Ciało sztywne umieszczone w przestrzeni może poruszać się ruchem złożonym, tj. ruchem będącym sumą dwóch niezależnych rodzajów ruchów, postępowego i obrotowego. Każdy z tych ruchów można zdekomponować na trzy składowe. Ruchy te mogą odbywać się zarówno wzdłuż osi x , y , z , jak i wokół tych osi. Każdy z tych ruchów jest niezależny.

Jeżeli na ciało sztywne nie zostały nałożone ograniczenia (tzw. **więzy**), to w przestrzeni ciało sztywne ma 6 stopni swobody. Stopnie swobody wskazują na **ruchliwość** ciała sztywnego [46].

Połączenie ze sobą dwóch ciał sztywnych z wykorzystaniem złącza tworzy wspomnianą już parę kinematyczną. Ruchliwość elementów pary kinematycznej względem siebie tzw. **ruchliwość wzajemna** ulega redukcji, tj. wyraża się liczbą stopni swobody mniejszą niż 6. Liczbę zredukowanych stopni swobody dla pary kinematycznej nazywamy **klasą pary kinematycznej**. Na przykład staw kolonowy jest połączeniem IV klasy, co oznacza, że jego konstrukcja nakłada cztery ograniczenia (więzy) [46].

Połączenie wielu par kinematycznych tworzy łańcuch kinematyczny. Przykładami łańcuchów kinematycznych w odniesieniu do ludzkiego ciała może być **kończyna górna** lub **dolna**. Takie łańcuchy nazywamy **łańcuchami prostymi**

11.3. Kinematyczna analiza układu ruchu człowieka

otwartymi, czyli mają strukturę szeregową. Na ruchliwość łańcucha kinematycznego wpływ ma zarówno liczba par kinematycznych, jak i sposób ich konfiguracji. Ruchliwość łańcucha kinematycznego to jego liczba stopni swobody, które zostały wyznaczone dla członów ruchomych łańcucha względem nieruchomej podstawy. Nieruchomą podstawą może być jedno z dowolnie wybranych ogniw łańcucha kinematycznego [46].

Ruchliwość, czyli liczbę stopni swobody łańcucha kinematycznego wyznaczamy z następującej zależności [46]:

$$\text{LSS} = 6 (N_{10} - 1) - \sum_{i=3}^5 \text{LP}_i i \quad (11.1)$$

gdzie i to klasa pary kinematycznej, LP_i liczba par i -tej klasy, N_{10} to liczba wszystkich ogniw łańcucha kinematycznego.

System ciał sztywnych złożony z kości i stawów tworzy **bierny układ ruchowy człowieka**. W przypadku układu ruchu człowieka nie występują złącza wykonujące ruchy postępowe. Wynika z tego, że ruchliwość ludzkich złączy (stawów) jest co najwyżej stopnia 3 i jest związana z typem budowy stawu oraz kształtem powierzchni stawowych [46]. Rozpatrzmy strukturę dwóch typów łańcuchów kinematycznych występujących w ludzkim układzie ruchu, tj. łańcuchy kinematyczne kończyny górnej i dolnej. W analizie uwzględnimy uproszczone modele tych łańcuchów kinematycznych. Łańcuch kinematyczny kończyny górnej człowieka składa się z 3 członów ruchomych, tj.

- ramię,
- przedramię,
- dłoń.

Człony te tworzą względem łopatki 3 pary kinematyczne, które połączone są stawami o różnym typie i o różnej liczbie stopni swobody:

- staw ramienny, tj. staw kulisty ma 3 stopnie swobody związane z: a) zginaniem i prostowaniem, b) odwodzeniem i przywodzeniem oraz c) ruchami obrotowymi ramienia do wewnątrz i na zewnątrz,
- staw łokciowy, tj. staw zawiasowo-obrotowy ma 2 stopnie swobody, które związane są: a) ze zginaniem i prostowaniem oraz b) z nawracaniem i odwracaniem przedramienia,
- staw promieniowo-nadgarstkowy, tj. staw typu eliptycznego ma 2 stopnie swobody, które odpowiedzialne są za a) zginanie grzbietowe i dłoniowe nadgarstka oraz b) odwodzenie i przywodzenie.

Wyznaczmy dla przykładu, korzystając z zależności (11.1), ruchliwość (liczbę stopni swobody) łańcucha kinematycznego kończyny górnej człowieka względem łopatki.

11. Kinematyka układu ruchu człowieka

Przykład 11.1 (Ruchliwość łańcucha kinematycznego kończyny górnej).

Dane:

- liczba członów ruchomych ($N_{10} - 1$), tj. bez podstawy: 3,
- liczba par klasy III to 1, tj. $LP_3 = 1$,
- liczba par klasy IV to 2, tj. $LP_4 = 2$.

Podstawiając te wartości do zależności (11.1), otrzymamy:

$$LSS = 6 \cdot 3 - (1 \cdot 3 + 2 \cdot 4) = 18 - 11 = 7 \quad (11.2)$$

Powtórzmy teraz rozważania dla łańcucha kinematycznego kończyny dolnej. Łańcuch ten składa się z następujących członów ruchomych:

- udo,
- podudzie,
- stopa.

Względem biodra wymienione człony tworzą 3 pary kinematyczne połączone następującymi stawami:

- staw biodrowy, tj. staw kulisty panewkowy ma 3 stopnie swobody, z którymi związane są następujące ruchy: a) prostowanie i zginanie, b) odwodzenie i przywodzenie oraz c) rotacja zewnętrzna i wewnętrzna;
- staw kolanowy, tj. staw zawiasowy ma 1 stopień swobody związany ze zginaniem i prostowaniem;
- staw skokowy górny, tj. staw zawiasowy ma 1 stopień swobody, z którym związane jest zginanie grzbietowe stopy oraz zginanie podszwowe stopy.

Przeprowadzając podobne obliczenia, określimy teraz ruchliwość łańcucha kinematycznego kończyny górnej.

Przykład 11.2 (Ruchliwość łańcucha kinematycznego kończyny dolnej).

Dane:

- liczba członów ruchomych ($N_{10} - 1$), tj. bez podstawy: 3,
- liczba par klasy III to 1, tj. $LP_3 = 1$,
- liczba par klasy V to 2, tj. $LP_4 = 1$.

Tym razem po podstawieniu określonych wartości do zależności (11.1), otrzymamy:

$$LSS = 6 \cdot 3 - (1 \cdot 3 + 2 \cdot 5) = 18 - 13 = 5 \quad (11.3)$$

Przedstawione łańcuchy kinematyczne kończyny górnej i dolnej są łańcuchami uproszczonymi. W pracy [46] omówiono dokładne łańcuchy tych kończyn.

Dysponując modelami kinematyki, np. kończyn górnych lub dolnych, możliwe jest przeprowadzenie analizy kinematyki ruchów tych kończyn podczas różnych aktywności człowieka, np. podczas biegania, wykonywania ćwiczeń fizycznych lub wykonywania ruchów typowych dla danej dyscypliny sportowej. Dokładne

11.4. Zadanie proste kinematyki w analizie układu ruchu człowieka

modele umożliwiają głębsze zrozumienie układu ruchu człowieka. Lepsze z kolei zrozumienie układu ruchu umożliwi opracowanie lepszych, z danego punktu widzenia, metod wspomagających np. trening lub rehabilitację (np. po kontuzji) sportowca.

Jednym z podejść, które zostało wykorzystane w książce, do modelowania kinematyki układu ruchu człowieka jest metoda, w której ciało ludzkie dzieli się na segmenty. Z kolei każdy z tych segmentów traktowany jest jako ciało sztywne, które może być połączone z innym ciałem sztywnym przez odpowiedni staw.

11.4. Zadanie proste kinematyki w analizie układu ruchu człowieka

Z przeprowadzonej w rozdziale 10 analizy wiadomo, że celem prostego zadania kinematyki jest wyznaczenie położenia i orientacji ostatniego ogniwa względem układu odniesienia.

Wyniki rozwiązania prostego zadania kinematyki mogą zostać wykorzystane m.in. do analizy aktywności ruchowej człowieka. Podejście to może być stosowane zarówno do oceny sprawności układu ruchu człowieka (trening sportowy), jak i do badań nad występującymi w nim patologiami (rehabilitacja).

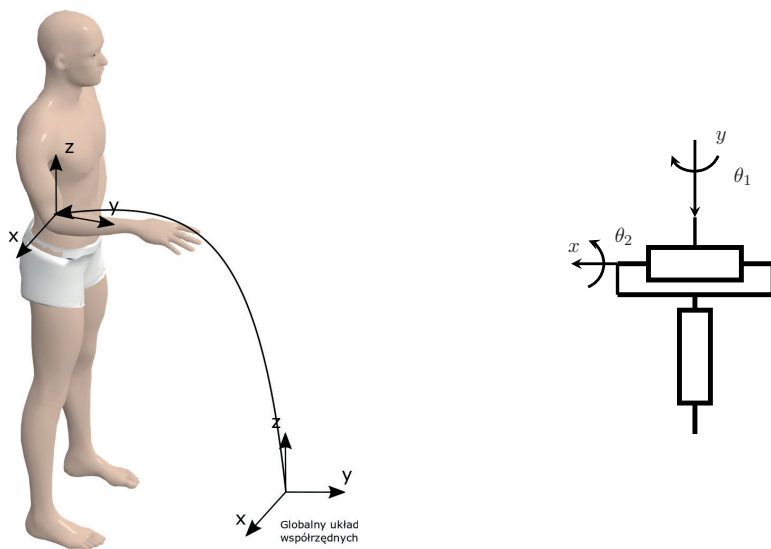
W tym podrozdziale omówiono podejście do opisu kinematyki dwóch stawów, tj. stawu łokciowego i kolanowego. Staw łokciowy człowieka (rys. 11.1 z lewej) może być modelowany jako podwójny staw zawiasowy (ang. *hinge joint*) [95, 119]. Model stawu łokciowego przedstawiono na rys. 11.1 (z prawej). W stawie tym wyróżniamy dwie osie obrotu (dwa stopnie swobody). Wokół jednej z osi obrotu możliwe jest zginanie i prostowanie, a wokół drugiej z nich odwracanie i nawracanie.

Z kolei staw kolanowy (rys. 11.2 z lewej) może być reprezentowany jako prosty staw zawiasowy z jedną osią obrotu. Model stawu kolanowego przedstawiono na rys. 11.2 (z prawej). Staw ten charakteryzuje się jednym stopniem swobody. Wokół tej osi możliwe jest jedynie zginanie i prostowanie.

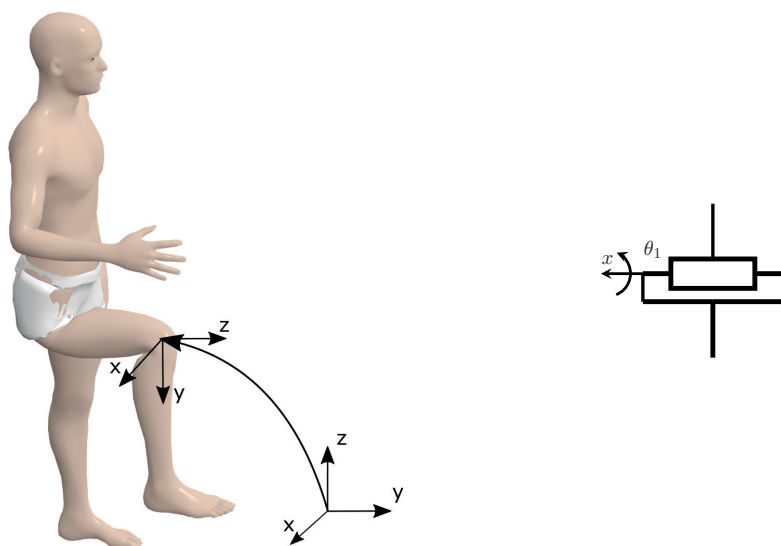
Dysponując modelami kinematyki stawu łokciowego oraz kolanowego możliwe jest ich rozszerzenie. W efekcie otrzymany (odpowiednio) modele kinematyki dla kończyny górnej i dolnej człowieka. W tym celu wprowadzimy dodatkowe opisy. I tak dla kończyny górnej będą to opisy kinematyki stawu barkowego i stawu promieniowo-nadgarstkowego. Z kolei dla kończyny dolnej wprowadzimy dodatkowy opis dla stawu biodrowego i skokowego górnego.

Ostatnim zagadnieniem omówionym w tym podrozdziale będzie zadanie modelowania kinematyki całego układu ruchu człowieka.

11. Kinematyka układu ruchu człowieka



Rys. 11.1. Staw łokciowy z zaznaczonym układem współrzędnych (z lewej) oraz jego model kinematyki (z prawej)

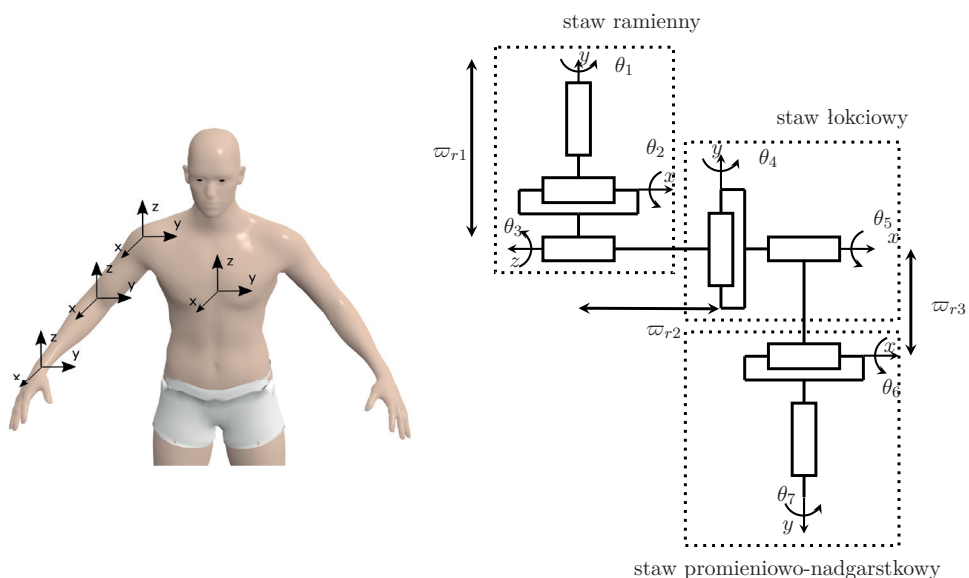


Rys. 11.2. Staw kolanowy z zaznaczonym układem współrzędnych (z lewej) oraz jego model kinematyki (z prawej)

11.4.1. Model kinematyki prostej kończyny górnej

Konstrukcję modeli kinematyki dla układu ruchu człowieka rozpoczniemy od modelu kończyny górnej. Model ten składa się z trzech ogniw, tj. ramienia, przedramienia i dłoni, oraz z trzech stawów, tj. ramiennego, łokciowego i promieniowo-nadgarstkowego. Na rysunku 11.3 przedstawiono łańcuch kinematyczny kończyny górnej człowieka składający się z wymienionych ogniw i stawów. Oznaczmy długości ogniw w tym łańcuchu w następujący sposób:

- ramię (długość ϖ_{r1}),
- przedramię (długość ϖ_{r2}),
- dłoń (długość ϖ_{r3}).



Rys. 11.3. Łańcuch kinematyczny kończyny górnej człowieka (z lewej) oraz jego model (z prawej)

W łańcuchu kinematycznym kończyny górnej człowieka wyróżniamy również trzy stawy. Staw ramienny jest stawem kulistym o trzech stopniach swobody (ang. *ball and socket joint*) [72]. Ruch wokół trzech osi tego stawu umożliwia: odwodzenie–przywodzenie, zgięcie–wyprostowanie oraz rotację wewnątrzno–zewnętrzno. Rozpocznijmy opis tego stawu od wyznaczenia skrętników dla tego stawu (rys. 11.3).

11. Kinematyka układu ruchu człowieka

Osie obrotu stawu ramiennego opisujemy następującymi wektorami:

$$\check{\omega}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \check{\omega}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \check{\omega}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (11.4)$$

Natomiast wektory przesunięć dla rozpatrywanego przypadku mają postać:

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ \varpi_{r1} \\ 0 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (11.5)$$

Biorąc pod uwagę ustalone osie obrotu oraz wektory przesunięć, skrętniki dla każdego ze stopni swobody są następujące:

$$\xi_1 = \begin{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{r1} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{r1} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.6a)$$

$$\xi_2 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (11.6b)$$

$$\xi_3 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (11.6c)$$

Przeprowadźmy teraz podobną analizę dla stawu łokciowego. W tym przypadku mamy dwa stopnie swobody. Staw łokciowy jest stawem przegubowym, w którym możliwe jest zginanie–prostowanie oraz odwracanie–nawracanie. Również i w tym przypadku, bazując na rys. 11.3, wyznaczmy skrętniki dla tego stawu. W tym

11.4. Zadanie proste kinematyki w analizie układu ruchu człowieka

przypadku osie obrotu są następujące:

$$\check{\omega}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \check{\omega}_5 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.7)$$

natomiast wektory przesunięć mają postać:

$$\mathbf{p}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{p}_5 = \begin{bmatrix} 0 \\ \varpi_{r1} + \varpi_{r2} \\ 0 \end{bmatrix} \quad (11.8)$$

Z kolei skrętniki dla tego stawu przedstawiają się następująco:

$$\xi_4 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (11.9a)$$

$$\xi_5 = \begin{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{r1} + \varpi_{r2} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{r1} - \varpi_{r2} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.9b)$$

Ostatni staw w tym łańcuchu kinematycznym to staw promieniowo-nadgarstkowy. Jest to staw kłyckiowy (ang. *condylar joint*, *ellipsoid joint*), o dwóch stopniach swobody, w którym możliwe jest a) odwodzenie-przywodzenie oraz b) zgięcie-wyprostowanie. Osie obrotu w tym przypadku zapisujemy w postaci:

$$\check{\omega}_6 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \check{\omega}_7 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (11.10)$$

a wektory przesunięcia mają następującą formę:

$$\mathbf{p}_6 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{p}_7 = \begin{bmatrix} 0 \\ \varpi_{r1} + \varpi_{r2} + \varpi_{r3} \\ 0 \end{bmatrix} \quad (11.11)$$

11. Kinematyka układu ruchu człowieka

Skretniki dla stawu promieniowo-nadgarstkowego mają postać:

$$\xi_6 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (11.12a)$$

$$\xi_7 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{r1} + \varpi_{r2} + \varpi_{r3} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{r1} - \varpi_{r2} - \varpi_{r3} \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (11.12b)$$

Biorąc pod uwagę wyznaczone skretniki dla poszczególnych stawów w łańcuchu kinematycznym kończyny górnej możemy zapisać zależność opisującą jego kinematykę.

$$\begin{aligned} \mathbf{g}_{\text{ram,dl}}(\theta) &= \left(e^{[\check{\xi}_1]\theta_1} e^{[\check{\xi}_2]\theta_2} e^{[\check{\xi}_3]\theta_3} e^{[\check{\xi}_4]\theta_4} e^{[\check{\xi}_5]\theta_5} e^{[\check{\xi}_6]\theta_6} e^{[\check{\xi}_7]\theta_7} \right) \mathbf{g}_{\text{ram,dl}}(0) = \\ &= \prod_{i=1}^7 e^{[\check{\xi}_i]\theta_i} \mathbf{g}_{\text{ram,dl}}(0) \quad (11.13) \end{aligned}$$

gdzie ram odnosi się do układu ramienia, natomiast dl do układu dłoni.

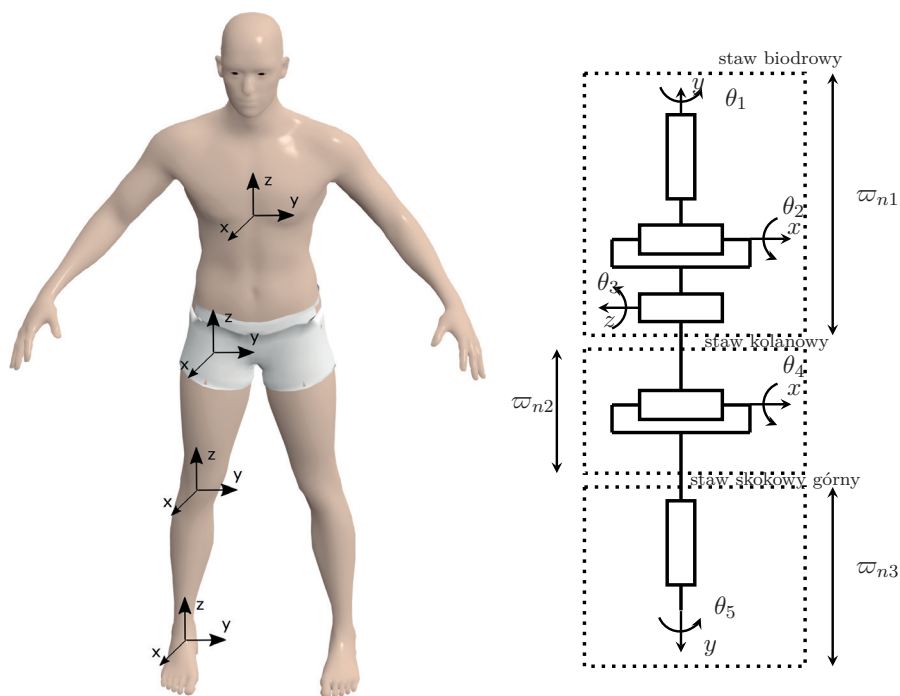
Podsumowanie parametrów dla łańcucha kinematycznego kończyny górnej przedstawiono w tabeli 11.1.

Tabela 11.1. Parametry dla łańcucha kinematycznego kończyny górnej prawej

| Staw | $\check{\omega}$ | \mathbf{p} | ξ |
|------------------|---------------------------------------|---|---|
| Ramienny | $\check{\omega}_1 = [1 \ 0 \ 0]^\top$ | $\mathbf{p}_1 = [0 \ \varpi_{r1} \ 0]^\top$ | $\xi_1 = [0 \ 0 \ -\varpi_{r1} \ 1 \ 0 \ 0]^\top$ |
| | $\check{\omega}_2 = [0 \ 1 \ 0]^\top$ | $\mathbf{p}_2 = [0 \ 0 \ 0]^\top$ | $\xi_2 = [0 \ 0 \ 0 \ 0 \ 1]^\top$ |
| | $\check{\omega}_3 = [0 \ 0 \ 1]^\top$ | $\mathbf{p}_3 = [0 \ 0 \ 0]^\top$ | $\xi_3 = [0 \ 0 \ 0 \ 0 \ 1]^\top$ |
| Łokciowy | $\check{\omega}_4 = [0 \ 1 \ 0]^\top$ | $\mathbf{p}_4 = [0 \ 0 \ 0]^\top$ | $\xi_4 = [0 \ 0 \ 0 \ 0 \ 1 \ 0]^\top$ |
| | $\check{\omega}_5 = [1 \ 0 \ 0]^\top$ | $\mathbf{p}_5 = [0 \ (\varpi_{r1} + \varpi_{r2}) \ 0]^\top$ | $\xi_5 = [0 \ 0 \ -(\varpi_{r1} + \varpi_{r2}) \ 1 \ 0 \ 0]^\top$ |
| Prom.- -nadg. | $\check{\omega}_6 = [0 \ 1 \ 0]^\top$ | $\mathbf{p}_6 = [0 \ 0 \ 0]^\top$ | $\xi_6 = [0 \ 0 \ 0 \ 0 \ 1 \ 0]^\top$ |
| | $\check{\omega}_7 = [0 \ 0 \ 1]^\top$ | $\mathbf{p}_7 = [0 \ (\varpi_{r1} + \varpi_{r2} + \varpi_{r3}) \ 0]^\top$ | $\xi_7 = [0 \ 0 \ -(\varpi_{r1} + \varpi_{r2} + \varpi_{r3}) \ 0 \ 0 \ 1]^\top$ |

11.4.2. Model kinematyki prostej kończyny dolnej

Przeprowadzimy teraz analogiczną analizę na potrzeby konstrukcji modelu kinematyki kończyny dolnej. Również w przypadku kończyny górnej jej łańcuch kinematyczny składa się z trzech ogniw oraz trzech stawów. Ogniwami tego łańcucha kinematycznego to udo, podudzie oraz stopa. Natomiast stawy w tym łańcuchu to: staw biodrowy, kolanowy oraz skokowy. Na rysunku 11.4 (z lewej) przedstawiono łańcuch kończyny dolnej człowieka. Z kolei z prawej strony tego rysunku zamieszczono model tej kończyny.



Rys. 11.4. Łańcuch kinematyczny kończyny dolnej człowieka (z lewej) oraz jego model (z prawej)

Dysponując modelem kończyny dolnej możliwe jest jego opisanie. Oznaczmy długości ogniw łańcucha kinematycznego dla tej kończyny w następujący sposób:

- udo (długość ϖ_{n1}),
- podudzie (długość ϖ_{n2}),
- stopa (długość ϖ_{n3}).

11. Kinematyka układu ruchu człowieka

Podobnie jak staw ramienny, staw biodrowy jest stawem kulistym (panewkowym) (ang. *ball and socket joint*). Osie obrotu w tym przypadku opisujemy następująco:

$$\check{\omega}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \check{\omega}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (11.14)$$

Z kolei wektory przesunięć w tym łańcuchu zapisujemy jako:

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ \varpi_{n1} \\ 0 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (11.15)$$

Bazując na (11.14) oraz (11.15), możemy również określić skrętniki dla tego stawu w następującej postaci:

$$\xi_1 = \begin{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{n1} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{n1} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.16a)$$

$$\xi_2 = \begin{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (11.16b)$$

Staw kolanowy jest natomiast stawem zawiasowym (ang. *synovial hinge joint*). W przypadku tego stawu mamy jeden stopień swobody, który opisujemy w następujący sposób:

$$\check{\omega}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.17)$$

gdzie wektor przesunięcia ma postać:

$$\mathbf{p}_3 = \begin{bmatrix} 0 \\ \varpi_{n1} + \varpi_{n2} \\ 0 \end{bmatrix} \quad (11.18)$$

11.4. Zadanie proste kinematyki w analizie układu ruchu człowieka

Bazując na analizie kinematycznej stawu kolanowego możemy teraz zapisać skrętnik dla tego stawu w formie:

$$\xi_3 = \begin{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{n1} + \varpi_{n2} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{n1} - \varpi_{n2} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.19)$$

Ostatni staw w tym łańcuchu to staw skokowy. Jest to staw typu zawiasowego. Ma on tylko jeden stopień swobody. Oś obrotu tego stawu zapisujemy w formie:

$$\check{\omega}_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.20)$$

wektor przesunięcie natomiast jako:

$$\mathbf{p}_4 = \begin{bmatrix} 0 \\ \varpi_{n1} + \varpi_{n2} + \varpi_{n3} \\ 0 \end{bmatrix} \quad (11.21)$$

W przypadku tego stawu, skrętnik ma postać:

$$\xi_4 = \begin{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \varpi_{n1} + \varpi_{n2} + \varpi_{n3} \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\varpi_{n1} - \varpi_{n2} - \varpi_{n3} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (11.22)$$

Biorąc pod uwagę poprzednie rozważania, możemy teraz zaproponować model kinematyki kończyny dolnej w postaci:

$$\mathbf{g}_{\text{bd,st}}(\theta) = \left(e^{[\check{\xi}_1]\theta_1} e^{[\check{\xi}_2]\theta_2} e^{[\check{\xi}_3]\theta_3} e^{[\check{\xi}_4]\theta_4} \right) \mathbf{g}_{\text{bd,st}}(0) = \prod_{i=1}^4 e^{[\check{\xi}_i]\theta_i} \mathbf{g}_{\text{bd,st}}(0) \quad (11.23)$$

gdzie bd odnosi się do układu biodra, st odnosi się do układu stopy.

Podsumowanie parametrów dla łańcucha kinematycznego kończyny dolnej prawej przedstawiono w tabeli 11.2.

11. Kinematyka układu ruchu człowieka

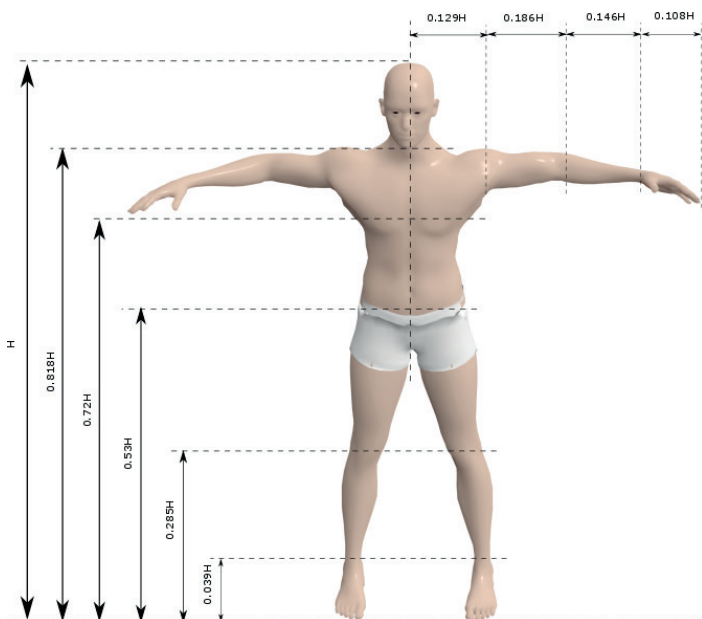
Tabela 11.2. Parametry dla łańcucha kinematycznego kończyny dolnej prawej

| Staw | $\dot{\omega}$ | \mathbf{p} | ξ |
|----------|--|--|---|
| Biodrowy | $\dot{\omega}_1 = [1 \ 0 \ 0]^T$ $\dot{\omega}_2 = [0 \ 1 \ 0]^T$ | $\mathbf{p}_1 = [0 \ \varpi_{n1} \ 0]^T$ $\mathbf{p}_2 = [0 \ 0 \ 0]^T$ | $\xi_1 = [0 \ 0 \ -\varpi_{n1} \ 1 \ 0 \ 0]^T$ $\xi_2 = [0 \ 0 \ 0 \ 0 \ 1 \ 0]^T$ |
| Kolanowy | $\dot{\omega}_3 = [1 \ 0 \ 0]^T$ | $\mathbf{p}_4 = [0 \ (\varpi_{n1} + \varpi_{n2}) \ 0]^T$ | $\xi_4 = [0 \ 0 \ -(\varpi_{n1} + \varpi_{n2}) \ 1 \ 0 \ 0]^T$ |
| Skokowy | $\dot{\omega}_4 = [1 \ 0 \ 0]^T$ | $\mathbf{p}_5 = [0 \ (\varpi_{n1} + \varpi_{n2} + \varpi_{n3}) \ 0]^T$ | $\xi_5 = [0 \ 0 \ -(\varpi_{n1} + \varpi_{n2} + \varpi_{n3}) \ 1 \ 0 \ 0]^T$ |

11.5. Antropometria

Konstruując modele kinematyki prostej dla układu ruchu człowieka, oprócz znajomości parametrów stawów, należy również uwzględnić wymiary danych ogniw, np. ramię lub podudzie. Wymiary te można pozyskać, dokonując właściwych pomiarów osoby, dla której model jest przygotowywany. Wadą tego podejścia jest to, że wyznaczone parametry będą prawdziwe jedynie dla konkretnej osoby. Warto zwrócić również uwagę na to, że metoda ta jest czasochłonna.

Alternatywą dla przedstawionego podejścia jest wykorzystanie antropometrii [442]. Antropometria zajmuje się pomiarami porównawczymi ciała ludzkiego.



Rys. 11.5. Wymiary wybranych segmentów ciała człowieka na podstawie pomiaru wysokości ciała (na podst. [442])

W ujęciu tematu pracy najważniejsza będzie osteometria (pomiar długości kości) oraz karpometria, czyli analiza pomiarów antropometrycznych w celu badania proporcji ciała.

Pomiarów antropometrycznych można dokonywać na całym ciele lub na jego części (np. nodze). Wyróżniamy trzy podstawowe typy pomiarów, tj. długościowe, szerokościowe i obwodowe [384].

W pracy [442] przedstawiono proporcje ciała ludzkiego z wykorzystaniem pomiaru wysokości ciała. Metoda wyznaczania długości poszczególnych ogniw (segmentów ciała) przedstawiono na rys. 11.5. Zauważmy, że znając wysokość człowieka (H), jesteśmy w stanie wyliczyć długości poszczególnych segmentów jego ciała.

Wyznaczone długości segmentów ludzkiego ciała mogą być następnie wykorzystane do określenia długości segmentów w modelach, które zostały omówione w tym rozdziale. Podejście do wyznaczania długości segmentów łańcucha kinematycznego na bazie antropometrii jest jednym z możliwych podejść. Zadanie to można rozwiązać również w inny sposób. Jednym z takich sposobów jest wyznaczanie tych wielkości na bazie przeprowadzonych, odpowiednio zaprojektowanych, pomiarów. Metoda ta została zaproponowana m.in. w [72, 95].

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

Podejście jest tą małą rzeczą, która robi dużą różnicę.

Winston S. Churchill

12.1. Wstęp

W rozdziale omówiono zagadnienia związane z estymacją orientacji ciała sztywnego w przestrzeni. Rozpatrzono dwa przypadki, tj. statyczny (gdy obiekt nie porusza się lub porusza się z małą prędkością) oraz dynamiczny (gdy obiekt porusza się z dużą prędkością). W rozdziale tym zamieszczono również przegląd algorytmów do estymacji orientacji obiektu w przestrzeni oraz śledzenia zmian orientacji na podstawie danych pomiarowych.

W dalszej części tego rozdziału sformułowaliśmy dwa kolejne zadania. Jedno z tych zadań dotyczy estymacji prędkości ciała sztywnego w przestrzeni na podstawie danych pomiarowych. Natomiast drugie z nich związane jest z problemem szacowania położenia obiektu w przestrzeni. W drugim z wymienionych przypadków w rozważaniach uwzględniono to, że dysponujemy pewnym zestawem danych pomiarowych. Warto zauważyć, że typ danych pomiarowych ma wpływ na to, w jaki sposób wymienione zadania będą rozwiązane. W naszych rozważaniach przyjmujemy, że do wyznaczenia orientacji obiektu, jego prędkości oraz położenia dysponujemy danymi z czujników pomiarowych, takich jak czujniki inercyjne (tj. akcelerometr i żyroskop) oraz z czujnik natężenia pola magnetycznego (tj. magnetometr). W wyniku wykorzystania w pomiarach czujników przyspieszenia uzyskujemy pomiary przyspieszeń poruszającego się obiektu. Z kolei żyroskop umożliwia pomiar prędkości kątowych obiektu. Ostatni z czujników, który wykorzystany zostanie w projektowanym systemie, to czujnik natężenia pola magnetycznego.

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

Rozdział ten zakończono przykładowymi zadaniami estymacji orientacji obiektu w przestrzeni z wykorzystaniem różnych (omówionych szczegółowo w tym rozdziale) algorytmów.

12.2. Estymacja orientacji ciała sztywnego

Położenie obiektu (ang. *attitude*) w lokalnym układzie określa się przez jego orientację (ang. *orientation*) względem pewnego układu odniesienia. Orientacja określa sekwencję rotacji, którą należy wykonać, by lokalny układ współrzędnych związany z ciałem sztywnym (tj. lokalnym układem współrzędnych) pokrył się z globalnym układem współrzędnych (tj. układem odniesienia).

Do reprezentacji orientacji wykorzystywane są kąty Eulera, kwaternion lub macierz rotacji. Metody te zostały omówione w rozdziale 9.

W problemie estymacji orientacji ciała sztywnego w przestrzeni wyróżniamy dwa przypadki, tj. statyczny i dynamiczny. W pierwszym przypadku na obiekt nie działają żadne siły oprócz siły grawitacji. W przypadku dynamicznym uwzględniamy to, że na obiekt działają dodatkowe siły, których wpływu m.in. na pomiary nie można pominąć. Dodatkowe siły działające na obiekt mogą np. wymuszać ruch obiektu.

Problem estymacji orientacji ciała sztywnego w przestrzeni dla obiektu nieruchomego (przypadek statyczny) w literaturze nazywany jest **problemem Wahba** [426].

Sformułujmy teraz ten problem. Niech zależność opisująca współrzędne pewnego punktu \mathbf{p}_b w układzie ciała $\{B\}$ względem układu odniesienia $\{A\}$ opisana jest następującą zależnością (12.1):

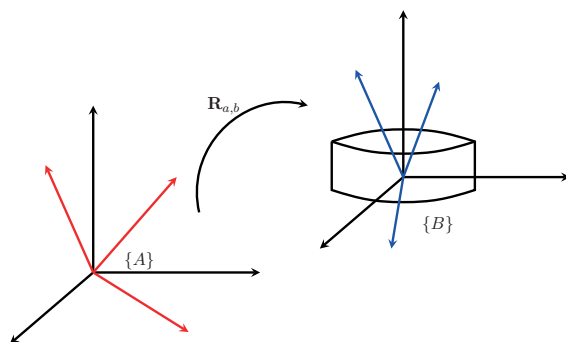
$$\mathbf{p}_a = \mathbf{R}_{a,b}\mathbf{p}_b \quad (12.1)$$

Problemem Wahba nazywamy zadanie optymalizacji, którego celem jest estymacja macierzy orientacji $\mathbf{R}_{a,b}$ przy danych współrzędnych punktów \mathbf{p}_a oraz \mathbf{p}_b i kwadratowego wskaźnika jakości. Na rysunku 12.1 zilustrowano istotę problemu Wahby.

12.2.1. Zadanie estymacji orientacji ciała sztywnego

Dysponując pomiarem współrzędnych punktu w układzie ciała, tj. \mathbf{p}_b i w układzie odniesienia \mathbf{p}_a , problem wyznaczania orientacji ciała sztywnego w przestrzeni można przedstawić następująco:

Problem 12.1 (Wyznaczenie orientacji ciała sztywnego w \mathbb{R}^3). Dla:



Rys. 12.1. Ilustracja problemu Wahby

- zbioru danych pomiarowych:

$$\mathbf{Y}_K = \{(\mathbf{p}_a(k), \mathbf{p}_b(k))\}_{k=1}^K$$

gdzie K to liczba pomiarów, \mathbf{p}_a to zbiór pomiarów w układzie ciała, a \mathbf{p}_b jest zbiorem pomiarów w układzie odniesienia.

- ustalonego wskaźnika jakości:

$$Q(\mathbf{Y}_K; \hat{\mathbf{R}}_{a,b}) = \|\mathbf{p}_a - \hat{\mathbf{R}}_{a,b}\mathbf{p}_b\|_2$$

należy znaleźć:

$$\hat{\mathbf{R}}_{a,b}^* \longrightarrow Q(\mathbf{Y}_K; \hat{\mathbf{R}}_{a,b}^*) = \min_{\hat{\mathbf{R}}_{a,b} \in \text{SO}(3)} Q(\mathbf{Y}_K; \hat{\mathbf{R}}_{a,b})$$

p.o.

$$\hat{\mathbf{R}}_{a,b} \hat{\mathbf{R}}_{a,b}^\top = \mathbf{I}$$

12.3. Algorytmy estymacji orientacji ciała sztywnego

W rozdziale omówiono wybrane algorytmy estymacji orientacji ciała sztywnego w przestrzeni \mathbb{R}^3 . Orientację obiektu w przestrzeni można wyznaczyć, rozwiązując problem Wahby. W ogólności algorytmy rozwiązujące problem Wahby dzieli się na algorytmy rozwiązujące deterministyczny oraz nondeterministyczny problem Wahby. W pierwszym przypadku przyjmuje się, że do rozwiązania tego zadania wykorzystane będą dane niezakłócone. W tym przypadku rozwiązanie zadania Wahby sprowadza się do rozwiązania układu równań nieliniowych [90]. W pierwszym przypadku do rozwiązania zadania należy wykorzystać obserwator

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

stanu. W drugim przypadku rozwiązanie zadania uzyskujemy przez zastosowanie odpowiedniego filtra lub ogólniej, estymatora stanu.

Jednym z pierwszych algorytmów rozwiązujący deterministyczny problem Wahby był algorytm TRIAD (ang. *TRI-axial Attitude Determination*) [43]. W problemie Wahby rozpatrujemy problem wyznaczania macierzy rotacji $\mathbf{R}_{b,a}$ taki, że:

$$\mathbf{p}_b = \mathbf{R}_{b,a} \mathbf{p}_a \quad (12.2)$$

gdzie \mathbf{p}_b jest wektorem w układzie lokalnym, \mathbf{p}_a jest wektorem w układzie odniesienia.

Przyjmijmy, że dysponujemy następującymi pomiarami $\mathbf{p}_a(1)$ i $\mathbf{p}_a(2)$ w układzie odniesienia oraz dwoma wektorami pomiarów $\mathbf{p}_b(1)$ i $\mathbf{p}_b(2)$ w układzie lokalnym.

Macierz rotacji $\mathbf{R}_{b,a}$ w algorytmie TRIAD wyznaczana jest z następującej zależności:

$$\mathbf{R}_{b,a} = \mathbf{R}_b \mathbf{R}_a^\top \quad (12.3)$$

gdzie

$$\mathbf{R}_b = \begin{bmatrix} \mathbf{p}_b(1) & \frac{\mathbf{p}_b(1) \times \mathbf{p}_b(2)}{|\mathbf{p}_b(1) \times \mathbf{p}_b(2)|} & \frac{(\mathbf{p}_b(1) \times (\mathbf{p}_b(1) \times \mathbf{p}_b(2)))}{|\mathbf{p}_b(1) \times \mathbf{p}_b(2)|} \end{bmatrix} \quad (12.4)$$

$$\mathbf{R}_a^\top = \begin{bmatrix} \mathbf{p}_a(1) & \frac{\mathbf{p}_a(1) \times \mathbf{p}_a(2)}{|\mathbf{p}_a(1) \times \mathbf{p}_a(2)|} & \frac{(\mathbf{p}_a(1) \times (\mathbf{p}_a(1) \times \mathbf{p}_a(2)))}{|\mathbf{p}_a(1) \times \mathbf{p}_a(2)|} \end{bmatrix}$$

W przypadku gdy pomiary są zakłócone i dysponujemy większą ich liczbą, do wyznaczenia orientacji obiektu w przestrzeni można wykorzystać algorytm QUEST (ang. *QUaternion ESTimator*). Algorytm bazuje na *q-metodzie Devenporta* służącej do wyznaczania optymalnej reprezentacji orientacji obiektu w przestrzeni w postaci kwaternionu [198, 415].

W przypadku większej liczby pomiarów problem Wahby (12.2) rozwiązywany jest przez poszukiwanie rozwiązania uprzednio sformułowanego zadania optymalizacyjnego. Dla tego zadania wskaźnik jakości oszacowania macierzy rotacji $\mathbf{R}_{b,a}$ ma postać:

$$Q(\mathbf{R}_{b,a}) = \frac{1}{2} \sum_{k=0}^K \mathbf{b}(k) \|\mathbf{p}_b(k) - \mathbf{R}_{b,a} \mathbf{p}_a(k)\|^2 \quad (12.5)$$

gdzie $\mathbf{p}_b(k)$ jest zbiorem pomiarów w układzie lokalnym, a $\mathbf{p}_a(k)$ jest zbiorem pomiarów w układzie odniesienia, natomiast $\mathbf{b}(k)$ jest zbiorem nieujemnych wag. Przepiszmy teraz wskaźnik jakości (12.5) w formie:

$$Q(\mathbf{R}_{b,a}) = \varkappa_0 - \text{trace}(\mathbf{R}_{b,a} \mathbf{R}^\top) \quad (12.6)$$

12.3. Algorytmy estymacji orientacji ciała sztywnego

gdzie

$$\varkappa_0 = \sum_{k=0}^K \mathbf{b}(k) \quad (12.7)$$

$$\mathbf{R} = \sum_{k=0}^K \mathbf{b}(k) \mathbf{p}_b(n) \mathbf{p}_a^T(k) \quad (12.8)$$

Ponieważ zależność pomiędzy macierzą rotacji $\mathbf{R}_{b,a}$ i kwaternionem \mathbf{q} ma postać [108]:

$$\mathbf{R}(\mathbf{q}) = \begin{bmatrix} \mathbf{q}_4 \mathbf{I}_{3 \times 3} + \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \end{bmatrix}^T \begin{bmatrix} \mathbf{q}_4 \mathbf{I}_{3 \times 3} - \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \end{bmatrix} \quad (12.9)$$

więc

$$\text{trace}(\mathbf{R}_{b,a} \mathbf{R}^T) = \mathbf{q}^T \mathbf{K} \mathbf{q} \quad (12.10)$$

gdzie [90]

$$\mathbf{K} = \begin{bmatrix} \mathbf{R} + \mathbf{R}^T - \text{trace}(\mathbf{R}) \mathbf{I}_{3 \times 3} & \sum_{k=0}^K \mathbf{b}(k) \mathbf{p}_b(k) \times \mathbf{p}_a(k) \\ \left(\sum_{k=0}^K \mathbf{b}(k) \mathbf{p}_b(k) \times \mathbf{p}_a(k) \right)^T & \text{trace}(\mathbf{R}) \end{bmatrix} \quad (12.11)$$

Co prowadzi do następującej modyfikacji wskaźnika jakości estymacji orientacji ciała sztywnego w przestrzeni (12.6):

$$Q(\mathbf{R}_{b,i}) = \varkappa_0 - \varkappa_{\max} \quad (12.12)$$

gdzie \varkappa_{\max} spełnia

$$\mathbf{K} \mathbf{q}^* = \varkappa_{\max} \mathbf{q}^* \quad (12.13)$$

Inne algorytmy należące do tej grupy to algorytm bazujący na dekompozycji SVD (ang. *Singular Value Decomposition*) [274], algorytm FOAM (ang. *Fast Optimal Matrix Algorithm*) [275], algorytmy ESOQ-1 i ESOQ-2 (ang. odpowiednio:

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

Estimator of the Opimal Quaternion i *Second Estimator of the Opimal Quaternion*) oraz algorytm EAA (ang. *Energy Approach Algorithm*) [294].

Do estymacji orientacji obiektu w przestrzeni można wykorzystać nie tylko pomiary pochodzące z czujników przyspieszenia i żyroskopu, ale również z magnetometru. Wykorzystanie dodatkowej wielkości, tj. natężenia pola magnetycznego wpływa na poprawę jakości oszacowania orientacji obiektu. Niestety takie rozwiązanie jest wrażliwe na obecność ferromagnetyków i/lub urządzeń podłączonych do sieci elektrycznej [345], [347]. W tym przypadku zakłócony pomiar z magnetometru wpływa na pogorszenie jakości estymacji orientacji. W pracy [462] zaproponowano algorytm FQA (ang. *Factored Quaternion Algorithm*), który bazując na pomiarach zarówno z czujników inercyjnych (tj. akcelerometru i żyroskopu), jak i z magnetometru, szacuje orientację obiektu w przestrzeni oraz minimalizuje wpływ zakłóceń magnetycznych. Cechą charakterystyczną zaproponowanego rozwiązania jest to, że wyliczenia wartości kwaternionu reprezentującego orientację obiektu w przestrzeni podzielono na dwa etapy. Na pierwszym etapie wyznaczane są dwa kąty Eulera (φ oraz ϑ). Kąt ψ wyznaczany jest w drugim kroku. Zaproponowane w algorytmie FQA etapowe wyznaczanie kątów Eulera, zmniejsza wpływ zakłóceń powodowanych np. przez ferromagnetyki na jakość ich estymacji.

Przyjmijmy, że do reprezentacji obiektu w przestrzeni korzystamy z kwaternionu:

$$\mathbf{q} = [q_0 \quad q_1 \quad q_2 \quad q_3]^\top \quad (12.14)$$

Dodatkowo dysponujemy pomiarami przyspieszenia z trójosiowego akcelerometru:

$$\mathbf{y}_a = [y_{a_x} \quad y_{a_y} \quad y_{a_z}]^\top \quad (12.15)$$

jak i natężenia pola magnetycznego z trójosiowego magnetometru:

$$\mathbf{y}_m = [y_{m_x} \quad y_{m_y} \quad y_{m_z}]^\top \quad (12.16)$$

Zależności szacujące kwaterniony związane z poszczególnymi kątami Eulera są następujące [462]:

- kąt φ :

$$\mathbf{q}_\varphi = \cos\left(\frac{\varphi}{2}\right) [1 \quad 0 \quad 0 \quad 0]^\top + \sin\left(\frac{\varphi}{2}\right) [0 \quad 0 \quad 1 \quad 0]^\top \quad (12.17)$$

- kąt ϑ :

$$\mathbf{q}_\vartheta = \cos\left(\frac{\vartheta}{2}\right) [1 \quad 0 \quad 0 \quad 0]^\top + \sin\left(\frac{\vartheta}{2}\right) [0 \quad 1 \quad 0 \quad 0]^\top \quad (12.18)$$

- kąt ψ :

$$\mathbf{q}_\psi = \cos\left(\frac{\psi}{2}\right) [1 \ 0 \ 0 \ 0]^\top + \sin\left(\frac{\psi}{2}\right) [0 \ 0 \ 0 \ 1]^\top \quad (12.19)$$

Opierając się na oszacowaniach poszczególnych kwaternionów ((12.17)–(12.19)), ostatecznie otrzymujemy:

$$\hat{\mathbf{q}} = \mathbf{q}_\varphi \mathbf{q}_\theta \mathbf{q}_\psi \quad (12.20)$$

Wadą algorytmu FQA jest to, że jest on dostosowany do szacowania orientacji obiektu w przestrzeni tylko wówczas, gdy obiekt nie porusza się lub porusza się bardzo powoli [280]. Oznacza to, że wyniki uzyskane przez zastosowanie algorytmu FQA dla obiektów będących w ruchu zwykle są niedokładne.

Wśród dostępnych pomiarów przyspieszenia i natężenia pola magnetycznego dostępne są również pomiary prędkości kątowych, które mierzone są z wykorzystaniem żyroskopu. Dalej omówiona zostanie klasa metod do szacowania orientacji obiektu w przestrzeni, które bazują na pomiarach przyspieszenia, natężenia pola magnetycznego oraz prędkości kątowej.

Oprócz omówionych już metod estymacji orientacji obiektu w przestrzeni, w literaturze zaproponowano wiele innych metod, które możemy podzielić na metody oparte na filtrze Kalmana [350, 351], nieliniowym obserwatorze stanu (ang. *nonlinear state observer*) [37, 38, 146, 267, 352] oraz filtrze komplementarnym (ang. *complementary filter*) [264, 266].

Klasyczny liniowy filtr Kalmana do szacowania orientacji obiektu w przestrzeni jest stosowany rzadko. Najczęściej stosowanym rozwiązaniem w tym przypadku jest rozszerzony filtr Kalmana [60, 328]. Filtr ten bazuje na rozwinięciu w szereg Taylora i szczegółowo został omówiony w podrozdziale 6.3.2. Filtr ten pomimo swoich wad i ograniczeń jest szeroko stosowanym rozwiązaniem w systemach do estymacji orientacji obiektu w przestrzeni, w których wykorzystywane są dane pomiarowe z różnych czujników.

Zastosowanie filtra EKF w zadaniu estymacji orientacji obiektu w przestrzeni wiąże się z kilkoma problemami, które należy uwzględnić, projektując algorytm estymacji orientacji bazujący na tym filtrze. Jednym z problemów jest **problem rozbieżności** (ang. *divergence problem*) [326, 333]. Problem rozbieżności związany jest z samą konstrukcją algorytmu, tj. filtr EKF pracuje w ustalonym punkcie pracy. W przypadku gdy model procesu lub model układu pomiarowego ewentualnie zarówno jeden, jak i drugi model są silnie nieliniowe, wówczas może wystąpić zjawisko rozbieżności.

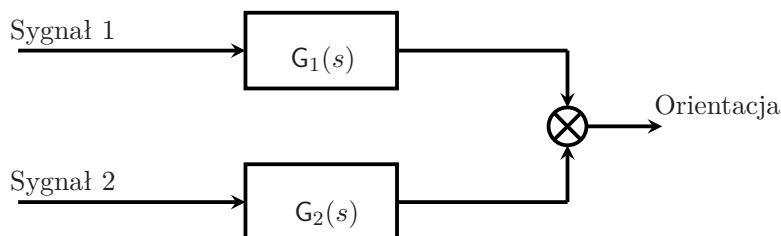
Inny problem, który związany jest z filtrem EKF, dotyczy modelu zakłóceń. W równaniach filtra EKF przyjmuje się *gaussowski* model szumu, zarówno dla

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

procesu jak i dla pomiaru. W rzeczywistości jest to warunek trudny do spełnienia, co również może doprowadzić do omówionego wcześniej problemu rozbieżności [138]. Kolejną wadą tego filtra jest to, że estymacja orientacji z wykorzystaniem EKF jest zadaniem kosztownym obliczeniowo.

Rozwiązaniem części z opisanych problemów związanych z filtrem EKF jest dwustopniowy filtr EKF (ang. *two-stage EKF*) [206, 282, 467]. Alternatywą dla filtra EKF jest filtr komplementarny (ang. *complementary filter*). Filtr komplementarny jest przykładem procesu fuzji danych na poziomie pierwszym, tj. poziomie danych, ponieważ dane z czujników są jednomodalne (por. z rozdz. 2). Kluczowym elementem filtracji komplementarnej jest dobór odpowiednich czujników do fuzji danych. Czujniki powinny zostać tak dobrane, by ich charakterystyki częstotliwościowe były względem siebie komplementarne. Jeden z czujników powinien być czuły na składowe wolnozmiennne (ang. *low frequency component*) w mierzonej wielkości. Z kolei drugi z nich powinien być czuły na składowe szybkozmiennne (ang. *high frequency component*).

Czujnikami czułymi na zmiany w niskich zakresach częstotliwości są akcelerometry i magnetometry, a w zakresach wysokich częstotliwości wykorzystywane są w tym celu żyroskopy. Przykład prostego filtra komplementarnego przedstawiono na rys. 12.2.



Rys. 12.2. Filtr komplementarny

Istotnym elementem filtra komplementarnego jest właściwy dobór transmitancji filtrów wchodzących w jego skład. Transmitancje tych filtrów powinny być dobierane tak, by:

$$\begin{aligned} G_1(s) &= 1 - G(s) \\ G_2(s) &= G(s) \end{aligned} \tag{12.21}$$

gdzie $G_1(s)$ jest transmitancją filtra 1, a $G_2(s)$ jest transmitancją filtra 2. Jeżeli $G(s)$ jest transmitancją filtra górnoprzepustowego, to $1 - G(s)$ jest transmitancją filtra dolnoprzepustowego. Wyznamy transformaty Laplace'a sygnałów przed-

12.3. Algorytmy estymacji orientacji ciała sztywnego

stawionych na rys. 12.2:

$$\begin{aligned}\mathcal{L}[s(t) + n_1(t)] &= S(s) + N_1(s) \\ \mathcal{L}[s(t) + n_2(t)] &= S(s) + N_2(s) \\ \mathcal{L}[y(t)] &= Y(s)\end{aligned}\tag{12.22}$$

gdzie \mathcal{L} to transformata Laplace'a. Uwzględniając zależności (12.21), transmitancję operatorową filtra komplementarnego możemy zapisać w sposób następujący:

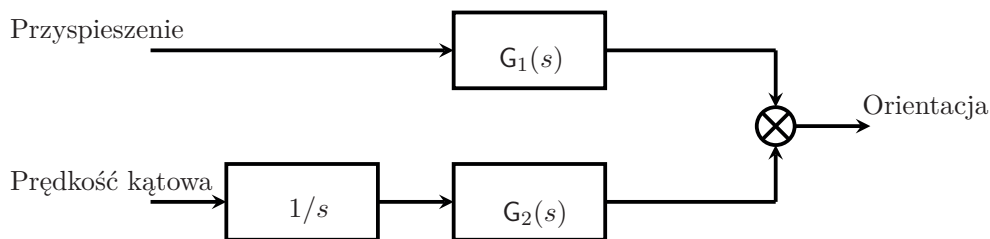
$$\begin{aligned}Y(s) &= (1 - G(s))(S(s) - N_1(s)) + G(s)(S(s) - N_2(s)) = \\ &= (1 - G(s))S(s) + (1 - G(s))N_1(s) + G(s)X(s) + G(s)N_2(s) = \\ &= S(s) + (1 - G(s))N_1(s) + G(s)N_2(s)\end{aligned}\tag{12.23}$$

Zauważmy (12.23), że właściwy dobór filtra komplementarnego umożliwia wyeliminowanie zakłóceń $n_1(t)$ oraz $n_2(t)$, tj. wpływu wolno- i szybkozmiennych składowych. W tym przypadku zachodzi $y(t) = s(t)$. Niestety w rzeczywistości idealne odfiltrowanie zakłóceń z sygnału pomiarowego jest niemożliwe. Oznacza to, że w ogólnym przypadku $y(t) \sim s(t)$.

Warto zwrócić uwagę na zjawisko, którego wystąpienie należy uwzględnić, projektując filtr komplementarny. Do filtrowania zakłóceń szybkozmiennych można zastosować filtr dolnoprzepustowy. Jednakże należy pamiętać, że taki filtr wprowadza opóźnienie do przetwarzanego sygnału.

Składową wolnozmienną (np. dryft) można usunąć, stosując filtr górnoprzepustowy. Należy jednak pamiętać, że, oprócz zakłóceń, mogą być wyeliminowane z sygnału również składowe użyteczne. Projektując filtr komplementarny, należy wziąć to pod uwagę.

Na rysunku 12.3 przedstawiono schemat filtra komplementarnego do szacowania orientacji obiektu w przestrzeni.



Rys. 12.3. Filtr komplementarny do szacowania orientacji

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

Kolejną grupą metod wykorzystywanych w zadaniach estymacji orientacji są algorytmy bazujące na nieliniowych obserwatorach stanu (ang. *nonlinear attitude observers*). Cechą charakterystyczną tych algorytmów jest to, że bazują one na macierzy rotacji. Jedną z wad tego rozwiązania jest duża wrażliwość algorytmów na zakłócenia w pomiarach. Inną wadą tych metod są duże koszty obliczeniowe, które związane są głównie z przetwarzaniem 9-elementowej macierzy rotacji.

Przykład filtra bazującego na nieliniowym obserwatorze stanu zastosowanego do szacowania orientacji ciała sztywnego w przestrzeni zaprezentowano w pracy [266]. Idea przedstawionego algorytmu filtra jest podobna do omówionego już filtra komplementarnego.

Niech $\mathbf{R} \in \mathbf{SO}(\mathbf{3})$ będzie macierzą rotacji, która reprezentuje orientację obiektu w układzie lokalnym względem układu odniesienia, oraz niech $\mathbf{y}_\omega = [y_{\omega_x} \ y_{\omega_y} \ y_{\omega_z}]^T$ będzie wektorem prędkości kątowej w układzie lokalnym. Przypomnijmy, że (por. (9.36)):

$$[\omega] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \quad (12.24)$$

jest macierzą antysymetryczną (skośnie symetryczną). Wiemy również, że równanie dynamiki obiektu ma postać (patrz podrozdz. 9.5):

$$\dot{\mathbf{R}} = \mathbf{R}\omega = [\mathbf{R}\omega]\mathbf{R} \quad (12.25)$$

Oznaczmy oszacowanie macierzy rotacji \mathbf{R} przez $\hat{\mathbf{R}}$. Celem zadania szacowania orientacji jest $\mathbf{R} \sim \hat{\mathbf{R}}$. Z kolei błąd oszacowania orientacji obiektu wyznaczamy z zależności:

$$\tilde{\mathbf{R}} = \hat{\mathbf{R}}\mathbf{R}^T \quad (12.26)$$

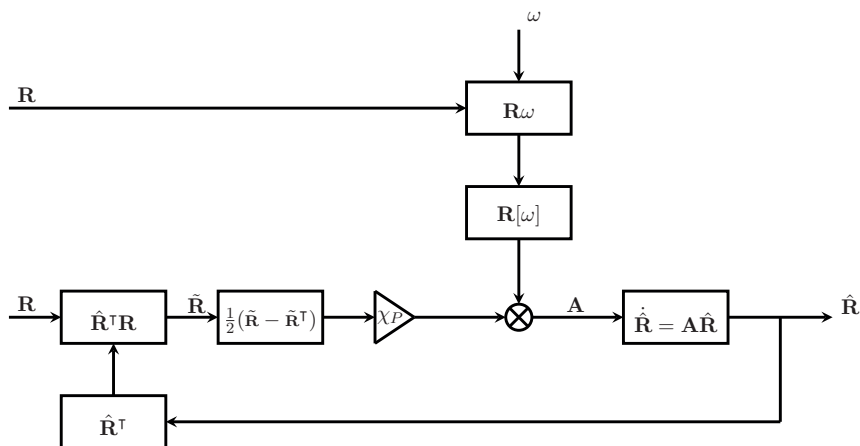
gdzie $\tilde{\mathbf{R}}$ jest błędem tego oszacowania w układzie odniesienia. Ogólne równanie dla nieliniowego obserwatora stanu ma postać:

$$\dot{\hat{\mathbf{R}}} = [\mathbf{R}\omega + \chi_P \hat{\mathbf{R}}\eta]\hat{\mathbf{R}} \quad (12.27)$$

gdzie χ_P jest dodatnim współczynnikiem skalującym, a η jest współczynnikiem korekcji.

Celem konstrukcji obserwatora stanu do szacowania orientacji jest taki dobór współczynnika korekcji η , że $\hat{\mathbf{R}} \rightarrow \mathbf{I}$. Schemat blokowy algorytmu przedstawiono na rys. 12.4. Algorytm ten w literaturze nazywany jest algorytmem Mahonego.

Tak jak wspomniano, wadą przedstawionego algorytmu oraz innych bazujących na nieliniowym obserwatorze stanu jest wrażliwość na zakłócenia pomiarowe.



Rys. 12.4. Filtr orientacji bazujący na nieliniowym obserwatorze stanu

Zakłócenia te mogą prowadzić do problemu rozbieżności w oszacowaniach orientacji obiektu w przestrzeni [366].

Algorytm zaproponowany w [263, 264] należy, podobnie jak przedstawiony wcześniej algorytm Mahonego, do grupy algorytmów z filtrem o stałym wzmocnieniu. W tym przypadku orientacja obiektu w przestrzeni jest reprezentowana przez kwaternion, dzięki czemu można uniknąć problemu niejednoznaczności (ang. *singularity problem*). Schemat ideowy algorytmu Madgwicka został przedstawiony na rys. 12.5.

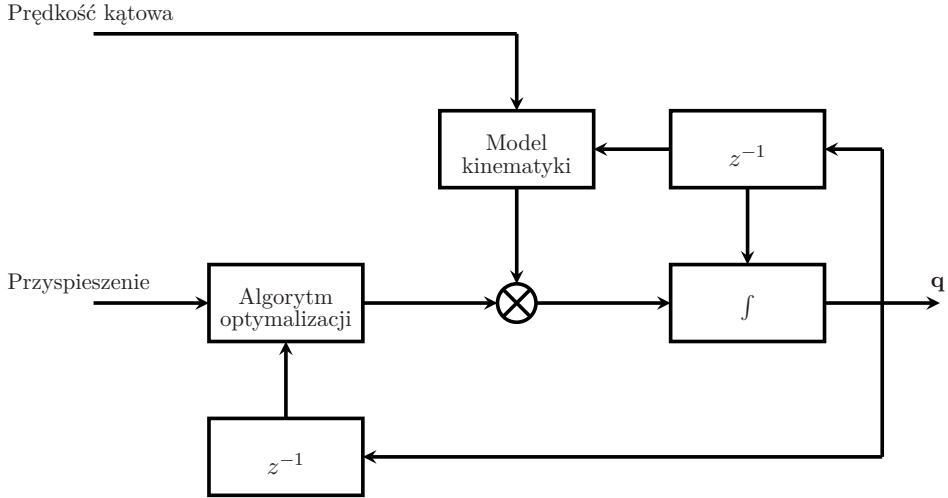
Pierwszy etap szacowania orientacji obiektu w przestrzeni to estymacja kwaternionu z wykorzystaniem danych pozyskanych z żyroskopu. Drugi etap to estymacja kwaternionu na podstawie pomiarów z akcelerometru i magnetometru. Następnie tak oszacowany kwaternion jest zastosowany do korekcy kwaternionu oszacowanego na podstawie danych z żyroskopu na pierwszym etapie. Etap korekty wykonany jest z wykorzystaniem metod optymalizacji numerycznej. W oryginalnym algorytmie zaproponowanym przez Madgwicka zastosowano metodę najszybszego spadku (ang. *gradient descent method*).

Ostatni etap związany jest z fuzją danych, której celem jest poprawa oszacowania orientacji wykonanej na wcześniejszych etapach. Ponieważ podczas szacowania orientacji obiektu mogą pojawić się zakłócenia zarówno wolnozmiennie, jak i szybkozmiennie, algorytm powinien umożliwiać filtrację tych zakłóceń.

Oznaczmy dane z czujnika żyroskopowego w następujący sposób

$$\mathbf{y}_\omega = [0 \quad y_{\omega_x} \quad y_{\omega_y} \quad y_{\omega_z}]^T$$

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej



Rys. 12.5. Idea filtra Madgwicka

Zapiszmy równanie wiążące zmiany wartości kwaternionu i prędkości kątowej:

$$\dot{\hat{\mathbf{q}}}_{a,b} = \frac{1}{2} \hat{\mathbf{q}}_{a,b} \otimes \mathbf{y}_\omega \quad (12.28)$$

Kwaternion w równaniu (12.28) można wyliczyć numerycznie (por. (9.36)):

$$\hat{\mathbf{q}}_{a,b}^{(i+1)} = \hat{\mathbf{q}}_{a,b}^{(i)} + \Delta t \left(\frac{1}{2} \hat{\mathbf{q}}_{a,b} \otimes \mathbf{y}_\omega \right) \quad (12.29)$$

Wyniki estymacji orientacji obiektu w przestrzeni bazujące jedynie na pomiarach z żyroskopu są niedokładne. W celu poprawy wyników szacowania do obliczeń w filtrze Madgwicka wykorzystano dodatkowe pomiary z czujników przyspieszenia oraz magnetometru. Oznaczmy pomiary przyspieszenia jako $\mathbf{y}_a = [y_{a_x} \ y_{a_y} \ y_{a_z}]^\top$, a pomiary z magnetometru następująco $\mathbf{y}_m = [y_{m_x} \ y_{m_y} \ y_{m_z}]^\top$. Wartości kwaternionu w opisywanym algorytmie wyznaczane są na bazie wymienionych pomiarów oraz przez rozwiązanie następującego zadania optymalizacyjnego:

$$\hat{\mathbf{q}}_{a,b} = \min_{\mathbf{q}_{a,b} \in \mathbb{R}^4} Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \tilde{\mathbf{h}}) \quad (12.30)$$

gdzie wykorzystany wskaźnik jakości ma postać:

$$Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \tilde{\mathbf{h}}) = \begin{bmatrix} Q_1(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g) \\ Q_2(\mathbf{q}_{a,b}, \mathbf{y}_m, \tilde{\mathbf{h}}) \end{bmatrix} \quad (12.31)$$

12.3. Algorytmy estymacji orientacji ciała sztywnego

Sformułowane zadanie optymalizacyjne (12.30) można rozwiązać metodami numerycznymi. Ogólna postać algorytmu numerycznego jest następująca:

$$\mathbf{q}_{a,b}^{(i+1)} = \mathbf{q}_{a,b}^{(i)} - \eta \frac{\nabla Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \hbar)}{\|\nabla Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \hbar)\|} \quad (12.32)$$

gdzie $\eta \in \mathbb{R}_+$, \mathbf{a}_g reprezentuje przyspieszenie związane z grawitacją, a \hbar jest wartością nachylenia magnetycznego w danej lokalizacji. Z kolei $\nabla Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \hbar)$ w zależności (12.32) wyznaczane jest jako:

$$\nabla Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \hbar) = \mathbf{J}^\top(\mathbf{q}_{a,b}, \hbar) Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{y}_m, \hbar) \quad (12.33)$$

gdzie $\mathbf{J}(\mathbf{q}_{a,b}, \hbar)$ jest macierzą Jakobiego, której ogólna postać w omawianym zadaniu ma postać:

$$\mathbf{J}(\mathbf{q}_{a,b}, \hbar) = \begin{bmatrix} \mathbf{J}_1^\top(\mathbf{q}_{a,b}) \\ \mathbf{J}_2^\top(\mathbf{q}_{a,b}, \hbar) \end{bmatrix} \quad (12.34)$$

Uwzględniając to, że dysponujemy pomiarami z czujnika przyspieszenia i magnetometru, odpowiednie wskaźniki jakości mają postać:

$$Q_1(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g) = \begin{bmatrix} 2(q_2q_4 - q_1q_3) - y_{a_x} \\ 2(q_1q_2 + q_3q_4) - y_{a_y} \\ 2\left(\frac{1}{2} - q_2^2 - q_3^2\right) - y_{a_z} \end{bmatrix} \quad (12.35)$$

$$Q_2(\mathbf{q}_{a,b}, \mathbf{y}_m, \hbar) = \begin{bmatrix} 2\hbar_x \left(\frac{1}{2} - q_3^2 - q_4^2\right) + 2\hbar_z (q_2q_4 - q_1q_3) - y_{m_y} \\ 2\hbar_x (q_2q_3 - q_1q_4) + 2\hbar_z (q_1q_2 + q_3q_4) - y_{m_y} \\ 2\hbar_x (q_1q_3 + q_2q_4) + 2\hbar_z \left(\frac{1}{2} - q_2^2 - q_3^2\right) - y_{m_z} \end{bmatrix} \quad (12.36)$$

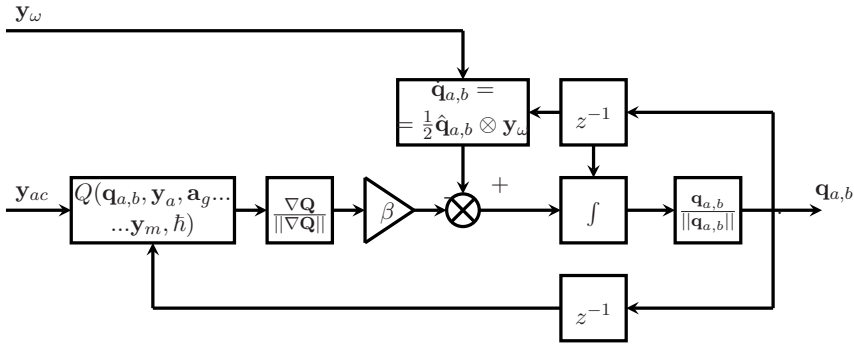
Z kolei macierze Jakobiego w tym zadaniu wyliczane są z zależności:

$$\mathbf{J}_1^\top(\mathbf{q}_{a,b}) = \begin{bmatrix} -2q_3 & 2q_4 & -2q_1 & 2q_2 \\ 2q_2 & 2q_1 & 2q_4 & 2q_3 \\ 0 & -4q_2 & -4q_3 & 0 \end{bmatrix} \quad (12.37)$$

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

$$\mathbf{J}_1^T(\mathbf{q}_{a,b}, \mathbf{h}) = \begin{bmatrix} -2\hbar_z q_3 & 2\hbar_z q_4 & -4\hbar_x q_3 - 2\hbar_z q_1 & -4\hbar_x q_4 + 2\hbar_z q_2 \\ -2\hbar_x q_4 + 2\hbar_x q_2 & 2\hbar_x q_3 + 2\hbar_z q_1 & 2\hbar_x q_2 + 2\hbar_z q_4 & -2\hbar_x q_1 + 2\hbar_z q_3 \\ 2\hbar_x q_3 & 2\hbar_x q_4 - 4\hbar_z q_2 & 2\hbar_x q_1 - 4\hbar_z q_3 & 2\hbar_x q_2 \end{bmatrix} \quad (12.38)$$

Schemat ogólny algorytmu Madgwicka przedstawiono na rys. 12.5. Na rysunku 12.6 przedstawiono szczegółowy schemat algorytmu Madgwicka bazujący na danych z czujników przyspieszenia i żyroskopowego, natomiast na rys. 12.7 przedstawiono schematy zmodyfikowanego algorytmu Madgwicka, w których dodatkowo wykorzystany został pomiar z magnetometru.



Rys. 12.6. Filtr Madgwicka z wykorzystaniem sygnału przyspieszenia i prędkości kątowej

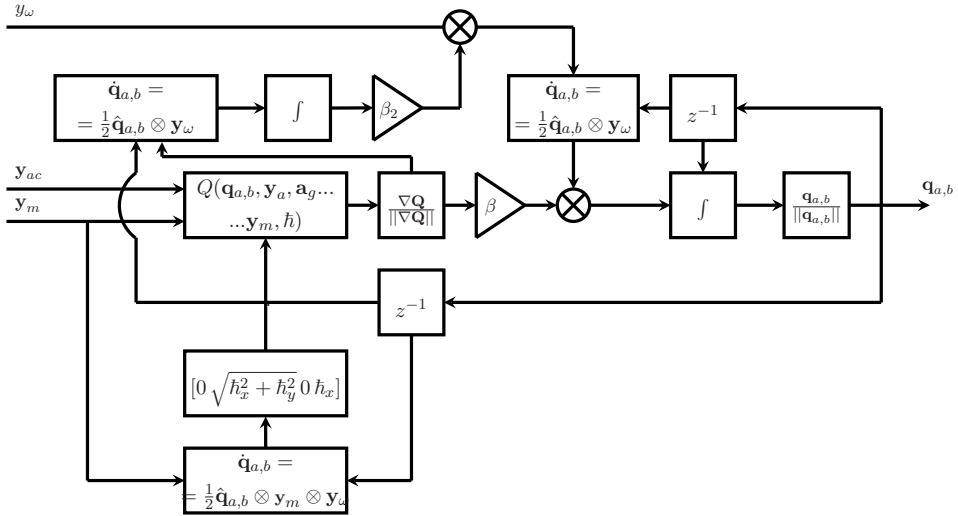
Przedstawiony algorytm Madgwicka charakteryzuje się stałym wzmocnieniem. W pracy [403] zaproponowano algorytm, w którym wzmocnienie jest wyznaczone adaptacyjnie. Sformułowanie problemu jest takie same jak dla poprzedniego zadania. Natomiast, zamiast metody najszybszego spadku do numerycznego rozwiązania zadania optymalizacji zaproponowano metodę Gaussa–Newtona:

$$\mathbf{q}_{a,b}^{(i+1)} = \mathbf{q}_{a,b}^{(i)} - \eta(\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T Q(\mathbf{q}_{a,b}, \mathbf{y}_a, \mathbf{a}_g, \mathbf{y}_m, \mathbf{h}) \quad (12.39)$$

W zależności (12.39) współczynnik η był stały. W zaproponowanym algorytmie współczynnik ten jest wyznaczany w każdej iteracji zgodnie z następującą zależnością:

$$\eta = \nu \|\mathbf{q}_{a,b}(k+1)\| \Delta t \quad (12.40)$$

gdzie $\nu > 1$. Schemat algorytmu przedstawiono na rys. 12.8.



Rys. 12.7. Filtr Madgwicka z wykorzystaniem sygnału przyśpieszenia, prędkości kątowej i natężenia pola magnetycznego

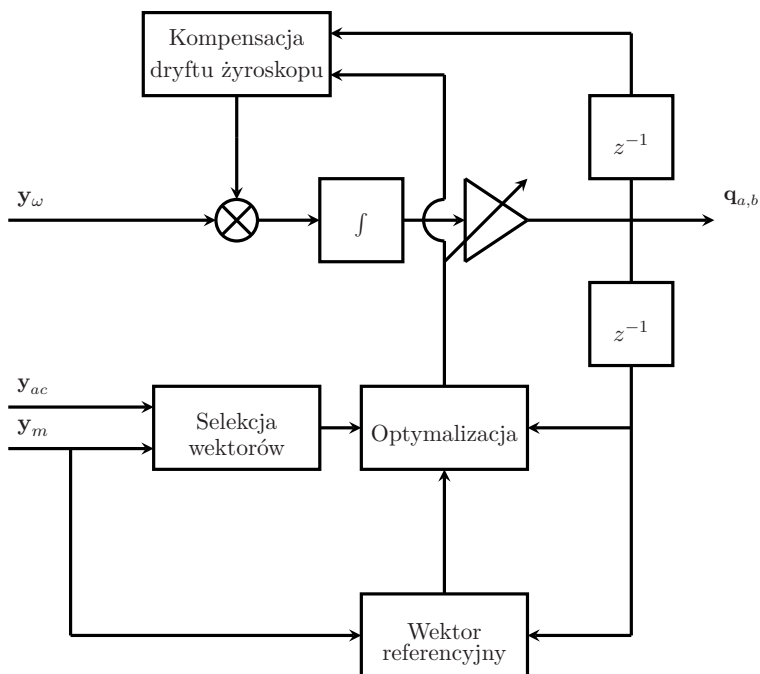
Przykład 12.1 (Porównanie algorytmów Mahonego i Madgwicka). *Jako podsumowanie tego rozdziału zostanie przedstawiony sposób wykorzystania algorytmów estymacji orientacji. W badaniach wykorzystano dane z czujników przyśpieszenia, prędkości kątowych oraz natężenia pola magnetycznego. Przykład został częściowo opracowany na bazie wyników przedstawionych w pracach [237, 238].*

Oznaczmy pomiary przyśpieszeń jako $y_a(k)$, prędkości kątowych jako $y_\omega(k)$, a natężenie pola magnetycznego $y_m(k)$, dla $k = 1, 2, \dots, K$. Wynik porównania algorytmów Madgwicka i Mahonego przedstawiono na rys. 12.9. Eksperyment wykonany został przy niewielkiej prędkości dokonywanych zmian w orientacji czujnika. W takich warunkach eliminuje się wpływ przyśpieszenia związanego z ruchem czujnika.

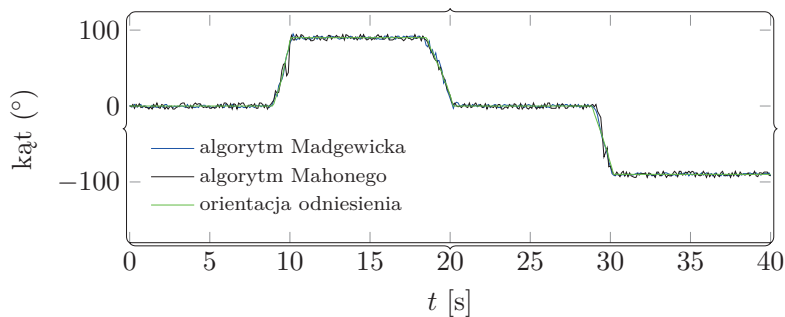
12.4. Wykorzystanie filtrów Kalmana do szacowania orientacji obiektu w przestrzeni

Na koniec tego rozdziału przedstawimy, w jaki sposób skonstruować algorytm estymacji orientacji obiektu w przestrzeni, który będzie bazował na filtrze Kalmana i rozszerzonym filtrze Kalmana. Sytuacja pomiarowa została przedstawiona na rys. 12.10. Do obiektu umieszczonego w przestrzeni przymocowano urządzenie po-

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej



Rys. 12.8. Filtr ze wzmocnieniem adaptacyjnym [403]

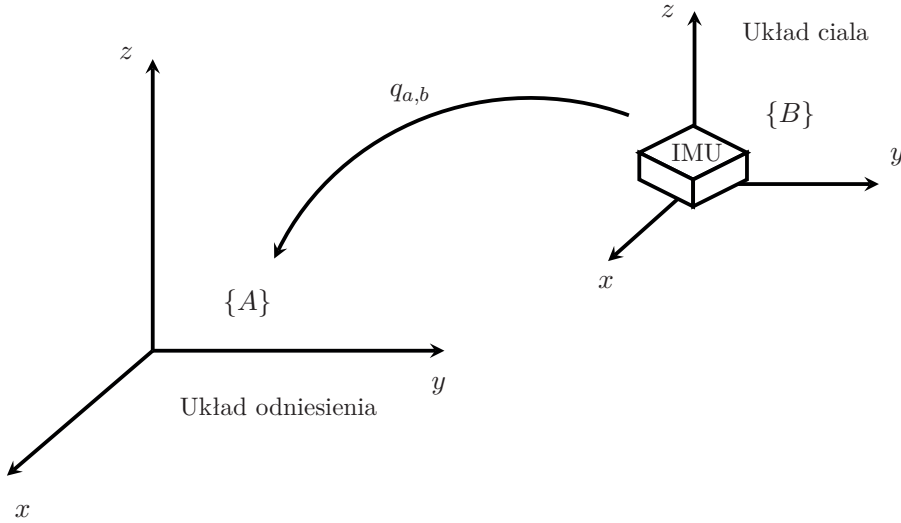


Rys. 12.9. Porównanie wyników estymacji orientacji czujnika z wykorzystaniem algorytmów Madgwicka i Mahonego

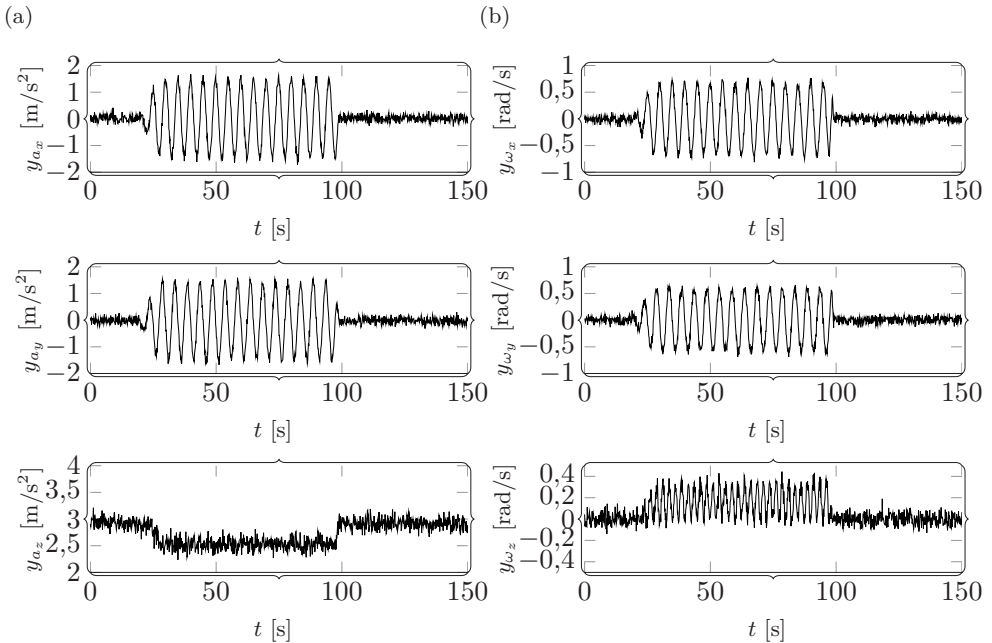
miarowe typu IMU, które umożliwia pomiar przyspieszeń (czujnik przyspieszenia) oraz prędkości kątowych (czujnik żyroskopowy).

Oznaczmy pomiary przyspieszeń jako $\mathbf{y}_a(k)$ (rys. 12.11a), a pomiary prędkości kątowych niech będą reprezentowane przez $\mathbf{y}_\omega(k)$ dla $k = 1, 2, \dots, K$ (rys. 12.11b).

12.4. Wykorzystanie filtrów Kalmana do szacowania orientacji...

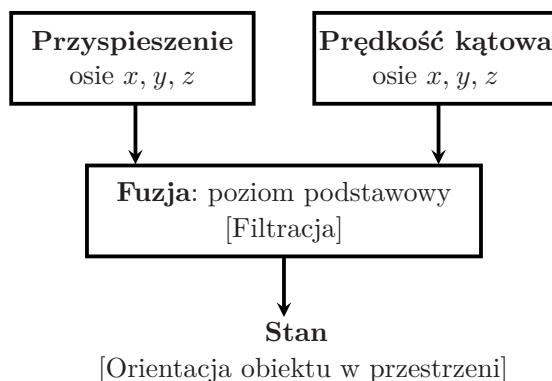


Rys. 12.10. Ilustracja sytuacji pomiarowej dla przykładu z podrozdziału 12.4



Rys. 12.11. Pomiary wykorzystane w zadaniu: a) przyśpieszenie dla osi x , y i z ,
b) prędkość kątowna dla osi x , y i z

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej



Rys. 12.12. Schemat fuzji danych dla estymacji orientacji obiektu w przestrzeni

Orientacja obiektu będzie reprezentowana przez kwaternion $(\mathbf{q}_{a,b})$. Indeks a związany jest z układem odniesienia, a indeks b z układem lokalnym. Model opisujący dynamikę systemu ma postać:

$$\dot{\mathbf{q}}_{a,b} = \frac{1}{2} \hat{\mathbf{q}}_{a,b} \otimes \omega_b \quad (12.41)$$

Przedstawimy dwa podejścia do rozwiązania zadania estymacji orientacji obiektu w przestrzeni z wykorzystaniem pomiarów z czujnika przyspieszenia oraz żyroskopu. Pierwsze z nich bazuje na filtrze Kalmana, a drugie na rozszerzonym filtrze Kalmana. Układ pomiarowy wraz z algorytmami wstępnego przetworzenia danych pomiarowych oraz algorytmem fuzji danych na poziomie podstawowym przedstawiono na rys. 12.12.

12.4.1. Rozwiązanie z wykorzystaniem filtra Kalmana

Wykorzystując omówiony już model systemu (12.41), ustalmy wektor stanu w postaci:

$$\mathbf{x} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (12.42)$$

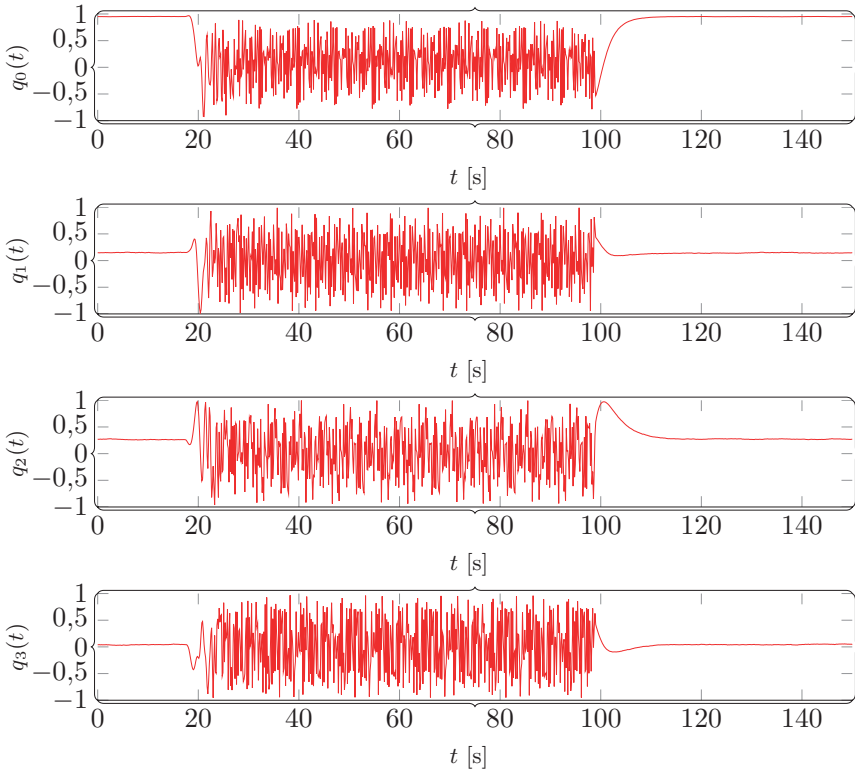
12.4. Wykorzystanie filtrów Kalmana do szacowania orientacji...

gdzie składowe wektora q_i są składowymi kwaternionu. Zapiszmy teraz równanie (12.41) w formie dyskretnej (por. (12.28)):

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} (i+1) = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} (i) + \frac{1}{2} \Delta t \begin{bmatrix} 0 & -y\omega_x & -y\omega_y & -y\omega_z \\ y\omega_x & 0 & y\omega_z & -y\omega_y \\ y\omega_y & y\omega_z & 0 & y\omega_x \\ y\omega_z & y\omega_y & -y\omega_x & 0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \quad (12.43)$$

Stąd macierz \mathbf{A} w filtrze Kalmana ma postać:

$$\mathbf{A} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} (i) + \frac{1}{2} \Delta t \begin{bmatrix} 0 & -y\omega_x & -y\omega_y & -y\omega_z \\ y\omega_x & 0 & y\omega_z & -y\omega_y \\ y\omega_y & y\omega_z & 0 & y\omega_x \\ y\omega_z & y\omega_y & -y\omega_x & 0 \end{bmatrix}. \quad (12.44)$$



Rys. 12.13. Wartości kwaternionu oszacowane z wykorzystaniem filtra Kalmana

12. Orientacja ciała sztywnego w przestrzeni trójwymiarowej

Zapiszmy teraz wektor pomiarów przyspieszeń jako:

$$\mathbf{y} = \begin{bmatrix} y_{a_x} \\ y_{a_y} \\ y_{a_z} \end{bmatrix} \quad (12.45)$$

Uwzględniając (12.45), macierz \mathbf{H} w modelu układu pomiarowego ma postać:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (12.46)$$

Macierze kowariancji \mathbf{Q} oraz \mathbf{R} należy dobrać eksperymentalnie. Dysponując podanymi zależnościami, możliwe jest wyznaczenie orientacji obiektu w przestrzeni, korzystając z algorytmu 1. Wynik estymacji kwaternionu reprezentującego orientację obiektu w przestrzeni z wykorzystaniem filtra Kalmana przedstawiono na rys. 12.13.

12.4.2. Rozwiązanie z wykorzystaniem rozszerzonego filtra Kalmana

Zadanie wyznaczenia orientacji obiektu w przestrzeni może być również rozwiązane z wykorzystaniem rozszerzonego filtra Kalmana. Dokonajmy linearyzacji równania stanu w modelu (6.37):

$$\Delta x = \frac{\partial f(x)}{\partial x} \Big|_{x=\hat{x}} \Delta x + n(k) \quad (12.47)$$

Zapiszmy teraz równanie (12.47) w postaci dyskretniej:

$$\Delta x(k+1) = \mathbf{F}(k)\Delta x(k) + n(k) \quad (12.48)$$

gdzie macierz \mathbf{F} ma postać:

$$\mathbf{F}(k) = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \quad (12.49)$$

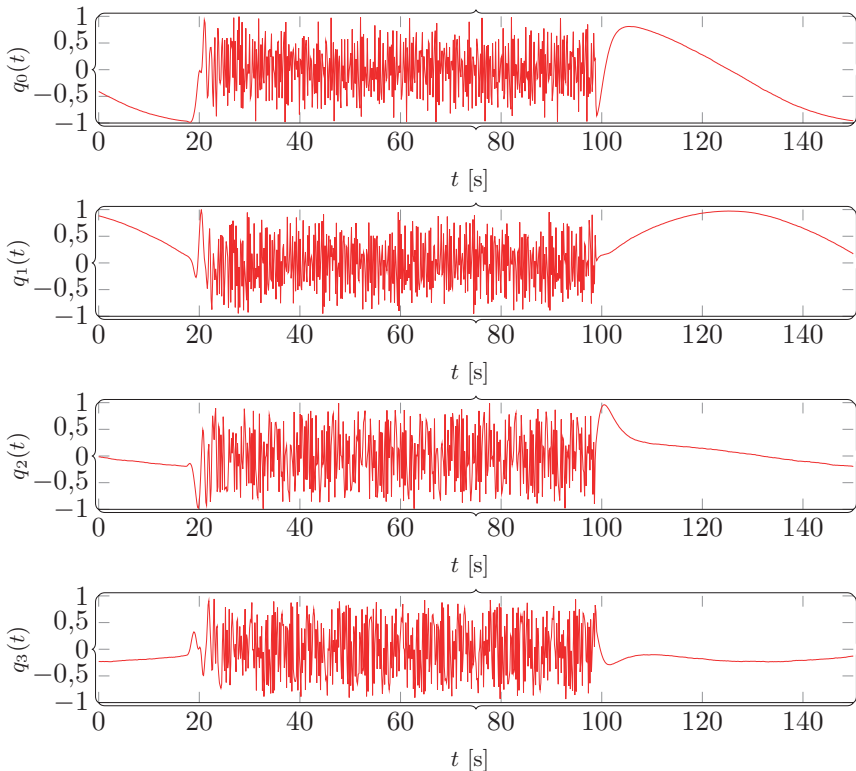
Przejdźmy teraz do modelu układu pomiarowego. W tym przypadku model nie wymaga linearyzacji, dlatego model pomiarowy zapiszemy w postaci:

$$\mathbf{y}(k) = \mathbf{H}(k)x(k) + v(k) \quad (12.50)$$

gdzie macierz \mathbf{H} ma postać taką jak w (12.46).

12.4. Wykorzystanie filtrów Kalmana do szacowania orientacji...

Macierze kowariancji dla modelu systemu \mathbf{Q} oraz układu pomiarowego \mathbf{R} wyznaczamy również eksperymentalnie. Wynik estymacji kwaternionu z wykorzystaniem rozszerzonego filtra Kalmana przedstawiono na rys. 12.14.



Rys. 12.14. Wartości kwaternionu oszacowane z wykorzystaniem rozszerzonego filtra Kalmana

13. Ilościowa analiza układu ruchu człowieka

A planta pedis usque verticem eius

Księga Hioba

13.1. Wstęp

Rozdział poświęcony jest praktycznym zagadnieniom związanym z modelowaniem kinematyki układu ruchu człowieka. W pierwszej części rozdziału omówiono zagadnienia związane z kalibracją czujników wchodzących w skład systemu pomiarowego. Następnie skupiono się na zadaniu estymacji parametrów modeli kinematyki układu ruchu człowieka. Rozpatrzono kilka przypadków, z których pierwszy dotyczy estymacji parametrów modelu kinematyki ludzkiego łokcia. Biorąc ten staw za przykład, omówiono istotne zagadnienia związane z estymacją parametrów modelu łańcucha kinematycznego z wykorzystaniem danych pomiarowych. W dalszej części tego rozdziału powtórzono rozważania dla łańcuchów kinematycznych kończyny górnej i dolnej.

Ostatnia część tego rozdziału została poświęcona estymacji kątów w stawach z wykorzystaniem modeli łańcuchów kinematycznych odpowiednich kończyn.

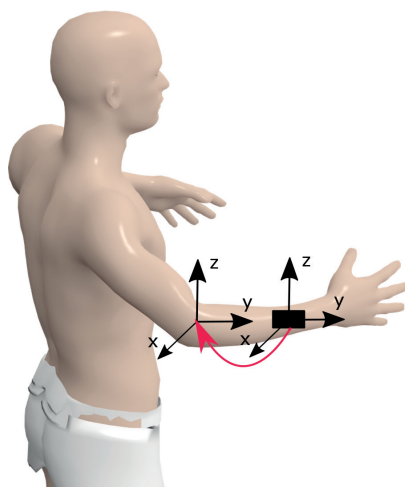
13.2. Główne zadania w ilościowej analizie układu ruchu człowieka

Każdy z omawianych łańcuchów kinematycznych składa się z ogniw i stawów. W celu ustalenia położenia danego segmentu w przestrzeni, jak również zależności pomiędzy segmentami, należy pozyskać dane pomiarowe, które umożliwią ich wyznaczenie. W tym celu należy przymocować do każdego segmentu analizowanego łańcucha kinematycznego jeden czujnik pomiarowy, który umożliwi wyznaczenie orientacji tego czujnika w przestrzeni. Dla uproszczenia rozważań przyjmijmy, że połączenie pomiędzy odpowiednim segmentem i czujnikiem jest sztywne. Jest to

istotne założenie, ponieważ wskazuje, że orientacja czujnika jest również orientacją segmentu.

13.2.1. Kalibracja

Ze względu na budowę anatomiczną człowieka nie jest możliwe umieszczenie czujnika pomiarowego w dowolnym miejscu danego segmentu. Wynika z tego, że układ współrzędnych czujnika nie pokrywa się z układem współrzędnych segmentu, z którym został połączony. Ilustruje to rys. 13.1. Do kompensacji tych różnic wykorzystuje się **kalibrację**, która musi zostać wykonana przed rozpoczęciem pomiarów.

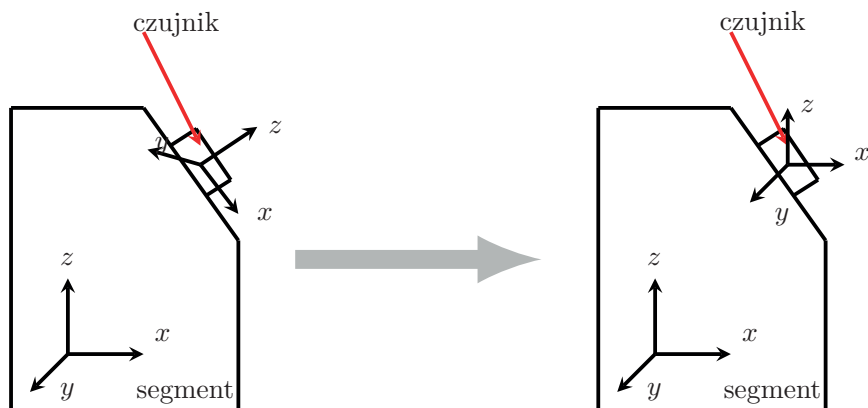


Rys. 13.1. Układ stawu łokciowego i układ czujnika pomiarowego

Kalibracja czujników pomiarowych to zestaw działań, które mają na celu przygotowanie czujnika lub zespołu czujników do wykonania właściwych pomiarów. Jednym z tych działań jest takie zorientowanie osi układu współrzędnych czujnika i stawu, by nakładały się one na siebie. Takie zadanie nazywane jest **osiowaniem czujnika** (ang. *sensor alignment*). Ilustracja operacji osiowania została przedstawiona na rys. 13.2.

Wyznaczenie orientacji pomiędzy układem stawu a układem czujnika odbywa się na podstawie wyznaczonej pozycji czujnika w przestrzeni, co oznaczmy przez $\mathbf{q}_{a,c}$, gdzie a oznacza układ odniesienia, a c układ czujnika. Znana jest również orientacja układu ciała względem układu odniesienia, tj. $\mathbf{q}_{a,b}$, gdzie b oznacza

13.2. Główne zadania w ilościowej analizie układu ruchu człowieka



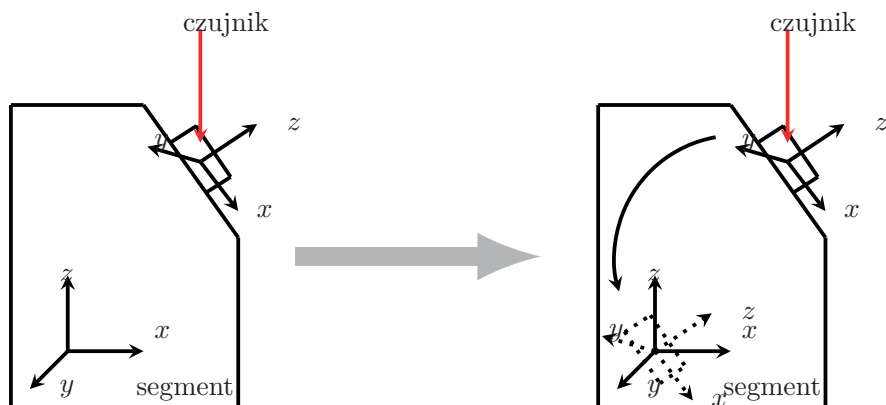
Rys. 13.2. Ilustracja osiowania układów stawu i czujnika

układ ciała. Korzystając teraz z zależności:

$$\mathbf{q}_{b,c} = \mathbf{q}_{b,a} \otimes \mathbf{q}_{a,c} \quad (13.1)$$

wyznamy kwaternion orientacji pomiędzy układem ciała $\{B\}$ a układem czujnika $\{C\}$.

Kolejnym zadaniem wykonywanym w ramach kalibracji jest przesunięcie punktu zerowego układu współrzędnych czujnika w taki sposób, by pokrywał się on z punktem zerowym układu stawu, z którym został on powiązany. Przypadek ten został przedstawiony na rys. 13.3.



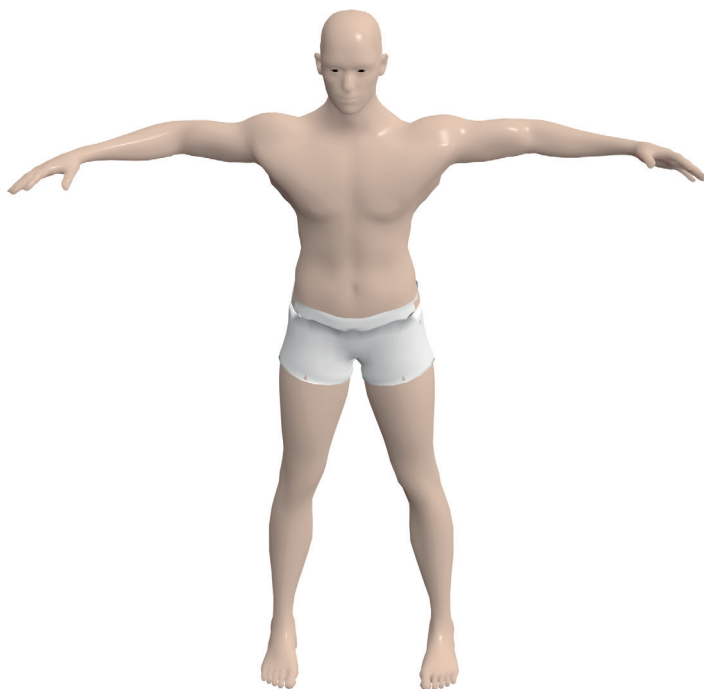
Rys. 13.3. Ilustracja przesuwania środka układu czujnika w celu nałożenia na środek układu stawu

13. Ilościowa analiza układu ruchu człowieka

Znając orientację stawu względem układu odniesienia ($\mathbf{q}_{b,a}$) oraz czujnika względem układu odniesienia ($\mathbf{q}_{a,c}$), można wyznaczyć wektor przesunięcia pomiędzy środkami stawu i czujnika, tj. $r_{b,c}$.

Podczas kalibracji określone są również wymiary ciała ludzkiego istotne z perspektywy rozwiązywanego zadania. Jeśli przedmiotem zainteresowania jest np. kończyna górna, to istotne okaże się wyznaczenie długości dłoni, przedramienia, ramienia oraz odległości między barkiem a klatką piersiową. Wyróżniamy trzy podejścia do wyznaczania wymaganych wielkości. Jednym z nich jest bezpośredni pomiar kończyn [480]. Innym podejściem jest wyznaczanie tych wielkości z wykorzystaniem modeli antropometrycznych [346] (patrz również 11.5). Możliwe jest również wyznaczanie tych wielkości z wykorzystaniem odpowiednich algorytmów [103].

W celu wykonania kalibracji należy wykonać zestaw pomiarów z wykorzystaniem czujników, umieszczonych na odpowiednich segmentach. Wiąże się to z przyjęciem przez badanego tzw. pozycji typu „T” (ang. *T-pose*). Pozycja typu „T” została przedstawiona na rys. 13.4. Wykonując pomiary w tej pozycji, z jednoczesną znajomością wymaganych wymiarów ciała ludzkiego możliwe jest wyznaczenie



Rys. 13.4. Pozycja typu „T”

13.2. Główne zadania w ilościowej analizie układu ruchu człowieka

wymaganych wartości związanych ze względnymi przesunięciami i orientacjami segmentów łańcucha i związanymi z nimi czujnikami [325, 340, 396].

Alternatywne podejście do kalibracji bazuje na pozycji typu „N”. Pozycję tę nazywa się pozycją neutralną. W tym przypadku ramiona trzymane są podczas procedury kalibracji wzdłuż ciała [325, 474].

Opracowano również metody, w których kalibracji dokonuje się, wykonując określone ruchy kończynami wokół ustalonych osi obrotu. Na przykład w pracy [102] opracowano zestaw ruchów wykonywanych kolejno wokół jednej określonej osi dla klatki piersiowej, ramienia, przedramienia oraz dłoni. W podejściu tym na jakość kalibracji duży wpływ ma to, z jaką dokładnością wykonywane są ustalone ruchy. W pracy [359] zaproponowano metodę, która podczas kalibracji uwzględnia ograniczenia anatomiczne ludzkiego ciała.

Procedura kalibracji jest zazwyczaj bardzo czasochłonna i co więcej może mieć znaczny wpływ na jakość uzyskanych wyników odtwarzania ruchów człowieka. Dlatego coraz częściej podejmowany jest problem automatycznej kalibracji [390]. Algorytmy automatycznej kalibracji charakteryzują się tym, że nie wymagają żadnych informacji odnośnie do tego, jakie ruchy będą wykonywane podczas wyznaczania wymaganych parametrów. Zagadnienie automatycznej kalibracji jest szczególnie istotne w przypadku rozwiązań przeznaczonych dla masowego użytkownika, dla którego przeprowadzenie procedury kalibracji może być kłopotliwe.

W dalszej części tego rozdziału będziemy przyjmowali, że procedura kalibracji została już wykonana.

13.2.2. Dryft

Jednym z głównych problemów związanych z estymacją ruchu człowieka bazującą na danych z czujników typu akcelerometry, żyroskopy i magnetometry jest zjawisko dryftu. Tak jak już wspomnieliśmy, dryft to wolnozmienna składowa mająca niekorzystny wpływ na jakość wyników estymacji. Co więcej występowanie dryftu ogranicza zakres stosowania tego podejścia. Dryft może bardzo szybko w znaczący sposób pogorszyć wyniki estymacji.

Negatywny wpływ dryftu na wyniki estymacji ruchu człowieka można eliminować m.in. przez jego estymację i późniejsze usuwanie z sygnału. Inne podejście do usuwania dryftu bazuje na osłabianiu zjawisk wpływających na pojawienie się dryftu. Takie podejście przedstawiono m.in. w pracy [457]. W przedstawionym rozwiązaniu zaproponowano algorytm do estymacji liniowego przyśpieszenia, które jest jedną ze składowych przyśpieszenia mierzonego z wykorzystaniem akcelerometru.

Inne z podejść, które przedstawiono m.in. w pracach [118, 119], opiera się na estymacji szumów powstających w żyroskopie. Rozwiązanie to bazuje na

13. Ilościowa analiza układu ruchu człowieka

obserwacji, że żyroskop jest jednym z głównych źródeł błędów, które skutkują dryftem w estymowanych sygnałach [350].

Do ograniczenia niekorzystnego wpływu dryftu na jakość estymacji ruchu człowieka, bazującej na pomiarach z czujników inercyjnych i magnetometrów, zaproponowano również podejście, w którym wykorzystuje się ograniczenia biomechaniczne [187, 325, 473].

Mało popularnym nurtem badań nad usuwaniem dryftu z estymat sygnałów związanych z ludzkim ruchem jest wykorzystanie metod usuwania składowych o niskich częstotliwościach. Zastosowanie tego podejścia wiąże się z użyciem metod uwzględniających złożoną naturę występującego w omawianych oszacowaniach sygnałów dryftu. Przykładowymi pracami wpisującymi się w ten nurt badań są, m.in. [49, 50].

13.2.3. Zakłócenia magnetyczne

Do szacowania położeń i orientacji stawów w łańcuchach kinematycznych, oprócz czujników inercyjnych (tj. akcelerometrów i żyroskopów), wykorzystywane są również, jak już wspominaliśmy, magnetometry. Dane pozyskane z tych czujników są bardzo przydatne do poprawy jakości estymacji orientacji obiektów w przestrzeni, a co jest z tym związane, również położeń. Jest to powodowane tym, że dane z magnetometrów poprawiają jakość estymacji m.in. kursu (ang. *heading*), który wyznaczany jest głównie z wykorzystaniem danych z żyroskopu [246]. Niestety magnetometry mają też swoje wady, z których najistotniejszą jest podatność na zakłócenia powodowane przez materiały ferromagnetyczne będące w ich otoczeniu. Typowe zakłócenia powodowane przez te materiały dzieli się na zakłócenia powodowane przez materiały magnetycznie twarde i miękkie [353]. Materiał magnetycznie twarde to taki, którego namagnesowanie jest trwałe. Z kolei w materiałach magnetycznie miękkich efekt namagnesowania nie jest trwały i zanika, gdy pole magnetyczne przestaje oddziaływać na materiał.

Jeżeli pole magnetyczne w otoczeniu czujnika magnetoelektrycznego nie zmienia się, niekorzystny jego wpływ na pomiary może być wyeliminowany przez kalibrację czujnika. Bardziej złożonym zagadnieniem jest eliminowanie zakłóceń wywołanych przez zmienne pole magnetyczne. W literaturze można wyróżnić kilka podejść do eliminowania niekorzystnego wpływu pola magnetycznego na czujnik magnetyczny. Jednym z nich jest analiza amplitudy mierzonego sygnału i bieżące decydowanie o tym, czy zaobserwowany pomiar jest wiarygodny czy też nie. Możliwe jest również inne rozwiązanie tego problemu, które zaproponowano np. w [45]. W pracy tej dane z magnetometru zastosowano jedynie do szacowania kursu.

Bardziej złożone podejście zaproponowano z kolei w pracy [347]. Autorzy zaproponowali model zakłóceń powodowanych przez pole magnetyczne. Opracowany

13.2. Główne zadania w ilościowej analizie układu ruchu człowieka

model został następnie wykorzystany do usuwania niechcianych składowych z pomiarów pola magnetycznego.

13.2.4. Kinematyka i ograniczenia biomechaniczne

W pracach dotyczących analizy układu ruchu człowieka zakłada się, że ludzkie kończyny są ciałami sztywnymi. W ujęciu kinematyki problem analizy redukuje się do problemu estymacji orientacji i położenia ludzkich kończyn, wchodzących w skład rozpatrywanego łańcucha kinematycznego [482]. Jeśli liczba segmentów wchodzących w skład łańcucha kinematycznego jest większa, to należy wziąć pod uwagę konstrukcję modelu dla tego łańcucha. Do parametryzacji łańcucha kinematycznego wykorzystuje się kąty Eulera [45], konwencję Denavita–Hartenberga [95, 325] lub mapy wykładnicze [207].

Liczba segmentów wchodzących w skład łańcucha kinematycznego nie jest jedyną wielkością, którą należy brać pod uwagę, wybierając metodę jego analizy. W wielu zadaniach istotne może okazać się uwzględnienie ograniczeń anatomicznych ludzkiego układu kostnego. Z jednej strony uwzględnienie ograniczeń anatomicznych może poprawić jakość estymacji poszczególnych składowych łańcucha kinematycznego, z drugiej natomiast, jak już wspomniano, jest to jedna z metod ograniczania niekorzystnego wpływu dryftu. Wykorzystanie modelowania łańcucha kinematycznego może ułatwić uwzględnienie ograniczeń anatomicznych ludzkiego ciała. Ograniczenia mogą być uwzględniane na etapie konstrukcji algorytmu estymacji orientacji [45, 325]. Ograniczenia kinematyczne mogą też zostać uwzględnione na późniejszym etapie, jak to zostało przedstawione m.in. w pracy [474].

W podejściu, w którym każdy segment rozpatrywany jest oddzielnie również możliwe jest uwzględnienie ograniczeń anatomicznych. Na przykład w pracy [213] zaproponowane algorytmy uwzględniają ograniczenia anatomiczne w budowie ludzkiego ciała.

13.2.5. Ocena wyników ilościowej analizy układu ruchu

Ważnym etapem analizy układu ruchu człowieka jest ocena jakości wykorzystywanych metod. Główne podejście bazuje na porównaniu uzyskanych wyników z danymi referencyjnymi. Do porównań najczęściej wykorzystuje się metryki (por. rozdz. 3). Na przykład w [347] wykorzystano błąd średniokwadratowy. Z kolei w pracy [247] oprócz błędu średniokwadratowego do porównania zastosowano również współczynnik korelacji.

W systemach analizy układu ruchu człowieka dane pochodzące z systemów typu *motion capture* są uznanym standardem i przez to są najczęściej wykorzystywane jako dane referencyjne. Dane zebrane z tych systemów umożliwiają weryfikację

13. Ilościowa analiza układu ruchu człowieka

zarówno algorytmów estymacji orientacji, jak i położenia obiektów w przestrzeni. Weryfikacja wyników estymacji, np. orientacji czy położenia, nie jest zadaniem prostym w przypadku wykorzystania pomiarów z optycznych systemów analizy układu ruchu człowieka. Jedną z przyczyn są trudności związane z dokładnym wyznaczaniem długości segmentów wchodzących w skład analizowanego łańcucha kinematycznego. Inną trudnością jest dopasowanie danych pochodzących z czujników inercyjnych i magnetometrów do tych danych zebranych z wykorzystaniem systemów *motion capture* (ang. *data alignment*).

Oprócz danych z systemów przechwytywania ruchu, do porównania wyników estymacji orientacji i położenia obiektów w przestrzeni wykorzystuje się również dane pochodzące z robotów lub manipulatorów [119, 405]. Uznany źródłem danych referencyjnych są również systemy pomiarowe firmy Xsens [346].

Warto również zaznaczyć, że wyniki oceny jakości estymacji orientacji i położenia obiektu w przestrzeni zależą od wybranej metody. Powoduje to, że trudno jest porównać wyniki prezentowane w różnych pracach.

13.3. Śledzenie ruchu w przestrzeni stawu łokciowego

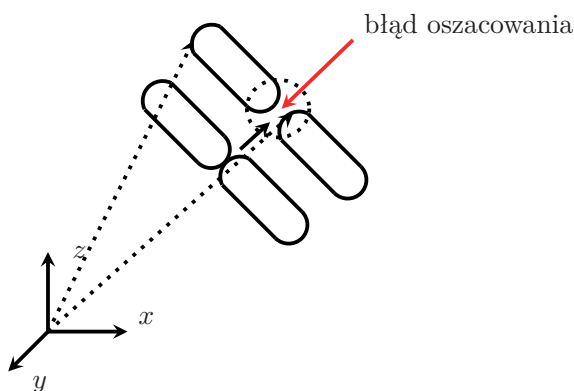
W podrozdziale przedstawimy, na przykładzie prostego stawu, znane podejścia do śledzenia ruchu oparte na danych z czujników typu akcelerometr, żyroskop i magnetometr. Przedstawione w tym podrozdziale podejścia można uogólnić na bardziej złożone przypadki, co zostanie przedstawione w kolejnym podrozdziale.

Wśród podejść do estymacji położenia segmentów łańcucha kinematycznego w przestrzeni możemy wyróżnić takie, w których orientacja i położenie każdego z segmentów wyznaczone są: a) oddzielnie, b) łącznie. W przypadku pierwszym wyróżniamy dwa sposoby rozwiązania zadania, tj. bez uwzględniania ograniczeń anatomicznych oraz z ich uwzględnianiem. W przypadku drugim warto zaznaczyć, że do rozwiązania wymagane jest przygotowanie odpowiedniego modelu dla rozpatrywanego łańcucha kinematycznego.

W pierwszej z wymienionych metod położenie wszystkich segmentów wyznaczone jest dla każdego z nich oddzielnie. Wariant bez uwzględniania ograniczeń anatomicznych jest rzadko wykorzystywany w zagadnieniach praktycznych. Związane jest to m.in. z tym, że podejście to jest nieodporne na zakłócenia pomiarowe. Zakłócenia w pomiarach sprawiają, że podczas obliczeń mogą pojawić się różnice w estymowanych położeniach segmentów. Prowadzi to do sytuacji, w której punkty przestają się ze sobą stykać pomimo przeprowadzonej na początku badań kalibracji. Zjawisko to zilustrowano na rys. 13.5.

Sposób, w którym orientacja każdego segmentu wyznaczana jest oddzielnie, zostało wykorzystane np. w pracy [461]. Autorzy pracy zaproponowali dwuetapową

13.3. Śledzenie ruchu w przestrzeni stawu łokciowego



Rys. 13.5. Ilustracja błędu w oszacowaniu położenia segmentów

metodę do estymacji każdego segmentu rozpatrywanego łańcucha kinematycznego. Na pierwszym etapie do szacowania orientacji segmentów wykorzystano algorytm QUEST, który bazował na danych z czujnika przyśpieszenia i magnetometru. Na drugim etapie wyniki estymacji z pierwszego etapu poprawiano przez zastosowanie filtra EKF, który w tym przypadku wykorzystano do fuzji wyników z pierwszego etapu i danych z żyroskopu. Z kolei w pracy [480] orientacja każdego z segmentów wyznaczana była jedynie przez filtr EKF. Natomiast w pracy [482] nie uwzględniono ograniczeń anatomicznych podczas estymacji orientacji każdego ze stawów. W tym przypadku do szacowania orientacji wykorzystano filtr Kalmana.

Obecność zakłóceń pomiarowych sprawia, że dokładność metod, w których nie uwzględnia się ograniczeń anatomicznych, jest mała. W celu zwiększenia jakości estymacji orientacji segmentów, a co za tym idzie również jakości estymacji ich wzajemnego położenia, należy uwzględnić podczas obliczeń ograniczenia nakładane przez anatomię człowieka. Na przykład autorzy [474] wykorzystali filtr cząsteczkowy do estymacji orientacji poszczególnych segmentów. Zaproponowany algorytm uwzględnia również ograniczenia anatomiczne. Podobne podejście zostało również zaproponowane w pracy [346]. Autorzy tej pracy zaprojektowali algorytm, który umożliwi estymację wszystkich segmentów zarówno kończyn górnych, jak i dolnych. Ograniczenia anatomiczne zostały również uwzględnione w pracach [213, 214].

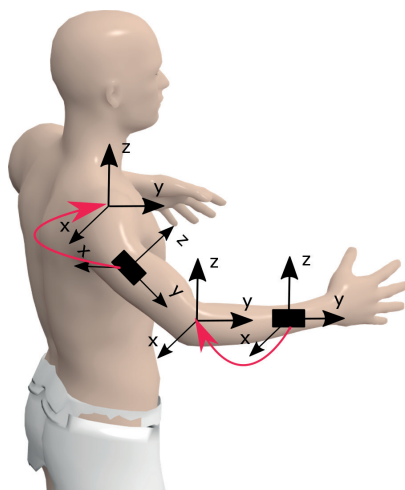
Drugie z podejść wykorzystuje nie tylko dane z czujników inercyjnych i magnetometru, ale również model kinematyki dla rozpatrywanej kończyny. W rozdziale 11 omówione zostały metody konstrukcji łańcuchów kinematycznych wraz z przykładami ich modelowania dla kończyny górnej i dolnej. W pracach wykorzystywane są

13. Ilościowa analiza układu ruchu człowieka

dwa podejścia do parametryzacji łańcucha kinematycznego, tj. metody Denavita–Hartenberga oraz mapy eksponencjalne.

Metoda Denavita–Hartenberga do parametryzacji łańcucha kinematycznego została wykorzystana na przykład w pracy [473]. W badaniach skupiono się na kończynie górnej. Z kolei do estymacji parametrów modelu wykorzystano algorytm UKF. Autorzy prac [118] oraz [119] również swoje rozwiązanie zastosowali do kończyny górnej. W pracach [285] oraz [286] autorzy skupili się na kończynie górnej. Parametryzacja łańcucha kinematycznego z wykorzystaniem map eksponencjalnych została z kolei wykorzystana m.in. w pracach [64, 186] oraz [273].

W dalszej części tego podrozdziału omówimy każde z wymienionych podejść na przykładzie stawu łokciowego (rys. 13.6). Podczas rozpatrywania stawu łokciowego przyjęto, że staw barkowy jest nieruchomy. Założenie to skupia się na istotnych zagadnieniach dotyczących modelowania pojedynczego stawu. Założenie to nie wpływa na ogólność przeprowadzonych rozważań.



Rys. 13.6. Schemat połączenia dwóch segmentów ciała w stawie łokciowym

Do szacowania położenia segmentów wchodzących w skład łańcucha stawu łokciowego wykorzystamy dwa czujniki. Czujniki umieszczone są na segmentach łańcucha kinematycznego (ramię i przedramię), tak jak to pokazano na rys. 13.6. Przyjmijmy również, że procedura kalibracji została wykonana, a stosowane wartości względne przesunięć oraz orientacji są znane.

13.3.1. Estymacja położenia segmentów bez uwzględniania ograniczeń anatomicznych

Pierwsze z zaprezentowanych podejść będzie podejściem, w którym każdy z segmentów rozpatrywany jest oddzielnie. Do reprezentacji orientacji segmentów wykorzystamy kwaterniony.

W pierwszym kroku, korzystając z procedur estymacji orientacji omówionych w podrozdziale 13.2.1, wyznaczamy wartości kwaternionów dla każdego z czujników, tj. jednego ulokowanego do ramienia i jednego przymocowanego do przedramienia. Do wyznaczenia kwaternionów wykorzystamy następujące równanie ruchu:

$$\dot{\mathbf{q}}_{a,c} = \frac{1}{2} \hat{\mathbf{q}}_{a,c} \otimes \omega_c \quad (13.2)$$

gdzie a oraz c to odpowiednio indeks związany z układem odniesienia $\{A\}$ i z układem czujnika $\{C\}$. Wielkość ω_c reprezentuje pomiar prędkości kątowych w układzie czujnika. Zapiszmy teraz to równanie dla wygody dalszych obliczeń w następującej formie:

$$\dot{\mathbf{q}}_{a,c} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & \omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \hat{\mathbf{q}}_{a,c} \quad (13.3)$$

Przyjmijmy

$$\Omega(\boldsymbol{\omega}) = \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & \omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \quad (13.4)$$

gdzie $\mathbf{y}_\omega = [y_{\omega_x} \ y_{\omega_y} \ y_{\omega_z}]^T$ to pomiary z czujnika żyroskopowego. Biorąc pod uwagę (13.4), zapiszmy teraz równanie (13.3) w postaci:

$$\dot{\mathbf{q}}_{a,c} = \frac{1}{2} \Omega(\boldsymbol{\omega}(t)) \mathbf{q}_{a,c} \quad (13.5)$$

Ponieważ pomiary żyroskopowe obarczone są błędami pomiarowymi, to uwzględnijmy je w zależności (13.4)

$$\dot{\mathbf{q}}_{a,c} = \frac{1}{2} \Omega(\boldsymbol{\omega}(t) + \mathbf{v}(t)) \mathbf{q}_{a,c} \quad (13.6)$$

13. Ilościowa analiza układu ruchu człowieka

Bazując na rozważaniach przedstawionych w pracy [407], zależność (13.6) przepisujemy do postaci

$$\dot{\mathbf{q}}_{a,c} = \frac{1}{2}\Omega(\boldsymbol{\omega}(t))\mathbf{q}_{a,c} + \Omega(\mathbf{q}_{a,c})v(t) \quad (13.7)$$

gdzie

$$\Omega(\mathbf{q}_{a,c}) = \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \quad (13.8)$$

Rozwiązanie zależności (13.4) dla czasu dyskretnego i bez uwzględniania zakłóceń pomiarowych przyjmuje postać:

$$\mathbf{q}_{a,c}(k) = \exp\left(\frac{1}{2}\Omega(\boldsymbol{\omega}(k)\Delta t)\right)\mathbf{q}_{a,c}(k-1) \quad (13.9)$$

gdzie Δt jest okresem próbkowania.

Ponieważ wyznaczenie eksponenty macierzy jest zadaniem złożonym, uprościmy teraz zależność (13.9). Biorąc pod uwagę to, że

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (13.10)$$

eksponentę macierzy z zależności (13.9) możemy zapisać w formie:

$$\exp\left(\frac{1}{2}\Omega(\boldsymbol{\omega}(k)\Delta t)\right) = \sum_{n=0}^{\infty} \left(\frac{\Delta t}{2}\right)^n \frac{1}{n!}\Omega(\boldsymbol{\omega}(k))^n \quad (13.11)$$

Bazując na równości (13.11), zależność (13.9) uproszczamy z kolei (por. z (9.46)) do postaci [407]:

$$\mathbf{q}_{a,c}(k) = \exp\left(\cos\left(\frac{\|\boldsymbol{\omega}(k)\|\Delta t}{2}\right)\mathbf{I} + \frac{\sin\left(\frac{\|\boldsymbol{\omega}(k)\|\Delta t}{2}\right)}{\|\boldsymbol{\omega}(k)\|}\Omega(\boldsymbol{\omega}(k))\right)\mathbf{q}_{a,c}(k-1) \quad (13.12)$$

Zależność (13.12) może zostać jeszcze uproszczona, jeśli przyjmiemy, że $\|\boldsymbol{\omega}(k)\|\Delta t$ przyjmuje małe wartości. Założenie to jest prawdziwe dla wysokich częstotliwości

13.3. Śledzenie ruchu w przestrzeni stawu łokciowego

próbki. Zapiszmy więc (13.12) w postaci:

$$\mathbf{q}_{a,c}(k) = \exp\left(\mathbf{I} + \frac{\Delta t}{2}\Omega(\boldsymbol{\omega}(k))\right)\mathbf{q}_{a,c}(k-1) \quad (13.13)$$

Dysponując zależnościami do szacowania orientacji obiektu na podstawie danych z żyroskopu, kolejnym krokiem jest konstrukcja modelu dynamiki. Oznaczmy teraz kwaterniony reprezentujące orientację czujników przymocowanych do ramienia (\mathbf{q}_{a,c_1}) oraz przedramienia (\mathbf{q}_{a,c_2}). Pomiary prędkości kątowych oznaczmy natomiast jako $y_{\omega_{c_1}}(k)$ (dla czujnika powiązanego z ramieniem) oraz $y_{\omega_{c_2}}(k)$ (przedramię). Biorąc pod uwagę dotychczasowe rozważania, model dynamiki dla rozpatrywanego zadania ma postać:

$$\begin{bmatrix} \mathbf{q}_{a,c_1}(k) \\ \mathbf{q}_{a,c_2}(k) \end{bmatrix} = \begin{bmatrix} \exp\left(\mathbf{I} + \frac{\Delta t}{2}\Omega(\boldsymbol{\omega}_{c_1}(k))\right) & 0 \\ 0 & \exp\left(\mathbf{I} + \frac{\Delta t}{2}\Omega(\boldsymbol{\omega}_{c_2}(k))\right) \end{bmatrix} \begin{bmatrix} \mathbf{q}_{a,c_1}(k-1) \\ \mathbf{q}_{a,c_2}(k-1) \end{bmatrix} \quad (13.14)$$

Przejdźmy teraz do konstrukcji modelu pomiarowego. W rozdziale 4 wprowadzono następujący model pomiarowy dla czujnika przyspieszenia (4.9):

$$\bar{\mathbf{y}}_a = \mathbf{y}_a + \mathbf{y}_{a_{bs}} + \mathbf{y}_{a_{wh}} = \mathbf{R}_{a,c}(\mathbf{y}_{a_F} - \mathbf{y}_{a_g}) + \mathbf{y}_{a_{bs}} + \mathbf{y}_{a_{wh}} = \mathbf{R}_{a,c}(\mathbf{q}) + \mathbf{y}_{a_F} + \mathbf{y}_{a_{bs}} + \mathbf{y}_{a_{wh}} \quad (13.15)$$

Również w tym rozdziale wprowadziliśmy modele dla żyroskopu (4.16)

$$\bar{\mathbf{y}}_{\boldsymbol{\omega}} = \mathbf{y}_{\boldsymbol{\omega}} + \mathbf{y}_{\boldsymbol{\omega}_{bs}} + \mathbf{y}_{\boldsymbol{\omega}_{wh}} \quad (13.16)$$

oraz magnetometru (4.21)

$$\bar{\mathbf{y}}_m = \mathbf{R}_{a,c}(\mathbf{q})h + \mathbf{y}_{m_{bs}} + \mathbf{y}_{m_{wh}} \quad (13.17)$$

Bazując na zależnościach (13.15)–(13.17) zapiszmy modele pomiarowe dla rozpatrywanego stawu łokciowego. Model dla czujnika przyspieszenia ma postać

$$\bar{\mathbf{y}}_a = \mathbf{R}_{a,c_1}(\mathbf{q})\mathbf{R}_{a,c_2}(\mathbf{q}) + \mathbf{y}_{a_F} + \mathbf{y}_{a_{bs}} + \mathbf{y}_{a_{wh}} \quad (13.18)$$

Pozostałe modele są następujące, tj. dla żyroskopu

$$\bar{\mathbf{y}}_{\boldsymbol{\omega}} = \mathbf{R}_{a,c_2}(\mathbf{q})y_{\omega_1} + \mathbf{y}_{\omega_2} + \mathbf{y}_{\boldsymbol{\omega}_{bs}} + \mathbf{y}_{\boldsymbol{\omega}_{wh}} \quad (13.19)$$

a dla magnetometru

$$\bar{\mathbf{y}}_m = \mathbf{R}_{a,c_1}(\mathbf{q})\mathbf{R}_{a,c_2}(\mathbf{q})h + \mathbf{y}_{m_{bs}} + \mathbf{y}_{m_{wh}} \quad (13.20)$$

13. Ilościowa analiza układu ruchu człowieka

Podsumowując poprzednie rozważania zapiszmy model pomiarowy dla zadania estymacji położenia dwóch czujników umieszczonych na ramieniu i przedramieniu

$$\begin{bmatrix} \bar{\mathbf{y}}_{a_1} \\ \bar{\mathbf{y}}_{\omega_1} \\ \bar{\mathbf{y}}_{m_1} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{a,c_1}(\mathbf{q}) \mathbf{R}_{a,c_2}(\mathbf{q}) \\ \mathbf{R}_{a,c_2}(\mathbf{q}) \\ \mathbf{R}_{a,c_1}(\mathbf{q}) \mathbf{R}_{a,c_2}(\mathbf{q}) \end{bmatrix} + \begin{bmatrix} \mathbf{y}_{a_{bs,1}} \\ \mathbf{y}_{\omega_{bs,1}} \\ \mathbf{y}_{m_{bs,1}} \end{bmatrix} \quad (13.21)$$

Po wyznaczeniu wartości kwaternionu dla każdego z segmentów możliwe jest wyznaczenie położenia jednego segmentu względem drugiego. W tym celu korzystamy z zależności

$$\mathbf{p}_{a,b_1}(k) = \mathbf{p}_{c,b_2}(k) - \mathbf{q}_{a,c_1}(k) \otimes \varpi_{b_1} \otimes (\mathbf{q}_{a,c_1}^*(k)) \quad (13.22)$$

gdzie ϖ_{b_1} jest długością pierwszego segmentu w rozpatrywanym łańcuchu.

Algorytm do wyznaczania położenia segmentu przedstawiono w formie pseudokodu (por. algorytm 5).

Algorytm 5: Algorytm do szacowania położenia segmentów w przestrzeni

Wejście : zestaw pomiarów z żyroskopu $\mathbf{y}_\omega = [y_{\omega_x}, y_{\omega_y}, y_{\omega_z}]^\top$,
częstotliwość próbkowania F_s

for $k = 1$ **to** K **do**

$$\mathbf{q}_{a,c}(k) = \exp\left(\frac{1}{2}\Omega(\boldsymbol{\omega}(k))\right)\mathbf{q}_{a,c}(k-1)$$

$$\mathbf{p}_{a,b_1}(k) = \mathbf{p}_{a,b_2}(k) - \mathbf{q}_{a,c}(k) \otimes \varpi_{b_1} \otimes (\mathbf{q}_{a,c}^*(k))$$

end

Wyjście : $\mathbf{q}_{a,c}$.

13.3.2. Estymacja położenia segmentów z uwzględnieniem ograniczeń anatomicznych

Rozwiązaniem problemu jest przeformułowanie zadania estymacji tak, aby uwzględniało ono ograniczenia biomechaniczne. Przykład konstrukcji ograniczeń biomechanicznych przedstawimy na przykładzie stawu łokciowego (rys. 13.6). W tym przypadku bierzemy pod uwagę dwie zależności. Jedna z zależności dotyczy segmentu pierwszego (13.22), a druga z tych zależności dotyczy segmentu drugiego

13.3. Śledzenie ruchu w przestrzeni stawu łokciowego

i ma następującą postać:

$$\mathbf{p}_{a,b_2}(k) = \mathbf{p}_{a,b_1}(k) - \mathbf{q}_{a,c_2}(k) \otimes \varpi_{b_2} \otimes \left(\mathbf{q}_{a,c_2}^*(k) \right) \quad (13.23)$$

gdzie ϖ_{b_2} reprezentuje długość drugiego segmentu.

Ponieważ segmenty 1 i 2 dla poprawnie wyznaczonych położeń p_{a,b_1} oraz p_{a,b_2} powinny się ze sobą stykać, wówczas z zależności ((13.22) oraz (13.23)) wynika, że:

$$0 = \mathbf{p}_{a,b_1}(k) + \mathbf{q}_{a,c_1}(k) \otimes \varpi_{b_1} \otimes \left(\mathbf{q}_{a,c_1}^*(k) \right) - \mathbf{p}_{a,b_2}(k) - \mathbf{q}_{a,c_2}(k) \otimes \varpi_{b_2} \otimes \left(\mathbf{q}_{a,c_2}^*(k) \right) \quad (13.24)$$

W tym zadaniu istotne jest, by estymowane położenia segmentów w przestrzeni stykały się ze sobą. Na podstawie przeprowadzonych rozważań możemy zaproponować następujące zadanie estymacji wzajemnego położenia segmentów w przestrzeni jako następujący problem optymalizacyjny:

$$\{\hat{\mathbf{q}}_{a,c_1}, \hat{\mathbf{q}}_{a,c_2}\} = \min_{\mathbf{q}_{a,c_1}, \mathbf{q}_{a,c_2}} Q(\mathbf{q}_{a,c_1}, \mathbf{q}_{a,c_2}) \quad (13.25)$$

gdzie

$$Q(\mathbf{q}_{a,c_1}, \mathbf{q}_{a,c_2}) = \mathbf{p}_{a,b_1}(k) + \mathbf{q}_{a,c_1}(k) \otimes l_{b_1} \otimes \left(\mathbf{q}_{a,c_1}^*(k) \right) + \\ - \mathbf{p}_{a,b_2}(k) - \mathbf{q}_{a,c_2}(k) \otimes l_{b_2} \otimes \left(\mathbf{q}_{a,c_2}^*(k) \right) \quad (13.26)$$

13.3.3. Estymacja położenia segmentów bazująca na modelu kinematyki

Omówimy teraz alternatywne podejście do estymacji orientacji i położenia segmentów, w którym wykorzystano model łańcucha kinematycznego. Z dwóch wymienionych metod parametryzacji wybrano metodę bazującą na mapach wykładniczych.

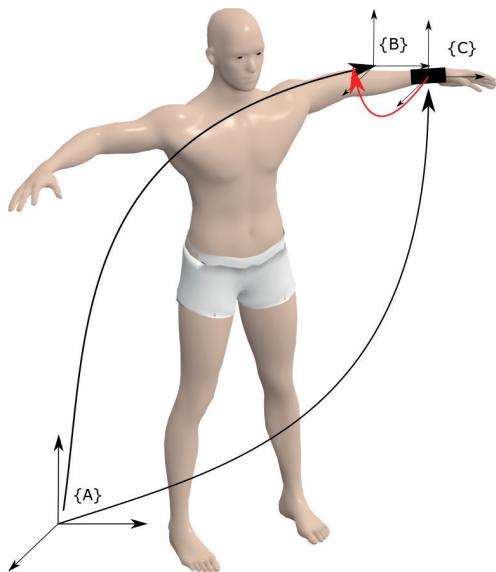
Przyjmując, że staw barkowy jest nieruchomy, łańcuch kinematyczny dla stawu łokciowego można zapisać jako:

$$\mathbf{g}_{\text{ram,przedr}}(\Theta) = \left(\underbrace{e^{[\xi_1]\theta_1} e^{[\xi_2]\theta_2} e^{[\xi_3]\theta_3}}_{\text{ram}} \overbrace{e^{[\xi_4]\theta_4} e^{[\xi_5]\theta_5}}^{\text{przedr}} \right) \mathbf{g}_{\text{ram,przedr}}(0) \quad (13.27)$$

gdzie Θ reprezentuje zbiór kątów wszystkich stawów wchodzących w skład łańcucha kinematycznego.

13. Ilościowa analiza układu ruchu człowieka

Wprowadźmy teraz układy współrzędnych, które będą niezbędne do opisu przedstawianego modelu. Niech układ współrzędnych czujników pomiarowych będzie oznaczony jako $\{C_i\}$, gdzie i jest indeksem i -tego czujnika pomiarowego. Z kolei globalny układ współrzędnych oznaczmy jako $\{A\}$, a układ współrzędnych ciała jako $\{B_i\}$, gdzie i jest indeksem i -tego segmentu. Niech teraz $\mathbf{q}_{a,b_i} : \{A\} \rightarrow \{B_i\}$, $\mathbf{q}_{a,c_i} : \{A\} \rightarrow \{C_i\}$, a $\mathbf{q}_{c_i,b_i} : \{C_i\} \rightarrow \{B_i\}$ (rys. 13.7).



Rys. 13.7. Definicja układu globalnego, czujnika i ciała

Ustalmy, że procedura kalibracji została wykonana i znane są wszystkie wartości \mathbf{q}_{a,c_i} . Celem teraz jest wyznaczenie \mathbf{q}_{a,b_i} . Operacja ta jest możliwa na podstawie pomiarów \mathbf{q}_{c_i,b_i} , tj.

$$\mathbf{q}_{a,b_i} = \mathbf{q}_{a,c_i} \otimes \mathbf{q}_{c_i,b_i} \quad (13.28)$$

Zauważmy, że odpowiednie wartości \mathbf{q}_{c_i,b_i} wyznaczamy, korzystając z algorytmów estymacji orientacji obiektu w przestrzeni na podstawie danych pomiarowych (por. rozdz. 12).

Kolejnym zadaniem jest wyznaczenie zestawu parametrów łańcucha kinematycznego (13.27). W tym celu rozwiązane zostanie następujące zadanie optymalizacyjne:

$$\Theta^* = \min_{\Theta} Q(\Theta, \mathbf{q}) \quad (13.29)$$

13.4. Estymacja położenia kończyny górnej i dolnej w przestrzeni

gdzie

$$Q(\Theta, \mathbf{q}) = \sum_{k=1}^K \sum_{i=1}^{N_{cz}} \left(\mathbf{q}_{c_i, b_i}(k-1) \otimes \mathbf{q} \left(\mathbf{g}_{\text{ram,przede}}(\Theta) \right) - \mathbf{q}_{c_i, b_i}(k) \right)^2 \quad (13.30)$$

gdzie N_{cz} jest liczbą wykorzystanych czujników pomiarowych.

Zapis $\mathbf{q} \left(\mathbf{g}_{\text{ram,przede}}(\Theta) \right)$ w zależności (13.30) rozumiemy jako operację wyznaczenia wartości odpowiednich kwaternionów (por. (9.25)):

$$\mathbf{q} = \begin{bmatrix} \cos \frac{\theta}{2} \\ \check{\omega} \cos \frac{\theta}{2} \end{bmatrix} \in \mathbb{R}^4 \quad (13.31)$$

Z kolei na podstawie wyznaczonych wartości kwaternionów możliwe jest wyznaczenie odpowiednich wartości (patrz (9.27)):

$$\theta = 2 \cos^{-1} q_0$$

$$\check{\omega} = \begin{cases} \frac{[q_1, q_2, q_3]^T}{\sin \left(\frac{\theta}{2} \right)} & \text{jeżeli } \theta \neq 0 \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (13.32)$$

Wyliczony na podstawie $\mathbf{q} \left(\mathbf{g}_{\text{ram,przede}}(\Theta) \right)$ kwaternion reprezentuje położenie ostatniego segmentu w łańcuchu kinematycznym.

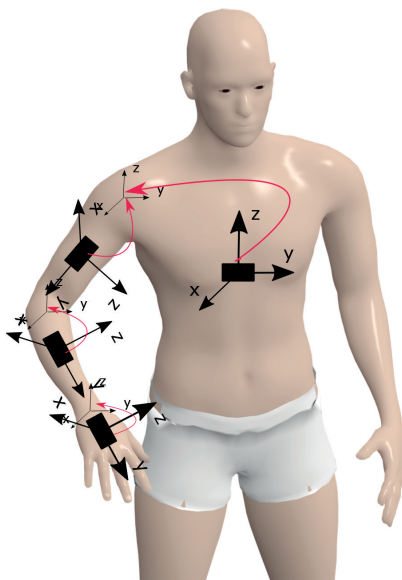
Znając postać funkcji, możliwe jest teraz rozwiązanie sformułowanego zadania optymalizacji (13.30). W tym celu można wykorzystać jedną z metod numerycznej optymalizacji. Na przykład w pracy [273] zaproponowano rozwiązanie bazujące na algorytmie Levenberga–Marquardta.

13.4. Estymacja położenia kończyny górnej i dolnej w przestrzeni

13.4.1. Kończyna górna

W tej części omówiony zostanie sposób szacowania położenia segmentów kończyn górnej i dolnej człowieka. Rozważania bazują na wynikach analizy przeprowadzonej w poprzednim podrozdziale dla prostego stawu. Na rysunku 13.8 przedstawiono schematycznie połączenie segmentów ciała dla kończyny górnej. W przypadku kończyny górnej mamy trzy segmenty b_1 (ramię), b_2 (przedramię) oraz b_3 (dłoń) oraz trzy stawy, tj. barkowy, łokciowy oraz promieniowo-nadgarstkowy. W stawie

13. Ilościowa analiza układu ruchu człowieka



Rys. 13.8. Schemat połączenia segmentów ciała dla kończyny górnej

barkowym wyróżniamy trzy osie obrotu, a w stawach łokciowym i promieniowo-nadgarstkowym po dwie osie obrotu.

Estymacja kwaternionów z trzech czujników pomiarowych umieszczonych na wymienionych segmentach rozpatrywanego łańcucha kinematycznego umożliwia wyznaczenie osi oraz kątów obrotu dla każdego ze stawu, tj.

$$\begin{aligned}
 \mathbf{q}_{\text{kpiers}} &\Rightarrow \{\theta_{\text{kpiers}}, \boldsymbol{\omega}_{\text{kpiers}}\} \\
 \mathbf{q}_{\text{ram}} &\Rightarrow \{\theta_{\text{ram}}, \boldsymbol{\omega}_{\text{ram}}\} \\
 \mathbf{q}_{\text{przedr}} &\Rightarrow \{\theta_{\text{przedr}}, \boldsymbol{\omega}_{\text{przedr}}\} \\
 \mathbf{q}_{\text{dl}} &\Rightarrow \{\theta_{\text{dl}}, \boldsymbol{\omega}_{\text{dl}}\}
 \end{aligned} \tag{13.33}$$

Biorąc to pod uwagę, zapiszmy teraz reprezentację łańcucha kinematycznego w postaci:

$$\begin{aligned}
 \mathbf{g}_{\text{kpiers,dl}}(\boldsymbol{\theta}) &= \\
 &= \left(\underbrace{e^{[\xi_{01}]\theta_{01}} e^{[\xi_{02}]\theta_{02}} e^{[\xi_{03}]\theta_{03}}}_{\text{kpiers}} \underbrace{e^{[\xi_1]\theta_1} e^{[\xi_2]\theta_2} e^{[\xi_3]\theta_3}}_{\text{ram}} \underbrace{e^{[\xi_4]\theta_4} e^{[\xi_5]\theta_5}}_{\text{przedr}} \underbrace{e^{[\xi_6]\theta_6} e^{[\xi_7]\theta_7}}_{\text{dl}} \right) \mathbf{g}_{\text{kpiers,dl}}(0)
 \end{aligned} \tag{13.34}$$

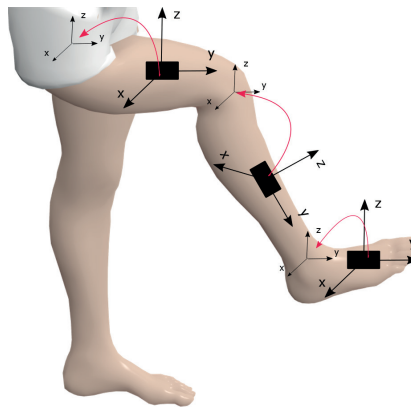
13.4. Estymacja położenia kończyny górnej i dolnej w przestrzeni

W zaproponowanym modelu dodany został dodatkowy segment, tj. klatka piersiowa. Na klatce piersiowej należy umieścić dodatkowy czujnik pomiarowy. Czujnik ten stanowi punkt odniesienia w sytuacji, gdy zmianie ulegnie położenie tułowia osoby. Jest to konieczne, ponieważ wraz ze zmianą położenia tułowia zmianie ulega położenie innych elementów łańcucha.

Przygotowany model, wraz z estymowanymi wartościami parametrów (patrz rozdz. 13.3.3), umożliwia wyznaczenie przesunięć i obrotów poszczególnych stawów względem pozycji początkowej. Wykorzystanie odpowiedniego łańcucha kinematycznego ułatwia sprowadzenie wyliczonych wartości do jednego układu współrzędnych, tj. do układu współrzędnych klatki piersiowej.

13.4.2. Kończyna dolna

Teraz przeprowadzimy podobne rozważania dla kończyny dolnej. Schemat połączeń segmentów ciała dla kończyny dolnej został przedstawiony na rys. 13.9.



Rys. 13.9. Schemat połączenia segmentów ciała dla kończyny dolnej

Również i w tym przypadku mamy trzy segmenty, tj. udo, podudzie i stopa. W skład tego łańcucha kinematycznego wchodzi również trzy stawy: biodrowy, kolanowy oraz skokowy. Do każdego segmentu przymocowano jeden czujnik. Dane z czujnika umożliwiają wyliczenie:

$$\begin{aligned} \mathbf{q}_{\text{biod}} &\Rightarrow \{\theta_{\text{biod}}, \boldsymbol{\omega}_{\text{kpiers}}\} \\ \mathbf{q}_{\text{udo}} &\Rightarrow \{\theta_{\text{udo}}, \boldsymbol{\omega}_{\text{udo}}\} \\ \mathbf{q}_{\text{podud}} &\Rightarrow \{\theta_{\text{podud}}, \boldsymbol{\omega}_{\text{podud}}\} \\ \mathbf{q}_{\text{st}} &\Rightarrow \{\theta_{\text{st}}, \boldsymbol{\omega}_{\text{st}}\} \end{aligned} \tag{13.35}$$

13. Ilościowa analiza układu ruchu człowieka

W tym przypadku łańcuch kinematyczny ma postać:

$$\mathbf{g}_{\text{biod,st}}(\theta) = \left(\overbrace{e^{[\xi_{01}]\theta_{01}} e^{[\xi_{02}]\theta_{02}} e^{[\xi_{03}]\theta_{03}}}^{\text{biod}} \underbrace{e^{[\xi_1]\theta_1} e^{[\xi_2]\theta_2}}_{\text{udo}} \overbrace{e^{[\xi_3]\theta_3}}^{\text{podud}} \underbrace{e^{[\xi_4]\theta_4}}_{\text{st}} \right) \mathbf{g}_{\text{biod,st}}(0) \quad (13.36)$$

13.5. Przykład ilościowej analizy układu ruchu człowieka

W podrozdziale przedstawimy przykład wykorzystania omówionego aparatu do estymacji kąta w stawie łokciowym na podstawie danych pomiarowych. Przykład został opracowany na bazie wyników przedstawionych w pracach [238, 237].

13.5.1. Dane pomiarowe

W badaniach wykorzystano czujniki umożliwiające pomiar przyśpieszeń, prędkości kątowych oraz natężenia pola magnetycznego. Pomiaru były wykonywane w trzech osiach x , y oraz z . Częstotliwość próbkowania wynosiła 512 Hz. Sposób umocowania czujników przedstawiono na rys. 13.10.

Badania wykonano w dwóch przypadkach. W pierwszym z nich ruch w stawie był wykonywany bardzo powoli. W drugim przypadku ruch był wykonywany z normalną prędkością.

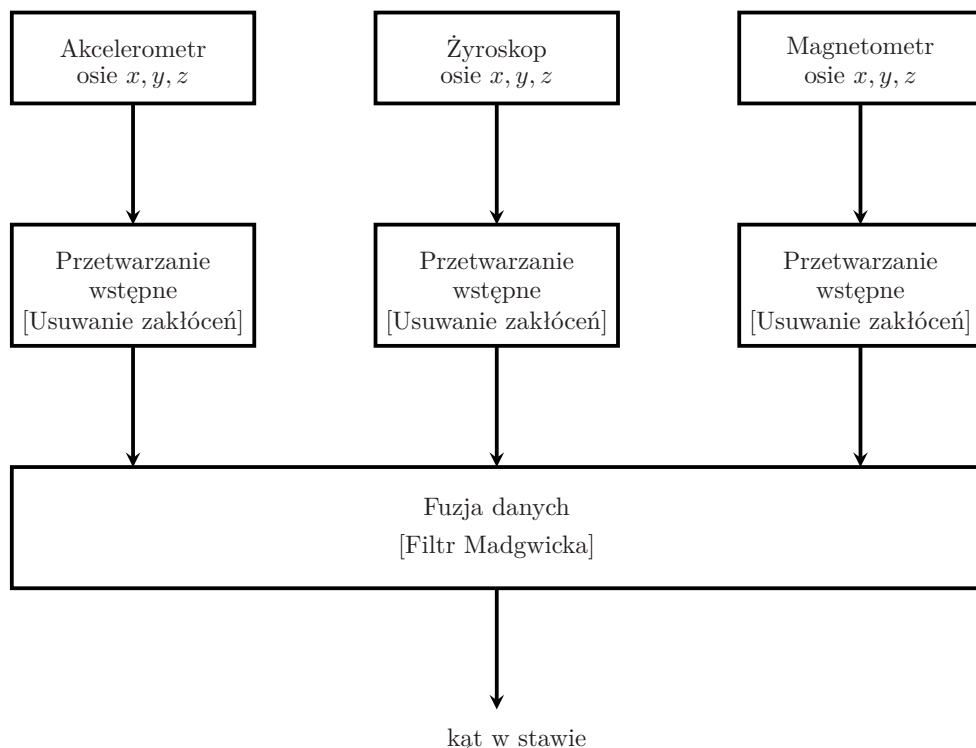


Rys. 13.10. Umocowanie czujników na ramieniu i przedramieniu [238]

Dane pomiarowe z czujników zostały zapisane. Oznaczmy pomiary przyśpieszeń jako $\mathbf{y}_a(k)$, prędkości kątowych jako $\mathbf{y}_\omega(k)$, a natężenie pola magnetycznego $\mathbf{y}_m(k)$, dla $k = 1, 2, \dots, K$.

13.5.2. Algorytm estymacji parametrów łańcucha kinematycznego kończyny górnej

W celu estymacji orientacji każdego z segmentów (w tym przypadku ramię i przedramię), do którego został dołączony czujnik, wykorzystano algorytm Madgwicka (patrz rys. 12.5). Następnie, z wykorzystaniem modelu łańcucha kinematycznego, wyznaczono kąt w stawie łokciowym. Schemat koncepcji rozwiązania zadania przedstawiono na rys. 13.11.

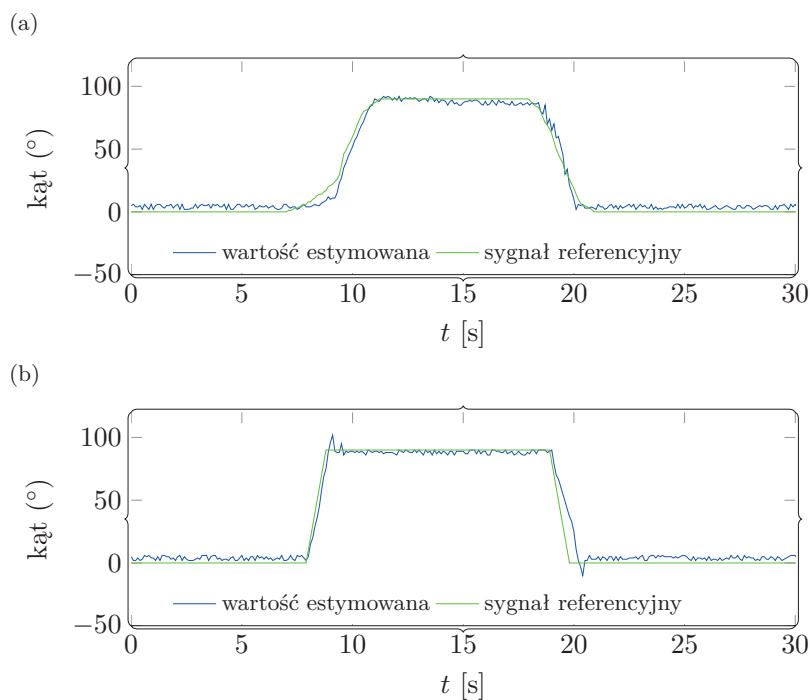


Rys. 13.11. Algorytm do estymacji kątów w stawie bazujący na fuzji danych

Mając oszacowane wartości kwaternionów dla każdego ze stawów, wyznaczaliśmy kąt w zginanym stawie łokciowym. W tym celu wykorzystaliśmy opracowany łańcuch kinematyczny (13.34) oraz zależności przedstawione w podrozdziale 13.4.

Wynik działania zaproponowanego algorytmu przedstawiono na rys. 13.12. Górny wykres przedstawia wynik eksperymentu w przypadku, gdy ruch w sta-

13. Ilościowa analiza układu ruchu człowieka



Rys. 13.12. Wynik estymacji kąta w stawie łokciowym z wykorzystaniem filtra Madgwicka: a) w przypadku wolnego ruchu w stawie oraz b) w normalnym ruchu

wie wykonywany był bardzo powoli. Na wykresie dolnym przedstawiono wynik eksperymentu, gdy ruch w stawie wykonywany był z normalną prędkością.

Część IV

Zastosowania

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym

14.1. Wstęp

Rozdział poświęcony jest problemowi usuwania zakłóceń o niskich częstotliwościach z sygnału będącego wynikiem przetwarzania pomiarów z czujników przyspieszenia, żyroskopu i magnetometru. Tym sygnałem jest prędkość poruszającej się osoby. Prędkość poruszającej się osoby jest istotną wielkością, która ma znaczenie, np. w medycynie (rehabilitacja) [81, 88, 272, 304] jak też w sporcie [252] (por. rozdz. 1.7). Wyniki przedstawione w tym rozdziale bazują na pracy [49].

Najpopularniejszymi metodami wykorzystywanymi w badaniach nad kinematyką ludzkiego ciała są systemy przechwytywania obrazu tzw. systemy typu *motion capture*. Systemy te są popularnymi rozwiązaniami wykorzystywanymi m.in. w medycynie i w sporcie. Warto jednak zaznaczyć, że rozwiązania te nie są pozbawione wad. Wadami tych rozwiązań jest ich wysoka cena oraz brak możliwości wykorzystania poza laboratorium. Brak możliwości wykorzystania tych systemów poza laboratorium sprawia, że nie jest możliwe np. dokonanie analizy kinematycznej skoczka narciarskiego lub pływaka w warunkach laboratoryjnych.

Rozwinięciem systemów przechwytywania ruchu *motion capture* są systemy wizyjne. Jedną z zalet tych rozwiązań jest to, że nie wymagają one dużych nakładów pracy przed wykonaniem badań, tak jak jest to w przypadku systemów *motion capture* (np. instalacja markerów na ciele osoby, która ma zostać poddana badaniu). Zaletą tych rozwiązań jest również to, że badania mogą być wykonywane w typowych dla danej dyscypliny miejscach, tj. na stadionie, korcie czy na stoku narciarskim. Technologia ta jest aktualnie intensywnie rozwijana, m.in. z powodu dostępności do jednostek obliczeniowych o znacznych mocach, które wykorzystywane są do przetwarzania zebranych danych [84, 91].

Alternatywnym podejściem do wymienionych systemów analizy kinematyki ciała ludzkiego jest wykorzystanie czujników takich, jak akcelerometry, żyroskopy i magnetometry [347, 363]. Wykorzystanie tych czujników w konstrukcji systemów do analizy kinematyki człowieka ułatwia usunięcie jednej z głównych wad systemów typu *motion capture*, tj. umożliwiają one wykonywanie badań poza laboratorium,

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym

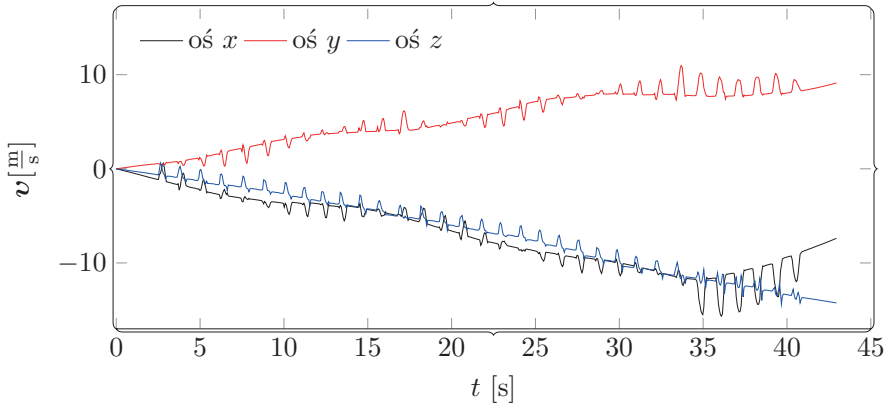
co w przypadku np. sportowców jest istotne, ponieważ właśnie poza laboratorium uprawiają swoje dyscypliny sportowe.

Jeden ze sposobów estymacji prędkości poruszającego się człowieka bazuje na modelowaniu matematycznym, w którym wykorzystuje się biomechaniczne modele ruchu człowieka [272] (por. rozdz. 13). Na przykład prędkość chodu można estymować z wykorzystaniem modelu kinematyki kończyn dolnych człowieka [364, 459] (por. rozdz. 1.7).

Drugie z podejść wykorzystuje metody znane z uczenia maszynowego [272]. Podejście to jest najmniej rozpowszechnione w tym obszarze badawczym. Jedną z głównych trudności w wykorzystaniu metod uczenia maszynowego do estymacji prędkości chodu człowieka jest problem z konstrukcją zbioru uczącego. Jakość estymacji prędkości chodu w przypadku niepoprawnie przygotowanego zbioru uczącego jest niewielka. Jest to głównie związane z tym, że trudno jest uwzględnić wszystkie możliwe przypadki, które mogą pojawić się podczas chodzenia. Dodatkową trudnością w stosowaniu metod uczenia maszynowego w estymacji prędkości chodu jest to, że dane pomiarowe, na podstawie których zadanie to jest rozwiązywane, są niestacjonarne.

Ostatni ze sposobów bazuje na metodach przetwarzania sygnałów. W podejściu tym prędkość chodu wyznaczana jest na podstawie pomiarów przyśpieszenia, prędkości kątowej i natężenia pola magnetycznego. Ponieważ nie jest możliwe uzyskanie wartości prędkości na podstawie pomiarów tylko przyśpieszeń, dane z akcelerometrów, żyroskopów i magnetometrów muszą zostać wstępnie przygotowane do dalszych obliczeń. Dopiero po ich wstępnym przetworzeniu możliwa jest estymacja prędkości. Niestety podejście bazujące na przetwarzaniu sygnałów z czujników takich, jak akcelerometr, żyroskop czy magnetometr ma swoje wady. Najpoważniejszą wadą tego podejścia jest niskoczęstotliwościowa składowa (tj. dryft), która jest składnikiem wyznaczonej tym sposobem prędkości chodu [380, 400]. Ilustracja dryftu w estymowanych wartościach prędkości chodu została przedstawiona na rys. 14.1.

Podstawową metodą do usuwania dryftu z estymowanej prędkości chodu jest dedykowana dla tego zadania metoda ZUPT (ang. *zero-velocity update*). W metodzie tej do usuwania dryftu wykorzystuje się to, że chód człowieka jest cykliczny [283]. Idea tej metody polega na zerowaniu prędkości w chwili, gdy stopa opiera się o podłoże i przez chwilę pozostaje bez ruchu, wówczas prędkość wynosi zero. Główną wadą tej metody jest to, że może być stosowana jedynie w przypadku chodu człowieka, ponieważ tylko podczas chodu występuje moment, gdy stopa nieruchomo opiera się na podłożu [310]. Ze względu na fakt, że metoda ZUPT została skonstruowana na bazie tej obserwacji, to ma ona ograniczony zakres stosowania. Inną wadą podejścia ZUPT jest również problem z wyznaczeniem momentu, gdy stopa nieruchomo opiera się na podło-



Rys. 14.1. Ilustracja estymacji prędkości z dryftem

żu. Precyzja w wyznaczaniu tego momentu ma wpływ na końcową dokładność omawianej metody.

Inny sposób do estymacji prędkości chodu przedstawiono w pracy [463], gdzie zaproponowano wstępne przetworzenie danych z akcelerometrów, żyroskopów i magnetometrów. Z kolei prędkość wyznaczana jest na podstawie opracowanych wcześniej danych. Jednak tak przygotowane dane nie gwarantują, że powstający na etapie estymacji prędkości dryft zostanie usunięty.

W pracach [128, 356] zaproponowano natomiast rozwiązanie, które bazuje na filtrach górnoprzepustowych. Niestety autorzy wykorzystali w swoich badaniach filtr Butterwortha, który nie umożliwia skutecznego usunięcia niskoczęstotliwościowych zakłóceń, które są trudne do usunięcia przez filtr tego typu.

W niniejszym podrozdziale zaproponowano metodę usuwania niskoczęstotliwościowych zakłóceń z wykorzystaniem metod modelowania matematycznego. Metody modelowania zostaną wykorzystane do estymacji dryftu. W prezentowanym przypadku zastosowano podejście do modelowania bazujące na modelach odcinkowo-liniowych (patrz podrozdz. 5.7.2). Podejście bazujące na funkcjach odcinkowo-liniowych do estymacji dryftu przedstawiono m.in. w [128, 356].

14.2. Algorytm estymacji prędkości chodu

14.2.1. Estymacja orientacji

Oznaczmy pomiary przyśpieszeń jako $\mathbf{y}_a(k)$, prędkości kątowych jako $\mathbf{y}_\omega(k)$, a natężenie pola magnetycznego $\mathbf{y}_m(k)$, dla $k = 1, 2, \dots, K$.

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym

Do reprezentacji orientacji obiektu w \mathbb{R}^3 wykorzystamy kwaternion $\mathbf{q}_{a,b}$. Zastosujemy następujące równanie ruchu (por. rozdz. 13):

$$\dot{\mathbf{q}}_{a,b} = \frac{1}{2}\Omega(\boldsymbol{\omega}(t))\mathbf{q}_{a,b} \quad (14.1)$$

gdzie indeks a oznacza powiązanie z układem odniesienia a b z układem lokalnym.

W podrozdziale 12.3 przedstawiliśmy różne metody do estymacji orientacji obiektu w przestrzeni, tj. do estymacji kwaternionu $\mathbf{q}_{a,b}$. Opracowany algorytm do estymacji prędkości chodu bazuje na omówionym w tym rozdziale algorytmie Madgwicka.

14.2.2. Estymacja prędkości

Kolejny krok w zaproponowanym rozwiązaniu to estymacja prędkości na podstawie pomiarów przyśpieszenia. W przypadku dysponowania idealnymi pomiarami zadanie sprowadziłoby się do:

$$\dot{\mathbf{v}}_a = \mathbf{a}_a \quad (14.2)$$

Niestety nie dysponujemy dokładnymi pomiarami przyśpieszenia \mathbf{a}_a w układzie odniesienia. Posiadamy natomiast pomiary przyśpieszenia wykonane w układzie ciała, tj. $\mathbf{y}_a(k)$. Dysponując tymi pomiarami oraz znając wartość kwaternionu $\mathbf{q}_{a,b}$, możliwe jest wyznaczenie przyśpieszenia w układzie odniesienia:

$$\hat{\mathbf{a}}_a = \hat{\mathbf{q}}_{a,b} \otimes \mathbf{y}_a \otimes \hat{\mathbf{q}}_{a,b} - [0, 0, g]^\top \quad (14.3)$$

gdzie $g = 9,81 \text{ m/s}^2$.

Teoretycznie całkując (numerycznie) wyznaczone wartości $\hat{\mathbf{a}}_a$, szacujemy prędkość:

$$\hat{\mathbf{v}}_a(t) = \int_0^{\Delta t} \hat{\mathbf{a}}_a(t) dt + \hat{\mathbf{v}}_b(t-1) \quad (14.4)$$

gdzie $[0, \Delta t]$ jest okresem próbkowania. Operacja całkowania w zależności (14.4) może być wykonana jedną ze znanych metod (np. metodą trapezów, metodą Simpson etc.).

14.2.3. Usuwanie dryftu

Kolejne zadanie to korekta estymowanej prędkości poruszającego się człowieka. Zadanie to składa się z dwóch części, tj. estymacji dryftu i usuwania dryftu z estymowanej wartości prędkości. Opiszmy te kroki.

14.2. Algorytm estymacji prędkości chodu

Przyjmijmy następujący model dla estymowanej wartości prędkości (równ. (14.4))

$$\bar{\mathbf{v}}_a(t) = \mathbf{v}_a(t) + \mathbf{v}_{a_{bs}}(t) \quad (14.5)$$

gdzie $\mathbf{v}_a(t)$ jest prawdziwą wartością prędkości chodu (14.2), a $\mathbf{v}_{a_{bs}}(t)$ jest dryftem w estymowanych wartościach prędkości. Z kolei naszym zadaniem jest estymować dryft na podstawie pomiarów:

$$\hat{\mathbf{v}}_{a_{bs}}(t) \approx \mathbf{v}_{a_{bs}}(t) \quad (14.6)$$

Dysponując oszacowaniem $\hat{\mathbf{v}}_{a_{bs}}(t)$, możliwe jest usunięcie dryftu z oszacowanej wartości prędkości $\hat{\mathbf{v}}_a(t)$:

$$\hat{\mathbf{v}}_a(t) = \bar{\mathbf{v}}_a(t) - \hat{\mathbf{v}}_{a_{bs}}(t) \quad (14.7)$$

gdzie $\hat{\mathbf{v}}_a(t)$ estymowaną prędkością chodu.

Algorytm estymacji dryftu

Dryft w estymowanej prędkości chodu jest niestacjonarny, przez co trudny w modelowaniu (por. rozdz. 13). Klasyczne metody bazujące na wielomianach nie uzyskują zadowalających wyników [324].

W rozdziale zaproponowaliśmy rozwiązanie, które bazuje na modelowaniu dryftu funkcjami odcinkowo-liniowymi (por. podrozdz 5.7.2). Metodę można przedstawić w następujący sposób, podzielić analizowany sygnał na N_S nienakładających się na siebie segmentów i estymuj dryft w każdym z nich oddzielnie. Na koniec dokonaj konkatenacji dryftu wyznaczonego w każdym segmencie. Schemat blokowy algorytmu przedstawiono na rys. 14.2.

Segmentacja sygnału

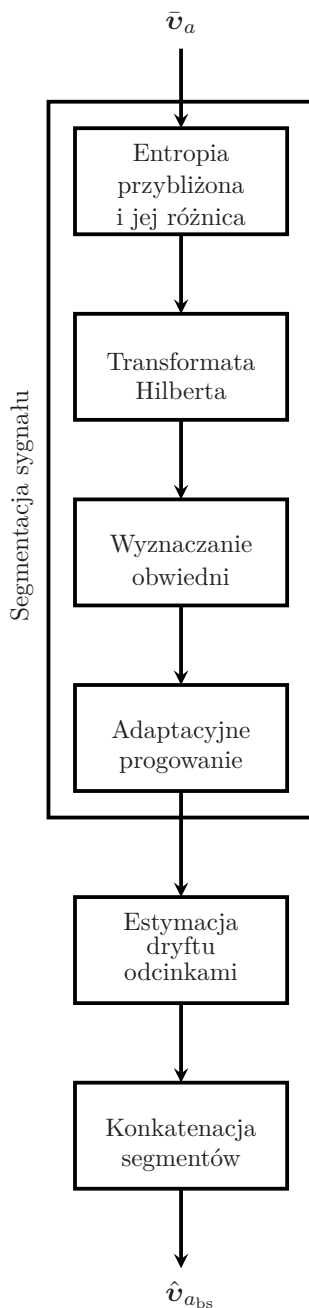
Analizując rysunki 14.1 i 14.3, zauważamy, że w estymowanej prędkości możemy wyróżnić dwa powtarzające się obszary: takie, w których wartość prędkości jest niewielka i takie, w których prędkość jest duża. Prędkość zostanie podzielona na dwa obszary, takie, w których wartość prędkości jest duża i takie, w których prędkość jest mała.

Podzielmy teraz sygnał prędkości chodu $\hat{\mathbf{v}}_a$ na N_S nienakładających się na siebie segmentów (por. podrozdz. 7.3):

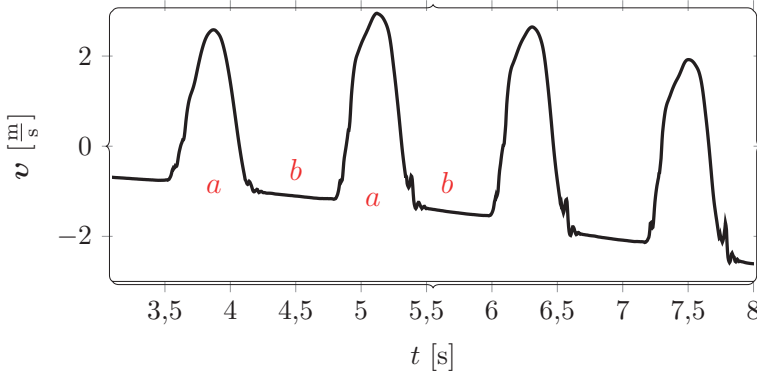
$$\mathbb{S} = (\mathbb{s}_1, \mathbb{s}_2, \dots, \mathbb{s}_{N_S}) \quad (14.8)$$

W zadaniu tym zaproponowano nowy algorytm segmentacji bazujący na nieliniowych metodach analizy sygnałów. W opisywanym przypadku wykorzystano

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym



Rys. 14.2. Schemat blokowy algorytmu estymacji dryftu



Rys. 14.3. Prędkość ruchu stopy: a) duża prędkość, b) mała prędkość

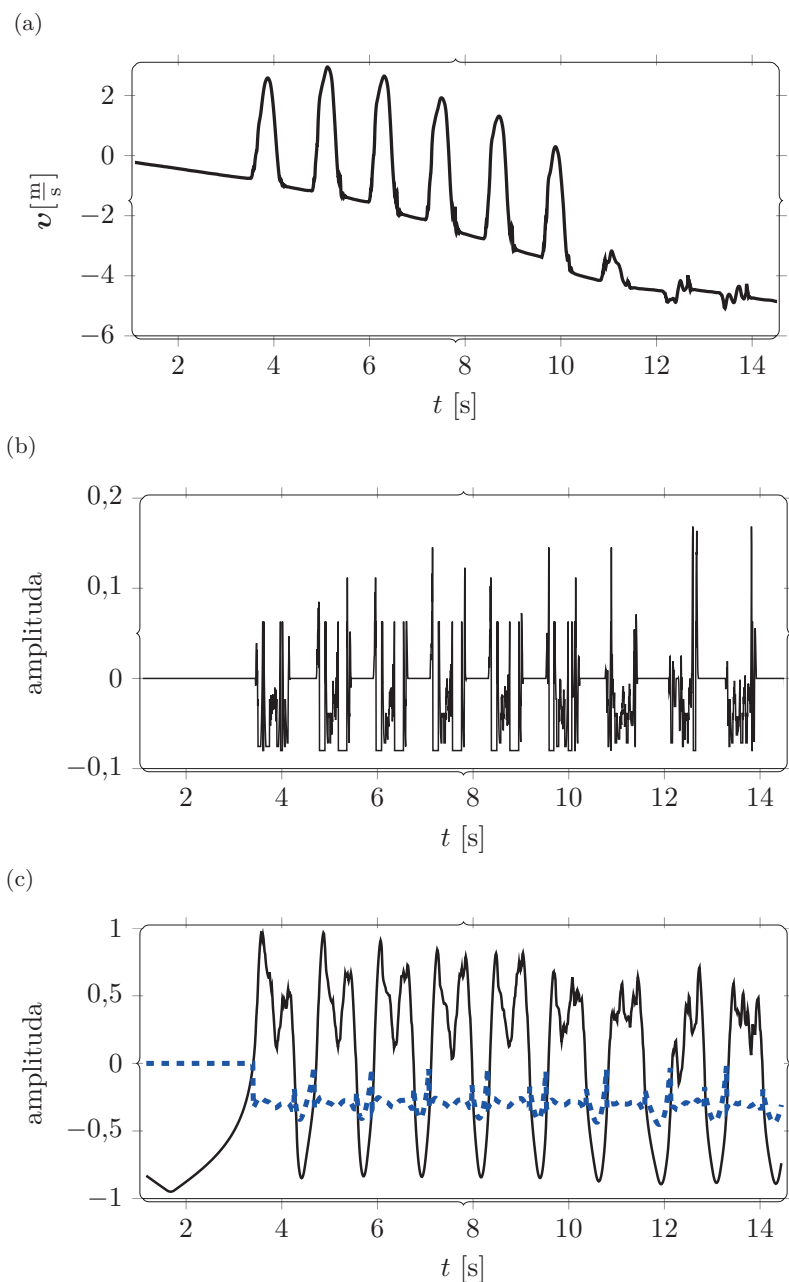
przybliżoną entropię (por. podrozdz. 7.5) jako narzędzie do przekształcenia sygnału prędkości chodu na wielkość, która umożliwi precyzyjną segmentację sygnału [154]. Przybliżona entropia umożliwi ustalenie w sygnale regionów o różnej dynamice [329, 339]. Ponieważ sygnał prędkości chodu ma takie regiony, przybliżona entropia ułatwi proces segmentacji, jeśli wykorzystana zostanie w pierwszym kroku do przekształcenia sygnału chodu. Dla ilustracji na rys. 14.4a przedstawiliśmy fragment estymowanej prędkości, a na rys. 14.4b pochodną z wyliczonej dla sygnału prędkości przybliżonej entropii. Z rysunku 14.4b widać, że możliwe jest wyraźne odseparowanie regionów o różnej dynamice zmian w sygnale.

Po wstępnym przygotowaniu sygnału następnym krokiem jest już związany z segmentacją. W zaproponowanym algorytmie wykorzystano podejście bazujące na transformacji Hilberta i adaptacyjnym progu [344]. W pierwszym kroku wyznaczamy transformację Hilberta sygnału (podrozdz. 7.4.2) i bazując na jej wyniku, wyznaczamy obwiednie przetworzonego sygnału (rys. 14.4c). Następnie, stosując adaptacyjny próg (rys. 14.4c) dla wyznaczonej obwiedni, wyznaczamy granice kolejnych segmentów.

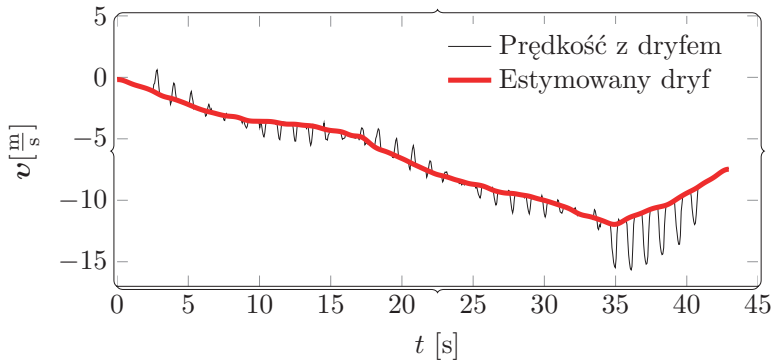
Estymacja dryftu

Przedstawimy teraz algorytm estymacji dryftu w sygnale \hat{v}_a , wykorzystując aproksymację odcinkowo-liniową. Wykorzystany model został przedstawiony w podrozdziale 5.7.2. Przykładowy wynik estymacji dryftu z prędkości chodu (dla osi x) przedstawiono na rys. 14.5.

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym



Rys. 14.4. Prędkość chodu dla jednej z osi z dryftem (a), pochodna przybliżonej entropii z sygnału (b), obwiednia (linia czarna) oraz adaptacyjny próg (niebieska linia – (c))

Rys. 14.5. Estymowany dryft z prędkości chodu dla osi x

14.3. Dane pomiarowe

W badaniach wykorzystano dwa zbiory danych. Pierwszy zbiór danych wykorzystany w badaniach został opisany w [264] i jest dostępny [262]. System pomiarowy wykorzystuje czujnik Xsens MTx [452]. W skład urządzenia wchodzi 16-bitowe trójosiowe czujniki przyspieszenia, prędkości kątowych i natężenia pola magnetycznego. Dane były próbkowane z częstotliwością 512 Hz.

W trakcie eksperymentu czujnik był przymocowany do buta badanej osoby (parz rys. 4.5 oraz 14.6). Badania wykonano dla scenariusza **chodzenie na spiralnych schodach** (zbiór ten oznaczyliśmy A).



Rys. 14.6. Ilustracja pomiaru w poszczególnych fazach chodu

Drugi zbiór testowy pochodzi z [14]. Z tego zbioru wybraliśmy badanie dotyczące **biegu** (który oznaczyliśmy jako B).

W tym przypadku dane zostały zarejestrowane z wykorzystaniem czujnika Xsens MTx-28A53G25 [452] z częstotliwością 100 Hz [15]. Zbiór ten zawiera dodatkowe pomiary z systemu *motion capture* (platforma Bonita). Umożliwi to porównanie zaproponowanego rozwiązania z inną metodą pomiarową, która nie opiera się na danych z czujników typu akcelerometr, żyroskop czy magnetometr.

14.4. Badania eksperymentalne

14.4.1. Kryteria oceny

Do oceny jakości estymacji dryftu z prędkości chodu wykorzystano zależność:

$$\varepsilon(k) = \mathbf{v}_{a_{\text{ref}}}(k) - \hat{\mathbf{v}}_a(k) \quad (14.9)$$

gdzie $\hat{\mathbf{v}}_a(k)$ (14.7) jest estymowaną prędkością chodu, $\mathbf{v}_{a_{\text{ref}}}(k)$ jest referencyjnym sygnałem pozyskanym z metody ZUPT (w przypadku zbioru A) lub systemu *motion capture* Bonita (dla zbioru B).

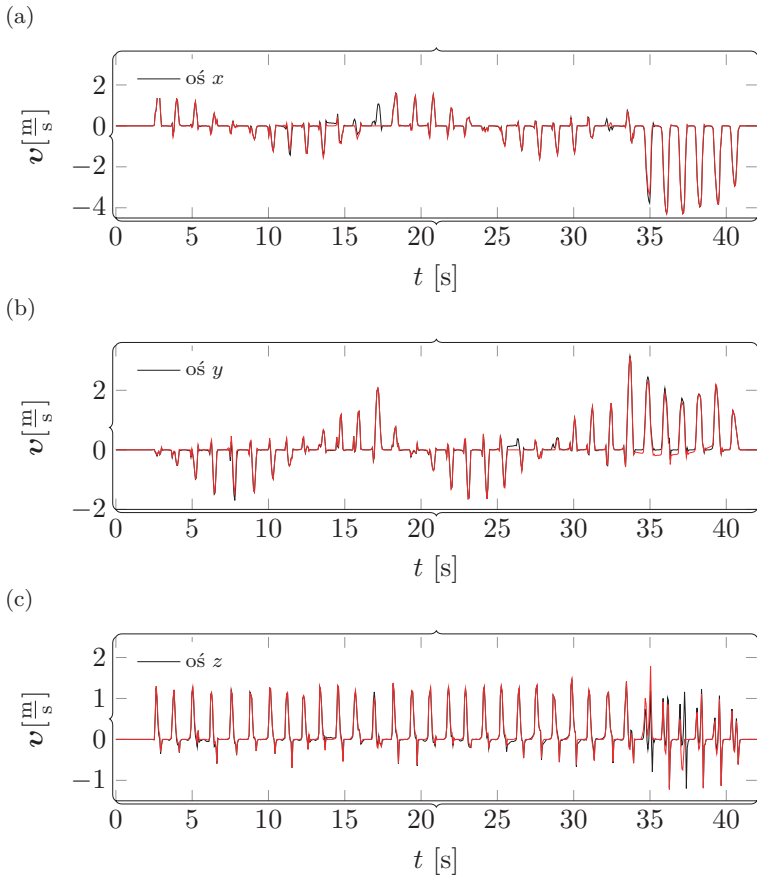
14.4.2. Porównanie z metodą ZUPT

Na rysunku 14.7 porównano zaproponowane podejście algorytmem ZUPT. W tych badaniach wykorzystano zbiór A (patrz podrozdz. 14.3). Reszty dla zaproponowanej metody przedstawiono z kolei na rys. 14.8 (reszty zostały zdefiniowane zależnością (14.9)).

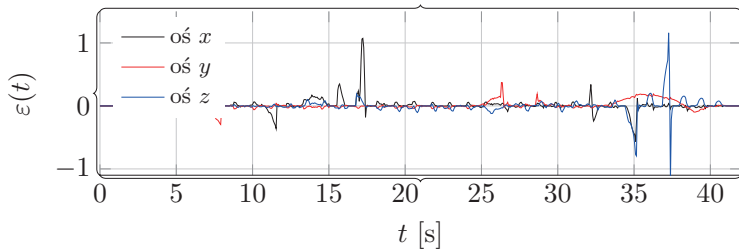
14.4.3. Porównanie z systemem *motion capture*

Wyniki uzyskane z wykorzystaniem zaproponowanej w tym rozdziale metody zostały również porównane z systemem typu *motion capture*. Zbiór przedstawiony w [14] zawiera dane zebrane nie tylko z czujników typu akcelerometr, żyroskop czy magnetometr, ale również z systemu do przechwytywania ruchu Bonita. Dane te zostały wykorzystane w badaniach jako dane referencyjne. Wykorzystanie danych z systemu typu Bonita umożliwia porównanie efektywności zaproponowanej metody z systemem, który estymuje prędkość chodu w inny sposób niż metody bazujące na danych typu przyspieszenie, prędkość kątowna lub natężenie pola magnetycznego. Drugi cel tych badań dotyczy oceny jakości estymacji prędkości poruszania się nie podczas chodu, lecz na przykład podczas biegu. Badanie to ułatwi ocenę, czy przedstawione podejście umożliwia estymację prędkości osoby biegnącej. Charakteryzując metodę ZUPT, zaznaczono, że nie wpływa ona na poprawną estymację prędkości biegu uczestnika badania.

W drugiej części badań wykorzystano zbiór B, tj. zbiór z zarejestrowanym biegiem uczestnika badania. Ponieważ metoda ZUPT nie jest w stanie poprawnie estymować prędkości w tych warunkach, dlatego wyniki uzyskane dla zaproponowanej metody porównano z danymi pozyskanymi z systemu do przechwytywania ruchu. Z eksperymentu wynika, że zaproponowana metoda umożliwia usunięcie niskoczęstotliwościowych zakłóceń z estymowanej prędkości biegu. Ze względu na wyższą intensywność ruchu, jaką jest bieg, jakość estymacji jest niższa (patrz rys. 14.9). Jednakże metoda umożliwia usunięcie zakłóceń w przypadku biegu, co

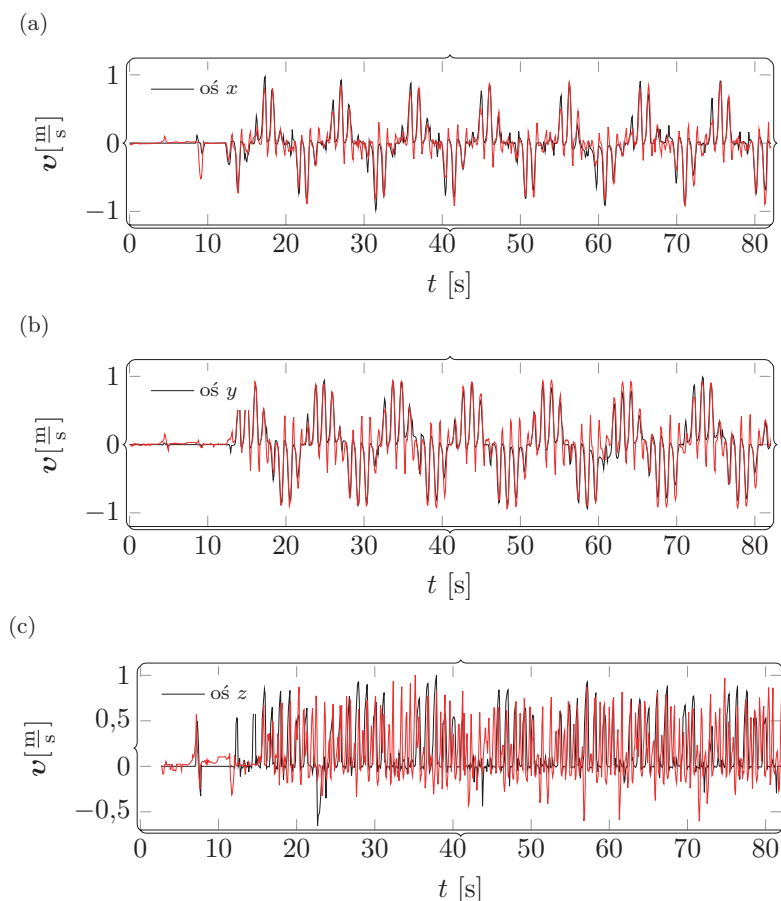


Rys. 14.7. Zbiór **chodzenie po spiralnych schodach**. Porównanie wyników dla metody zaproponowanej (linia czarna) z metodą ZUPT (linia czerwona) dla osi: a) x , b) y , c) z



Rys. 14.8. Zbiór **chodzenie po spiralnych schodach**: reszty dla zaproponowanej metody

14. Odszumianie bazujące na modelowaniu odcinkowo-liniowym



Rys. 14.9. Zbiór **bieg**. Porównanie uzyskanych wyników (linia czarna) z danymi z systemu *motion capture* (linia czerwona) dla osi: a) x , b) y , c) z

jest niemożliwe w przypadku metody ZUPT. Warto podkreślić, że jakość estymacji może zostać podwyższona przez zwiększenie jakości segmentacji.

14.5. Uwagi końcowe

W rozdziale przedstawiono algorytm do modelowania i usuwania z sygnału składowych o niskich częstotliwościach. Zaproponowane podejście charakteryzuje się tym, że do estymacji nie wymaga żadnych dodatkowych elementów infrastruktury (np. systemu GPS) i jest ogólniejsze od metod uznanych za typowe w tym zakresie (np. metoda ZUPT).

W badaniach wykorzystane zostały dwa zbiory danych, które pochodzą z różnych badań i zostały przygotowane przez różne zespoły. Zaproponowane podejście uzyskało bardzo dobre wyniki dla każdego z nich. Wskazuje to m.in. na uniwersalność metody.

Algorytm został również porównany z danymi uzyskanymi z systemów przechwytywania ruchu tzw. systemów *motion capture*. Również i w tym przypadku uzyskane wyniki wskazują na dużą precyzję zaproponowanego podejścia. Z przeprowadzonych badań wynika również, że jakość estymacji prędkości może zostać podwyższona przez opracowanie bardziej precyzyjnych algorytmów segmentacji. To właśnie jakość segmentacji została wskazana podczas badań jako jeden z kluczowych elementów wpływających na jakość estymacji prędkości.

Zaletą zaproponowanego algorytmu jest również to, że w porównaniu do metod, takich jak ZUPT, umożliwia ona estymację prędkości również podczas biegu. Cecha ta została potwierdzona wynikami badań ilościowych, które przedstawiliśmy w tym rozdziale. Możliwość estymacji prędkości poruszającej się osoby jest cechą odróżniającą zaproponowane podejście od innych znanych metod.

15. Odszumianie bazujące na reprezentacji rzadkiej

15.1. Wstęp

W rozdziale przedstawiono alternatywne podejście do usuwania niskoczęstotliwościowych zakłóceń z estymowanej prędkości chodu. Rozważania w tym rozdziale zostały rozszerzone o systemy wspomagające nawigację pieszych. Wprowadzenie dodatkowego zagadnienia będzie wiązało się z dodatkowymi zadaniami związanymi z wyznaczeniem pozycji użytkownika w przestrzeni. Rozdział ten powstał na bazie artykułu [50].

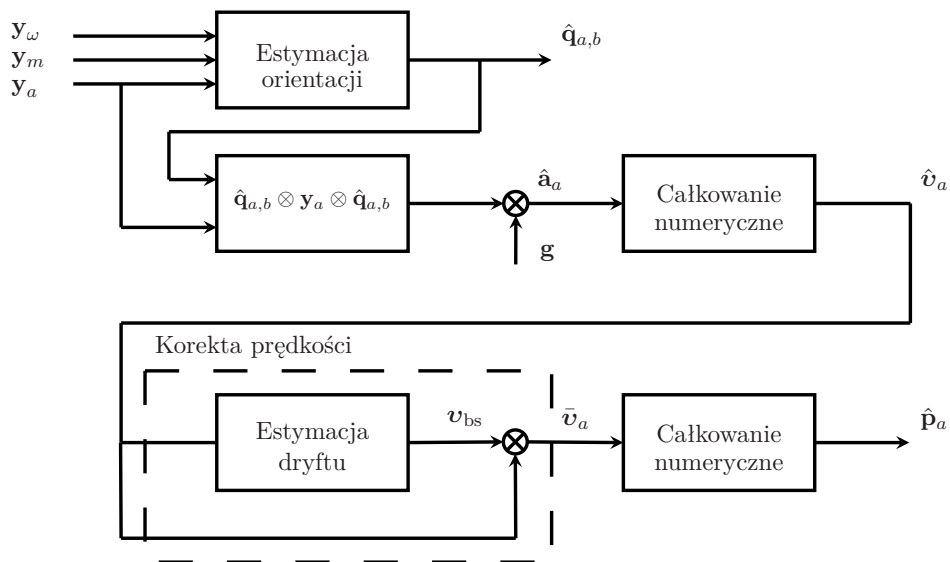
Systemy wspomagające nawigację pieszych (ang. *Pedestrian Navigation Systems*, PNS) można podzielić na systemy, które wykorzystują dodatkową infrastrukturę oraz na takie, które nie korzystają z dodatkowej infrastruktury [314].

Do pierwszej kategorii zaliczamy m.in. najpopularniejsze systemy nawigacji pieszych, które wspomagane są przez systemy nawigacji satelitarnej. Inne, mniej popularne, ale zyskujące popularność, to systemy wykorzystujące do nawigacji internetowe punkty dostępne (ang. *Internet access points*) czy techniki radiokomunikacyjne takie, jak UWB (ang. *Ultra Wide Band*). Celem dodatkowej infrastruktury jest zapewnienie informacji o punktach, których lokalizacje są znane.

Do drugiej kategorii zaliczamy metody, które wykorzystują pomiary z takich czujników, jak akcelerometr, żyroskop czy magnetometr [31]. Tak jak pokazaliśmy w poprzednim rozdziale, odpowiednie wykorzystanie danych z czujników ułatwia estymację prędkości poruszającej się osoby. W systemach nawigacji pieszych konieczne jest dwukrotne wykonanie operacji całkowania numerycznego. Wykonanie tej operacji dla sygnału przyspieszenia umożliwia, przynajmniej teoretycznie, wyznaczenie pozycji pieszego. Jednym z głównych problemów w tym podejściu jest to, że podwójne całkowanie takiego sygnału prowadzi nie tylko do wyznaczenia prędkości. Skutkiem ubocznym całkowania zaszumionego sygnału pomiarowego jest to, że w sygnale wynikowym pojawi się niepożądana składowa w postaci dryftu.

15.2. Algorytm estymacji prędkości chodu oraz pozycji

W rozdziale przedstawimy algorytm do estymacji prędkości i pozycji na podstawie danych pomiarowych z czujników przyspieszenia, prędkości kątowych oraz natężenia pola magnetycznego. Schemat blokowy dla zaproponowanego algorytmu przedstawiono na rys. 15.1.



Rys. 15.1. Schemat blokowy algorytmu do estymacji prędkości i pozycji na podstawie danych z czujników przyspieszenia, prędkości kątowych oraz natężenia pola magnetycznego

Pierwszy etap w algorytmie estymacji prędkości i pozycji dotyczy estymacji orientacji obiektu w przestrzeni. Kolejne kroki na tym etapie są takie jak te, opisane w podrozdziale 14.2.1.

Również i w tych badaniach do estymacji orientacji obiektu w przestrzeni wykorzystano algorytm Madgwicka [264].

Kolejny etap to wyznaczenie na podstawie przygotowanych pomiarów przyspieszenia prędkości oraz pozycji. W tym celu wykorzystano wyznaczony na wcześniejszym etapie kwaternion. Wówczas liniowe przyspieszenie wyznaczamy z zależności (14.3).

Następnie całkujemy przyspieszenie $\hat{\mathbf{a}}_a$, korzystając z zależności (14.4). Również i tym razem do całkowania numerycznego można wykorzystać jedną ze znanych metod (np. metodę trapezów, metodę Simpsona itd.).

Następnie, chcąc wyznaczyć pozycję obiektu na podstawie oszacowanej prędkości \hat{v}_a , należy powtórzyć operację całkowania:

$$\hat{p}_a(t) = \int_0^{\Delta t} \hat{v}_a(t) dt + \hat{p}_a(t-1) \quad (15.1)$$

Tak, jak to już zostało omówione w poprzednim rozdziale, estymacja prędkości (a w tym przykładzie również pozycji) przez numeryczne całkowanie zakłóconych pomiarów sprawia, że w estymowanym sygnale pojawi się niepożądany składnik, tj. dryft [283]. Przedstawimy teraz algorytm do estymacji dryftu w sygnałach, który bazuje na rzadkiej reprezentacji sygnałów (por. podrozdz. 5.7.7).

15.3. Estymacja dryftu

Rozwiązanie bazujące na rzadkiej reprezentacji ułatwia konstruować, np. filtry dolnoprzepustowe umożliwiające usuwanie z sygnału niestacjonarnych zakłóceń [127, 468]. Zaproponowane podejście bazuje na algorytmie wykorzystującym rzadką reprezentację, tj. na metodzie TARA (ang. *Transient Artifact Reduction Algorithm*) [361].

Ustalmy model dla estymowanej prędkości $\bar{v}_a(k)$ w formie:

$$\bar{v}_a(k) = \mathbf{v}_{bs}(k) + \boldsymbol{\kappa}_1(k) + \boldsymbol{\kappa}_2(k) + \mathbf{v}(k) \quad (15.2)$$

gdzie $\mathbf{v}_{bs}(k)$ jest dryftem, $\boldsymbol{\kappa}_1(k)$ oraz $\boldsymbol{\kappa}_2(k)$ to dwa typy zakłóceń (typ 1 oraz typ 2, por. podrozdz. 5.7.7) oraz $\mathbf{v}(k)$ jest zakłóceniem *gaussowskim*. Zakłócenie typu 1 jest sygnałem rzadkim mającym rzadką pierwszą pochodną, z kolei zakłócenie typu 2 jest sygnałem o rzadkiej pierwszej pochodnej [361].

Wykorzystanie zakłócenia typu 1 wiąże się z modelowaniem krótkich zakłóceń w postaci fali, natomiast zakłócenie typu 2 wykorzystano do modelowania zakłóceń w postaci funkcji schodkowej [361].

W dalszej części opisu składowe modelu (15.2) będziemy przedstawiać w formie wektorów. Niech model ma postać:

$$\bar{\mathbf{v}}_a = \mathbf{v}_{bs} + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2 + \mathbf{v} \quad (15.3)$$

gdzie $\bar{\mathbf{v}}_a \in \mathbb{R}^K$, $\mathbf{v}_{bs} \in \mathbb{R}^K$, $\boldsymbol{\kappa}_1 \in \mathbb{R}^K$, $\boldsymbol{\kappa}_2 \in \mathbb{R}^K$, a $\mathbf{v} \in \mathbb{R}^K$.

Zakłócenia $\boldsymbol{\kappa}_1$ oraz $\boldsymbol{\kappa}_2$ są sygnałami o rzadkiej pierwszej pochodnej. Do przybliżenia pochodnej zastosowano dyskretny operator różnicowy pierwszego rzędu (ang. *the discrete-time first-order difference operator*) \mathbf{D} tj. $[\mathbf{D}\boldsymbol{\kappa}]_k = [\boldsymbol{\kappa}]_{k+1} - [\boldsymbol{\kappa}]_k$.

15. Odszumianie bazujące na reprezentacji rzadkiej

Macierz \mathbf{D} ma następującą postać (por. (5.96)):

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & -1 & 1 \end{bmatrix} \quad (15.4)$$

gdzie $[\boldsymbol{\kappa}]_n$ oznacza k -ty składnik wektora \mathbf{x} .

Problem estymacji zakłóceń $\boldsymbol{\kappa}_1$ oraz $\boldsymbol{\kappa}_2$ można przedstawić jako następujące zadanie optymalizacji:

$$\{\hat{\boldsymbol{\kappa}}_1, \hat{\boldsymbol{\kappa}}_2\} = \arg \min_{\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2} \left\{ \frac{1}{2} \|\mathbf{F}_H(\bar{\mathbf{v}}_a - \boldsymbol{\kappa}_1 - \boldsymbol{\kappa}_2)\|_2^2 + \lambda_0 \sum_{k=1}^K \phi_0([\boldsymbol{\kappa}_1]_k) + \right. \\ \left. + \lambda_1 \sum_{k=1}^K \phi_1([\mathbf{D}\boldsymbol{\kappa}_1]_k) + \lambda_2 \sum_{k=1}^K \phi_2([\mathbf{D}\boldsymbol{\kappa}_2]_k) \right\} \quad (15.5)$$

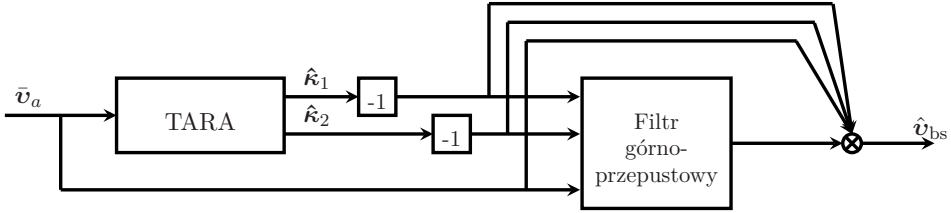
gdzie $\mathcal{F}_H \equiv \mathbf{F}_H$ jest filtrem górnoprzepustowym, ϕ_0 , ϕ_1 oraz ϕ_2 są takimi funkcjami kary, że $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$ generują sygnał rzadki, λ_0 , λ_1 i λ_2 są dodatnimi parametrami [361]. Przykładem funkcji, która generuje sygnał o takich właściwościach jest np. $\phi_i(x) = |x|$. Listę innych funkcji, które można w tym zadaniu wykorzystać, przedstawiono m.in. w [361].

Po wyznaczeniu zakłóceń $\boldsymbol{\kappa}_1$ oraz $\boldsymbol{\kappa}_2$, dryft $\hat{\mathbf{v}}_{\text{bs}}$ w prędkości (15.3) wyznaczamy z zależności:

$$\hat{\mathbf{v}}_{\text{bs}} = \mathbf{F}_L(\bar{\mathbf{v}}_a - \hat{\boldsymbol{\kappa}}_1 - \hat{\boldsymbol{\kappa}}_2) = \bar{\mathbf{v}}_a - \hat{\boldsymbol{\kappa}}_1 - \hat{\boldsymbol{\kappa}}_2 - \mathbf{F}_H(\bar{\mathbf{v}}_a - \hat{\boldsymbol{\kappa}}_1 - \hat{\boldsymbol{\kappa}}_2) \quad (15.6)$$

gdzie \mathbf{F}_L jest filtrem dolnoprzepustowym, takim że $\mathbf{F}_L = \mathbf{I} - \mathbf{F}_H$, w którym \mathbf{I} to macierz jednostkowa. Natomiast filtr górnoprzepustowy \mathbf{F}_H definiujemy jako $\mathbf{F}_H = \mathbf{B}\mathbf{A}^{-1}$, gdzie \mathbf{A} oraz \mathbf{B} są macierzami wstęgowymi (ang. *banded matrices*) [362]. Przyjmujemy również, że \mathbf{B} można przedstawić w formie $\mathbf{B} = \mathbf{B}_1\mathbf{D}$. W pracy [362] pokazano, że macierze \mathbf{D} oraz \mathbf{A}^{-1} komutują. Stąd problem optymalizacyjny (15.5) można przedstawić w formie:

$$\{\hat{\boldsymbol{\kappa}}_1, \hat{\boldsymbol{\kappa}}_2\} = \arg \min_{\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2} \left\{ \frac{1}{2} \|\mathbf{F}_H\bar{\mathbf{v}}_a - \mathbf{B}\mathbf{A}^{-1}\boldsymbol{\kappa}_1 - \mathbf{B}_1\mathbf{A}^{-1}\mathbf{D}\boldsymbol{\kappa}_2\|_2^2 + \lambda_0 \sum_k \phi_0([\boldsymbol{\kappa}_1]_k) + \right. \\ \left. + \lambda_1 \sum_n \phi_1([\mathbf{D}\boldsymbol{\kappa}_1]_k) + \lambda_2 \sum_k \phi_2([\mathbf{D}\boldsymbol{\kappa}_2]_k) \right\} \quad (15.7)$$



Rys. 15.2. Schemat blokowy algorytmu bazującego na metodzie TARA

Na rysunku 15.2 przedstawiono kluczowe elementy zaproponowanego algorytmu, które służą do:

Krok 1 Estymacja zakłóceń typu 1 (κ_1) oraz typu 2 (κ_2) (15.7),

Krok 2 Estymacja dryftu \hat{v}_{bs} , korzystając z zależności (15.6),

Krok 3 Estymacja prędkości na podstawie estymowanych wielkości: κ_1 , κ_2 oraz \hat{v}_{bs} .

15.3.1. Dobór parametrów algorytmu TARA

Przygotowano dwa warianty algorytmu TARA z następującymi funkcjami:

$$\phi_i(x) = |x| \quad (15.8)$$

oraz

$$\phi_i(x) = \frac{2}{a_i\sqrt{3}} \left(\tan^{-1} \left(\frac{1 + 2a_i\sqrt{x^2 + \epsilon}}{\sqrt{3}} \right) - \frac{\pi}{6} \right) \quad (15.9)$$

gdzie $i = 0, 1, 2$, wielkość ϵ w zależności (15.9) przyjmuje wartość z przedziału $10^{-8} \leq \epsilon \leq 10^{-10}$ [361].

Algorytm TARA wykorzystujący pierwszą z tych funkcji jako funkcje kary (15.8) oznaczyliśmy jako TARA/ ℓ_1 . Natomiast algorytm oparty na funkcji kary (15.9) nazwaliśmy w naszych badaniach TARA/ATAN.

W przypadku pierwszej z metod TARA/ ℓ_1 oszacować należy parametry $\lambda_0, \lambda_1, \lambda_2$. Sposób wyznaczenia tych parametrów został przedstawiony w pracy [361] w rozdziale III-B.

Dla algorytmu TARA/ATAN szacujemy parametry $\lambda_0, \lambda_1, \lambda_2$, tak jak w przypadku metody TARA/ ℓ_1 . Dodatkowo należy wyznaczyć wielkości takie, jak a_0, a_1, a_2 (por. (15.9)). Do wyznaczenia tych parametrów stosuje się odpowiedź impulsową zaprojektowanego wcześniej filtra górnoprzepustowego:

$$a_0 = \frac{1}{2} \frac{\|\mathbf{k}_{\text{imp}}\|_2^2}{\lambda_0}, \quad a_1 = \frac{1}{2} \frac{\|\mathbf{k}_{\text{imp1}}\|_2^2}{\lambda_1}, \quad a_2 = \frac{1}{2} \frac{\|\mathbf{k}_{\text{imp1}}\|_2^2}{\lambda_2} \quad (15.10)$$

15. Odszumianie bazujące na reprezentacji rzadkiej

gdzie \mathbf{k}_{imp} jest odpowiedzią impulsową filtra $\mathbf{F}_H = \mathbf{B}\mathbf{A}^{-1}$, $\mathbf{k}_{\text{imp}1}$ jest odpowiedzią impulsową filtra $\mathbf{F}_{H_1} = \mathbf{B}_1\mathbf{A}^{-1}$ (patrz (15.7)).

Alternatywną metodą estymacji parametrów $\lambda_0, \lambda_1, \lambda_2$ oraz a_0, a_1, a_2 jest ich wyznaczenie z wykorzystaniem algorytmów numerycznej optymalizacji. Warto jednak pamiętać, że czas uzyskania rozwiązania może znacznie wydłużyć się.

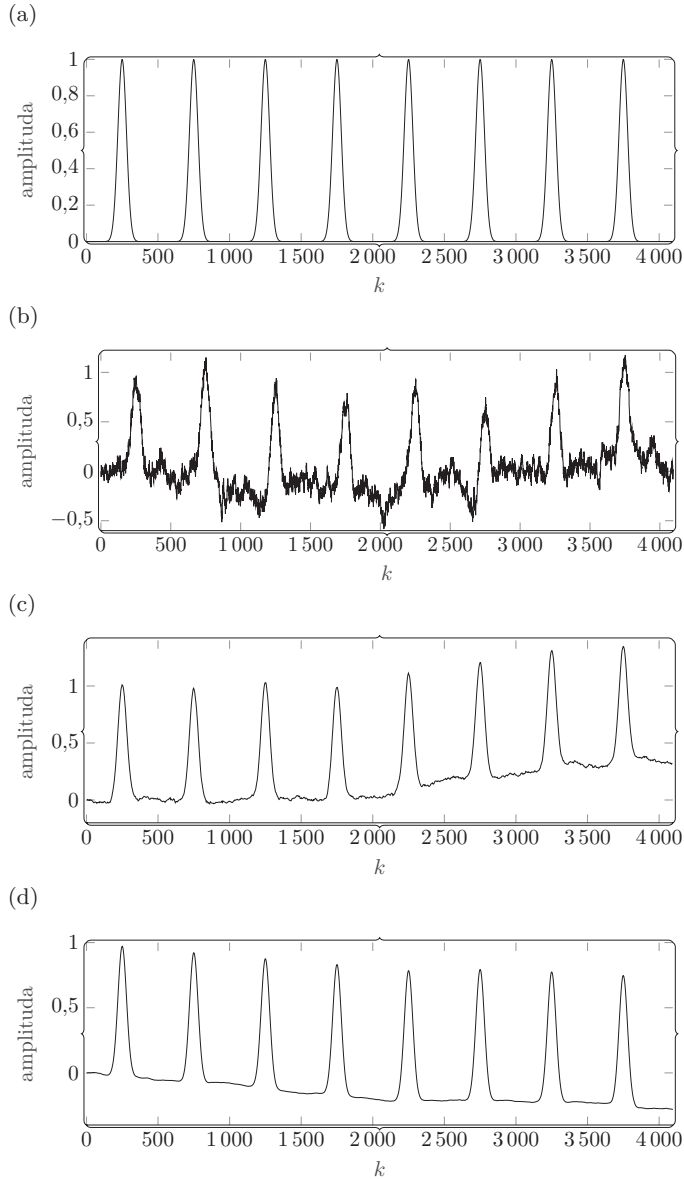
15.4. Analiza algorytmu

W celu oceny efektywności podejścia bazującego na algorytmie TARA przeprowadzone zostały zarówno badania z sygnałami syntetycznymi, jak i rzeczywistymi. W tym rozdziale zajmujemy się sygnałami syntetycznymi. Ponieważ analizowane sygnały prędkości chodu są sygnałami, w których można również wyróżnić obszary o małych i dużych wartościach przyspieszeń (patrz rys. 14.1), do badań wybraliśmy sygnał typu *pulse* (rys. 15.3a).

Oryginalny algorytm TARA został opracowany dla sygnałów zakłóconych *gaussowskim* szumem. Niestety, sygnały takie jak analizowane sygnały prędkości chodu są obciążone szumami *niegaussowskimi*. Dlatego zasadne jest zbadanie efektywności podejścia bazującego na metodzie TARA w przypadku sygnałów z szumami *niegaussowskimi*.

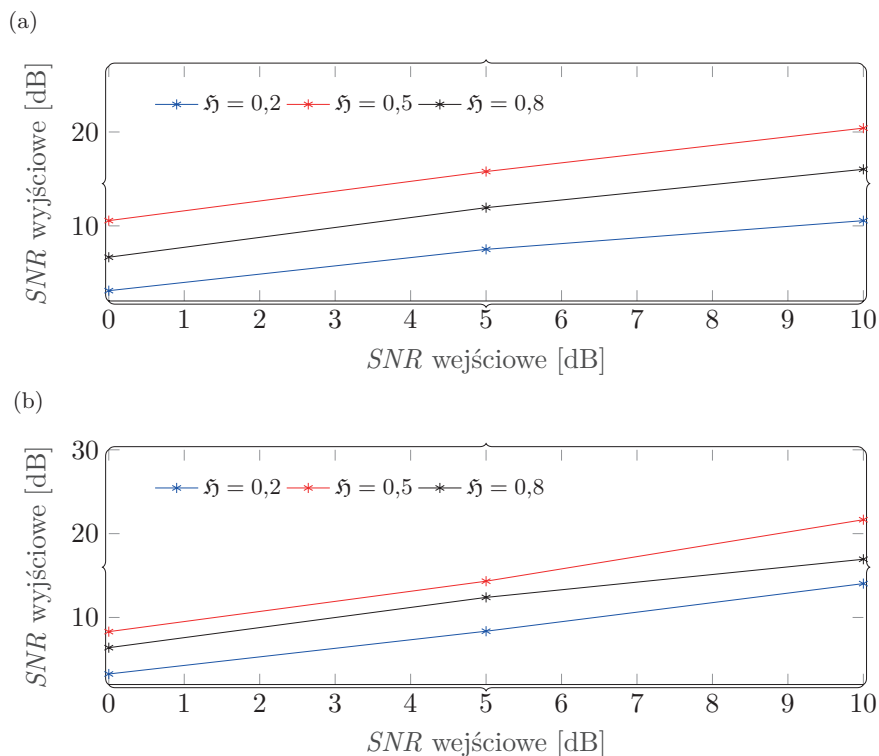
W badaniach wykorzystano sygnał z rys. 15.3a, który został zakłócony szumem niskoczęstotliwościowym o różnej wartości *SNR*. Niskoczęstotliwościowy szum wykorzystany w badaniach opisaliśmy z wykorzystaniem wykładnika Hursta (ułankowy szum *gaussowski*). W badaniach wykorzystaliśmy szumy z trzema jego wartościami, tj. $\mathfrak{H} = 0,2$, $\mathfrak{H} = 0,5$, i $\mathfrak{H} = 0,8$ oraz o trzech różnych poziomach współczynnika *SNR*, tj. $SNR = 0$ dB, $SNR = 5$ dB, $SNR = 10$ dB). Przykład sygnału dla $\mathfrak{H} = 0,2$, $\mathfrak{H} = 0,5$, $\mathfrak{H} = 0,8$ oraz $SNR = 5$ dB został przedstawiony na rys. 15.3.

Wynik zastosowania algorytmu TARA dla sygnałów z rys. 15.3 przedstawiono na rysunkach 15.4a i b. *SNR* wejściowy jest związany z poziomem szumów w sygnale wejściowym, natomiast *SNR* wyjściowy związany jest z sygnałem wyjściowym, po usunięciu z niego niskoczęstotliwościowych szumów. Wyniki dla algorytmu TARA/ ℓ_1 przedstawiono na rys. 15.4a, natomiast dla metody TARA/ATAN przedstawione zostały na rys. 15.4a. Najlepsze wyniki uzyskano w przypadku, gdy sygnał był zakłócony szumem charakteryzowanym przez współczynnik Hursta $\mathfrak{H} = 0,5$. W tym przypadku szum ma właściwości szumu *gaussowskiego*. W przypadku, gdy szum nie miał właściwości szumu *gaussowskiego*, efektywność podejścia bazującego na algorytmie TARA spadła. Co nie jest zaskoczeniem, ponieważ algorytm TARA został zaprojektowany dla sygnałów zakłóconych szumem *gaussowskim*. Warto jednak podkreślić, że spadek efektywności podejścia bazującego na algorytmie



Rys. 15.3. Wykorzystany w badaniach sygnał syntetyczny: a) sygnał niezakłócony, b) sygnał zakłócony ($SNR = 5$ dB, $\xi = 0,2$), c) sygnał zakłócony ($SNR = 5$ dB, $\xi = 0,5$), d) sygnał zakłócony ($SNR = 5$ dB, $\xi = 0,8$)

15. Odszumianie bazujące na reprezentacji rzadkiej



Rys. 15.4. Jakość filtracji algorytmu TARA zależność między SNR wejściowym a SNR wyjściowym: a) TARA/ ℓ_1 , b) TARA/ATAN

TARA nie jest duży. Dlatego warto sprawdzić jego efektywność w przypadku sygnałów, które zakłócone są szumem o *niegaussowskiej* naturze.

15.5. Badania eksperymentalne

W badaniach eksperymentalnych dla rzeczywistych danych pomiarowych wykorzystano zbiór (oznaczony jako A), który został opisany w podrozdziale 14.3, natomiast do oceny jakości działania metod wykorzystano miary przedstawione w podrozdziale 3.6.

15.6. Wyniki

Ostatnią część tego rozdziału poświęcimy ocenie jakości usuwania niskoczęstotliwościowych zakłóceń z estymowanych wartości prędkości chodu (rys. 15.1). Efektywność zaproponowanego podejścia została zbadana dla dwóch wersji algorytmu TARA:

- TARA/ ℓ_1 z ℓ_1 jako funkcja kary,
- TARA/ATAN, tj. ATAN jako funkcja kary.

W celu oceny jakości opracowanego podejścia wykorzystano zbiór danych opracowany na bazie metody ZUPT. W badaniach metoda ZUPT była metodą referencyjną.

Wyniki badań dla przygotowanych zbiorów danych (patrz podrozdz. 15.5) przedstawiono w tabeli 15.1. Wartości wskaźników *SNR* i *RMSE* zostały wyznaczone dla każdej z osi, tj. x , y , z . Wyniki wskazują na to, że zaproponowane podejście wykorzystujące algorytm TARA umożliwia efektywnie usunąć zakłócenia o niskich częstotliwościach.

W tabeli 15.1 zamieszczono wyniki uzyskane dla trzech podzbiorów ze zbioru oznaczonego jako A. Podzbiory te oznaczono jako A1, A2 oraz A3. Z analizy wynika, że podejście wykorzystujące algorytm TARA/ATAN ułatwia uzyskanie lepszych wyników niż podejście bazujące na metodzie TARA/ ℓ_1 . Ponadto, w większości przypadków wyniki dla osi z są znacznie gorsze zarówno w pierwszym, jak i drugim algorytmie. Jest to związane z mniejszym *SNR* dla sygnałów pomiarowych mierzonych dla tej osi.

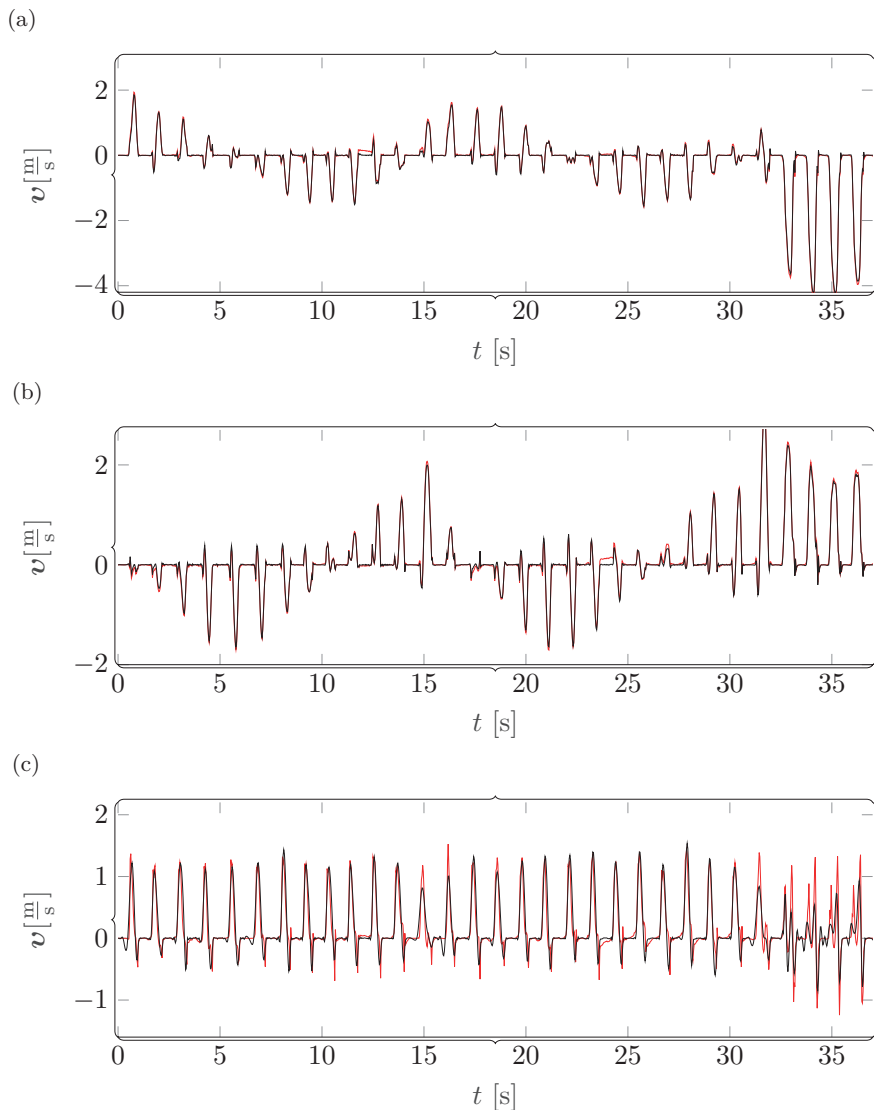
Oceniając zbiór danych, w którym zarejestrowano badanie z chodem po linii prostej, możemy zauważyć, że uzyskane wyniki są znacznie lepsze niż dla innych

Tabela 15.1. Porównanie jakości estymacji dryftu dla zbioru A

| Podzbiór | Miara | TARA/ ℓ_1 | | | TARA/ATAN | | |
|----------|-------------|----------------|-------|-------|-----------|-------|-------|
| | | x | y | z | x | y | z |
| A1 | <i>RMSE</i> | 0,03 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 |
| | <i>SNR</i> | 13,51 | 11,19 | 17,79 | 21,79 | 19,00 | 19,81 |
| A2 | <i>RMSE</i> | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 |
| | <i>SNR</i> | 21,17 | 17,96 | 15,61 | 34,27 | 24,59 | 15,17 |
| A3 | <i>RMSE</i> | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 |
| | <i>SNR</i> | 19,69 | 20,48 | 16,88 | 23,38 | 26,38 | 16,83 |

15. Odszumianie bazujące na reprezentacji rzadkiej

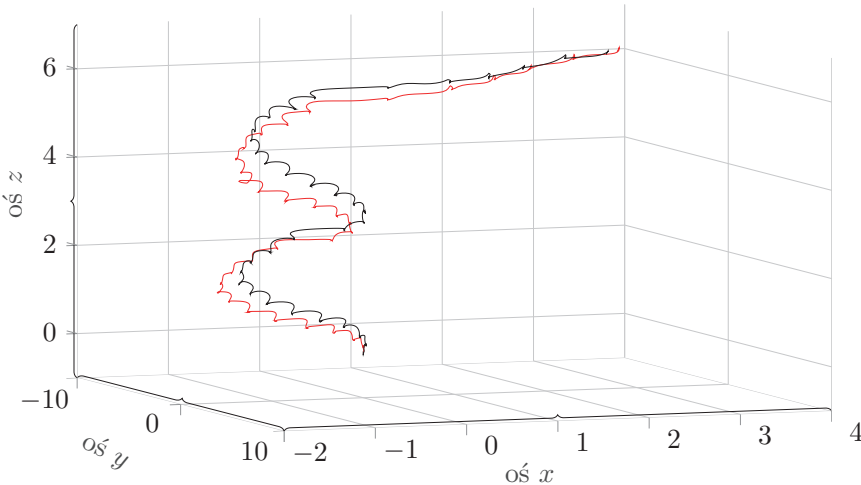
podzbiorów (patrz A1 w tab. 15.1). Uwaga jest zbieżna z tym, na co w swoich pracach zwracają inni badacze [472].



Rys. 15.5. Zbiór A1: Estymowana prędkość z wykorzystaniem metody ZUPT (linia czerwona) i metody bazującej na algorytmie TARA/ATAN (linia czarna) dla osi: a) x , b) y , c) z

Wyniki dla tego podzbioru zilustrowano na rys. 15.5. Wykresy wskazują na to, że niskoczęstotliwościowe zakłócenia w sygnałach pomiarowych zostały znacznie ograniczone.

Ponieważ opracowane rozwiązanie ma również wspomagać nawigację osób, istotne jest sprawdzenie, jaka jest jakość estymacji pozycji na podstawie estymowanej wartości prędkości chodu. Na rysunku 15.6 przeprowadziliśmy dodatkową analizę dla podzbioru A1. Wykorzystując estymowaną prędkość chodu, wyznaczyliśmy trajektorię poruszającej się osoby dla osi x, y, z (patrz rys. 15.6). Obserwując wyznaczoną trajektorię, można zauważyć, że pojawiają się pewne różnice pomiędzy trajektorią wyznaczoną dla zaproponowanego podejścia a tą uzyskaną z wykorzystaniem metody ZUPT. Można jednak również zauważyć, że na początku i na końcu trajektorii wyniki są bardzo do siebie zbliżone. Na rysunku 15.6 przedstawiliśmy te same wyniki dla podzbioru A1, lecz jako trajektorię w przestrzeni.



Rys. 15.6. Zbiór A1: Estymowana trajektoria w \mathbb{R}^3 z wykorzystaniem metody ZUPT (linia czerwona) i z wykorzystaniem metody bazującej na algorytmie TARA/ATAN (linia czarna)

Przeprowadzone badania eksperymentalne wskazują, że wykorzystanie algorytmu TARA w systemach do estymacji prędkości oraz pozycji (tzw. systemy nawigacji) umożliwia uzyskanie bardzo dobrych wyników bez wykorzystywania dodatkowych elementów infrastruktury.

15.7. Uwagi końcowe

W rozdziale przedstawione zostało podejście do estymacji niskoczęstotliwościowych zakłóceń z sygnału prędkości chodu człowieka. Zaproponowane podejście bazuje na algorytmie TARA. TARA jest przykładem metody wykorzystującej rzadką reprezentację sygnałów.

Większość metod, które wykorzystywane są w tym celu bazują na dodatkowej infrastrukturze (np. systemy GPS) lub wykorzystujące fakt, że część ruchów ludzkich, np. chód, ma naturę cykliczną. W tym drugim przypadku popularnym przykładem jest metoda ZUPT. Zaproponowane podejście jest alternatywą dla tych, które zostały wymienione. W tym przypadku do usuwania niskoczęstotliwościowych zakłóceń z estymowanego prędkości chodu wykorzystano jedynie algorytmy filtracji.

Zaproponowane podejście bazuje na algorytmie TARA – metodzie wykorzystującej rzadką reprezentację sygnałów. Opracowana metodyka została zbadana zarówno dla sygnałów syntetycznych, jak i rzeczywistych. Strata efektywności w usuwaniu niskoczęstotliwościowych zakłóceń z sygnału syntetycznego wskazała w przypadku *niegaussowskich* zakłóceń, że możliwe jest uzyskanie dobrych wyników w przypadku, gdy zakłócenia będą *niegaussowskie*. Było to istotne spostrzeżenie, ponieważ zakłócenia w estymowanej prędkości chodu mają naturę *niegaussowską*. W badaniach wykorzystano dwa typy zakłóceń *niegaussowskich*, które zostały scharakteryzowane wykładnikiem Hursta. Jeden z szumów był scharakteryzowany wykładnikiem Hursta $\mathfrak{H} = 0,2$, a drugi $\mathfrak{H} = 0,8$. Wykładnik Hursta z przedziału $0,5 < \mathfrak{H} \leq 1$ oznacza szum ze wzmocnionym trendem, czyli składowymi o niskiej częstotliwości. Taki szum charakteryzuje się pamięcią długoterminową i ma tendencję do powielania bieżącego trendu. Wzmocnienie trendu w szumach tego typu jest tym większe im wartość współczynnika Hursta zbliża się do 1.

W przypadku szumów, które charakteryzowane są współczynnikiem Hursta w przedziale $0,0 \leq \mathfrak{H} < 0,5$, istnieje z kolei silna tendencja do odwracania trendu, czyli do powracania do średniej. Oznacza to, że jeżeli w danej chwili trend wykazuje tendencję do wzrostu, to w chwilach następnych będzie miał tendencję do powrotu do średniej. W szumach tego typu jest to spowodowane długoterminowymi korelacjami ujemnymi. Siła do odwracania trendu jest tym większa, im wartość współczynnika jest bliższa 0.

W badaniach wykorzystane zostały również sygnały rzeczywiste, które były estymowanymi prędkościami chodu. Do estymacji prędkości chodu wykorzystano pomiary z czujników przyspieszenia, żyroskopów oraz magnetometrów. W ogólności estymacja niskoczęstotliwościowych zakłóceń z tych sygnałów pomiarowych nie jest zadaniem trywialnym. A zastosowanie technik filtracji w tym sygnale nie było

do tej pory wykorzystywane oprócz dwóch prac, które zostały wspomniane na początku tego rozdziału.

Przyszłe prace związane z przedstawionym podejściem będą związane z badaniami nad innymi typami aktywności ruchowej człowieka (np. bieganie, skakanie czy pływanie) i mającą inną niż cykliczną naturę. Wyniki tych badań wskazują na znalezienie odpowiedzi, czy możliwe jest zaprojektowanie systemu do nawigacji osobistej, który będzie niezależny od dodatkowej infrastruktury czy specyficznej aktywności człowieka.

Drugi możliwy kierunek badań związany jest z opracowaniem algorytmu bazującego na metodzie TARA, który będzie mniej wrażliwy na typ zakłóceń. Co oznacza, że jego efektywność nie będzie spadała lub będzie spadała w niewielkim stopniu w przypadku *niegaussowskich* zakłóceń pomiarowych.

16. Odszumianie bazujące na reprezentacji czasowo-częstotliwościowej

16.1. Wstęp

W rozdziałach 14 oraz 15 przedstawione zostały zagadnienia związane z drugim z wymienionych zagadnień. W tym rozdziale przedstawiony zostanie algorytm do usuwania składowych o wysokich częstotliwościach z sygnału pomiarowego. Testowym sygnałem pomiarowym będzie sygnał z żyroskopu.

W ogólności metody do usuwania składowych o wysokich częstotliwościach z sygnałów pomiarowych można podzielić na metody liniowe i nieliniowe. Ponieważ większość rzeczywistych sygnałów pomiarowych to sygnały, które mają naturę sygnałów niestacjonarnych, liniowe metody zazwyczaj prowadzą do niesatysfakcjonujących wyników. Problemem jest również to, że trudno wskazać zakres częstotliwości dla takiego sygnału (patrz podrozdz. 5.8).

Popularnymi metodami wykorzystywanymi w analizie sygnałów niestacjonarnych są krótkookresowa transformacja Fouriera albo falkowa. Trudności związane z zastosowaniem transformacji falkowej w usuwaniu wysokoczęstotliwościowych składowych związane są m.in. z ustaleniem funkcji bazowej oraz poziomu dekompozycji. Alternatywą dla tych metod jest podejście bazujące na algorytmie EMD [169]. W metodzie analizowany sygnał podlega dekompozycji na wiele funkcji składowych IMF (mod). W przeciwieństwie do transformacji falkowej, algorytm oparty na metodzie EMD nie wymaga ustalenia postaci funkcji bazowej. Jest to przykład podejścia całkowicie bazującego na danych (ang. *data-driven*). Wadą tej metody jest natomiast to, że nie ma dla niej opracowanej jeszcze teorii ani ścisłego opisu matematycznego. Utrudnia to formalną analizę algorytmu.

Podejście to natomiast znalazło szerokie zastosowanie w zagadnieniach praktycznych, takich jak analiza sygnałów biomedycznych [44, 408, 437], akustycznych [68, 202, 203] i sygnałów z czujników inercyjnych, takich jak akcelerometry, żyroskopy czy magnetometry [93, 147, 335].

W podrozdziale 5.8.2 wskazaliśmy, że są dwa główne podejścia do konstrukcji algorytmów filtracji bazujących na metodzie EMD. Pierwsze podejście wykorzystuje estymację sygnału biorąc pod uwagę tylko części z mod, które wyznaczono

16. Odszumianie bazujące na reprezentacji czasowo-częstotliwościowej

w wyniku dekompozycji sygnału wejściowego. Drugie podejście bazuje na wykorzystaniu operatora eliminacji progowej (por. podrozdz. 5.8.1). Druga technika jest inspirowana metodą oryginalnie opracowaną dla transformacji falkowej [217].

W podrozdziale przedstawione zostanie podejście do konstrukcji algorytmu do usuwania wysokoczęstotliwościowych niepożądanych składowych z sygnałów pomiarowych z wykorzystaniem algorytmu EMD, rzadkiej reprezentacji oraz technik rzadkiej optymalizacji. Zaproponowane podejście jest rozszerzeniem metody bazującej na operatorze eliminacji progowej.

Wykorzystanie reprezentacji rzadkiej jest bardzo popularnym obszarem badań [195, 336, 362, 410]. Jak wskazano m.in. w pracach [195, 311, 362], jest to podejście, które może być również wykorzystane do usuwania z sygnałów niepożądanych składowych o wysokich częstotliwościach.

Reprezentacja rzadka może zostać wykorzystana również w analizie sygnałów bazującej na transformacji falkowej [109]. Podejście to zostało również omówione w pracy [195]. Charakteryzuje się ono tym, że jest odporne na pseudoefekt Gibbsa [109].

W pracy [195] transformacja falkowa została wykorzystana wraz z techniką nazywaną TV (ang. *Total Variation*) do usuwania wysokoczęstotliwościowych składowych z sygnałów pomiarowych. Połączenie tych dwóch technik, na co wskazano w pracy, uzyskało bardzo dobre wyniki, przewyższające te, które uzyskano dla metod referencyjnych.

Wyniki przedstawione w tym rozdziale bazują na pracy [51].

16.2. Metodyka

16.2.1. Algorytm EMD

W podrozdziale 5.7.8 omówiliśmy metodę EMD. Algorytm ten służy do dekompozycji sygnału na skończony zbiór sygnałów o prostszych przebiegach. Podejście to ma cechy, które ułatwiają analizę sygnałów niestacjonarnych [169]. Cecha ta została wykorzystana w zaproponowanym podejściu.

16.2.2. Usuwanie zakłóceń z wykorzystaniem metody EMD

Przyjmijmy następujący model sygnału pomiarowego

$$y(k) = s(k) + v(k), \quad k = 1, 2, \dots, K \quad (16.1)$$

gdzie $y(k)$ jest sygnałem pomiarowym, $v(k)$ reprezentuje zakłócenie w sygnale, $s(k)$ jest nieznanym sygnałem, a K to liczba punktów pomiarowych. Zadaniem jest estymacja nieznanego sygnału $s(k)$ na podstawie sygnału pomiarowego $y(k)$.

16.3. Algorytm usuwania zakłóceń z wykorzystaniem metody EMD

16.3.1. Sformułowanie problemu

Przyjmijmy, że model sygnału pomiarowego ma postać (16.1). Sygnały reprezentujemy jako wektory, tj. $\mathbf{y} \in \mathbb{R}^K$, $\mathbf{s} \in \mathbb{R}^K$, $\mathbf{v} \in \mathbb{R}^K$.

Zadanie usuwania niepożądanych składowych o wysokich częstotliwościach polega na tym, że dla sygnału pomiarowego \mathbf{y} należy opracować algorytm usuwający zakłócenia \mathbf{v} , który możliwie najlepiej estymuje sygnał \mathbf{s} . W tym celu należy minimalizować różnicę $\boldsymbol{\epsilon} = \hat{\mathbf{s}} - \mathbf{s}$. Przy czym wielkość $\hat{\mathbf{s}}$ jest oszacowaniem sygnału \mathbf{s} .

16.3.2. Algorytm TVD

Jedną z najczęściej wykorzystywanych funkcji regularyzujących jest TV. Przedstawmy metodę estymacji sygnału \mathbf{s} na bazie sygnałów IMF $H_{N,K}$ wyznaczonych z sygnału pomiarowego \mathbf{y} .

Problem optymalizacji z normą TV [156] ma postać:

$$\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}} = \arg \min_{\mathbf{H}_{\text{imf}N_{\text{imf}}}} \left\{ \frac{1}{2} \|\mathcal{E}(\mathbf{Y}_K) - \mathbf{H}_{\text{imf}N_{\text{imf}}}\|_2^2 + \lambda_0 \left\| D\mathcal{E}^{-1}(\mathbf{H}_{\text{imf}N_{\text{imf}}}) \right\|_1 \right\} \quad (16.2)$$

gdzie λ_0 parametrem związanym z funkcją kary, D jest macierzą postaci (patrz (5.96)), a $\mathbf{Y}_K = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$.

Człon regularyzujący w zależności (16.2), tj. $\left\| D\mathcal{E}^{-1}(\mathbf{H}_{\text{imf}N_{\text{imf}}}) \right\|_1$ służy do ustalania liczby bloków w sygnale [334]. Rozwiązaniem zadania (16.2) jest [362]

$$\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}} = \text{tvd}(\mathbf{H}_{\text{imf}N_{\text{imf}}}, \lambda_0) \quad (16.3)$$

gdzie tvd jest algorytmem TVD (ang. *Total Variation Deonoising*). Należy podkreślić, że nie ma bezpośredniego algorytmu do rozwiązania zadania TVD. W tym celu można natomiast wykorzystać rozwiązanie pośrednie zadania (16.2), przedstawione w pracy [85].

Sygnał \mathbf{s} może zostać oszacowany na podstawie zależności (5.85) na podstawie wyznaczonych sygnałów $\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}}$. Ponieważ zaproponowany algorytm wykorzystany do konstrukcji metody usuwania zakłóceń z sygnału bazuje na algorytmie EMD oraz metodzie TVD, dlatego nową propozycję nazwiemy EMD-TVD. Algorytm przedstawiliśmy również w formie pseudokodu (algorytm 6).

Algorytm 6: EMD-TVD

```

Wejście : sygnał pomiarowy  $\mathbf{Y}_K = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$ , parametry:  $\lambda_0$ 
/* Dekompozycja sygnału z wykorzystaniem EMD */
 $\mathbf{H}_{\text{imf}N_{\text{imf}}} = \mathcal{E}(\mathbf{Y}_K)$ 
/* Algorytm TVD */
for  $i = 1$  to  $N_{\text{imf}}$  do
|  $\tilde{\mathbf{H}}_{\text{imf}i} = \text{tvd}(\mathbf{H}_{\text{imf}i}, \lambda_0)$ 
end
/* Odtworzenie sygnału z wykorzystaniem EMD */
 $\hat{\mathbf{s}}_K = \mathcal{E}^{-1}(\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}})$ 
Wyjście : sygnał bez zakłóceń  $\hat{\mathbf{s}}_K$ 

```

16.3.3. Algorytm FLSA

Podójście przedstawione w poprzednim podrozdziale, które bazuje na algorytmie TVD, może zostać rozszerzone przez dodanie dodatkowego wyrażenia w części z funkcją kary. Nowy problem optymalizacyjny ma postać (16.2):

$$\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}} = \arg \min_H \left\{ \frac{1}{2} \|\mathcal{E}(\mathbf{Y}_K) - \mathbf{H}_{\text{imf}N_{\text{imf}}}\|_2^2 + \lambda_1 \|\mathbf{H}_{\text{imf}N_{\text{imf}}}\|_1 + \lambda_0 \left\| D\mathcal{E}^{-1}(\mathbf{H}_{\text{imf}N_{\text{imf}}}) \right\|_1 \right\} \quad (16.4)$$

gdzie λ_0 oraz λ_1 są parametrami związanymi z funkcjami kary. Dodatkowy człon regularyzujący w zależności (16.4), tj. $\|\mathbf{H}_{\text{imf}N_{\text{imf}}}\|_1$ służy do ustalania liczby stopnia rzadkości sygnału przez dobór liczby zer w sygnale [334].

Rozwiązanie dla (16.4) [130] można znaleźć, stosując podejście F-LASSO (ang. *fused-lasso*) [404]

$$\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}} = \text{soft}(\text{tvd}(\mathbf{H}_{\text{imf}N_{\text{imf}}}, \lambda_0), \lambda_1) \quad (16.5)$$

gdzie *soft* jest operatorem eliminacji miękkiej. W przypadku sygnałów zdekomponowanych algorytmem EMD do wykonania operacji miękkiego progowania należy wykorzystać zależność (5.89). Algorytm opisany tą zależnością nazywamy algorytmem FLSA (ang. *Fused Lasso Signal Approximator*).

Algorytm do usuwania zakłóceń z sygnału bazujący na (16.5) nazwiemy EMD-FLS. Metoda w postaci pseudokodu została przedstawiona jako algorytm 7. Do dekompozycji sygnału EMD w prezentowanym rozwiązaniu wykorzystano metodę EMD.

Algorytm 7: EMD-FLS

```

Wejście : sygnał pomiarowy  $\mathbf{Y}_K = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K)$ , parametry:  $\lambda_0, \lambda_1$ 
/* Dekompozycja sygnału z wykorzystaniem EMD */
 $\mathbf{H}_{\text{imf}N_{\text{imf}}} = \mathcal{E}(\mathbf{Y}_K)$ 
/* Algorytm FLSA */
for  $i = 1$  to  $N_{\text{imf}}$  do
|  $\tilde{\mathbf{H}}_{\text{imf}i} = \text{soft}(\text{tvd}(\mathbf{H}_{\text{imf}i}, \lambda_0), \lambda_1)$ 
end
/* Odtworzenie sygnału z wykorzystaniem EMD */
 $\hat{\mathbf{s}}_K = \mathcal{E}^{-1}(\tilde{\mathbf{H}}_{\text{imf}N_{\text{imf}}})$ 
Wyjście : sygnał bez zakłóceń  $\hat{\mathbf{s}}_K$ 

```

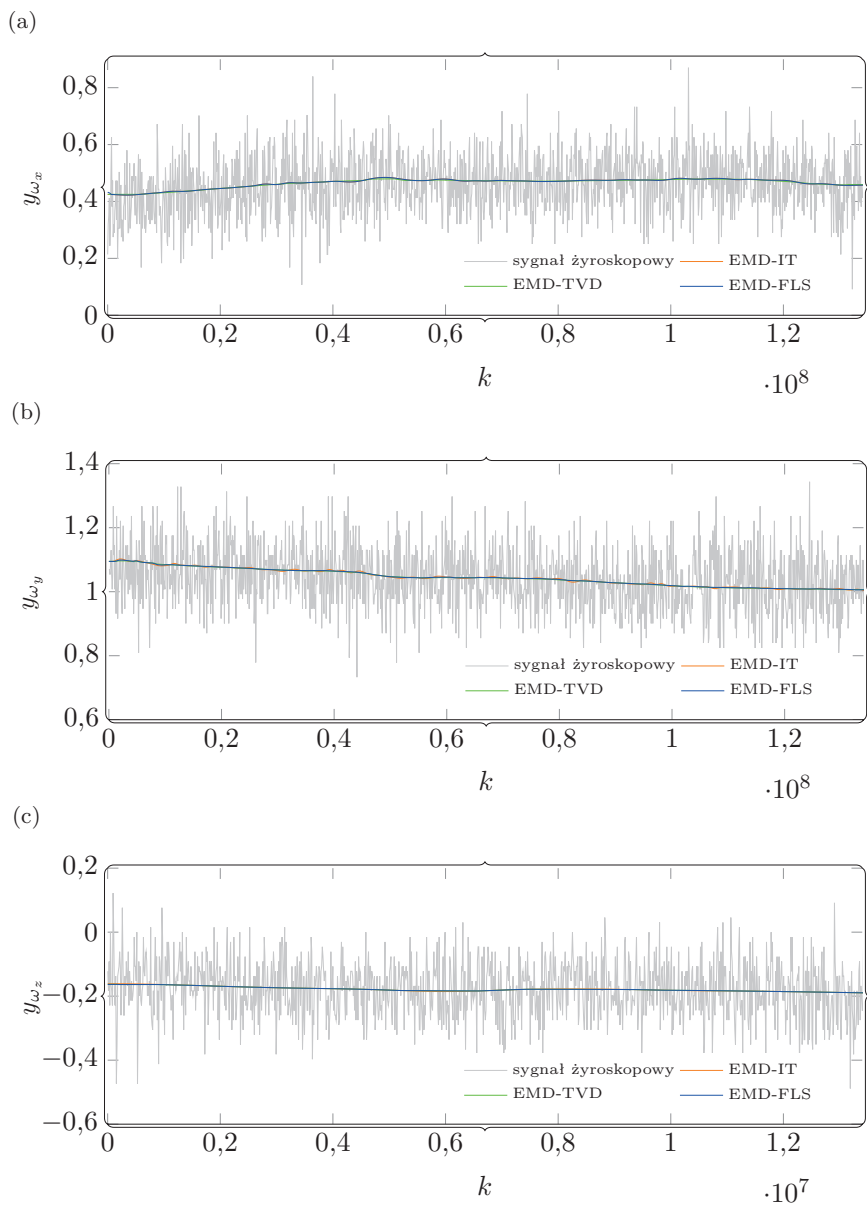
16.4. Badania eksperymentalne

W badaniach numerycznych wykorzystamy miarę SNR (3.9). Zaproponowane algorytmy zostały przetestowane dla zadania usuwania wysokoczęstotliwościowych składników z sygnałów pomiarowych. Pierwszym z badanych algorytmów jest algorytm EMD-TVD (patrz podrozdz. 16.3.2), drugim EMD-FLS (podrozdz. 16.3.3). Zaproponowane algorytmy bazują na szybkiej implementacji algorytmu EMD [122]. Zaproponowane metody zostały porównane z klasycznym algorytmem wykorzystującym algorytm EMD, tj. EMD-IT (ang. *interval thresholding algorithm*) zaproponowanym w pracy [217].

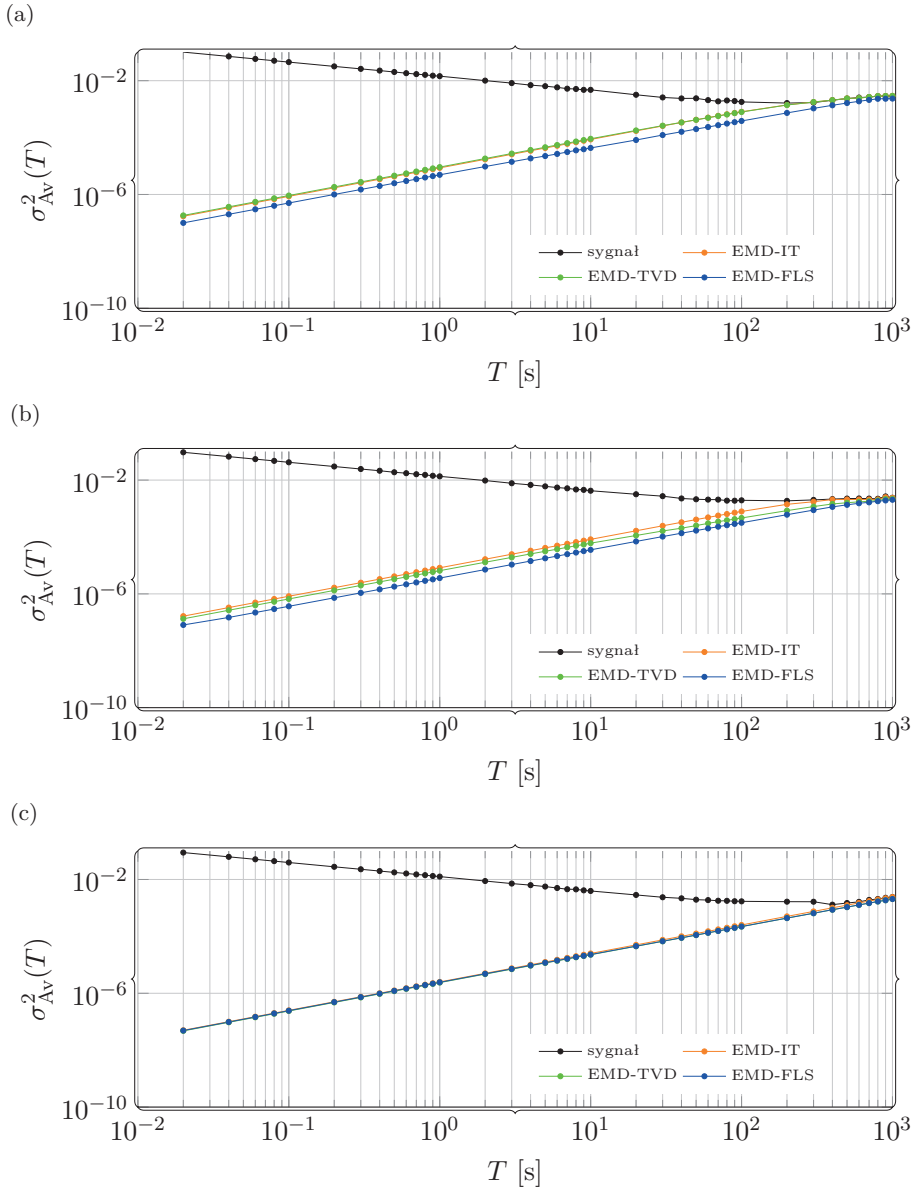
Charakterystyka sygnałów wykorzystanych w badaniach

W badaniach nad efektywnością zaproponowanych algorytmów usuwania niepożądanych składowych z pomiarów wykorzystamy dane z czujnika żyroskopowego. Żyroskopy wykonane w technologii MEMS mają wady, które wpływają na jakość pomiarów. Do najważniejszych wad tych urządzeń należy wymienić podatność na wpływ temperatury, wibracji oraz ciśnienia [455] (por. podrozdz. 4.1). Zjawiska te przejawiają się tym, że w pomiarach z tych urządzeń pojawia się zakłócenie niskoczęstotliwościowe (dryft) oraz wysokoczęstotliwościowe. Zakłócenia te mają zarówno naturę szumu białego, jak i szumu fraktalnego [40]. Istotnym pytaniem badawczym jest to, w jaki sposób usunąć z pomiarów niepożądane składowe zarówno te o niskich, jak i wysokich częstotliwościach. Jest to o tyle istotna kwestia, że usunięcie tych niepożądanych składowych wpływa na jakość rozwiązań bazujących na pomiarach żyroskopowych. Przykładem są systemy nawigacji ludzi niewidomych itp.

16. Odszumianie bazujące na reprezentacji czasowo-częstotliwościowej



Rys. 16.1. Sygnał z żyroskopu po usunięciu wysokoczęstotliwościowych zakłóceń dla osi:
a) x , b) y , c) z żyroskopu



Rys. 16.2. Wariancja Allana sygnału oryginalnego oraz sygnału oczyszczonego dla osi: a) x , b) y , c) z żyroskopu

16. Odszumianie bazujące na reprezentacji czasowo-częstotliwościowej

Badania wykonano dla unieruchomionego czujnika w temperaturze otoczenia równej 22 °C, z częstotliwością próbkowania równą 50 Hz. W badaniach wykorzystano czujnik Shimmer3. Pomiary prędkości kątowych dla żyroskopu wykonano w trzech osiach x , y oraz z .

Wyniki usuwania niepożądanych składowych z danych pomiarowych zostały przeanalizowane z wykorzystaniem wariancji Allana (patrz rozdz. 4).

Wyniki badań numerycznych

Na rysunku 16.1a–c przedstawiono efekty usunięcia niepożądanych składowych z pomiarów żyroskopowych. W kolorze szarym zaznaczono oryginalny sygnał z czujnika. Wyniki przedstawiono dla metody EMD-IT oraz zaproponowanych algorytmów EMD-TVD oraz EMD-FLS. W celach porównawczych wykorzystano również wyniki uzyskane dla algorytmu bazującego na falkach.

Na rysunku 16.2a–c przedstawiono wyniki analizy ilościowej uzyskanych wyników. W analizie wykorzystano wariancję Allana. Zaproponowane algorytmy, tj. EMD-TVD oraz EMD-FLS dały lepsze wyniki dla osi pomiarowych x oraz y , w porównaniu do algorytmu EMD-IT. Natomiast wyniki uzyskane dla osi z algorytmów EMD-TVD oraz EMD-FLS są porównywalne z tymi, które uzyskano dla metody EMD-IT.

Bibliografia

- [1] Abdelgawad A., Bayoumi M., *Data fusion in WSN*, [w:] *Resource-aware data fusion algorithms for wireless sensor networks*, Springer, 2012, 17–35.
- [2] Abry P., Veitch D., *Wavelet analysis of long-range-dependent traffic*, *IEEE Transactions on Information Theory*, 44, 1, 1998, 2–15.
- [3] Acharya U.R., Fujita H., Sudarshan V.K., Bhat S., Koh J.E., *Application of entropies for automated diagnosis of epilepsy using EEG signals: A review*, *Knowledge-Based Systems*, 88, 2015, 85–96.
- [4] Aggarwal C.C., *Data classification: algorithms and applications*, CRC Press, Boca Raton 2014.
- [5] Aggarwal J.K., Cai Q., *Human motion analysis: A review*, [w:] *Proceedings of Nonrigid and Articulated Motion Workshop*, IEEE, 1997, 90–102.
- [6] Aghabozorgi S., Shirkhorshidi A.S., Wah T.Y., *Time-series clustering – A decade review*, *Information Systems*, 53, 2015, 16–38.
- [7] Ahmadi A., Mitchell E., Destelle F., Gowing M., O'Connor N.E., Richter C., Moran K., *Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors*, [w:] *11th International Conference on Wearable and Implantable Body Sensor Networks*, IEEE, 2014, 98–103.
- [8] Ahmadi A., Mitchell E., Richter C., Destelle F., Gowing M., O'Connor N.E., Moran K., *Toward automatic activity classification and movement assessment during a sports training session*, *IEEE Internet of Things Journal*, 2, 1, 2015, 23–32.
- [9] Ali Z., Talha M., *Innovative method for unsupervised voice activity detection and classification of audio segments*, *IEEE Access*, 6, 2018, 15494–15504.
- [10] Allan D.W., *Should the classical variance be used as a basic measure in standards metrology?*, *IEEE Transactions on Instrumentation and Measurement*, 1001, 2, 1987, 646–654.
- [11] Allan D.W., *Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators*, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 34, 6, 1987, 647–654.
- [12] Allan D.W., Levine J., *A historical perspective on the development of the Allan variances and their strengths and weaknesses*, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 63, 4, 2016, 513–519.
- [13] Amsüss S., Goebel P.M., Jiang N., Graimann B., Paredes L., Farina D., *Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control*, *IEEE Transactions on Biomedical Engineering*, 61, 4, 2014, 1167–1176.
- [14] Angermann M., Robertson P., Kemptner T., Khider M., *A high precision reference data set for pedestrian navigation using foot-mounted inertial sensors*, [w:] *Interna-*

- tional Conference on Indoor Positioning and Indoor Navigation*, Zurich, 2010, URL http://www.dlr.de/kn/desktopdefault.aspx/tabid-8498/14560_read-36505.
- [15] Angermann M., Robertson P., Kemptner T., Khider M., *A high precision reference data set for pedestrian navigation using foot-mounted inertial sensors*, [w:] *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, 2010, 1–6.
- [16] Antczak K., *Uczenie głębokie w diagnostyce medycznej*, Symulacja w badaniach i rozwoju, 7, 3–4, 2016.
- [17] Arasu A., Babu S., Widom J., *An abstract semantics and concrete language for continuous queries over streams and relations*, Rap. tech. 2002-57, Stanford InfoLab, 2002.
- [18] Arteche J., Garcia-Enriquez J., *Singular spectrum analysis for signal extraction in stochastic volatility models*, *Econometrics and Statistics*, 1, 2017, 85–98.
- [19] Arulampalam M.S., Maskell S., Gordon N., Clapp T., *A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking*, *IEEE Transactions on Signal Processing*, 50, 2, 2002, 174–188.
- [20] Arunkumar N., Ramkumar K., Venkatraman V., Abdulhay E., Fernandes S.L., Kadry S., Segal S., *Classification of focal and non focal EEG using entropies*, *Pattern Recognition Letters*, 94, 2017, 112–117.
- [21] Arvind D., Bates A., *The speckled golfer*, [w:] *Proceedings of the ICST 3rd international conference on Body area networks*, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, 1–7.
- [22] Aydemir G.A., Saranlı A., *Characterization and calibration of mems inertial sensors for state and parameter estimation applications*, *Measurement*, 45, 5, 2012, 1210–1225.
- [23] Azami H., Hassanpour H., Escudero J., Sanei S., *An intelligent approach for variable size segmentation of non-stationary signals*, *Journal of Advanced Research*, 6, 5, 2015, 687–698.
- [24] Baek W.S., Kim D.M., Bashir F., Pyun J.Y., *Real life applicable fall detection system based on wireless body area network*, [w:] *IEEE 10th Consumer Communications and Networking Conference (CCNC)*, IEEE, 2013, 62–67.
- [25] Bajaj V., Pachori R.B., *Automatic classification of sleep stages based on the time-frequency image of EEG signals*, *Computer Methods and Programs in Biomedicine*, 112, 3, 2013, 320–328.
- [26] Barański P., Strumiłło P., *Układ nawigacji zliczeniowej do korekcji odczytów GPS w terenie miejskim*, *Przegląd Elektrotechniczny*, 85, 11, 2009, 160–163.
- [27] Barański P., Polańczyk M., Strumiłło P., *A remote guidance system for the blind*, [w:] *The 12th IEEE International Conference on e-Health Networking, Applications and Services*, IEEE, 2010, 386–390.
- [28] Bardou D., Zhang K., Ahmad S.M., *Lung sounds classification using convolutional neural networks*, *Artificial Intelligence in Medicine*, 2018.
- [29] Barunik J., Kristoufek L., *On hurst exponent estimation under heavy-tailed distributions*, *Physica A: Statistical Mechanics and its Applications*, 389, 18, 2010, 3844–3855.

- [30] Basseville M., Nikiforov I.V., et al., *Detection of abrupt changes: Theory and application*, 104, Prentice Hall, Englewood Cliffs 1993.
- [31] Basso M., Galanti M., Innocenti G., Miceli D., *Pedestrian dead reckoning based on frequency self-synchronization and body kinematics*, IEEE Sensors Journal, 17, 2, 2017, 534–545.
- [32] Beenamol M., Prabavathy S., Mohanalin J., *Wavelet based seismic signal de-noising using shannon and tsallis entropy*, Computers and Mathematics with Applications, 64, 11, 2012, 3580–3593.
- [33] Bendat J.S., Piersol A.G., *Metody analizy i pomiaru sygnałów losowych*, Wydawnictwo Naukowe PWN, Warszawa 1976.
- [34] Benedetti C., Buscemi F., Bordone P., Paris M.G., *Quantum probes for the spectral properties of a classical environment*, Physical Review A, 89, 3, 2014, 1–8.
- [35] Bengio Y., Delalleau O., *Justifying and generalizing contrastive divergence*, Neural Computation, 21, 6, 2009, 1601–1621.
- [36] Bennett T.R., Wu J., Kehtarnavaz N., Jafari R., *Inertial measurement unit-based wearable computers for assisted living applications: A signal processing perspective*, IEEE Signal Processing Magazine, 33, 2, 2016, 28–35.
- [37] Berkane S., Abdessameud A., Tayebi A., *Hybrid attitude and gyro-bias observer design on $so(3)$* , IEEE Transactions on Automatic Control, 62, 11, 2017, 6044–6050.
- [38] Berkane S., Tayebi A., *On the design of attitude observers on the special orthogonal group $so(3)$* , arXiv preprint arXiv:1606.06208, 2016.
- [39] Białasiewicz J., *Falki i aproksymacje*, Wydawnictwo Naukowo-Techniczne, Warszawa 2000.
- [40] Bielas M.S., *Stochastic and dynamic modeling of fiber gyros*, [w:] *Fiber Optic and Laser Sensors XII*, 2292, 1994, 240–254.
- [41] Billings S.A., *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*, John Wiley and Sons, Hoboken 2013.
- [42] Bishop C.M., *Pattern recognition and machine learning*, Springer, New York 2006.
- [43] Black H.D., *A passive system for determining the attitude of a satellite*, AIAA Journal, 2, 7, 1964, 1350–1351.
- [44] Blanco-Velasco M., Weng B., Barner K.E., *ECG signal denoising and baseline wander correction based on the empirical mode decomposition*, Computers in Biology and Medicine, 38, 1, 2008, 1–13.
- [45] Bleser G., Hendeby G., Miezal M., *Using egocentric vision to achieve robust inertial body tracking under magnetic disturbances*, [w:] *10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, 2011, 103–109.
- [46] Bober T., Zawadzki J., *Biomechanika układu ruchu człowieka*, Wydawnictwo BK, Wrocław 2006.
- [47] Bouvier B., Duprey S., Claudon L., Dumas R., Savescu A., *Upper limb kinematics using inertial and magnetic sensors: Comparison of sensor-to-segment calibrations*, Sensors, 15, 8, 2015, 18813–18833.
- [48] Brooks R.R., Iyengar S.S., *Multi-sensor fusion: fundamentals and applications with software*, Prentice-Hall, Inc., Englewood Cliffs 1998.
- [49] Brzostowski K., *Novel approach to human walking speed enhancement based on drift estimation*, Biomedical Signal Processing and Control, 42, 1, 2018, 18–29.

Bibliografia

- [50] Brzostowski K., *Toward the unaided estimation of human walking speed based on sparse modeling*, IEEE Transactions on Instrumentation and Measurement, 67, 6, 2018, 1389–1398.
- [51] Brzostowski K., Swiatek J., *Empirical mode decomposition based denoising algorithm for fibre optical gyroscope measurement*, [w:] *25th International Conference on Systems Engineering (ICSEng)*, IEEE, 2017, 225–230.
- [52] Brzostowski K., Szwach P., *Data fusion in ubiquitous sports training: Methodology and application*, Wireless Communications and Mobile Computing, 2018, 1, 2018, 1–14.
- [53] Brzozowska-Rup K., Dawidowicz A.L., *Metoda filtru cząsteczkowego*, Matematyka stosowana: matematyka dla społeczeństwa, 10, 51, 2009, 69–107.
- [54] Buffa C., *MEMS Lorentz Force Magnetometers: From Specifications to Product*, Springer, Cham 2017.
- [55] Błaszczyk Ł., *Podstawy teorii oszczędnego próbkowania*, praca doktorska, Politechnika Warszawska, Warszawa 2014.
- [56] Caesarendra W., Kosasih B., Tieu A.K., Moodie C.A., *Application of the largest Lyapunov exponent algorithm for feature extraction in low speed slew bearing condition monitoring*, Mechanical Systems and Signal Processing, 50, 2015, 116–138.
- [57] Callejas-Cuervo M., Gutierrez R.M., Hernandez A.I., *Joint amplitude MEMS based measurement platform for low cost and high accessibility telerehabilitation: Elbow case study*, Journal of Bodywork and Movement Therapies, 2016.
- [58] de Carvalho M., Rua A., *Real-time nowcasting the US output gap: Singular spectrum analysis at work*, International Journal of Forecasting, 33, 1, 2017, 185–198.
- [59] Castanedo F., *A review of data fusion techniques*, The Scientific World Journal, 2013.
- [60] Castellanos J.G., Lesecq S., Marchand N., Delamare J., *A low-cost air data attitude heading reference system for the tourism airplane applications*, [w:] *IEEE Sensors*, IEEE, 2005, 1388–1391.
- [61] Castells F., Laguna P., Sörnmo L., Bollmann A., Roig J.M., *Principal component analysis in ECG signal processing*, EURASIP Journal on Advances in Signal Processing, 2007, 1, 2007, 1–21.
- [62] Castiglioni P., *Letter to the editor: What is wrong in Katz’s method? Comments on: A note on fractal dimensions of biomedical waveforms*, Computers in Biology and Medicine, 40, 11–12, 2010, 950–952.
- [63] Cereatti A., Trojaniello D., Della Croce U., *Accurately measuring human movement using magneto-inertial sensors: techniques and challenges*, [w:] *IEEE International Symposium on Inertial Sensors and Systems*, 2015, 1–4.
- [64] Česić J., Joukov V., Petrović I., Kulić D., *Full body human motion estimation on lie groups using 3D marker position measurements*, [w:] *IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, IEEE, 2016, 826–833.
- [65] Chamoli A., Bansal A.R., Dimri V., *Wavelet and rescaled range approach for the Hurst coefficient for short and long time series*, Computers and Geosciences, 33, 1, 2007, 83–93.

- [66] Chang H., Xue L., Jiang C., Kraft M., Yuan W., *Combining numerous uncorrelated MEMS gyroscopes for accuracy improvement based on an optimal Kalman filter*, IEEE Transactions on Instrumentation and Measurement, 61, 11, 2012, 3084–3093.
- [67] Charte D., Charte F., García S., del Jesus M.J., Herrera F., *A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines*, Information Fusion, 44, 2018, 78–96.
- [68] Chatlani N., Soraghan J.J., *EMD-based filtering (EMDF) of low-frequency noise for speech enhancement*, IEEE Transactions on Audio, Speech, and Language Processing, 20, 4, 2012, 1158–1166.
- [69] Chen J., Li G., *Tsallis wavelet entropy and its application in power signal analysis*, Entropy, 16, 6, 2014, 3009–3025.
- [70] Chen W., Xu Y., Wang J., Zhang J., *Kinematic analysis of human gait based on wearable sensor system for gait rehabilitation*, Journal of Medical and Biological Engineering, 2016, 1–14.
- [71] Chen X., Yin C., He W., *Feature extraction of gearbox vibration signals based on EEMD and sample entropy*, [w:] *10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, IEEE, 2013, 811–815.
- [72] Chen X., Zhang J., Hamel W.R., Tan J., *An inertial-based human motion tracking system with twists and exponential maps*, [w:] *IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2014, 5665–5670.
- [73] Chen Y., Feng J., *Spatial analysis of cities using Rényi entropy and fractal parameters*, Chaos, Solitons and Fractals, 105, 2017, 279–287.
- [74] Cheng J., Chen X., Shen M., *A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals*, IEEE Journal of Biomedical and Health Informatics, 17, 1, 2013, 38–45.
- [75] Cheng J., Wei F., Li C., Liu Y., Liu A., Chen X., *Position-independent gesture recognition using sEMG signals via canonical correlation analysis*, Computers in Biology and Medicine, 103, 2018, 44–54.
- [76] Chi Y.M., Cauwenberghs G., *Wireless non-contact EEG/ECG electrodes for body sensor networks*, [w:] *International Conference on Body Sensor Networks*, IEEE, 2010, 297–301.
- [77] Chithra P., Aparna R., *Performance analysis of windowing techniques in automatic speech signal segmentation*, Indian Journal of Science and Technology, 8, 29, 2015.
- [78] Chou J.C., *Quaternion kinematic and dynamic differential equations*, IEEE Transactions on Robotics and Automation, 8, 1, 1992, 53–64.
- [79] Chu D., Liao L.Z., Ng M.K.P., Wang X., *Incremental linear discriminant analysis: A fast algorithm and comparisons*, IEEE Transactions on Neural Networks and Learning Systems, 26, 11, 2015, 2716–2735.
- [80] Chua K.C., Chandran V., Acharya U.R., Lim C.M., *Application of higher order statistics/spectra in biomedical signals—a review*, Medical Engineering and Physics, 32, 7, 2010, 679–689.
- [81] Cimolin V., Galli M., *Summary measures for clinical gait analysis: A literature review*, Gait and Posture, 39, 4, 2014, 1005–1010.

Bibliografia

- [82] Coelho A.L., Lima C.A., *Assessing fractal dimension methods as feature extractors for EMG signal classification*, Engineering Applications of Artificial Intelligence, 36, 2014, 81–98.
- [83] Cohen M.X., *Analyzing neural time series data: Theory and practice*, MIT Press, Cambridge 2014.
- [84] Colyer S.L., Evans M., Cosker D.P., Salo A.I., *A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system*, Sports Medicine – Open, 4, 1, 2018, 24.
- [85] Condat L., *A direct algorithm for 1-D total variation denoising*, IEEE Signal Processing Letters, 20, 11, 2013, 1054–1057.
- [86] Corley G., Mooney R., Quinlan L.Q., Ó Laighin G., *Application of video-based methods for competitive swimming analysis: A systematic review*, Sports and Exercise Medicine, 2015.
- [87] Cover T., Hart P., *Nearest neighbor pattern classification*, IEEE Transactions on Information Theory, 13, 1, 1967, 21–27.
- [88] Cowan R.E., Fregly B.J., Boninger M.L., Chan L., Rodgers M.M., Reinkensmeyer D.J., et al., *Recent trends in assistive technology for mobility*, Journal of Neuroengineering and Rehabilitation, 9, 1, 2012, 1–20.
- [89] Craig J.J., Knapczyk J., *Wprowadzenie do robotyki: mechanika i sterowanie*, Wydawnictwa Naukowo-Techniczne, Warszawa 1993.
- [90] Crassidis J.L., Markley F.L., Cheng Y., *Survey of nonlinear attitude estimation methods*, Journal of Guidance, Control, and Dynamics, 30, 1, 2007, 12–28.
- [91] Cronin N.J., Rantalainen T., Ahtiainen J.P., Hynynen E., Waller B., *Markerless 2D kinematic analysis of underwater running: A deep learning approach*, Journal of Biomechanics, 2019.
- [92] Cuesta-Vargas A.I., Galán-Mercant A., Williams J.M., *The use of inertial sensors system for human motion analysis*, Physical Therapy Reviews, 15, 6, 2010, 462–473.
- [93] Cui B., Chen X., *Improved hybrid filter for fiber optic gyroscope signal denoising based on EMD and forward linear prediction*, Sensors and Actuators A: Physical, 230, 2015, 150–155.
- [94] Cuong N.V., Dinh V., Ho L.S.T., *Mel-frequency cepstral coefficients for eye movement identification*, [w:] *IEEE 24th International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, 2012, 253–260.
- [95] Cutti A.G., Giovanardi A., Rocchi L., Davalli A., Sacchetti R., *Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors*, Medical and Biological Engineering and Computing, 46, 2, 2008, 169–178.
- [96] Czabanowski R., *Sensory i systemy pomiarowe*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2010.
- [97] Daponte P., De Vito L., Picariello F., Riccio M., *State of the art and future developments of the augmented reality for measurement applications*, Measurement, 57, 2014, 53–70.
- [98] Daponte P., De Vito L., Riccio M., Sementa C., *Design and validation of a motion-tracking system for ROM measurements in home rehabilitation*, Measurement, 55, 2014, 82–96.

- [99] Darooneh A.H., Naeimi G., Mehri A., Sadeghi P., *Tsallis entropy, escort probability and the incomplete information theory*, Entropy, 12, 12, 2010, 2497–2503.
- [100] Dasarathy B.V., *Sensor fusion potential exploitation-innovative architectures and illustrative applications*, Proceedings of the IEEE, 85, 1, 1997, 24–38.
- [101] Dash S.K., Rao G.S., *Arrhythmia detection using Wigner–Ville distribution based neural network*, Procedia Computer Science, 85, 2016, 806–811.
- [102] De Vries W., Veeger H., Cutti A., Baten C., Van Der Helm F., *Functionally interpretable local coordinate systems for the upper extremity using inertial and magnetic measurement systems*, Journal of Biomechanics, 43, 10, 2010, 1983–1988.
- [103] Dejnabadi H., Jolles B.M., Casanova E., Fua P., Aminian K., *Estimation and visualization of sagittal kinematics of lower limbs orientation using body-fixed sensors*, IEEE Transactions on Biomedical Engineering, 53, 7, 2006, 1385–1393.
- [104] Denavit J., Hartenberg R.S., *A kinematic notation for lower-pair mechanisms based on matrices*, Journal of Applied Mechanics, 22, 1955, 215–221.
- [105] Deng J.J., Leung C.H.C., *Dynamic time warping for music retrieval using time series modeling of musical emotions*, IEEE Transactions on Affective Computing, 6, 2, 2015, 137–151.
- [106] Desai N., Seghouane A.K., Palaniswami M., *Algorithms for two dimensional multi set canonical correlation analysis*, Pattern Recognition Letters, 111, 2018, 101–108.
- [107] Dias M., Madson L.D., Neto A.R.R., *Training soft margin support vector machines by simulated annealing: A dual approach*, Expert Systems with Applications, 2017.
- [108] Diebel J., *Representing attitude: Euler angles, unit quaternions, and rotation vectors*, Matrix, 58, 2006, 15–16.
- [109] Ding Y., Selesnick I., *Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization*, IEEE Signal Processing Letters, 22, 9, 2015, 1364–1368.
- [110] Diniz P., *Adaptive Filtering: Algorithms and Practical Implementation*, Springer, New York 2008.
- [111] Djurić Z., *Mechanisms of noise sources in microelectromechanical systems*, Microelectronics Reliability, 40, 6, 2000, 919–932.
- [112] Donoho D.L., Johnstone I.M., *Ideal spatial adaptation by wavelet shrinkage*, Biometrika, 1994, 425–455.
- [113] Doyle J.A., Evans A.C., *What colour is neural noise?*, arXiv preprint arXiv:1806.03704, 2018.
- [114] Drabik L., Stankiewicz A., Sobol E., *Słownik języka polskiego PWN*, Wydawnictwo Naukowe PWN, Warszawa 2011.
- [115] Drózdź K., *Estymacja zmiennych stanu i parametrów układu dwumasowego przy pomocy rozmytych filtrów Kalmana*, praca doktorska, Politechnika Poznańska, Poznań 2017.
- [116] Durrant-Whyte H.F., *Sensor models and multisensor integration*, The International Journal of Robotics Research, 7, 6, 1988, 97–113.
- [117] Eke A., Herman P., Kocsis L., Kozak L., *Fractal characterization of complexity in temporal physiological signals*, Physiological Measurement, 23, 1, 2002, 1–38.
- [118] El-Gohary M., McNames J., *Shoulder and elbow joint angle tracking with inertial sensors*, IEEE Transactions on Biomedical Engineering, 59, 9, 2012, 2635–2641.

Bibliografia

- [119] El-Gohary M., McNames J., *Human joint angle estimation with inertial sensors and validation with a robot arm*, IEEE Transactions on Biomedical Engineering, 62, 7, 2015, 1759–1767.
- [120] El-Sheimy N., Hou H., Niu X., *Analysis and modeling of inertial sensors using Allan variance*, IEEE Transactions on Instrumentation and Measurement, 57, 1, 2008, 140–149.
- [121] Esfandyari J., De Nuccio R., Xu G., *Introduction to MEMS gyroscopes*, URL <http://www.findmems.com/wikimems-learn/introduction-to-mems-gyroscopes> 2014.
- [122] *Fast EEMD*, URL <http://rcada.ncu.edu.tw/FEEMD.rar>.
- [123] Favre J., Aissaoui R., Jolles B., de Guise J., Aminian K., *Functional calibration procedure for 3D knee joint angle description using inertial sensors*, Journal of Biomechanics, 42, 14, 2009, 2330–2335.
- [124] Fidali M., *Zastosowanie widm wyższych rzędów do analizy sygnałów napięcia spawania*, Przegląd Spawalnictwa – Welding Technology Review, 89, 5, 2017.
- [125] Fister Jr I., Ljubić K., Suganthan P.N., Perc M., Fister I., *Computational intelligence in sports: challenges and opportunities within a new research domain*, Applied Mathematics and Computation, 262, 2015, 178–186.
- [126] Flandrin P., Rilling G., Goncalves P., *Empirical mode decomposition as a filter bank*, IEEE Signal Processing Letters, 11, 2, 2004, 112–114.
- [127] Flenniken W., Wall J., Bevely D., *Characterization of various IMU error sources and the effect on navigation performance*, [w:] *ION GNSS*, 2005, 967–978.
- [128] Floor-Westerdijk M.J., Schepers H.M., Veltink P.H., van Asseldonk E.H.F., Buurke J.H., *Use of inertial sensors for ambulatory assessment of center-of-mass displacements during walking*, IEEE Transactions on Biomedical Engineering, 59, 7, 2012, 2080–2084.
- [129] Fojt O., Holcik J., *Applying nonlinear dynamics to ECG signal processing*, IEEE Engineering in Medicine and Biology Magazine, 17, 2, 1998, 96–101.
- [130] Friedman J., Hastie T., Höfling H., Tibshirani R., *Pathwise coordinate optimization*, The Annals of Applied Statistics, 1, 2, 2007, 302–332.
- [131] Gama J., *Knowledge discovery from data streams*, Chapman and Hall/CRC, Boca Raton 2010.
- [132] Gao Y., Li H., Luo Y., *An empirical study of wearable technology acceptance in healthcare*, Industrial Management and Data Systems, 115, 9, 2015, 1704–1723.
- [133] Gebre-Egziabher D., Elkaim G.H., David Powell J., Parkinson B.W., *Calibration of strapdown magnetometers in magnetic field domain*, Journal of Aerospace Engineering, 19, 2, 2006, 87–102.
- [134] Gelb A., *Applied optimal estimation*, MIT Press, Cambridge 1974.
- [135] Ghasemzadeh H., Loseu V., Jafari R., *Wearable coach for sport training: A quantitative model to evaluate wrist-rotation in golf*, Journal of Ambient Intelligence and Smart Environments, 1, 2, 2009, 173–184.
- [136] Giannakopoulos T., Pikrakis A., *Introduction to Audio Analysis: A MATLAB® Approach*, Academic Press, Cambridge 2014.
- [137] Gianniotis N., Kügler S.D., Tiño P., Polsterer K.L., *Model-coupled autoencoder for time series visualisation*, Neurocomputing, 192, 2016, 139–146.

- [138] Gning A., Bonnifait P., *Constraints propagation techniques on intervals for a guaranteed localization using redundant data*, Automatica, 42, 7, 2006, 1167–1175.
- [139] Golshani L., Pasha E., Yari G., *Some properties of Rényi entropy and Rényi entropy rate*, Information Sciences, 179, 14, 2009, 2426–2433.
- [140] Gómez-Clapers J., Casanella R., *A fast and easy-to-use ECG acquisition and heart rate monitoring system using a wireless steering wheel*, IEEE Sensors Journal, 12, 3, 2012, 610–616.
- [141] Goodfellow I., Bengio Y., Courville A., *Deep learning*, MIT Press, Cambridge 2016.
- [142] Górecka J., *Detection of ocular artifacts in EEG data using the hurst exponent*, [w:] *20th International Conference on Methods and Models in Automation and Robotics (MMAR)*, IEEE, 2015, 931–933.
- [143] Graff B., Graff G., Kolesiak A., *Entropia w badaniach zaburzeń rytmu serca*, Matematyka stosowana: matematyka dla społeczeństwa, 9, 1, 2008, 46–56.
- [144] Gravina R., Alinia P., Ghasemzadeh H., Fortino G., *Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges*, Information Fusion, 35, 2017, 68–80.
- [145] Gucma M., Montewka J., *Podstawy morskiej nawigacji inercyjnej*, Akademia Morska, Gdynia 2006.
- [146] Guerrero-Castellanos J.F., Madrigal-Sastre H., Durand S., Torres L., Muñoz-Hernández G.A., *A robust nonlinear observer for real-time attitude estimation using low-cost MEMS inertial sensors*, Sensors, 13, 11, 2013, 15138–15158.
- [147] Guo X., Sun C., Wang P., Huang L., *Hybrid methods for MEMS gyro signal noise reduction with fast convergence rate and small steady-state error*, Sensors and Actuators A: Physical, 2017.
- [148] Gustafsson F., *Statistical sensor fusion*, Studentlitteratur, Lund 2010.
- [149] Hadjidj A., Souil M., Bouabdallah A., Challal Y., Owen H., *Wireless sensor networks for rehabilitation applications: Challenges and opportunities*, Journal of Network and Computer Applications, 36, 1, 2013, 1–15.
- [150] Hall D.L., Llinas J., *An introduction to multisensor data fusion*, Proceedings of the IEEE, 85, 1, 1997, 6–23.
- [151] Hall D.L., Llinas J., *Multisensor data fusion*, CRC Press, Boca Raton 2001.
- [152] Halvorsen K., Olsson F., *Pose estimation of cyclic movement using inertial sensor data*, [w:] *Statistical Signal Processing Workshop (SSP)*, IEEE, 2016, 1–5.
- [153] Hammersley J., *Monte Carlo methods*, Springer Science and Business Media, Berlin 2013.
- [154] Hassan M., Terrien J., Marque C., Karlsson B., *Comparison between approximate entropy, correntropy and time reversibility: Application to uterine electromyogram signals*, Medical Engineering and Physics, 33, 8, 2011, 980–986.
- [155] Hasse L., Spiralski L., *Szumy elementów i układów elektronicznych*, Wydawnictwo Naukowo-Techniczne, Warszawa 1981.
- [156] Hastie T., Tibshirani R., Wainwright M., *Statistical learning with sparsity: the lasso and generalizations*, CRC Press, Boca Raton 2015.
- [157] Haykin S., *Signal processing in a nonlinear, nongaussian, and nonstationary world*, [w:] *Nonlinear Speech Modeling and Applications*, Springer, 2005, 43–53.
- [158] Haykin S.O., *Adaptive filter theory*, Pearson Higher Ed., New York 2013.

Bibliografia

- [159] He F., Billings S.A., Wei H.L., Sarrigiannis P.G., Zhao Y., *Spectral analysis for nonstationary and nonlinear systems: A discrete-time-model-based approach*, IEEE Transactions on Biomedical Engineering, 60, 8, 2013, 2233–2241.
- [160] Muro-de-la Herran A., Garcia-Zapirain B., Mendez-Zorrilla A., *Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications*, Sensors, 14, 2, 2014, 3362–3394.
- [161] Hinton G.E., Sejnowski T.J., *Learning and relearning in Boltzmann machines*, Parallel Distributed Processing, 1, 1986.
- [162] Ho T.K., Hull J.J., Srihari S.N., *Decision combination in multiple classifier systems*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16, 1, 1994, 66–75.
- [163] Hood L., Auffray C., *Participatory medicine: A driving force for revolutionizing healthcare*, Genome Medicine, 5, 1–4, 2013, 110.
- [164] Horzyk A., Starzyk J.A., Graham J., *Integration of semantic and episodic memories*, IEEE Transactions on Neural Networks and Learning Systems, 28, 12, 2017, 3084–3095.
- [165] Hou S., Riley C., *Is uncorrelated linear discriminant analysis really a new method?*, Chemometrics and Intelligent Laboratory Systems, 142, 2015, 49–53.
- [166] Hsu S.H., Mullen T.R., Jung T.P., Cauwenberghs G., *Real-time adaptive EEG source separation using online recursive independent component analysis*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24, 3, 2016, 309–319.
- [167] Hsu Y.L., Chen Y.T., Chou P.H., Kou Y.C., Chen Y.C., Su H.Y., *Golf swing motion detection using an inertial-sensor-based portable instrument*, [w:] *IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, IEEE, 2016, 1–2.
- [168] Huang M., Lin R., Huang S., Xing T., *A novel approach for precipitation forecast via improved k-nearest neighbor algorithm*, Advanced Engineering Informatics, 33, 2017, 89–95.
- [169] Huang N.E., Shen Z., Long S.R., Wu M.C., Shih H.H., Zheng Q., Yen N.C., Tung C.C., Liu H.H., *The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454, 1971, 1998, 903–995.
- [170] Huang X., Ye Y., Xiong L., Lau R.Y., Jiang N., Wang S., *Time series k-means: A new k-means type smooth subspace clustering for time series data*, Information Sciences, 367, 2016, 1–13.
- [171] Huang Y., Wang W., Wang L., *Unconstrained multimodal multi-label learning*, IEEE Transactions on Multimedia, 17, 11, 2015, 1923–1935.
- [172] Huang Y.S., Suen C.Y., *The behavior-knowledge space method for combination of multiple classifiers*, [w:] *IEEE computer society conference on computer vision and pattern recognition*, Institute of Electrical Engineers, Inc., 1993, 347–347.
- [173] *Standard for a Smart Transducer Interface for Sensors and Actuators – Common Functions, Communication Protocols, and Transducer Electronic Data Sheet (TEDS) Formats*, IEEE Standard 1451.2, IEEE, 1997.
- [174] *Standard Specification Format Guide and Test Procedure for Single-axis Interferometric Fiber Optic Gyros*, IEEE Standard 952, IEEE, 1998.

- [175] *Standard Specification Format Guide and Test Procedure for Linear, Single-axis, Nongyroscopic Accelerometers*, IEEE Standard 1293, IEEE, 1999.
- [176] Isermann R., Münchhof M., *Identification of dynamic systems: an introduction with applications*, Springer Science and Business Media, Berlin 2010.
- [177] Ivanov P.C., Amaral L.A.N., Goldberger A.L., Havlin S., Rosenblum M.G., Struzik Z.R., Stanley H.E., *Multifractality in human heartbeat dynamics*, Nature, 399, 6735, 1999, 461–465.
- [178] Jagannathan S., Lewis F.L., *Identification of nonlinear dynamical systems using multilayered neural networks*, Automatica, 32, 12, 1996, 1707–1712.
- [179] Jagodnik K.M., Thomas P.S., van den Bogert A.J., Branicky M.S., Kirsch R.F., *Training an actor-critic reinforcement learning controller for arm movement using human-generated rewards*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25, 10, 2017, 1892–1905.
- [180] Jagodziński J., *Metody planowania ruchu układów bezdryfowych bazujące na algorytmie Lafferriera–Sussmanna*, praca doktorska, Politechnika Wrocławska, Wrocław 2011.
- [181] Janocha H., *Microactuators-principles, applications, trends*, [w:] *Proceeding of MICRO, VDE World Microtechnology Congress*, 2000, 25–27.
- [182] Jerath K., Brennan S., Lagoa C., *Bridging the gap between sensor noise modeling and sensor characterization*, Measurement, 116, 2018, 350–366.
- [183] Jevtić D.R., Ivić M.L.A., Reljin I.S., Reljin B.D., Plavec G.I., Petrović S.D., Mijjin D.Ž., *Diagnostic spectroscopic and computer-aided evaluation of malignancy from UV/VIS spectra of clear pleural effusions*, Physica A: Statistical Mechanics and its Applications, 403, 2014, 206–216.
- [184] Jezierski E., *Dynamika robotów*, Wydawnictwa Naukowo-Techniczne, Warszawa 2006.
- [185] Ji S., Ye J., *Generalized linear discriminant analysis: A unified framework and efficient model selection*, IEEE Transactions on Neural Networks, 19, 10, 2008, 1768–1782.
- [186] Joukov V., Česić J., Westermann K., Marković I., Kulić D., Petrović I., *Human motion estimation on Lie groups using IMU measurements*, [w:] *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2017, 1965–1972.
- [187] Joukov V., Karg M., Kulic D., *Online tracking of the lower body joint angles using IMUs for gait rehabilitation*, [w:] *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2014, 2310–2313.
- [188] Jovanov E., Milenkovic A., Otto C., De Groen P.C., *A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation*, Journal of NeuroEngineering and Rehabilitation, 2, 1, 2005, 1–10.
- [189] Julier S.J., Uhlmann J.K., *A general method for approximating nonlinear transformations of probability distributions*, Rap. tech., Department of Engineering Science, University of Oxford, 1996.
- [190] Julier S.J., Uhlmann J.K., *Unscented filtering and nonlinear estimation*, Proceedings of the IEEE, 92, 3, 2004, 401–422.

Bibliografia

- [191] Kaczorek T., Dzieliński A., Dąbrowski W., Łopatka R., *Podstawy teorii sterowania*, Wydawnictwo WNT, Warszawa 2014.
- [192] Kaiser E.B., Lawo M., *Wearable navigation system for the visually impaired and blind people*, [w:] *IEEE/ACIS 11th International Conference on Computer and Information Science (ICIS)*, IEEE, 2012, 230–233.
- [193] Kaiser J.F., *On a simple algorithm to calculate the “energy” of a signal*, [w:] *International Conference on Acoustics, Speech, and Signal Processing*, IEEE, 1990, 381–384.
- [194] Kaleem M.F., Sugavaneswaran L., Guergachi A., Krishnan S., *Application of empirical mode decomposition and teager energy operator to EEG signals for mental task classification*, [w:] *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2010, 4590–4593.
- [195] Kamilov U., Bostan E., Unser M., *Wavelet shrinkage with consistent cycle spinning generalizes total variation denoising*, *IEEE Signal Processing Letters*, 19, 4, 2012, 187–190.
- [196] Kannadasan K., Edla D.R., Kuppili V., *Type 2 diabetes data classification using stacked autoencoders in deep neural networks*, *Clinical Epidemiology and Global Health*, 2018.
- [197] Kantelhardt J.W., Koscielny-Bunde E., Rego H.H., Havlin S., Bunde A., *Detecting long-range correlations with detrended fluctuation analysis*, *Physica A: Statistical Mechanics and its Applications*, 295, 3–4, 2001, 441–454.
- [198] Keat J.E., *Analysis of least-squares attitude determination routine DOAOP*, Rap. tech. CSC/TM-77/6034, Computer Science Corporation, 1977.
- [199] Keesman K.J., *System identification: an introduction*, Springer Science and Business Media, Berlin 2011.
- [200] Kenfack L.T., Martin T., Fai L.C., *Classification of classical non-gaussian noises with respect to their detrimental effects on the evolution of entanglement using a system of three-qubit as probe* 2017, URL <https://hal.archives-ouvertes.fr/hal-01612776>.
- [201] Kesić S., Spasić S.Z., *Application of Higuchi’s fractal dimension from basic to clinical neurophysiology: A review*, *Computer Methods and Programs in Biomedicine*, 133, 2016, 55–70.
- [202] Khaldi K., Alouane M.T.H., Boudraa A.O., *A new EMD denoising approach dedicated to voiced speech signals*, [w:] *2nd International Conference on Signals, Circuits and Systems*, IEEE, 2008, 1–5.
- [203] Khaldi K., Boudraa A.O., Bouchikhi A., Alouane M.T.H., *Speech enhancement via EMD*, *EURASIP Journal on Advances in Signal Processing*, 2008, 1, 2008, 873204.
- [204] Khaleghi B., Khamis A., Karray F.O., Razavi S.N., *Multisensor data fusion: A review of the state-of-the-art*, *Information Fusion*, 14, 1, 2013, 28–44.
- [205] Kiani M., Pourtakdoust S.H., Sheikhy A.A., *Consistent calibration of magnetometers for nonlinear attitude determination*, *Measurement*, 73, 2015, 180–190.
- [206] Kim K.H., Lee J.G., Park C.G., *Adaptive two-stage extended Kalman filter for a fault-tolerant INS-GPS loosely coupled system*, *IEEE Transactions on Aerospace and Electronic Systems*, 45, 1, 2009, 125–137.
- [207] Kim M., Lee D., *Wearable inertial sensor based parametric calibration of lower-limb kinematics*, *Sensors and Actuators A: Physical*, 265, 2017, 280–296.

- [208] Kim P., Huh L., *Kalman filter for beginners: with MATLAB[®] examples*, CreateSpace, Scotts Valley 2011.
- [209] Kim S.J., Koh K., Boyd S., Gorinevsky D., *ell_1 trend filtering*, SIAM Review, 51, 2, 2009, 339–360.
- [210] Kiranyaz S., Ince T., Zabihi M., Ince D., *Automated patient-specific classification of long-term electroencephalography*, Journal of Biomedical Informatics, 49, 2014, 16–31.
- [211] Klamra M., *Podstawowe idee próbkowania oszczędnego*, Czasopismo Techniczne. Nauki Podstawowe, 109, 1-NP, 2012, 85–100.
- [212] Kohavi R., et al., *A study of cross-validation and bootstrap for accuracy estimation and model selection*, [w:] *International Joint Conference on Artificial Intelligence*, 2, 1995, 1137–1143.
- [213] Kok M., Hol J., Schön T., *An optimization-based approach to human body motion capture using inertial sensors*, [w:] *19th World Congress of the International Federation of Automatic Control (IFAC)*, International Federation of Automatic Control, 2014, 79–85.
- [214] Kok M., Pakazad S.K., Schön T.B., Hansson A., Hol J.D., *A scalable and distributed solution to the inertial motion capture problem*, arXiv preprint arXiv:1603.06443, 2016.
- [215] Kolen J.F., Pollack J.B., *Back propagation is sensitive to initial conditions*, [w:] *Advances in neural information processing systems*, 1991, 860–867.
- [216] Komaty A., Boudraa A.Q., Augier B., Daré-Emzivat D., *EMD-based filtering using similarity measure between probability density functions of IMFs*, IEEE Transactions on Instrumentation and Measurement, 63, 1, 2014, 27–34.
- [217] Kopsinis Y., McLaughlin S., *Development of EMD-based denoising methods inspired by wavelet thresholding*, IEEE Transactions on Signal Processing, 57, 4, 2009, 1351–1362.
- [218] Kostrikin A.I., Trzeciak J., *Wstęp do algebry: Podstawy algebry*, Wydawnictwo Naukowe PWN, Warszawa 2011.
- [219] Kozłowski K., Dutkiewicz P., Wróblewski W., *Modelowanie i sterowanie robotów*, Wydawnictwo Naukowe PWN, Warszawa 2012.
- [220] Kranz M., Möller A., Hammerla N., Diewald S., Plötz T., Olivier P., Roalter L., *The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices*, Pervasive and Mobile Computing, 9, 2, 2013, 203–215.
- [221] Krawczyk B., Woźniak M., *Incremental weighted one-class classifier for mining stationary data streams*, Journal of Computational Science, 9, 2015, 19–25.
- [222] Kuang W., Ling B.W.K., Yang Z., *Reconstructing signal from quantized signal based on singular spectral analysis*, Digital Signal Processing, 2018.
- [223] Kulczycki P., Korbicz J., Kacprzyk J., *Automatyka, robotyka i przetwarzanie informacji*, PWN, Warszawa 2020.
- [224] Kuncheva L.I., *Combining pattern classifiers: methods and algorithms*, John Wiley and Sons, Hoboken 2004.
- [225] Lai D.T., Begg R.K., Palaniswami M., *Computational intelligence in gait research: A perspective on current applications and future challenges*, IEEE Transactions on Information Technology in Biomedicine, 13, 5, 2009, 687–702.

Bibliografia

- [226] Långkvist M., Karlsson L., Loutfi A., *A review of unsupervised feature learning and deep learning for time-series modeling*, Pattern Recognition Letters, 42, 2014, 11–24.
- [227] Laudanski A., Brouwer B., Li Q., *Measurement of lower limb joint kinematics using inertial sensors during stair ascent and descent in healthy older adults and stroke survivors*, Journal of Healthcare Engineering, 4, 4, 2013, 555–576.
- [228] Lee J.K., Park E.J., Robinovitch S.N., *Estimation of attitude and external acceleration using inertial sensor measurement during various dynamic conditions*, IEEE Transactions on Instrumentation and Measurement, 61, 8, 2012, 2262–2273.
- [229] Lee J.K., Robinovitch S.N., Park E.J., *Inertial sensing-based pre-impact detection of falls involving near-fall scenarios*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23, 2, 2015, 258–266.
- [230] Lee J.M., Kim D.J., Kim I.Y., Park K.S., Kim S.I., *Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data*, Computers in Biology and Medicine, 32, 1, 2002, 37–47.
- [231] Leland R.P., *Mechanical-thermal noise in mems gyroscopes*, IEEE Sensors Journal, 5, 3, 2005, 493–500.
- [232] Leles M., Sansao J., Mozelli L., Guimaraes H., *Improving reconstruction of time-series based in singular spectrum analysis: A segmentation approach*, Digital Signal Processing, 77, 2018, 63–76.
- [233] Lenz J., Edelstein A.S., *Magnetic sensors and their applications*, Sensors Journal, 6, 3, 2006, 631–649.
- [234] Leonowicz Z., *Analiza sygnałów okresowych metodami statystyk wyższych rzędów*, praca magisterska, Politechnika Wroclawska, Wrocław 1997.
- [235] Leonowicz Z., *Zaawansowane metody analizy widmowej sygnałów elektrycznych*, praca doktorska, Instytut Podstaw Elektrotechniki i Elektrotechnologii Politechniki Wroclawskiej, Wrocław 2000.
- [236] Leontaritis I., Billings S., of Control Engineering; S.U.D., *Identification of non-linear systems using parameter estimation techniques*, [w:] *Proceedings of IEEE Conference of Control and Applications*, 1981, 183–190.
- [237] Leśniczek Ł., Brzostowski K., *Application of the sensor fusion for the system to track human upper limb motion*, [w:] *International Conference on Multimedia and Network Information System*, Springer, 2018, 45–55.
- [238] Leśniczek Ł., *Opracowanie algorytmu estymacji prędkości i kątów w stawach kończyny górnej z wykorzystaniem bezprzewodowych czujników pomiarowych*, praca magisterska, Politechnika Wroclawska, Wrocław 2015.
- [239] Li J., Besada J.A., Bernardos A.M., Tarrío P., Casar J.R., *A novel system for object pose estimation using fused vision and inertial data*, Information Fusion, 33, 2017, 15–28.
- [240] Li J., Fang J., *Sliding average allan variance for inertial sensor stochastic error analysis*, IEEE Transactions on Instrumentation and Measurement, 62, 12, 2013, 3291–3300.
- [241] Li M., Chen W., Zhang T., *Automatic epilepsy detection using wavelet-based nonlinear analysis and optimized SVM*, Biocybernetics and Biomedical Engineering, 36, 4, 2016, 708–718.

- [242] Li M., Rouf V.T., Thompson M.J., Horsley D.A., *Three-axis Lorentz-force magnetic sensor for electronic compass applications*, Journal of Microelectromechanical Systems, 21, 4, 2012, 1002–1010.
- [243] Li R., Wang J., Chen Y., *Effect of the signal filtering on detrended fluctuation analysis*, Physica A: Statistical Mechanics and its Applications, 494, 2018, 446–453.
- [244] Li Y., Lei M.Y., Guo Y., Hu Z., Wei H.L., *Time-varying nonlinear causality detection using regularized orthogonal least squares and multi-wavelets with applications to eeg*, IEEE Access, 6, 2018, 17826–17840.
- [245] Li Y., Wei H.L., Billings S.A., *Identification of time-varying systems using multi-wavelet basis functions*, IEEE Transactions on Control Systems Technology, 19, 3, 2011, 656–663.
- [246] Ligorio G., Sabatini A.M., *Dealing with magnetic disturbances in human motion capture: A survey of techniques*, Micromachines, 7, 3, 2016, 1–17.
- [247] Liu K., Liu T., Shibata K., Inoue Y., *Ambulatory measurement and analysis of the lower limb 3D posture using wearable sensor system*, [w:] *International Conference on Mechatronics and Automation*, IEEE, 2009, 3065–3069.
- [248] Liu T., Inoue Y., Shibata K., *Development of a wearable sensor system for quantitative gait analysis*, Measurement, 42, 7, 2009, 978–988.
- [249] Liu W., Wang Z., Liu X., Zeng N., Liu Y., Alsaadi F.E., *A survey of deep neural network architectures and their applications*, Neurocomputing, 234, 2017, 11–26.
- [250] Liu Y., Liu Y., Chan K.C., Hua K.A., *Hybrid manifold embedding*, IEEE Transactions on Neural Networks and Learning Systems, 25, 12, 2014, 2295–2302.
- [251] Liu Z., Sun J., Zhang Y., Rolfe P., *Sleep staging from the EEG signal using multi-domain feature extraction*, Biomedical Signal Processing and Control, 30, 2016, 86–97.
- [252] Lockie R.G., Murphy A.J., Schultz A.B., Knight T.J., de Jonge X.A.K.J., *The effects of different speed training protocols on sprint acceleration kinematics and muscle strength and power in field sport athletes*, The Journal of Strength and Conditioning Research, 26, 6, 2012, 1539–1550.
- [253] Luo R.C., Chang C.C., Lai C.C., *Multisensor fusion and integration: Theories, applications, and its perspectives*, IEEE Sensors Journal, 11, 12, 2011, 3122–3138.
- [254] Luo R.C., Yih C.C., Su K.L., *Multisensor fusion and integration: approaches, applications, and future research directions*, IEEE Sensors Journal, 2, 2, 2002, 107–119.
- [255] Luo Y., Huang Y., *A new combined approach on hurst exponent estimate and its applications in realized volatility*, Physica A: Statistical Mechanics and its Applications, 492, 2018, 1364–1372.
- [256] Lv H., Zhang L., Wang D., Wu J., *An optimization iterative algorithm based on nonnegative constraint with application to Allan variance analysis technique*, Advances in Space Research, 53, 5, 2014, 836–844.
- [257] Lv Z., Wang Y., Zhang C., Gao X., Wu X., *An ICA-based spatial filtering approach to saccadic EOG signal recognition*, Biomedical Signal Processing and Control, 43, 2018, 9–17.
- [258] Lyons R.G., Zarzycki J., Szymbor J., i Łączności W.K., *Wprowadzenie do cyfrowego przetwarzania sygnałów*, Wydawnictwa Komunikacji i Łączności, Warszawa 2010.

Bibliografia

- [259] Ma G., Yang X., Zhang B., Shi Z., *Multi-feature fusion deep networks*, Neurocomputing, 218, 2016, 164–171.
- [260] Maddirala A.K., Shaik R.A., *Removal of EOG artifacts from single channel EEG signals using combined singular spectrum analysis and adaptive noise canceler*, IEEE Sensors Journal, 16, 23, 2016, 8279–8287.
- [261] Maddirala A.K., Shaik R.A., *Separation of sources from single-channel EEG signals using independent component analysis*, IEEE Transactions on Instrumentation and Measurement, 67, 2, 2018, 382–393.
- [262] Madgwick S., *Open source IMU and AHRS algorithms*, URL <http://www.x-io.co.uk/gait-tracking-with-x-imu>.
- [263] Madgwick S., *An efficient orientation filter for inertial and inertial/magnetic sensor arrays*, Rap. tech., x-io Technologies and University of Bristol, 2010.
- [264] Madgwick S.O., Harrison A.J., Vaidyanathan R., *Estimation of IMU and MARG orientation using a gradient descent algorithm*, [w:] *IEEE International Conference on Rehabilitation Robotics (ICORR)*, IEEE, 2011, 1–7.
- [265] Madrigal J.A.B., Cardiel E., Rogeli P., Salas L.L., Guerrero R.M., *Evaluation of suitability of a micro-processing unit of motion analysis for upper limb tracking*, Medical Engineering and Physics, 2016.
- [266] Mahony R., Hamel T., Pfimlin J.M., *Nonlinear complementary filters on the special orthogonal group*, IEEE Transactions on Automatic Control, 53, 5, 2008, 1203–1218.
- [267] Mahony R., Hamel T., Trumpf J., Lageman C., *Nonlinear attitude observers on so(3) for complementary and compatible measurements: A theoretical study*, [w:] *Proceedings of the 48th IEEE Conference on Decision and Control*, IEEE, 2009, 6407–6412.
- [268] Mak J.N., Hu Y., Luk K.D., *An automated ECG-artifact removal method for trunk muscle surface EMG recordings*, Medical Engineering and Physics, 32, 8, 2010, 840–848.
- [269] Mandic D.P., ur Rehman N., Wu Z., Huang N.E., *Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis*, IEEE Signal Processing Magazine, 30, 6, 2013, 74–86.
- [270] Mangai U.G., Samanta S., Das S., Chowdhury P.R., *A survey of decision fusion and feature fusion strategies for pattern classification*, IETE Technical Review, 27, 4, 2010, 293–307.
- [271] Manis G., *Fast computation of approximate entropy*, Computer Methods and Programs in Biomedicine, 91, 1, 2008, 48–54.
- [272] Mannini A., Sabatini A.M., *Walking speed estimation using foot-mounted inertial sensors: Comparing machine learning and strap-down integration methods*, Medical Engineering and Physics, 36, 10, 2014, 1312–1321.
- [273] von Marcard T., Rosenhahn B., Black M.J., Pons-Moll G., *Sparse inertial poser: Automatic 3D human pose estimation from sparse IMUs*, Computer Graphics Forum, 36, 2, 2017, 349–360.
- [274] Markley F.L., *Attitude determination using vector observations and the singular value decomposition*, The Journal of the Astronautical Sciences, 36, 3, 1988, 245–258.
- [275] Markley F.L., *Attitude determination using vector observations: A fast optimal matrix algorithm*, Journal of the Astronautical Sciences, 41, 2, 1993, 261–280.

- [276] Marquardt D.W., *An algorithm for least-squares estimation of nonlinear parameters*, Journal of the Society for Industrial and Applied Mathematics, 11, 2, 1963, 431–441.
- [277] Mateos I., Patton B., Zhivun E., Budker D., Wurm D., Ramos-Castro J., *Noise characterization of an atomic magnetometer at sub-millihertz frequencies*, Sensors and Actuators A: Physical, 224, 2015, 147–155.
- [278] Maugis P.A.G., *Big data uncertainties*, Journal of Forensic and Legal Medicine, 2016, 7–11.
- [279] McAfee A., Brynjolfsson E., *Big data: The management revolution*, Harvard Business Review, 90, 10, 2012, 61–67.
- [280] McGhee R.B., *The factored quaternion algorithm for orientation estimation from measured earth gravity and magnetic field*, MOVES Inst., Naval Postgraduate School, Monterey, CA, 2004.
- [281] Meditch J.S., Nahorski Z., Siekierski T., *Estymacja i sterowanie statystycznie optymalne w układach liniowych*, Wydawnictwa Naukowo-Techniczne, Warszawa 1975.
- [282] Mendel J.M., *Extension of Friedland's bias filtering technique to a class of nonlinear systems*, IEEE Transactions on Automatic Control, 21, 2, 1976, 296–298.
- [283] Meng X., Zhang Z.Q., Wu J.K., Wong W.C., Yu H., *Self-contained pedestrian tracking during normal walking using an inertial/magnetic sensor module*, IEEE Transactions on Biomedical Engineering, 61, 3, 2014, 892–899.
- [284] Mert A., Akan A., *Detrended fluctuation thresholding for empirical mode decomposition based denoising*, Digital Signal Processing, 32, 2014, 48–56.
- [285] Miezal M., Taetz B., Bleser G., *On inertial body tracking in the presence of model calibration errors*, Sensors, 16, 7, 2016, 1–34.
- [286] Mihelj M., *Inverse kinematics of human arm based on multisensor data integration*, Journal of Intelligent and Robotic Systems, 47, 2, 2006, 139–153.
- [287] Mikołajewska E., Mikołajewski D., *Możliwości wykorzystania egzoszkieletu medycznego jako nowoczesnej formy terapii*, Niepełnosprawność – zagadnienia, problemy, rozwiązania, 4, 2012, 121–141.
- [288] Milenković A., Otto C., Jovanov E., *Wireless sensor networks for personal health monitoring: Issues and an implementation*, Computer Communications, 29, 13, 2006, 2521–2533.
- [289] Mitchell H.B., *Multi-sensor data fusion: an introduction*, Springer Science and Business Media, Berlin 2007.
- [290] Moghtaderi A., Flandrin P., Borgnat P., *Trend filtering via empirical mode decompositions*, Computational Statistics and Data Analysis, 58, 2013, 114–126.
- [291] Mohd-Yasin F., Nagel D., Korman C., *Noise in MEMS*, Measurement Science and Technology, 21, 1, 2009, 1–24.
- [292] Molino-Minero-Re E., Garcia-Nocetti F., Benitez-Perez H., *Application of a time-scale local hurst exponent analysis to time series*, Digital Signal Processing, 37, 2015, 92–99.
- [293] Morel M., Achard C., Kulpa R., Dubuisson S., *Time-series averaging using constrained dynamic time warping with tolerance*, Pattern Recognition, 74, 2018, 77–89.
- [294] Mortari D., *Energy approach algorithm for attitude determination from vector observations*, Journal of the Astronautical Sciences, 45, 1, 1997, 41–55.

Bibliografia

- [295] Moukadem A., Dieterlen A., Hueber N., Brandt C., *A robust heart sounds segmentation module based on S-transform*, Biomedical Signal Processing and Control, 8, 3, 2013, 273–281.
- [296] Mounier E., Fraux R., Robin L., Bonnabel A., *Inertial MEMS manufacturing trends*, Yole Developpement, 1, 2014, 2.
- [297] Mubashir M., Shao L., Seed L., *A survey on fall detection: Principles and approaches*, Neurocomputing, 100, 2013, 144–152.
- [298] Murphy K.P., *Machine learning: A probabilistic perspective*, MIT Press, Cambridge 2012.
- [299] Murphy R.R., *Biological and cognitive foundations of intelligent sensor fusion*, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 26, 1, 1996, 42–51.
- [300] Murray R.M., Li Z., Sastry S.S., Sastry S.S., *A mathematical introduction to robotic manipulation*, CRC Press, Boca Raton 1994.
- [301] Muthukrishnan S., et al., *Data streams: Algorithms and applications*, Foundations and Trends in Theoretical Computer Science, 1, 2, 2005, 117–236.
- [302] Mutlu A.Y., *Detection of epileptic dysfunctions in EEG signals using hilbert vibration decomposition*, Biomedical Signal Processing and Control, 40, 2018, 33–40.
- [303] Nam C.N.K., Kang H.J., Suh Y.S., *Golf swing motion tracking using inertial sensors and a stereo camera*, IEEE Transactions on Instrumentation and Measurement, 63, 4, 2014, 943–952.
- [304] Nascimento L.R., de Oliveira C.Q., Ada L., Michaelsen S.M., Teixeira-Salmela L.F., *Walking training with cueing of cadence improves walking speed and stride length after stroke more than walking training alone: A systematic review*, Journal of Physiotherapy, 61, 1, 2015, 10–15.
- [305] Naves R., Barbosa B.H., Ferreira D.D., *Classification of lung sounds using higher-order statistics: A divide-and-conquer approach*, Computer Methods and Programs in Biomedicine, 129, 2016, 12–20.
- [306] Nelles O., *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*, Springer Science and Business Media, Berlin 2013.
- [307] Nguyen H.N., Guillemin F., *On process noise covariance estimation*, [w:] *25th Mediterranean Conference on Control and Automation (MED)*, IEEE, 2017, 1345–1348.
- [308] Niedźwiedzińska H., *Od medycznych zastosowań EDI do telemedycyny*, Acta Universitatis Lodzianis. Folia Oeconomica, 2004.
- [309] Niewinski M., *Nieliniowa dynamika szeregów czasowych w badaniach ekonomicznych*, Akademia Ekonomiczna we Wrocławiu, Wrocław 2007.
- [310] Nilsson J.O., Skog I., Händel P., *A note on the limitations of ZUPTs and the implications on sensor error modeling*, [w:] *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, 1–4.
- [311] Ning X., Selesnick I., *ECG enhancement and QRS detection based on sparse derivatives*, Biomedical Signal Processing and Control, 8, 6, 2013, 713–723.
- [312] *National Geophysical Data Center*, URL <http://www.ngdc.noaa.gov> 2014.
- [313] Noga K.M., *Transmisja danych w kanale radiowym – wybrane zagadnienia w środowisku Vissim*, Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej, Gdańsk, 2014, 77–80.

- [314] Norrdine A., Kasmi Z., Blankenbach J., *Step detection for ZUPT-aided inertial pedestrian navigation system using foot-mounted permanent magnet*, IEEE Sensors Journal, 16, 17, 2016, 6766–6773.
- [315] Nweke H.F., Teh Y.W., Mujtaba G., Al-garadi M.A., *Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions*, Information Fusion, 46, 2019, 147–170.
- [316] Ocak H., *Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy*, Expert Systems with Applications, 36, 2, 2009, 2027–2036.
- [317] Ogata K., Romicki S., Amborski K., *Metody przestrzeni stanów w teorii sterowania*, Wydawnictwa Naukowo-Techniczne, Warszawa 1974.
- [318] Ogunfunmi T., *Adaptive nonlinear system identification: the Volterra and Wiener model approaches*, Springer Science and Business Media, Berlin 2007.
- [319] Owis M.I., Abou-Zied A.H., Youssef A.B., Kadah Y.M., *Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification*, IEEE transactions on Biomedical Engineering, 49, 7, 2002, 733–736.
- [320] Pamuła G., *Multifrakalne tło skończonych sygnałów perzystentnych oraz ich transformacji nieliniowych wraz z zastosowaniami interdyscyplinarnymi*, praca doktorska, Uniwersytet Wrocławski, Wrocław 2013.
- [321] Pang S., Ozawa S., Kasabov N., *Incremental linear discriminant analysis for classification of data streams*, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 35, 5, 2005, 905–914.
- [322] Parent R., *Animacja komputerowa. Algorytmy i techniki*, Wydawnictwo Naukowe PWN, Warszawa 2012.
- [323] Patil R.A., Patil A.D., *Use of information technology in healthcare sector for improving outcomes*, International Journal of Basic and Clinical Pharmacology, 3, 2, 2017, 269–271.
- [324] Peng J., Peng S., Xie Q., Wei J., *Baseline correction combined partial least squares algorithm and its application in on-line Fourier transform infrared quantitative analysis*, Analytica Chimica Acta, 690, 2, 2011, 162–168.
- [325] Peppoloni L., Filippeschi A., Ruffaldi E., Avizzano C.A., *A novel 7 degrees of freedom model for upper limb kinematic reconstruction based on wearable sensors*, [w:] *IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY)*, IEEE, 2013, 105–110.
- [326] Perea L., How J., Breger L., Elosegui P., *Nonlinearity in sensor fusion: Divergence issues in EKF, modified truncated SOF, and UKF*, [w:] *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2007, 1–16.
- [327] Petkov P., Slavov T., *Stochastic modeling of MEMS inertial sensors*, Cybernetics and Information Technologies, 10, 2, 2010, 31–40.
- [328] Pierro P., Monje C.A., Mansard N., Soueres P., Balaguer C., *Open solution for humanoid attitude estimation through sensory integration and extended Kalman filtering*, Automatika – Journal for Control, Measurement, Electronics, Computing and Communications, 56, 1, 2015, 9–20.
- [329] Pincus S.M., *Approximate entropy as a measure of system complexity*, Proceedings of the National Academy of Sciences, 88, 6, 1991, 2297–2301.

Bibliografia

- [330] Polak E., Kulasa J., Vences Brito A., Castro M.A., Fernandes O., *Motion analysis systems as optimization training tools in combat sports and martial arts*, Revista de Artes Marciales Asiáticas, 10, 2, 2016, 105–123.
- [331] Polikar R., *Ensemble based systems in decision making*, IEEE Circuits and Systems Magazine, 6, 3, 2006, 21–45.
- [332] Pons-Moll G., Baak A., Helten T., Müller M., Seidel H.P., Rosenhahn B., *Multisensor-fusion for 3D full-body human motion capture*, [w:] *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 2010, 663–670.
- [333] Price C.F., *An analysis of the divergence problem in the Kalman filter*, IEEE Transactions on Automatic Control, 13, 6, 1968, 699–702.
- [334] Qian J., Jia J., *On stepwise pattern recovery of the fused LASSO*, Computational Statistics and Data Analysis, 94, 2016, 221–237.
- [335] Qian L., Xu G., Tian W., Wang J., *A novel hybrid EMD-based drift denoising method for a dynamically tuned gyroscope (DTG)*, Measurement, 42, 6, 2009, 927–932.
- [336] Rao B., Engan K., Cotter S.F., Palmer J., Kreutz-Delgado K., *Subset selection in noise based on diversity measure minimization*, IEEE Transactions on Signal Processing, 51, 3, 2003, 760–770.
- [337] Raol J.R., *Multi-Sensor Data Fusion with MATLAB®*, CRC Press, Boca Raton 2009.
- [338] Renaudin V., Afzal M.H., Lachapelle G., *Complete triaxis magnetometer calibration in the magnetic domain*, Journal of Sensors, 2010.
- [339] Restrepo J.F., Schlotthauer G., Torres M.E., *Maximum approximate entropy and r threshold: A new approach for regularity changes detection*, Physica A: Statistical Mechanics and its Applications, 409, 2014, 97–109.
- [340] Robert-Lachaine X., Mecheri H., Larue C., Plamondon A., *Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis*, Medical and Biological Engineering and Computing, 55, 4, 2017, 609–619.
- [341] Rocha C., Tonetto C., Dias A., *A comparison between the Denavit–Hartenberg and the screw-based methods used in kinematic modeling of robot manipulators*, Robotics and Computer-Integrated Manufacturing, 27, 4, 2011, 723–728.
- [342] Rodriguez P.H., Alonso J.B., Ferrer M.A., Travieso C.M., *Application of the teager–kaiser energy operator in bearing fault diagnosis*, ISA Transactions, 52, 2, 2013, 278–284.
- [343] Rodriguez R., Mexicano A., Bila J., Cervantes S., Ponce R., *Feature extraction of electrocardiogram signals by applying adaptive threshold and principal component analysis*, Journal of Applied Research and Technology, 13, 2, 2015, 261–269.
- [344] Rodríguez R., Mexicano A., Bila J., Ponce R., Cervantes S., *Adaptive threshold and principal component analysis for features extraction of electrocardiogram signals*, [w:] *International Symposium on Computer, Consumer and Control (IS3C)*, IEEE, 2014, 1253–1258.
- [345] Roetenberg D., *Inertial and magnetic sensing of human motion*, praca doktorska, University of Twente, Twente 2006.
- [346] Roetenberg D., Luinge H., Slycke P., *Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors*, Rap. tech., Xsens Motion Technologies BV, 2009.

- [347] Roetenberg D., Luinge H.J., Baten C.T., Veltink P.H., *Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 13, 3, 2005, 395–405.
- [348] de la Rosa E., Yu W., *Randomized algorithms for nonlinear system identification with deep learning modification*, Information Sciences, 364, 2016, 197–212.
- [349] Rosenstein M.T., Collins J.J., De Luca C.J., *A practical method for calculating largest Lyapunov exponents from small data sets*, Physica D: Nonlinear Phenomena, 65, 1–2, 1993, 117–134.
- [350] Sabatini A.M., *Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing*, IEEE Transactions on Biomedical Engineering, 53, 7, 2006, 1346–1356.
- [351] Sabatini A.M., *Kalman-filter-based orientation determination using inertial/magnetic sensors: Observability analysis and performance evaluation*, Sensors, 11, 10, 2011, 9182–9206.
- [352] Salcudean S., *A globally convergent angular velocity observer for rigid body motion*, IEEE Transactions on Automatic Control, 36, 12, 1991, 1493–1497.
- [353] Salehi S., Mostofi N., Bleser G., *A practical in-field magnetometer calibration method for IMUs*, [w:] *Proceedings of the IROS Workshop on Cognitive Assistive Systems: Closing the Action-Perception Loop*, 2012, 39–44.
- [354] Särkkä S., *Bayesian filtering and smoothing*, 3, Cambridge University Press, Cambridge 2013.
- [355] Scharf L.L., Demeure C., *Statistical signal processing: detection, estimation, and time series analysis*, 63, Addison–Wesley, Reading 1991.
- [356] Schepers H.M., Asseldonk E.H.F.V., Buurke J.H., Veltink P.H., *Ambulatory estimation of center of mass displacement during walking*, IEEE Transactions on Biomedical Engineering, 56, 4, 2009, 1189–1195.
- [357] Schopp P., Klingbeil L., Peters C., Buhmann A., Manoli Y., *Sensor fusion algorithm and calibration for a gyroscope-free IMU*, Procedia Chemistry, 1, 1, 2009, 1323–1326.
- [358] Secer G., Barshan B., *Improvements in deterministic error modeling and calibration of inertial sensors and magnetometers*, Sensors and Actuators A: Physical, 247, 2016, 522–538.
- [359] Seel T., Raisch J., Schauer T., *IMU-based joint angle measurement for gait analysis*, Sensors, 14, 4, 2014, 6891–6909.
- [360] Şeker M., Özerdem M.S., *Classification of EEG data sets with Hilbert transform*, [w:] *24th Signal Processing and Communication Application Conference (SIU)*, IEEE, 2016, 1949–1952.
- [361] Selesnick I.W., Graber H.L., Ding Y., Zhang T., Barbour R.L., *Transient artifact reduction algorithm (TARA) based on sparse optimization*, IEEE Transactions on Signal Processing, 62, 24, 2014, 6596–6611.
- [362] Selesnick I.W., Graber H.L., Pfeil D.S., Barbour R.L., *Simultaneous low-pass filtering and total variation denoising*, IEEE Transactions on Signal Processing, 62, 5, 2014, 1109–1124.
- [363] Shaeffer D.K., *MEMS inertial sensors: A tutorial overview*, IEEE Communications Magazine, 51, 4, 2013, 100–109.

Bibliografia

- [364] Sharp I., Yu K., *Sensor-based dead-reckoning for indoor positioning*, Physical Communication, 13, 2014, 4–16.
- [365] Sheinker A., Shkalim A., Salomonski N., Ginzburg B., Frumkis L., Kaplan B.Z., *Processing of a scalar magnetometer signal contaminated by $1/f$ α noise*, Sensors and Actuators A: Physical, 138, 1, 2007, 105–111.
- [366] Shim H., Seo J.H., Teel A.R., *Nonlinear observer design via passivation of error dynamics*, Automatica, 39, 5, 2003, 885–892.
- [367] *Shimmer*, URL <http://www.shimmersensing.com> 2014.
- [368] Shirdel A.H., Böling J.M., Toivonen H.T., *System identification in the presence of trends and outliers using sparse optimization*, Journal of Process Control, 44, 2016, 120–133.
- [369] Siciliano B., Khatib O., *Springer handbook of robotics*, Springer, Berlin 2016.
- [370] Simon D., *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*, John Wiley and Sons, Hoboken 2006.
- [371] Simonsen I., Hansen A., Nes O.M., *Determination of the Hurst exponent by use of wavelet transforms*, Physical Review E, 58, 3, 1998, 1–22.
- [372] Smalec Z., *Wstęp do mechatroniki*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2010.
- [373] Smolka L., Ochmann B., *A novel method of anaerobic performance assessment in swimming*, The Journal of Strength and Conditioning Research, 27, 2, 2013, 533–539.
- [374] Snitkowska E., Kasprzak W., *Procedura analizy składowych niezależnych i jej wykorzystanie do opisu tekstur w obrazach cyfrowych*, [w:] *Postępy Robotyki. Sterowanie robotów z percepcją otoczenia*, K. Tchoń (red.), Wydawnictwa Komunikacji i Łączności, Warszawa, 2005, URL http://www.ia.pw.edu.pl/~wkasprza/PAP/PostepyRobotyki_2005_105.pdf.
- [375] Solin A., Kok M., Wahlström N., Schön T.B., Särkkä S., *Modeling and interpolation of the ambient magnetic field by gaussian processes*, IEEE Transactions on Robotics, 34, 4, 2018, 1112–1127.
- [376] Sozański H., *Podstawy teorii treningu*, Resortowe Centrum Metodyczno-Szkoleniowe Kultury Fizycznej i Sportu, Warszawa 1993.
- [377] Spong M.W., Hutchinson S., Vidyasagar M., *Robot modeling and control*, 3, Wiley, New York 2006.
- [378] Sprdlik O., *Detection and Estimation of Human Movement Using Inertial Sensors: Applications in Neurology*, praca doktorska, Czech Technical University in Prague, Prague 2012.
- [379] Stam C., Montez T., Jones B., Rombouts S., Van Der Made Y., Pijnenburg Y., Scheltens P., *Disturbed fluctuations of resting state EEG synchronization in alzheimer's disease*, Clinical Neurophysiology, 116, 3, 2005, 708–715.
- [380] Stiros S.C., *Errors in velocities and displacements deduced from accelerographs: an approach based on the theory of error propagation*, Soil Dynamics and Earthquake Engineering, 28, 5, 2008, 415–420.
- [381] Stockwell R.G., Mansinha L., Lowe R., *Localization of the complex spectrum: the S transform*, IEEE Transactions on Signal Processing, 44, 4, 1996, 998–1001.

- [382] Struber L., Courvoisier A., Cinquin P., Nougier V., *Development of a method and software for human motion training based on an inertial measurement units system*, [w:] *IEEE International Conference on Virtual Rehabilitation Proceedings (ICVR)*, IEEE, 2015, 251–257.
- [383] Strumillo P., Skulimowski P., Bujacz M., Obuchowicz M., Borowiecka I., Holak M., Radecki A., *A mobile application for interactive sonification of images for the visually impaired*, [w:] *11th International Conference on Human System Interaction (HSI)*, IEEE, 2018, 391–397.
- [384] Stupnicki R., *Pomiary i normy w antropometrii*, Zeszyty Naukowe WSKFiT, 11, 2016, 13–20.
- [385] Sugiyama J., Miura J., *A wearable robot control interface based on measurement of human body motion using a camera and inertial sensors*, [w:] *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, IEEE, 2011, 565–570.
- [386] *ZEUS robotic surgical system*, URL <http://spinoff.nasa.gov/spinoff2000/hm1.htm> 2014.
- [387] Szabatin J., *Przetwarzanie sygnałów*, Politechnika Warszawska, Ośrodek Kształcenia na Odległość OKNO, Warszawa 2004.
- [388] Szwach P., *Klasyfikacja typu odbić w grze w tenisa z wykorzystaniem czujników bezprzewodowych*, praca magisterska, Politechnika Wroclawska, Wrocław 2015.
- [389] Tadano S., Takeda R., Sasaki K., Fujisawa T., Tohyama H., *Gait characterization for osteoarthritis patients using wearable gait sensors (h-gait systems)*, *Journal of Biomechanics*, 49, 5, 2016, 684–690.
- [390] Taetz B., Bleser G., Miezal M., *Towards self-calibrating inertial body motion capture*, [w:] *19th International Conference on Information Fusion (FUSION)*, IEEE, 2016, 1751–1759.
- [391] Tan C.M., Lim S.Y., *Application of Wigner–Ville distribution in electromigration noise analysis*, *IEEE Transactions on Device and Materials Reliability*, 2, 2, 2002, 30–35.
- [392] Tan S., *Neighbor-weighted k-nearest neighbor for unbalanced text corpus*, *Expert Systems with Applications*, 28, 4, 2005, 667–671.
- [393] Tang D., Xu C., Yue Q., Wu W., Shi Z., Feng J., *Single point mooring system modal parameter identification based on empirical mode decomposition and time-varying autoregressive model*, *Applied Ocean Research*, 53, 2015, 250–256.
- [394] Tang F., Adam L., Si B., *Group feature selection with multiclass support vector machine*, *Neurocomputing*, 317, 2018, 42–49.
- [395] Tao W., Liu T., Zheng R., Feng H., *Gait analysis using wearable sensors*, *Sensors*, 12, 2, 2012, 2255–2283.
- [396] Taunyzov T., Omarali B., Shintemirov A., *A novel low-cost 4-DOF wireless human arm motion tracker*, [w:] *6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, IEEE, 2016, 157–162.
- [397] Tavenard Z.Z.R., Bailly A., Tang X., Tang P., Corpetti T., *Dynamic time warping under limited warping path length*, *Information Sciences*, 393, 2017, 91–107.
- [398] Tchoń K., *Manipulatory i roboty mobilne: modele, planowanie ruchu, sterowanie*, Akademicka Oficyna Wydawnicza PLJ, Warszawa 2000.

Bibliografia

- [399] Terrades O.R., Valveny E., Tabbone S., *Optimal classifier fusion in a non-bayesian probabilistic framework*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 31, 9, 2009, 1630–1644.
- [400] Thong Y.K., Woolfson M.S., Crowe J.A., Hayes-Gill B.R., Jones D.A., *Numerical double integration of acceleration measurements in noise*, Measurement, 36, 1, 2004, 73–92.
- [401] Tian X., Chen J., Han Y., Shang J., Li N., *A novel zero velocity interval detection algorithm for self-contained pedestrian navigation system with inertial sensors*, Sensors, 16, 10, 2016, 1–18.
- [402] Tian Y., Meng X., Tao D., Liu D., Feng C., *Upper limb motion tracking with the integration of IMU and kinect*, Neurocomputing, 159, 2015, 207–218.
- [403] Tian Y., Wei H., Tan J., *An adaptive-gain complementary filter for real-time human motion tracking with MARG sensors in free-living environments*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 21, 2, 2013, 254–264.
- [404] Tibshirani R., Saunders M., Rosset S., Zhu J., Knight K., *Sparsity and smoothness via the fused lasso*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67, 1, 2005, 91–108.
- [405] To G., Mahfouz M.R., *Quaternionic attitude estimation for robotic and human motion tracking using sequential Monte Carlo methods with von mises-fisher and nonuniform densities simulations*, IEEE Transactions on Biomedical Engineering, 60, 11, 2013, 3046–3059.
- [406] Tolstikov A., Hong X., Biswas J., Nugent C., Chen L., Parente G., *Comparison of fusion methods based on dst and dbn in human activity recognition*, Journal of Control Theory and Applications, 9, 1, 2011, 18–27.
- [407] Törnqvist D., *Estimation and detection with applications to navigation*, praca doktorska, Linköping University, Department of Electrical Engineering, Automatic Control, Linköping 2008.
- [408] Torres M.E., Colominas M.A., Schlotthauer G., Flandrin P., *A complete ensemble empirical mode decomposition with adaptive noise*, [w:] *IEEE international conference on Acoustics, speech and signal processing (ICASSP)*, IEEE, 2011, 4144–4147.
- [409] Trawny N., Roumeliotis S.I., *Indirect Kalman filter for 3D attitude estimation*, Rap. tech., University of Minnesota, Dept. of Comp. Sci. and Eng., Tech. Rep., 2005.
- [410] Tropp J.A., *Just relax: Convex programming methods for identifying sparse signals in noise*, IEEE Transactions on Information Theory, 52, 3, 2006, 1030–1051.
- [411] Tsai A.C., Luh J.J., Lin T.T., *A novel STFT-ranking feature of multi-channel EMG for motion pattern recognition*, Expert Systems with Applications, 42, 7, 2015, 3327–3341.
- [412] Ugbohue U.C., Papi E., Kaliarntas K.T., Kerr A., Earl L., Pomeroy V.M., Rowe P.J., *The evaluation of an inexpensive, 2D, video based gait assessment system for clinical use*, Gait and Posture, 38, 3, 2013, 483–489.
- [413] Vakman D., *On the analytic signal, the Teager–Kaiser energy algorithm, and other methods for defining amplitude and frequency*, IEEE Transactions on Signal Processing, 44, 4, 1996, 791–797.

- [414] Valchinov E., Antoniou A., Rotas K., Pallikarakis N., *Wearable ECG system for health and sports monitoring*, [w:] *EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth)*, IEEE, 2014, 63–66.
- [415] Valenti R.G., Dryanovski I., Xiao J., *Keeping a good attitude: A quaternion-based orientation filter for IMUs and MARGs*, *Sensors*, 15, 8, 2015, 19302–19330.
- [416] Van Der Heijden F., Duin R., De Ridder D., Tax D.M., *Classification, parameter estimation and state estimation: an engineering approach using MATLAB®*, John Wiley and Sons, Hoboken 2005.
- [417] Vaseghi S.V., *Advanced digital signal processing and noise reduction*, John Wiley and Sons, Hoboken 2008.
- [418] Vasilescu G., *Electronic noise and interfering signals: principles and applications*, Springer Science and Business Media, Berlin 2006.
- [419] Vega C.F., Noel J., *Parameters analyzed of Higuchi's fractal dimension for EEG brain signals*, [w:] *Signal Processing Symposium (SPSymposium)*, IEEE, 2015, 1–5.
- [420] Veitch D., Abry P., *A wavelet-based joint estimator of the parameters of long-range dependence*, *IEEE Transactions on Information Theory*, 45, 3, 1999, 878–897.
- [421] Venkatesan C., Karthigaikumar P., Paul A., Satheeskumaran S., Kumar R., *ECG signal preprocessing and svm classifier-based abnormality detection in remote health-care applications*, *IEEE Access*, 6, 2018, 9767–9773.
- [422] Vergin R., O'Shaughnessy D., Farhat A., *Generalized mel frequency cepstral coefficients for large-vocabulary speaker-independent continuous-speech recognition*, *IEEE Transactions on Speech and Audio Processing*, 7, 5, 1999, 525–532.
- [423] Vig J.R., *Definitions of physical quantities for fundamental frequency and time metrology – random instabilities*, *IEEE Standard 1139–1999*, IEEE, 1999.
- [424] da Vinci Surgery, URL <http://www.davincisurgery.com/da-vinci-surgery/da-vinci-surgical-system/> 2014.
- [425] Vorobyov S., Cichocki A., *Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis*, *Biological Cybernetics*, 86, 4, 2002, 293–303.
- [426] Wahba G., *A least squares estimate of satellite attitude*, *SIAM Review*, 7, 3, 1965, 409–409.
- [427] Wallner P.E., Kanski A., *The impact of technology on health care cost and policy development*, *Seminars in Radiation Oncology*, 18, 3, 2008, 194–200, *Health Policy and Health Care Economics*.
- [428] Wan E.A., Van Der Merwe R., *The unscented Kalman filter for nonlinear estimation*, [w:] *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, IEEE, 2000, 153–158.
- [429] Wang L., Hu W., Tan T., *Recent developments in human motion analysis*, *Pattern Recognition*, 36, 3, 2003, 585–601.
- [430] Wang Q., Zhang L., *Least squares online linear discriminant analysis*, *Expert Systems with Applications*, 39, 1, 2012, 1510–1517.
- [431] Wang X., Huang F., Cheng Y., *Computational performance optimization of support vector machine based on support vectors*, *Neurocomputing*, 211, 2016, 66–71.

Bibliografia

- [432] Wang X., Yu W., Qi X., Deng Y., Liu Y., *Radiofrequency interference suppression in synthetic aperture radar based on singular spectrum analysis with extended-FAPI subspace tracking*, IET Radar, Sonar and Navigation, 6, 9, 2012, 881–890.
- [433] Webb A.R., *Statistical pattern recognition*, John Wiley and Sons, Hoboken 2003.
- [434] Wei H., Billings S., Liu J., *Term and variable selection for non-linear system identification*, International Journal of Control, 77, 1, 2004, 86–110.
- [435] Wei H.L., Billings S., et al., *Feature subset selection and ranking for data dimensionality reduction*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29, 1, 2007, 162–166.
- [436] Wei Y., Fei Q., He L., *Sports motion analysis based on mobile sensing technology*, [w:] *International Conference on Global Economy, Finance and Humanities Research (GEFHR)*, Atlantis Press, 2014, 74–76.
- [437] Weng B., Blanco-Velasco M., Barner K.E., *ECG denoising based on the empirical mode decomposition*, [w:] *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006, 1–4.
- [438] Westwick D.T., Kearney R.E., *Identification of nonlinear physiological systems*, 7, John Wiley and Sons, Hoboken 2003.
- [439] Whelan D., O'Reilly M., Ward T., Delahunt E., Caulfield B., *Evaluating performance of the lunge exercise with multiple and individual inertial measurement units*, [w:] *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare*, ICST (Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), 2016, 101–108.
- [440] Wiak S. (red.), *Mechatronika: Podstawy, Metody, Modele Komputerowe, Technologie*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2009.
- [441] Wiener N., *Cybernetyka: czyli sterowanie i komunikacja w zwierzeu u maszynie*, Państwowe Wydawnictwo Naukowe, Warszawa 1971.
- [442] Winter D.A., *Biomechanics and motor control of human movement*, John Wiley and Sons, Hoboken 2009.
- [443] Wojnowski J. (red.), *Wielka Encyklopedia PWN*, Wydawnictwo Naukowe PWN, Warszawa 2001–2005.
- [444] Wong C.C.L., *Fusion of wearable and visual sensors for human motion analysis*, praca doktorska, Imperial College, London 2015.
- [445] Woodman O.J., *An introduction to inertial navigation*, Rap. tech., University of Cambridge, Computer Laboratory, 2007.
- [446] Wu J., Ma J., *Rényi entropy penalized learning algorithm for gaussian mixture with automated model selection*, [w:] *9th International Conference on Signal Processing*, IEEE, 2008, 1561–1564.
- [447] Wu X., Liu X., Dai S., *The reliability of big data*, [w:] *IEEE 7th Joint International Information Technology and Artificial Intelligence Conference*, IEEE, 2014, 295–299.
- [448] Wu Z., Huang N.E., *A study of the characteristics of white noise using the empirical mode decomposition method*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 460, 2046, 2004, 1597–1611.
- [449] Wu Z., Huang N.E., *Ensemble empirical mode decomposition: A noise-assisted data analysis method*, Advances in Adaptive Data Analysis, 1, 1, 2009, 1–41.

- [450] Xia B., Bao C., *Wiener filtering based speech enhancement with weighted denoising auto-encoder and noise classification*, *Speech Communication*, 60, 2014, 13–29.
- [451] Xiong P., Wang H., Liu M., Zhou S., Hou Z., Liu X., *ECG signal enhancement based on improved denoising auto-encoder*, *Engineering Applications of Artificial Intelligence*, 52, 2016, 194–202.
- [452] *Xsens*, URL <https://www.xsens.com>.
- [453] Yan R., Liu Y., Gao R.X., *Correlation dimension analysis: A non-linear time series analysis for data processing*, *IEEE Instrumentation and Measurement Magazine*, 13, 6, 2010.
- [454] Yang B., Dong Y., Yu C., Hou Z., *Singular spectrum analysis window length selection in processing capacitive captured biopotential signals*, *IEEE Sensors Journal*, 16, 19, 2016, 7183–7193.
- [455] Yang G., Liu Y., Wang Y., Zhu Z., *EMD interval thresholding denoising based on similarity measure to select relevant modes*, *Signal Processing*, 109, 2015, 95–109.
- [456] Yang G.Z., Yacoub M., *Body sensor networks*, Springer Verlag, London 2006.
- [457] Young A.D., *Use of body model constraints to improve accuracy of inertial motion capture*, [w:] *International Conference on Body Sensor Networks (BSN)*, IEEE, 2010, 180–186.
- [458] Yu L., Chen J., Ding G., Tu Y., Yang J., Sun J., *Spectrum prediction based on taguchi method in deep learning with long short-term memory*, *IEEE Access*, 6, 2018, 45923–45933.
- [459] Yuan Q., Chen I., et al., *Localization and velocity tracking of human via 3 IMU sensors*, *Sensors and Actuators A: Physical*, 212, 2014, 25–33.
- [460] Yuan X.C., Pun C.M., Chen C.P., *Robust mel-frequency cepstral coefficients feature detection and dual-tree complex wavelet transform for digital audio watermarking*, *Information Sciences*, 298, 2015, 159–179.
- [461] Yun X., Bachmann E.R., *Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking*, *IEEE Transactions on Robotics*, 22, 6, 2006, 1216–1227.
- [462] Yun X., Bachmann E.R., McGhee R.B., *A simplified quaternion-based algorithm for orientation estimation from earth gravity and magnetic field measurements*, *IEEE Transactions on Instrumentation and Measurement*, 57, 3, 2008, 638–650.
- [463] Yun X., Calusdian J., Bachmann E.R., McGhee R.B., *Estimation of human foot motion during normal walking using inertial and magnetic sensor measurements*, *IEEE Transactions on Instrumentation and Measurement*, 61, 7, 2012, 2059–2072.
- [464] Yuvaraj R., Murugappan M., Ibrahim N.M., Sundaraj K., Omar M.I., Mohamad K., Palaniappan R., *Detection of emotions in parkinson's disease using higher order spectral features from brain's electrical activity*, *Biomedical Signal Processing and Control*, 14, 2014, 108–116.
- [465] Zahra A., Kanwal N., ur Rehman N., Ehsan S., McDonald-Maier K.D., *Seizure detection from EEG signals using multivariate empirical mode decomposition*, *Computers in Biology and Medicine*, 88, 2017, 132–141.
- [466] Zhai G., Chen J., Li C., Wang G., *Pattern recognition approach to identify loose particle material based on modified MFCC and HMMs*, *Neurocomputing*, 155, 2015, 135–145.

Bibliografia

- [467] Zhang J., Welch G., Bishop G., Huang Z., *A two-stage Kalman filter approach for robust and real-time power system state estimation*, IEEE Transactions on Sustainable Energy, 5, 2, 2014, 629–636.
- [468] Zhang M., Hol J.D., Slot L., Luinge H., *Second order nonlinear uncertainty modeling in strapdown integration using MEMS IMUs*, [w:] *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, IEEE, 2011, 1–7.
- [469] Zhang N., Ding S., Zhang J., Xue Y., *An overview on restricted Boltzmann machines*, Neurocomputing, 275, 2018, 1186–1199.
- [470] Zhang Y., Chen K., Yi J., Liu T., Pan Q., *Whole-body pose estimation in human bicycle riding using a small set of wearable sensors*, IEEE/ASME Transactions on Mechatronics, 21, 1, 2016, 163–174.
- [471] Zhang Z., Xu Y., Yang J., Li X., Zhang D., *A survey of sparse representation: algorithms and applications*, IEEE Access, 3, 2015, 490–530.
- [472] Zhang Z.O., Meng X., *Use of an inertial/magnetic sensor module for pedestrian tracking during normal walking*, IEEE Transactions on Instrumentation and Measurement, 64, 3, 2015, 776–783.
- [473] Zhang Z.Q., Wong W.C., Wu J.K., *Ubiquitous human upper-limb motion estimation using wearable sensors*, IEEE Transactions on Information Technology in Biomedicine, 15, 4, 2011, 513–521.
- [474] Zhang Z.Q., Wu J.K., *A novel hierarchical information fusion method for three-dimensional upper limb motion estimation*, IEEE Transactions on Instrumentation and Measurement, 60, 11, 2011, 3709–3719.
- [475] Zhao J., Philip L., Shi L., Li S., *Separable linear discriminant analysis*, Computational Statistics and Data Analysis, 56, 12, 2012, 4290–4300.
- [476] Zheng C., Tian X., Yang Z., Wang S., Fan Z., *Application of S-transform threshold filtering in anhui experiment airgun sounding data de-noising*, Geodesy and Geodynamics, 9, 4, 2018, 320–327.
- [477] Zheng J., Cheng J., Yang Y., *Partly ensemble empirical mode decomposition: an improved noise-assisted method for eliminating mode mixing*, Signal Processing, 96, 2014, 362–374.
- [478] Zhong F., Zhang J., *Linear discriminant analysis based on L1-norm maximization*, IEEE Transactions on Image Processing, 22, 8, 2013, 3018–3027.
- [479] Zhou H., Hu H., *Human motion tracking for rehabilitation – a survey*, Biomedical Signal Processing and Control, 3, 1, 2008, 1–18.
- [480] Zhou H., Hu H., *Reducing drifts in the inertial measurements of wrist and elbow positions*, IEEE Transactions on Instrumentation and Measurement, 59, 3, 2010, 575–585.
- [481] Zhou H., Stone T., Hu H., Harris N., *Use of multiple wearable inertial sensors in upper limb motion tracking*, Medical Engineering and Physics, 30, 1, 2008, 123–133.
- [482] Zhu R., Zhou Z., *A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 12, 2, 2004, 295–302.
- [483] Zidelmal Z., Amirou A., Ould-Abdeslam D., Moukadem A., Dieterlen A., *QRS detection using S-transform and Shannon energy*, Computer Methods and Programs in Biomedicine, 116, 1, 2014, 1–9.

- [484] Zieliński T.P., *Cyfrowe przetwarzanie sygnałów: od teorii do zastosowań*, Wydawnictwa Komunikacji i Łączności, Warszawa 2007.
- [485] Zou R., Wang H., Chon K.H., *A robust time-varying identification algorithm using basis functions*, *Annals of Biomedical Engineering*, 31, 7, 2003, 840–853.
- [486] Zunino L., Pérez D., Kowalski A., Martín M., Garavaglia M., Plastino A., Rosso O., *Fractional Brownian motion, fractional Gaussian noise, and Tsallis permutation entropy*, *Physica A: Statistical Mechanics and its Applications*, 387, 24, 2008, 6057–6068.
- [487] Zygarlicka M., *Wybrane metody przetwarzania obrazów w analizach czasowo-częstotliwościowych na przykładzie zakłóceń w sieciach elektroenergetycznych*, praca doktorska, Politechnika Opolska, Opole 2011.

Skorowidz

- agregacja danych, 34
- aktuator, *zob.* urządzenie wykonawcze
- akwizycja danych, 73
- Algorytm
 - Sieci Neuronowe
 - Głębokie, 310
 - Warunkowa ograniczona maszyna Boltzmann, 311
- algorytm
 - k*-Najbliższych Sąsiadów, 306
 - analizy głównych składowych, 284
 - analizy korelacji kanonicznych, 290
 - analizy składowych niezależnych, 285
 - autoenkodera, 293
 - BKS, 332
 - Borda, 334
 - CCA, *zob.* algorytm analizy korelacji kanonicznych
 - Dempstera–Shafera, 336
 - EMD, 168
 - estymacji orientacji ciała sztywnego, 389
 - LDA, *zob.* algorytm liniowej analizy dyskryminacyjnej
 - liniowej analizy dyskryminacyjnej, 287
 - maszyna wektorów nośnych, 309
 - PCA, *zob.* algorytm analizy głównych składowych, *zob.* algorytm analizy składowych niezależnych
 - regresja logistyczna, 309
 - rozkładu według wartości osobliwych, 284
 - SSA, *zob.* algorytm widma osobliwego
 - SVD, *zob.* algorytm rozkładu według wartości osobliwych
 - ślepa separacja źródeł, 285
 - widma osobliwego, 291
- analiza
 - spektralna, *zob.* analiza widmowa
- analiza widmowa, 243
- architektura
 - centralna, 61, 324
 - hybrydowa, 61
 - rozproszona, 61, 324
- autoenkoder, *zob.* algorytm autoenkodera
- autoenkoder odsumiający, 182
- bezprzewodowa sieć czujnikowa, 23
- blokada przegubów, 348
- cech
 - ekstrakcja, 56, 224
 - generowanie, 56, 224, 235
 - konstrukcja wektora, 56
 - redukcja, 57, 224, 283
 - selekcja, 56, 57, 224, 280
 - wektor, 57, 223
- chwilowa
 - amplituda, 257
 - częstotliwość, 257
 - częstotliwość fazowa, 257
 - faza, 257
- ciało człowieka
 - antropometria, 384
- ciało ludzkie
 - kość, 372
 - staw, 372
- ciało sztywne, 341

Skorowidz

- obrót, 343
- orientacja, 342
- położenie, 342
- przemieszczenie, 356
- prędkość kątowna ruchu, 360
- prędkość ruchu, 360
- ruch, 341, 355
- Coriolisa
 - efekt, 111
 - siła, 111
- czas, 341
- czujnik
 - IMU, 31
 - magnetometr, 116
 - magnetometryczny, 31
 - MARG, 31
 - przyśpieszenia, 31, 105
 - żyroskopowy, 31, 111
- częstotliwość próbkowania, 54
- dane
 - niekompletne, 302
 - niezbalansowane, 302
 - pomiarowe
 - jednomodalne, 53
 - strumieniowe, 35
 - wrażliwe, 22
- detekcja zmian, 226, 231
- dryft, 41, 413
- dyskretyzacja, 216
- dziedzina
 - czasowo-częstotliwościowa, 56
 - czasu, 56
 - częstotliwości, 56
- eliminacja
 - progowa, 179
- Empirical Mode Decomposition, *zob.* algorytm EMD
- entropia
 - przybliżona, 269
 - próbkowana, 269
 - Rényi, 270
 - Tsallisa, 241
- estymacja
 - dryft, 436
 - orientacja, 435
 - orientacji ciała sztywnego, 388
 - prędkość, 436
- filtr
 - Bayesa, 197
 - bezsładowy Kalmana, 208
 - cząsteczkowy, 214
 - Kalmana, 55, 199
 - Kalmana, rozszerzony, 204
 - nieskończonej odpowiedzi impulsowej, 141
 - skończonej odpowiedzi impulsowej, 140
 - Wienera, 138
- Filtr Kalmana
 - rozszerzony, 205
- filtracja
 - optymalna, 190
 - progowa
 - mięka, 180
 - twarda, 180
- fuzja
 - bepośrednia, 53, 189
 - cech, 53
 - danych, 33, 53
 - modele, 64
 - strumieniowych, 35
 - decyzji, 53
 - informacji, 34
 - komplementarna, 60
 - wielosensoryczna, 34
- fuzja danych, 33
- głosowanie
 - większościowe, 330
 - większościowe ważone, 331
- integracja wielosensoryczna, 34
- kalibracja, 410
- kinematyka, 363, 369
 - model kończyny dolnej, 381
 - model kończyny górnej, 377
 - proste zadanie, 375

- zadanie odwrotne, 369
- zadanie proste, 369
- zadanie proste, rozwiązanie, 370
- zadanie proste, sformułowanie, 369
- klasyfikacja, 56, 298
- konfiguracja
 - komplementarna, 59
 - nadmiarowa, 59
 - współpracy, 59
- konwolucja, *zob.* spłot
- kończyna
 - dolna, 372
 - górna, 372
- kwaternion, 350, 436, 448
 - jednostkowy, 350
- kąty
 - Eulera, 345
 - Eulera, RPY, 348
 - Eulera, ZYZ, 345, 346
- łańcuch kinematyczny, 364
 - prosty otwarty, 364
 - prosty zamknięty, 364
 - złożony, 364
- macierz
 - antysymetryczna, 353
 - kowariancji, 198
 - odległości, 307
 - rotacji, 343, 366
 - transformacji jednorodnej, 357, 366, 368
- Markowa
 - warunek, 193
- medycyna zaangażowana, 23
- metoda
 - Blackmana–Tukeya, 244
 - Monte Carlo, 214
 - MUSIC, 246
 - PCA, 284
 - podprzestrzeni, 246
 - Welcha, 244
- metody
 - nieparametryczne, 58
 - parametryczne, 58
- mikrosystem, *zob.* mikrouządzenie²³
- mikrouządzenie
 - elektromechaniczne, 23
- mnożenie
 - skalarane, 341
 - wektorowe, 341
- mnożnik Lagrange’a, 181
- model
 - blokowo-zorientowany, 153
 - Boyda, 67
 - Hammersteina, 153
 - JDL, 64
 - odcinkowo-liniowy, 151
 - przestrzeń stanu, 175
 - systemu, 192
 - typu wodospad, 69
 - typu zmodyfikowany wodospad, 69
 - układu pomiarowego, 193
 - w przestrzeni stanu, 194
 - Wienera, 153
- normalizacja, 57, 134
- ograniczenia
 - biomechaniczne, 415
- okno
 - przesuwne, 300
 - ważone, 300
 - z punktem początkowym, 299
 - zmiennej skali czasowej, 300
- operator Teagera–Kaisera, 239
- orientacja
 - ciała sztywnego, 388
- osobiste sieci czujnikowe, 24
- para kinematyczna, 364
 - klasa, 372
- periodogram, 243
 - uśredniony, 244
 - zmodyfikowany, 243
- pętla Boyda, *zob.* model, Boyda⁶⁷
- pomiar
 - synchronizacja, 54
- potencjał ruchowy człowieka, 371
- położenie

- ciała sztywnego, 388
- predykcja, 190
- problem
 - Wahba, 388
- profil decyzyjny, 334
- przeciek widma, 244
- przegub, 363
 - obrotowy, 363
 - przesuwny, 363
- przekształcenie bezśladowe, 208, 210
- przekształcenie sztywne, 343
- przestrzeń, 341
 - konfiguracyjna, 342
- przestrzeń fizyczna, *zob.* przestrzeń
- przestrzeń stanu, 191
 - model, 192
- przesunięcie liniowe, 364
- przetwarzanie
 - blokowe, 300
 - przyrostowe, 300
- przetwarzanie wstępne, 55
- punkt
 - materialny, 342
- reprezentacja
 - Denavita–Hartenberga, 365
 - Denavita–Hartenberga, zmodyfikowana, 365
 - oś-kąt, 344
 - rzadka, 162
 - wykładnicza, 365, 366
- rotacja
 - elementarna, 345
- rozmaitość
 - gładka, 356
- ruch
 - obrotowy, 355
 - postępowy, 355
- ruch ciała ludzkiego
 - analiza struktury dynamicznej, 371
 - analiza struktury kinematycznej, 371
- ruch śrubowy, 367
- ruchliwość
 - wzajemna, 372
- równanie
 - oscylatora, 106
- samopodobieństwo, 262
- segmentacja, 56, 224, 226, 437
- sieć
 - głęboka, 182
- siła
 - wewnętrzna, 371
 - zewnątrzna, 371
- skrzętnik, 359
- specjalna grupa euklidesowa, 355
 - algebra Liego, 359
- specjalna grupa obrotów
 - algebra Liego, 359
- specjalna grupa ortogonalna, 343
 - algebra Liego, 343
- splot, 141
- stan
 - estymacja, 189, 190
 - układu fizycznego, 342
 - wektor, 192
- standaryzacja, 134
- sygnał, 74
 - analityczny, 257
 - ciągły, 76
 - czasowy, 74
 - dyskretny, 76
 - nieokresowy, 75
 - niestacjonarny, 75, 76
 - niezdecydowany, 75
 - okresowy, 75
 - przestrzenno-czasowy, 74
 - przestrzenny, 74
 - rzadka reprezentacja, 449
 - rzadki, 162
 - stacjonarny, 75
 - zdecydowany, 75
 - zespolony, 257
- synergia, 33
- system
 - ciał sztywnych, 363
 - rozproszony, 32
 - scentralizowany, 32
- system, wspomagający nawigację pieszych, 447

- szablony decyzyjne, 336
- szereg
 - Taylora, 354
 - Volterry, 152
- szum
 - biały, *gaussowski*, 89
 - Browna, *zob.* szum czerwony
 - cieplny, 92
 - czerwony, 94
 - filoletowy, 98
 - kanałowy, 92
 - lawinowy, 92
 - niebieski, 97
 - różowy, 95
 - śrutowy, 92
 - wewnętrzny, 74
 - wybuchowy, 92
 - zewewnętrzny, 74, 90
- szum kwantyzacji, 123
- tablica pomyłek, 312
- telemedycyna, 26
- transformacja
 - beźśladowa, *zob.* przekształcenie beźśladowe
 - falkowa, 83, 147, 164, 259
 - Fouriera, 146, 259
 - Gabora, 83, 147
 - Hilberta, 167
 - Hilberta–Huanga, 83, 146, 167, 259
 - jednorodna, 356
 - krótkoczasowa, Fouriera, 83
 - S, 259
 - Wignera–Ville, 259
- transformacja Stockwella
 - seealsotransformacja S, 259
- trening sportowy, 25
- uczenie
 - bez nadzoru, 304
 - częściowo nadzorowane, 305
 - nadzorowane, 303
- ukryty model Markowa, 216
- układ
 - fizyczny, 342
 - odniesienia, 342
- układ ruchu człowieka
 - kinematyka, 372, 375, 409
 - metody analizy, 39
- układ współrzędnych
 - kartezyjski, 365
- urządzenie
 - pomiarowe, *zob.* czujnik
 - wykonawcze, 103
- wariancja
 - Allana, 121
- wektor
 - położenia, 343, 366
- widmowa gęstość mocy, 121
- współczynnik
 - MAE, 100
 - PSNR, 100
 - RMSE, 100
 - SNR, 100
- współrzędne
 - jednorodne, 356
 - wykładnicze, 353, 357
- wygładzanie, 190
- wykładnik
 - Hursta, 273
 - Lapunowa, 270
- wykładnik Hursta, 178, 184
- wymiar
 - fraktalny, 262
 - korelacyjny, 267
- wzór
 - Rodrigueza, 354
- zakłócenia, *zob.* szum zewnętrzny
 - magnetyczne, 414
 - w czujnikach pomiarowych, analiza, 120
- zmienna
 - złączowa, 364
- złącze, *zob.* przegub

Application of signal processing in the data stream fusion

Data fusion (DF) is the process of combining data from a number of different data sources in order to provide a complete description of the object of interest. An important attribute of data fusion is that it increases the overall description of the object compared to the description produced by classical approaches. It is possible since DF provides algorithms and methods for creating a synergy of the data acquired from multiple sources. The synergy results in better accuracy and robustness of the object's description. These attributes made data fusion a special significance and importance in many applications where a large amount of data must be processed and fused intelligently. In this book, we focus on the issues of data fusion in the context of signal processing. The first part of the work is related to fundamental concepts of data fusion. We discussed in details basic architectures of data fusion systems and algorithms. Additionally, in this part of the book, the taxonomy of the DF methods is provided. To this end, a typical classification based on the abstraction levels is introduced, i.e., low-level (fusion of data), middle-level (fusion of features) and high-level (fusion of decision).

This part of the book approaches a question related to signal representations, methods for the analysis of noise types and the modeling of data sources. Signal representations are considered in the viewpoint of three typical domains, i.e., time (T), frequency (F) and time-frequency (TF). TF means that the signal is represented over both time and frequency. Some parts of the book were also related to a modern approach of signal representation which is based on the sparse representation. In general, the sparse representation of signals is the most compact representation of a signal in terms of a linear combination of atoms in an overcomplete dictionary. One of the popular and widely-used dictionaries applied for sparse representation is based on the wavelet transform. In the book, we also consider other methods based on Empirical Mode Decomposition which is one of the original contributions in signal processing-based data fusion.

One of the chapters of this part of the book is also related to the methods of noise analysis. In this section, we focus on classical Spectral density and Allan Variance methods. The last section of this part is connected to the issues of data sources modeling. Since the main problems of this book are related to the problems of processing the data from IMU and MARG sensors, we discussed the problems of modeling data acquired from accelerometers, gyroscopes, and magnetometers. The second part of the book is devoted to the methods of signal processing-based data fusion. The algorithms in this part were divided into pre-processing methods and low-, middle-, and high-level processing methods. The pre-processing methods used to remove unwanted components of raw signals in order to improve the signal-to-noise ratio. The main algorithms presented in this part

Application of signal processing in the data stream fusion

are the methods of signal denoising. We also discuss the problem of normalization and standardization of acquired signals.

At the low-level processing, we mostly focus on methods based on Kalman filter and its modifications, i.e., Extended Kalman filter, Unscented Kalman filter, and particle filtering. Middle-level data fusion is related to the problem of feature sets extracted from multiple data sources in order to create a new feature vector that represents the input data. Thus, we start this part of the work with a review of the relevant methods of feature extraction algorithms. We concentrated on the methods of nonlinear methods extraction. At the end of this part, we discuss the problem of feature fusion which is based on classification methods. The last group of problems considered in the book's fragment is connected to decision fusion. Generally, decision-level fusion is a unique decision made based on local decisions of multiple data sources. It is an important phase of various data fusion systems since it allows the combination of the heterogeneous signals acquired from various data sources.

The last part of the work is related to the problem of human motion tracking and estimation based on the data acquired from IMU and MARG sensors. At the beginning of this book's fragment, the fundamental problems of human motion modeling are considered. This part of the work mostly based on a theory known from robotics. Subsequently, we focus on the methods of human motion tracking and estimation based on the data gathered via sensors. At the end of the work, we present some original results related to the data fusion of streaming data acquired from IMU and MARG sensors.

W książce *Zastosowanie przetwarzania sygnałów w fuzji danych strumieniowych* przedstawiono zagadnienia dotyczące fuzji danych strumieniowych zarówno od strony teoretycznej, jak i praktycznej. Pojęcie fuzji danych strumieniowych obejmuje teorie, metody, techniki i narzędzia wykorzystywane do łączenia danych pochodzących z różnych źródeł (np. czujników pomiarowych) w celu uzyskania lepszego opisu obserwowanego obiektu w porównaniu z opisem uzyskanym z wykorzystaniem tradycyjnych podejść, w których dane z wielu źródeł przetwarzane są oddzielnie.

Konstruowane algorytmy fuzji danych strumieniowych można traktować jako zestaw odpowiednio dobranych i przystosowanych metod, które zaczerpnięto z innych obszarów, takich jak np. przetwarzanie sygnałów, teoria estymacji stanu, rozpoznawanie obrazów, identyfikacja systemów czy podejmowanie decyzji. W monografii omówiono więc zagadnienia związane m.in. z filtrem Kalmana, algorytmami klasyfikacji, a także z metodami wstępnego przetwarzania sygnałów. Szczególną uwagę zwrócono na algorytmy, w których wykorzystuje się różne metody reprezentacji sygnałów, w tym metody oparte na reprezentacjach czasowo-częstotliwościowych oraz rzadkich.

W części książki dotyczącej praktycznych aspektów fuzji danych strumieniowych skupiono się na zagadnieniach związanych z analizą ruchu człowieka. Zamieszczono również wiele przykładów wykorzystania metod i algorytmów fuzji danych strumieniowych w zadaniach śledzenia oraz rozpoznawania aktywności fizycznej osób. W monografii zaprezentowano także możliwości wykorzystania fuzji danych w problemie kinematycznej analizy układu ruchu człowieka.

Aktualność przedstawionej w monografii problematyki sprawia, że może być ona traktowana jako spójne podejście do konstrukcji algorytmów przetwarzania danych pochodzących z różnych źródeł w nowoczesnych systemach informatycznych.



Wydawnictwa Politechniki Wrocławskiej są do nabycia w księgarni
ul. C.K. Norwida 9, 50-374 Wrocław, tel. 71 328 08 95
Prowadzimy sprzedaż wysyłkową: zamawianie.ksiazek@pwr.edu.pl

ISBN 978-83-7493-099-4