Optica Applicata, Vol. XLII, No. 4, 2012
DOI: 10.5277/oa120418

Person identification system using an identikit picture of the suspect

JANUSZ BOBULSKI*, MARIUSZ KUBANEK

Institute of Computer and Information Sciences, Częstochowa University of Technology, Dąbrowskiego 73, 42-201 Częstochowa, Poland

*Corresponding author: januszb@icis.pcz.pl

The article presents a person identification system, which may work with an identikit picture. The identikit picture (sketch) is often used in practice as an investigative tool to search for the perpetrators of an unknown identity. With a portrait of the perpetrator of a crime, one may identify the criminal. When the face database for comparisons is large, this is labour-absorbing. With the help of a computer system of face identification, this process becomes quick and easy.

Keywords: person identification, face recognition, biometrics, identikit picture.

1. Introduction

An identikit picture is often used in practice as an investigative tool to search for the perpetrators of an unknown identity. A contemporary identikit picture has two forms: descriptive description, and the image that is a form of visual looks of the person's face [1].

The identikit picture is an important part of an investigation, because in the absence of a real image it allows to identify the suspect. This helps the police to find the trail through which the offender may be covered. In the past, portraits were made in the form of an identikit picture drawing based on the testimony of witnesses or victims. It is currently being established by a computer, which adjusts the relevant elements of the face such as a nose, eyes, lips, eyebrows, face shape by using ready patterns from a database, and then generates the resulting facial image (Fig. 1).

It is possible to make the identification with a portrait of the perpetrator. When the face database for comparisons is large, this is becoming tedious. With the help of a computer system of face identification, this process becomes quick and easy [2, 3].

Computer faces identification systems use mathematical tools to compare the pattern. Such systems operate in two modes: training and testing. The first mode is used to

Fig. 1. Sample of identikit picture.

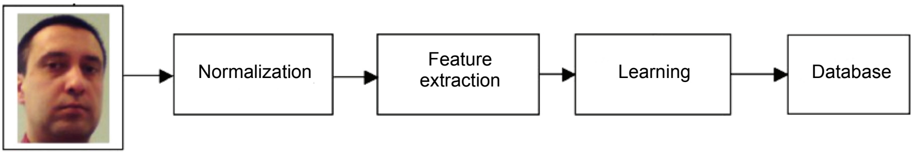


Fig. 2. Scheme of identification system in learning mode.

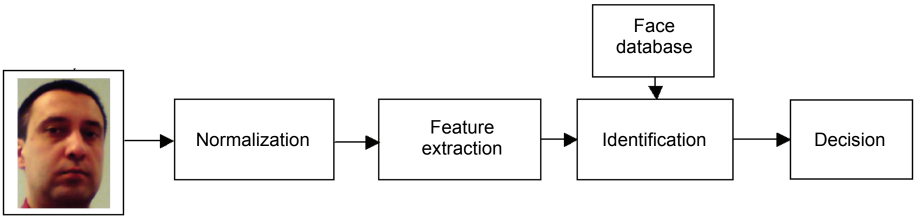


Fig. 3. Scheme of identification system in testing mode.

enter the pattern to the base faces (Fig. 2). However, the second mode is used to identify the suspect, that is to prepare a comparison portrait of the perpetrator with the images from the database and point the most similar (Fig. 3).

2. User identification

The user identification process works on the base of the frontal face image. The fusion of wavelet transformation (WT) and hidden Markov models (HMM) is used for processing the three face parts (eyes, nose, mouth). The likelihood of observation generating is calculated for each face part. The decision is made on the basis of the sum maximalization of likelihood.

The most popular method of face identification is principal component analysis (PCA) [2, 4]. Other popular methods use WT [2, 5] or HMM [2, 6]. The analysis of the existing solutions revealed their defects, which caused their weak effectiveness. The disadvantages of these methods are as follows:

– In the case of the new user's registration, the process of learning and addition of his/her facial image to a database, requires repeated learning of the whole system;
– They work with the whole face;
– They are computationally very expensive.

The proposed method is a combination of two mathematical tools, WT and HMM. Both were mainly used for speech recognition. Here, WT is used for features extraction, and HMM for identification. This system works in two modes, learning and testing. These modes differ from each other. The algorithm of this method consists of four main parts:

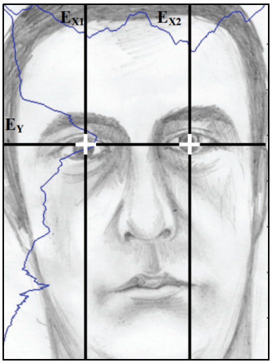
– Pre-processing: normalization and face division into three parts;
– Features extraction: WT of the face image;
– Learning: generating and learning HMM for each part of the face,
  Testing: testing models from the database;
– Learning: saving to database the learned models of the face,
  Testing: making a decision – maximum likelihood of the model.

2.1. Pre-processing

The normalization consists of fixing the eyes centres, and then of respective face scaling so that the distance between them equals 120 pixels [7]. This is a necessary process by which it is possible to compare objects, such as in this case faces. The idea of standardization is the appropriate scaling of facial images, so that they could be compared. The first step in the standardization process is to determine the eyes centres. This is done by calculating the gradient of the whole image, and then by summing these values for each row and column (Fig. 4).

$$\begin{cases} W_1 = \sum_{i=1}^{i-1} \left| m_{ij} - m_{i+1,j+1} \right| \\ W_2 = \sum_{j=1}^{j-1} \left| m_{ij} - m_{i+1,j+1} \right| \end{cases} \tag{1}$$

$$\begin{cases} E_Y = \max(W_1) \\ E_{X1} = \max\left(W_2\left(1 : \dfrac{j}{2}\right)\right) \\ E_{X2} = \max\left(W_2\left(\dfrac{j}{2} : j\right)\right) \end{cases} \tag{2}$$



◀ Fig. 4. Determine the eyes centres.

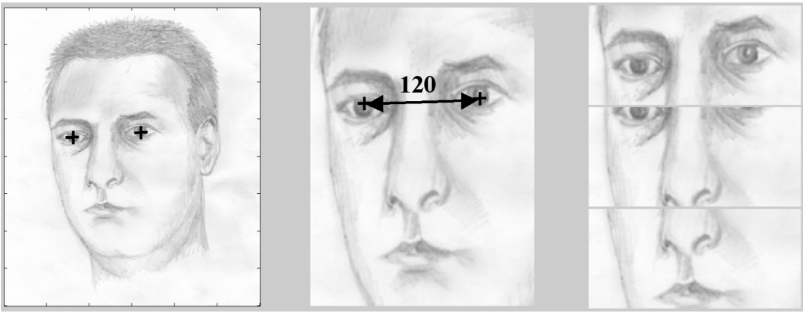


Fig. 5. Pre-processing of the face image.

The second part of this process is a division of the normalized face into three parts: the area of eyes, nose, and mouth (Fig. 5).

2.2. Features extraction

WT is used for features extraction. Using 2D WT (Fig. 6), the face image is decomposed into four sub-images via the high-pass and low-pass filtering. The image is decomposed along the column direction into sub-images to the high-pass frequency band H and the low-pass frequency band L. Assuming that the input image is a matrix of $m \times n$ pixels, the resulting sub-images become $m/2 \times n$ matrices. At the second step, the images H and L are decomposed along the row vector direction and respectively produce the high and low frequency band HH and HL for H, and LH and LL for L. The four output images become the matrices of $m/2 \times n/2$ pixels. The low frequency sub-image LL (A_1) possesses high energy, and is a smaller copy of the original images (A_0). The remaining sub-images LH, HL, and HH respectively extract the changing components in the horizontal (D_{11}), vertical (D_{12}), and diagonal (D_{13}) direction [8].

Wavelet transform of the second level (Fig. 7) is used for features extraction in the proposed technique. After the first level wavelet decomposition, the output images

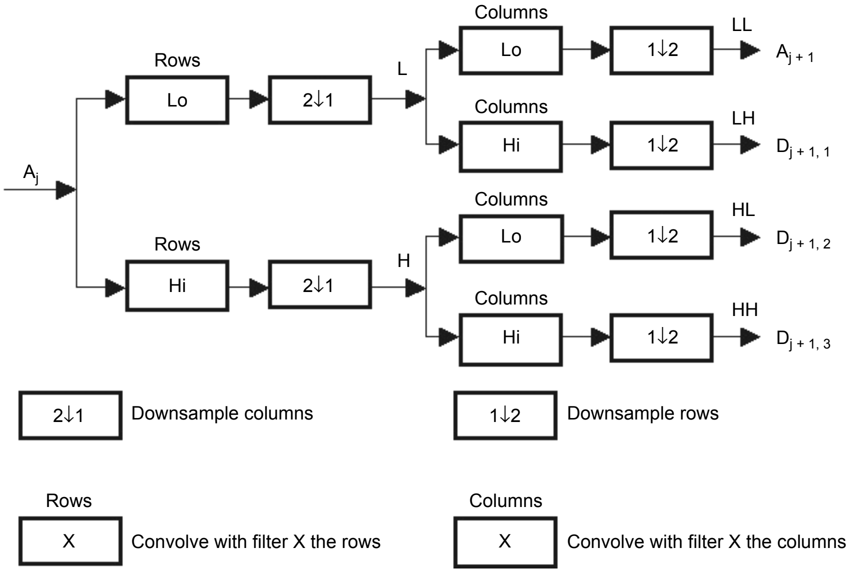


Fig. 6. Scheme of one-level two-dimensional wavelet transform [12].

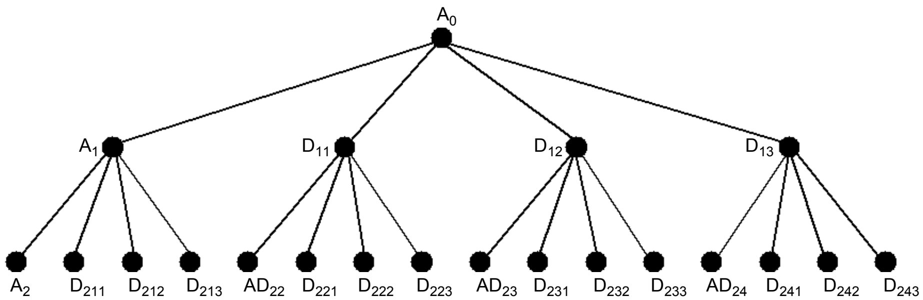


Fig. 7. The wavelet decomposition tree.

become the input images of the second level decomposition. The results of two-level 2D WT are coded in this way, so that they can be applied in HMM (Fig. 8). One of the simplest methods of reduction and information coding is to calculate the standard deviation or mean value. Each face part is transformed separately by discrete wavelet transform (Fig. 9). The bank filters' selection is an important thing in this transformation. It guarantees a good recognition rate [9].

2.3. Training system

HMM is used for the identification process. The HMM is a double stochastic process with an underlying stochastic process that is not observable (hidden), but may be

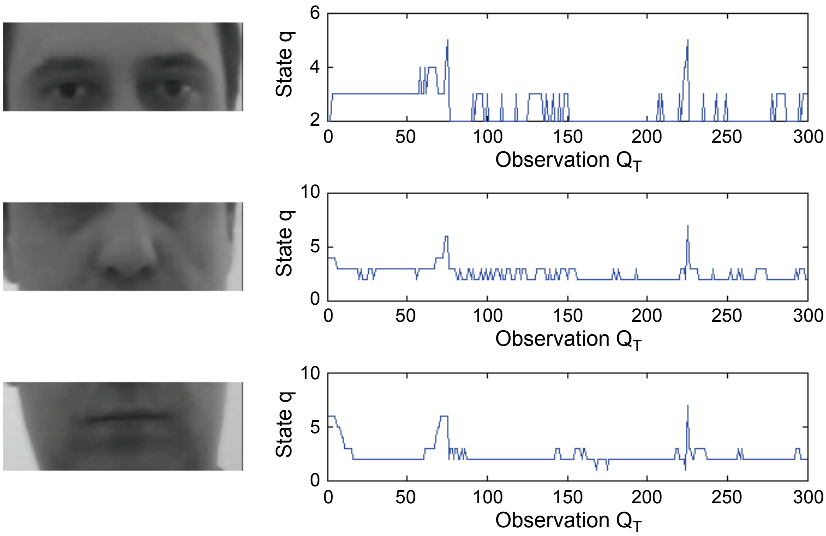


Fig. 8. Face part and corresponding sequences of observation.

Fig. 9. Example of 2nd level of the wavelet decomposition of eyes area image.

observed through another set of stochastic processes that produce a sequence of observation.

Let $O = \{O_1, ..., O_T\}$ be the sequence of observation of feature vectors, where T is the total number of feature vectors in the sequence. The statistical parameters of the model may be defined as follows [10]:

– The number of states of the model N;

– The transition probabilities of underlying Markov chain, $A = \{a_{ij}\}$, $1 \leq i, j \leq N$, where a_{ij} is the probability of transition from state i to state j subject to the constraint $\sum_{j=1}^{N} a_{ij} = 1$;

– The observation probabilities, $B = \{b_j(O_T)\}$, $1 \leq j \leq N$, $1 \leq t \leq T$ which represents the probability of t-th observation conditioned on j-th state;

– The initial probability vector, $\Pi = \{\pi_i\}$, $1 \leq i \leq N$.

Hence, the HMM requires three probability measures to be defined, A, B, π and the notation $\lambda = (A, B, \pi)$ is often used to indicate the set of parameters of the model. In the proposed method, one model is made for each part of the face. The parameters of the model are generated at random at the beginning. Then they are estimated with Baum–Welch algorithm, which is based on the forward–backward algorithm. The forward algorithm calculates the coefficient $\alpha_t(i)$ (probability of observing the partial sequence $(o_1, ..., o_t)$ such that state q_t is i). The backward algorithm calculates the coefficient $\beta_t(i)$ (probability of observing the partial sequence $(o_{t+1}, ..., o_T)$ such that state q_t is i). The Baum–Welch algorithm, which computes λ, can be described as follows [11]:

1. Let initial model be λ_0;
2. Compute new λ based on λ_0 and observation O;
3. If $\log[P(O|\lambda) - \log(P(O)|\lambda_0] <$ DELTA stop;
4. Else set $\lambda_0 \leftarrow \lambda$ and go to step 2.

The parameters of the new model λ, based on λ_0 and observation O, are estimated from the equation of Baum–Welch algorithm [11], and then are recorded to the database.

2.4. Testing system

The testing process consists of computing the probability of observation generated by the models saved in the database and choosing the model with a maximal likelihood. In the proposed method, probabilities are calculated separately for each of the three models representing face parts, and then they are added. The face, for which the sum of probability is maximum, is chosen as the correct face. The probability of generating sequences of observations is computed from the following equations [11]:

$$P(O|\lambda) = \sum_q P(O|q, \lambda) P(q|\lambda) \tag{3}$$

$$P(O|q, \lambda) = \prod_{i=1}^{T} P(o_t|q_t, \lambda) = b_{q1}(o_1) b_{q2}(o_2) \ldots b_{qT}(o_T) \tag{4}$$

$$P(q|\lambda) = \pi_{q1}\, a_{q1q2}\, a_{q2q3} \ldots a_{q(T-1)qT} \tag{5}$$

$$PF = \sum_{i=1}^{3} P(O_i|\lambda_i) \tag{6}$$

3. Experiments

The criminals usually have one image in the police face database. The services like police do not have many images of suspects. They make the identikit picture during the investigation, which may be compared with a single image from the database. This

Table 1. Comparison recognition rate.

Method	Face dataset	No. of testing images	Top one match score [%]
2DPCA [13]	AR	600	74.8
1D-DHMM [14]	AR	1440	89.8
(PC)2A [15]	FERET	200	83.5
E(PC)2A [16]	FERET	200	85.5
Modular FLDA [17]	FERET	200	86.5
Component LDA [18]	FERET	350	78.6
EBGM [19]	FERET	1196	95.0
Discriminant PCA [20]	FERET	914	72.0
Proposed method	Forensic	1563	90.0

situation makes the identification of a suspect difficult. The problem has been described in details [2].

The experiments testing the presented method were carried on the face database of the Forensic Laboratory of the Regional Police Headquarters in Katowice (Poland). There was used one image per person for learning and one for testing. There were 1563 persons' images (size 300×400 pixels). The recognition rate of achieved results equalled 90%. The error rate (FAR) was 10%. The comparison to other methods was shown in Table 1.

4. Conclusions

The article presents a person identification system, which may work with identikit pictures. The recognition rate of the system is 90% and is comparable to other methods.

The system is sensitive to face rotation. Future works will concentrate on the elimination of this problem through the detection and the rotation of the face about the X axis, and the addition of other pose of face to eliminate the face rotatation about the axes X and Z.

On the basis of experimental research it was stated that the area of eyes contained the most useful information for the persons' identification, and it could be successfully applied in specific methods of identification or detection. The proposed method of user identification is characterized by the usage of the three face areas and creating for each of them one independent HMM (which can be used separately or together). This procedure makes possible a short calculation request and permits to obtain a good recognition rate.

References

[1] KOZIEŁ T., DĘBIŃSKI Z., *Portret obrazowy w identyfikacji i poszukiwaniu osób*, Problemy Kryminalistyki, No. 197–198, 1992, p. 10, (in Polish).

[2] XIAOYANG TAN, SONGCAN CHEN, ZHI-HUA ZHOU, FUYAN ZHANG, *Face recognition from a single image per person: A survey*, Pattern Recognition **39**(9), 2006, pp. 1725–1745.

[3] SONG CHEN, TAO ZHANG, CHENGPU ZHANG, YU CHENG, *A real-time face detection and recognition system for a mobile robot in a complex background*, Artificial Life and Robotics **15**(4), 2010, pp. 439–443.

[4] KIRBY M., SIROVICH L., *Application of the Karhunen–Loeve procedure for the characterization of human faces*, IEEE Transactions on Pattern Analysis and Machine Intelligence **12**(1), 1990, pp. 103–108.

[5] JEN-TZUNG CHIEN, CHIA-CHEN WU, *Discriminant waveletface and nearest feature classifiers for face recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(12), 2002, pp. 1644–1649.

[6] SAMARIA F, YOUNG S., *HMM-based architecture for face identification*, Image and Vision Computing **12**(8), 1994, pp. 537–543.

[7] Norm ISO/IEC JTC 1/SC 37 Biometric Data Interchange Formats – Part 5: Face Image Data.

[8] GARCIA C., ZIKOS G., TZIRITAS G., *Wavelet pocket analysis for face recognition*, Image and Vision Computing **18**(4), 2000, pp. 289–297.

[9] KOZUANI A.Z. HE F. SAMMUT K., *Wavelet packet face representation and recognition*, IEEE International Conference on Systems, Man and Cybernetics, Vol. 2, 1997, pp. 1614–1619.

[10] RABINER L.R., *A tutorial on hidden Markov models and selected applications in speech recognition*, Proceedings of the IEEE **77**(2), 1989, pp. 257–286.

[11] KANUNGO T., *Hidden Markov Model Tutorial*, www.cfar.umd.edu/kanungo

[12] *Wavelet Toolbox User's Guide, MatLab 5.0*, The MathWorks.

[13] JIAN YANG, ZHANG D., FRANGI A.F., JING-YU YANG, *Two-dimensional PCA: a new approach to appearance-based face representation and recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence **26**(1), 2004, pp. 131–137.

[14] LE H.-S., LI H., *Recognizing frontal face images using hidden Markov models with one training image per person*, Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), Vol. 1, 2004, pp. 318–321.

[15] JIANXIN WU, ZHI-HUA ZHOU, *Face recognition with one training image per person*, Pattern Recognition Letters **23**(14), 2002, pp. 1711–1719.

[16] SONGCAN CHEN, DAOQIANG ZHANG, ZHI-HUA ZHOU, *Enhanced (PC)2A for face recognition with one training image per person*, Pattern Recognition Letters **25**(10), 2004, pp. 1173–1181.

[17] SONGCAN CHEN, JUN LIU, ZHI-HUA ZHOU, *Making FLDA applicable to face recognition with one sample per person*, Pattern Recognition **37**(7), 2004, pp. 1553–1555.

[18] JIAN HUANG, YUEN P.C., WEN-SHENG CHEN, LAI J.H., *Component-based LDA method for face recognition with one training sample*, IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2003), 2003, pp. 120–126.

[19] WISKOTT L., FELLOUS J.-M., KRÜGER N., VON DER MALSBURG C., *Face recognition by elastic bunch graph matching*, IEEE Transactions on Pattern Analysis and Machine Intelligence **19**(7), 1997, pp. 775–779.

[20] JIE WANG, PLATANIOTIS K.N., VENETSANOPOULOS A.N., *Selecting discriminant eigenfaces for face recognition*, Pattern Recognition Letters **26**(10), 2005, pp. 1470–1482.

Received June 14, 2011
in revised form March 8, 2012