PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1064 2005
Pozyskiwanie wiedzy i zarządzanie wiedzą

Janusz Świerzowicz
Technical University of Rzeszów

MULTIMEDIA DATA MINING – PAST, PRESENT, AND FUTURE

1. Introduction

Development of information technology is particularly noticeable in methods and techniques of data acquisition. Data can be stored in many forms of digital media e.g. still images taken by digital camera, songs, mpeg, videos from desktop or cell phones video cameras. The capacity of digital storage data is doubled every nine month with respect to the price. Such data exceeds the total cumulative handwriting and printing during all of recorded human history.

According to current analysis carried out by IBM Almaden Research data volumes are growing at different speeds. The fastest one is Internet resources growth – it will achieve the digital online threshold of exabytes within few years [26]. In these fast growing volumes of data environment, restrictions are connected with a human's low data complexity and dimensionality analysis.

This paper deals with the multimedia data mining. It presents a short overview of application domains and practical applications of multimedia data mining multimedia data mining goals and methods used in knowledge acquisition. It focuses on supervised and unsupervised classification, uncovering interesting rules, decision trees, artificial neural networks, and rough-neural computing. It presents main multimedia data mining advantages, and discusses some critical issues as well as future trends.

One of the results of the inexorable growth of multimedia data volumes and complexity is a data overload problem. It is impossible to solve the data overload issue in a human manner; it takes strong effort to use intelligent and automatic software tools for turning rough data into valuable information and information into knowledge.

Data Mining is one of the central activities associated with understanding, navigating and exploiting the world of digital data. It is an intelligent and automatic process of identifying and discovering useful structures in data such as patterns, models and relations. We can consider data mining as a part of the overall Knowledge Discover in Data process. Kantardzic [15] defines Data Mining as "a process of discovering various models, summaries, and derived values from a given collection of data". It should be iterative and carefully planed process of using properly analytic techniques, to extract hidden, valuable information.

Data mining is essential as we struggle to solve data overload and complexity issues. With the fastest acceleration of offline data resources in the Internet, the WWW is a natural area for using data mining techniques to automatically discover and extract actionable information from Web documents and services. Those techniques are named Web Mining. We also consider Text Mining as a data mining task that help us summarize, cluster, classify and find similar text documents in a set of documents. Due to advances in informational technology and high performance computing, very large sets of images such as digital or digitalized photographs, medical images, satellite images, digital sky surveys, images from computer simulations and images generated in many scientific disciplines are becoming available. The method that deals with the extraction of implicit knowledge, image data relationship, and other patterns not explicitly stored in the image databases is called Image Mining [33]. A main issue of the image mining is to deal with relative data, implicit spatial information and multiple interpretations of the same visual patterns. We can consider application oriented functional approach and image driven approach. In the latter one the following hierarchical layers are established [34]: the lower layer that consists of pixel and object information and the higher layer that takes into consideration domain knowledge to generate semantic concepts from the lower layer and incorporates them with related alphanumerical data to discover domain knowledge.

The main aim of the multimedia data mining is an extraction of the interesting knowledge and understanding semantic captured in multimedia data that contain correlated images, audio, video and text.

Multimedia data bases, containing combinations of various data types could be first integrated via Distributed Multimedia Processors and then mined or one could apply data mining tools on the homogenous databases and then combine the results of the various data miners [28].

In the paper, after a short introduction to data mining we consider different kind of data, both structured, semistructured and unstructured with emphasize on the special role of multimedia. The remainder of the paper is organized as follows. The second section presents a short overview of multimedia data mining history. The next section discusses examples of practical and successful applications in different application domains, goals, methods and techniques used in multimedia data mining. This section focuses the brief discussion on supervised and

unsupervised classification, uncovering interesting rules, decision trees, artificial neural networks, and rough-neural computing. The next section presents future views of multimedia data mining and suggests how to overcome data mining critical issues. Finally, we conclude in the next section, together with a discussion of future work.

2. Past of the Multimedia Data Mining

Investigations on combining different media data, multimedia, into one application have begun as early as 1960's, when text and images were combined in a document. During the research and development process audio, video, and animation were synchronized using a time-line to specify when they should be played [24]. Since the middle 1990's problems of multimedia data capture, storage, transmission and presentation have extensively been investigated. Over the past few years research on multimedia standards (e.g. MPEG-4, X3D, MPEG-7), has continued to growth. These standards are adapted to represent very complex multimedia data sets, can transparently handle sound, images, videos, 3-D objects combined with events, synchronization, scripting languages and can describe the content of any multimedia object.

Different algorithms need to be used in multimedia distribution and multimedia database applications. As an example is given an image database, that stores pictures of birds, and a sound database that stores recording of birds [16]. The distributed query that asks for "top ten different kinds of birds that have black feathers and a high voice" is described in [6].

3. Multimedia Data Mining State of Art

This section presents multimedia data mining application domains and examples, basic multimedia data mining goals, methods some techniques and tools.

3.1. Application Domains and Examples

Multimedia data mining is successfully used in many application domains. It is applied for example in analysis of audio data, medical images, multimedia data available on the Internet, anthropometry data for apparel and transportation industries, movies data, pattern recognition, satellite images, security, spatiotemporal multimedia streams data, text extracting, segmenting and recognizing from multimedia data, and TV data [27].The most characteristic of these domains and some representative examples are presented below:

Biometrics

One of the most promising applications of multimedia data mining is biometrics. It refers to the automatic identification of an individual by using certain physiological or behavioural traits associated with the person [11] and combines

many human's traits of hand (hand geometry, fingerprints or palmprints), eye (iris, retina), face image or facial thermogram, ear, voice, gait, and signature for identify of an unknown user or verify a claimed identity.

Biometrical systems must solve numerous problems of noise in biometric data, modification of sensor characteristics, spoof and replay - attacks in various real-life applications.

A major area of research within biometric signal processing is face recognition. A face detection system that works with edge features of greyscale still images and the modified Hausdorff distance, as described in [14]. It is used as a similarity measure between a general face model and possible instances of the object within the image. The face detection module is a part of the multi modal biometric authentication system "BioID" described in [7]. Using multimedia data mining in multibiometric systems makes them more reliable due to the presence of multiple, an independent piece of human's traits information.

Medicine

An example of practical multimedia data mining application for medical images data mining is presented in [19]. They have used image data mining approach to formulate recommendation rules that help physicians to recognize gastroenterological diseases during medical examinations. A parallel environment for image data mining contains a pattern database. Each pattern in the database, considered as representative case, contains formalised text, numeric values and endoscopy image. During a patient examination, the automatic classification of the examined case is performed.

TV Data Mining

TV data mining is used in monitoring TV news, retrieve interesting stories, extracting face sequence from video sequence, extracting of soccer goal events [27]. Multimedia data mining can be applied for discovering structures in video news to extract topic of a sequence or persons involved in video. A basic approach of multimedia data mining presented in [4], is to separate the visual, audio and text media channels. The separated multimedia data include features extracted from the video stream e.g. visual spatial content (color, texture, sketch, shape), visual temporal content (camera or object motion), from the audio stream (loudness, frequency, timbre) and from the text information appearing on the screen. They focused on key-frame color mining in order to notice the appearance of important information on the screen and on discovering the presence of inlays in a key-frame. A key frame is a representative image of each shot i.e. a sequence of images on which there is no change of camera.

Chen et all [2] propose a framework that uses data mining combined with multi-modal processing in extracting the soccer goal events from soccer videos. It is composed of three major components, namely video parsing, data pre-filtering, and data mining. The integration of data mining and multimodal processing of

video is a powerful approach for effective and efficient extraction of soccer goal events.

Apparel and Transportation Industries

Mining anthropometry data for apparel and transportation industries is presented in [9]. A multimedia data base project, called CAESAR™ (i.e. Civilian American and European Surface Anthropometry Resource Project), consists of anthropometrical and statistical databases that contain data about worldwide population, 3-D scans of individuals' bodies, consumer habits, livestyle ect. In a multimedia data mining system, called Cleopatra, a clustering data mining technique is used to find similar individuals within the population, based on an archetype, i.e. a typical, real individual within the cluster.

3.2. Goals and Methods

In this section we present the most popular multimedia data mining goals and methods. In multimedia data mining, classification is mainly interpreted as object recognition. Object models (e.g. letters or digits) are known a priori, and automatic recognition system finds letters or digits from handwritten or scanned documents. Other examples are identification of images or scenarios on the base of sets visual data from photos, satellite or aero observations, finding common patterns in a set of images, identification of speakers and words in speech recognition. Image association rule mining is used for finding associations between structures and functions of human brain.

Dissecting a Set of Objects

One of the most popular goals in data mining is ordering or dissecting a set of objects described by high-dimensional data into small comprehensive units, classes, substructures, or parts. These substructures give better understanding and control, and can assign a new situation to one of these classes base on suitable information; which can be classified as supervised or unsupervised. In the former classification, each object originates from one of predefined classes and is described by a data vector [1]. But it is unknown to which class the object belongs, and this class must be reconstructed from the data vector. In unsupervised classification (clustering), a new object is classified into cluster of objects according to the object content without a priori knowledge. It is often used in the early stages of the multimedia data mining processes.

Uncovering Rules

If a goal of multimedia data mining can be expressed as uncovering interesting rules an association rule method is used. An association rule takes a form of an implication $X \rightarrow Y$, where X denotes antecedent of the rule, Y denotes consequent of the rule X, Y belong to the set of objects (itemset) I, $X3Y=\Phi$, and D denotes a set of cases [33]. We can determine two parameters named support s and confidence c. The rule $X \rightarrow Y$ has support s in D, where s% of the data cases in D

contains both X and Y and the rule holds confidence c in D, where c% of the data cases in D that support X also support Y. Association rule mining selects rules that have support greater than some user specified minimum support threshold (typically around 10^{-2} - 10^{-4}) and the confidence of the rule is at least a given (from 0 to 1) confidence threshold [17]. A typical association rule mining algorithms works in two steps. The first step finds all large itemsets that meet the minimum support constraint. The second step generates rules from all large itemsets that satisfy the minimum confidence constraints.

Decision Tree

One of a natural structure of knowledge is a decision tree. Each node in such a tree is associated with a test on the values on an attribute, each edge from a node is labelled with a particular value of the attribute, and each leaf of the tree is associated with a value of the class [23]. However, when the values of attributes for the description change slightly, the decision associated with the previous description can vary greatly. It is a reason to introduce fuzziness in decision trees to obtain fuzzy decision trees [18]. A fuzzy decision trees method, equivalent to a set of fuzzy rules "if then" represents a natural and understandable knowledge [4].

Pattern Recognition or Trend Prediction

In case goal of multimedia data mining is pattern recognition or trend prediction with limited domain knowledge, the Artificial Neural Network approach can be applied to construct a model of the data. Artificial Neural Networks can be viewed as a highly distributed parallel computing systems consisting of a large number of simple processors (similar to neurons) with many weighted interconnections. Neural network models attempt to use some organizational principles (such as learning, generalization, adaptively, fault tolerance and distributed representation and computation) in a network of weighted directed graphs in which the nodes are artificial neurons and directed edges (with weights) are connections between neuron outputs and neuron inputs [11, 12]. The main differences between neural networks and the other approaches to pattern recognition are that these networks have the ability to learn complex non-linear input-output relationships, and use sequential training procedures.

Dimensionality Reduction

In the needs of dimensionality reduction, the principal component analysis (PCA) often is performed [9, 25]. In this method, the square covariance matrix, that characterizes the training data set, is computed. The next, the eigenvalues are evaluated and arranged in decreasing order, with corresponding eigenvectors. Then the optimal linear transformation is provided to transform from the n dimensional space into m dimensional space, where $m \leq n$, and m is the number of most dominant, principal eigenvalues, that corresponds to the importance of each dimension. Dimensions corresponding to the smallest eigenvalues are neglected.

The optimal transformation matrix minimizes the mean last square reconstruction error. In addition to PCA, rough set theory can be applied for choose eligible principal components, that describe all concepts in a data set, for classification. An appropriate algorithm of feature extraction/selection using PCA and rough sets is presented by [25].

4. Future of Multimedia Data Mining

The investigations of multimedia data mining methods, algorithms, frameworks and standards should have an impact on the future research in this promising field of Information Technology. Moreover, multimedia data mining is currently still immature. As said Zhang [33] "The current images association rule mining are far from mature and perfection" Multimedia data are mostly mining separately [4]. Even if some standards used for multimedia data look very promising, it is too early to draw a conclusion about their usefulness in data mining.

In the future, the author expects that the framework would be made more robust and scalable to a distributed multimedia environment. Other interesting future work concerns the multimedia data mining standardization [20, 21].

In multimedia data, rare objects are often of great interest. These objects are much harder to identify than common objects. Weiss in [29] states, that the most of data mining algorithms have a great deal of difficulty dealing with rarity. Also the systems need to be evaluated against mining with rarity and testing of appropriate evaluation metrics. Finally the multimedia data mining implementations need to be integrated with the intelligent user interfaces.

Multimedia data mining can open new threats to informational privacy and information security, if not used properly. These activities can give occasion of new types of privacy invasion that may be achieved through the use of cyberspace technology for such things as dataveillance, e.i. surveillance by tracking data shadows which are left behind as individuals undertake their various electronic transactions, [13]. Further invasion can also be occasioned by secondary usage of data which individuals are highly unlikely to be aware of.

5. Concluding Remarks

This paper deals with some important issues of multimedia data mining. It presents a short overview of data mining goals and methods used in, advantages offered by multimedia data mining, application domains and examples of practical applications, critical issues and main multimedia data mining disadvantages that should be overcome in the future.

Research on text or images mining, carried on separately cannot be considered as multimedia data mining unless these media are combined. Multimedia research during the past decade has focused an audio and video media, but now, wider using

of multimodal interfaces and collection smart devices with embedded computers should generate flood of multimedia data, from which knowledge will be extracted using multimedia data mining methods. The social impact of multimedia data mining is also very important. New threats to informational privacy and information security can occur if these tools are not used properly.

Future works must concentrate on overcome multimedia data mining disadvantages such as lack of mature and perfection of current algorithms and frameworks, limited success in specific applications, lack of multimedia data mining standards, and on difficulty in dealing with data rarity.

References

[1] Bock H., *The Goal of Classification*, [in:] *Handbook of Data Mining and Knowledge Discovery*, Klosgen W., Zytkow J.M. (ed.), Oxford University Press, 2002, 254-258.

[2] Chen S.C., Shyu M.L., Chen M., Chengcui Z., *A Decision Tree-based Multimodal Data Mining Framework for Soccer Goal Detection*, IEEE International Conference on Multimedia and Expo (ICME 2004), Taipei, Taiwan.

[3] Ciaccia P., Patella M., *Searching in Metric Spaces with User-Defined and Approximate Distances*, ACM Transactions on Database Systems, 2002, 27(4), 398-437.

[4] Detyniecki M., Marsala C., *Fuzzy Multimedia Mining Applied to Video News*, IPMU 2002, July 1-5,2002, Annency, France, 1001-1008.

[5] Doorn M., de Vries A., *The Psychology of Multimedia Databases*, Proceedings of the fifth ACM Conference on Digital Libraries, 2000, 1-9.

[6] Fagin R., Wimmers E., *Incorporating User Preferences in Multimedia Queries*, Proceedings of the International Conference on Database Theory (ICDT), Volume 1186 of Lecture Notes in Computer Science (LNCS) Springer-Verlag, 1997, 247-261.

[7] Frischholz R.W., Werner A., *Avoiding Replay-Attacs in a Face Recognition System using Head-Pose Estimation*, Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03), 2003, 1-2.

[8] Han J., Kamber M., *Data Mining. Concepts and Techniques*, San Mateo, CA. Morgan Kaufmann., 2001

[9] Herena V., Paquet E., le Roux G., *Cooperative Learning and Virtual Reality-Based Visualization for Data Mining*, Data Mining. Opportunities and Challenges, J. Wang (ed.), IGP, 2003, Hershey, PA, 55-79.

[10] Hsu J., *Critical and Future Trends in Data Mining*, A Review of Key Data Mining Technologies/Applications. Data Mining Opportunities and Challenges, J. Wang (ed.), IGP, 2003, Hershey, PA, 437-452.

[11] Jain A.K., Ross A., *Multibiometric Systems*, CACM, 2004, 47(1), 34-40.

[12] Jain A.K., Duin R.P., Mao J., *Statistical Pattern Recognition*, A Review, MSU-CSE-00-5, 2000.

[13] Jefferies P., *Multimedia, Cyberspace & Ethics*, Proceedings of the IEEE International Conference on Information Visualization (IV'00), London, 2000, 99-104.

[14] Jesorsky O., Kirchberg K.J., Frischholz R.W., *Robust Face Detection Using the Hausdorff Distance*, Proc. Third International Conference on Audio- and Video-based Biometric Person Authentication. Volume 2091 of Lecture Notes in Computer Science (LNCS) Springer-Verlag, 2001, 90-95.

[15] Kantardzic M., *Data Mining. Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, 2003.

[16] Kossmann D., *The State of the Art in Distributed Query Processing*, ACM Computing Surveys, 2000, 32(4), 422-469.

[17] Mannila H., *Association Rules*. [in:] *Handbook of Data Mining and Knowledge Discovery*, Klosgen W., Zytkow J.M. (ed.), Oxford University Press, 2002, 344-348.

[18] Marsala C., *Fuzzy Decision Trees to Help Flexible Querying*, Kybernetika, 2000, 36(6), 689-705.

[19] Mazurkiewicz A., Krawczyk H., *A Parallel Environment for Image Data Mining*, Proceedings of the International Conference on Parallel Computing in Electrical Engineering (PARELEC'02), 2002.

[20] Melton J., Eisenberg A., *SQL Multimedia and Application Packages (SQL/MM)*, SIGMOD Record, 2001, 30(4), 97-102.

[21] Noirhomme-Fraiture M., *Multimedia Support for Complex Multidimensional Data Mining*, Procceedings of the First International Workshop on Multimedia Data Mining (MDM/KDD' 2000), in conjunction with Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD 2000, Boston, MA. ACM Press.

[22] Oliviera S.R.M., Zaiane O.R., *Toward Standardization In Privacy-Preserving Data Mining*, Proceedings of the 2nd International Workshop on Data Mining Standards, Services and Platforms, Grossman R.(ed.),KDD 2004, Seattle, 7-17.

[23] Quinlan J.R., *Introduction of decision trees. Machine Learning*, 1986,1(1), 86-106.

[24] Rowe L.A., Jain R., *ACM SIGMM Retreat Report on Future Directions In Multimedia Research*, ACM SIGMM, 2004, 1-12.

[25] Skowron A., Swinarski R.W., *Information Granulation and Pattern Recognition*, [in:] Rough-Neural Computing, Pal S.K., Polkowski L., Skowron A. (eds.), Springer-Verlag, 2004, 599-636.

[26] Świerzowicz J., *Decision Support System for Data and Web Mining Tools Selection. Issues and Trends of Information Technology Management in Contemporary Organizations*, Khosrow-Pour M. (ed). IGP, Hershey, London, 2002, 1118-1120.

[27] Świerzowicz J., Multimedia *Data Mining Concept: an Overview*, to Appear in Encyclopeda of Multimedia Technology and Networking, Pagani M. (ed). IGP, Hershey, London, 2005.

[28] Thuraisingham B., *XML Databases and the Semantic Web*, Auerbach Publications, 2002.

[29] Weiss G.M., *Mining with Rarity*, A Unifying Framework, SIGKDD Explorations, 2004,6(1), 7-19.

[30] Wijesekera D., Barbara D., *Multimedia applications*, [in:] *Handbook of Data Mining and Knowledge Discovery*, Klosgen W., Zytkow J.M. (ed.), Oxford University Press, 2002, 758-769.

[31] Zaiane O.R., Han J., Li Z., Hou J., *Mining MultiMedia Data*, In Proceedings of Meeting of Minds, CASCON'98, Toronto, Canada, 1-18.

[32] Zaiane O.R., Han J., Zhu H., *Mining Recurrent Items in Multimedia with Progressive Resolution Refinement*, [in:] Proceedings of the International Conference on Data Engineering ICDE'00, San Diego, CA, IEEE, 15-28.

[33] Zhang J., Hsu W., Li Lee M., *Image Mining. Issues, Frameworks and Techniques*, Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD' 2001) in conjunction with Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD 2001, Zaiane O.,R, Simoff S.J.(eds.) San Francisco, USA, ACM Press, 13-21.

[34] Zhang J., Hsu W., Li Lee M., *An Information driven Framework for Image Mining*, Proceedings of 12th International Conference on Database and Expert Systems Applications (DEXA), Munich, Germany, 2001.

MULTIMEDIALNE BAZY DANYCH – PRZESZŁOŚĆ, TERAŹNIEJSZOŚĆ I PRZYSZŁOŚĆ

Streszczenie

Artykuł stanowi przegląd dziedzin i praktycznych zastosowań multimedialnych baz danych, celów i metod wykorzystywanych w zdobywaniu wiedzy. Głównym celem multimedialnych baz danych jest pozyskiwanie interesującej wiedzy i rozumienie znaczeń utrwalonych w danych multimedialnych, które zawierają współzależne obrazy, przekazy audio i wideo oraz tekst.

Artykuł pokrótce opisuje nadzorowaną i nienadzorowaną klasyfikację, pokazując m.in. interesujące reguły i schematy decyzyjne. Przedstawia główne korzyści multimedialnych baz danych oraz dyskutuje pewne decydujące kwestie i przyszłe trendy.