Multimedia Systems (2005) 10(6): 457–463
DOI 10.1007/s00530-005-0174-7

REGULAR PAPER

Chabane Djeraba · **Nicu Sebe** · **Michael S. Lew**

Systems and architectures for multimedia information retrieval

Published online: 30 September 2005

Abstract In this paper, we provide a brief survey on multimedia information retrieval and we introduce some ideas investigated in the special issue. We hope that the contributions of this issue will stimulate the readers to tackle the current challenges and problems in this highly important research direction. Such contributions are the basis of tomorrow's multimedia information systems. Our aims are to clarify some notions raised by this new technology by reviewing its current capabilities and potential usefulness to users in various areas. The research and development issues cover a wide range of fields, many of which are shared with media processing, signal processing, database technologies, and data mining.

Categories and subject descriptors
H.3.1, H.3.3, H.3.7

General terms
Multimedia, Algorithms, Design, Experimentation, Human factors

Keywords Multimedia information · Indexing · Retrieval · Architecture · System

1 Introduction

The last ACM SIGMM retreat report [7] presented observations and future directions in multimedia research. According to this report, a grand challenge that the community should focus on solving is "to make capturing, storing, finding, and using digital media an everyday occurrence in our computing environment." Our special issue lies at the heart of this challenge.

The ongoing expansion of the information superhighway and the associated increase in multimedia content in databases, broadcasts, streaming media, etc. has generated new requirements for more effective access to these giant global information repositories. Content extraction, indexing, and retrieval of multimedia data continue to be among the most challenging and fastest-growing research areas. A consequence of the growing consumer demand for multimedia information is that sophisticated technology is needed for representing, modeling, indexing, and retrieving multimedia data. In particular, we need robust techniques to index/retrieve and compress multimedia information, new scalable browsing algorithms that allow access to very large multimedia databases, and semantic visual interfaces that integrate the above components into a unified multimedia browsing and retrieval system. Applications should be investigated extensively to obtain real-world feedback necessary to achieve robust solutions. These solutions should consider different media and sources and use context and user specific information to improve application performance. This special issue directly addresses these recent developments in this challenging research area.

Not too long ago, we published two special issues [3][8] on the new challenges of multimedia indexing and retrieval. The core idea behind those special issue articles was the gap between primitive and semantics features and how to bridge the two levels for multimedia information retrieval. Our conclusion remains valid today. Multimedia information retrieval is still a young field, and more and more users have yet to integrate it into their everyday activities. Several commercial enterprises specialized in this domain have even shut down.

Even though promising research systems exist for content-based multimedia information retrieval [4], they have not reached the popularity of text-based systems such as Google or Lycos. These systems became very popular

C. Djeraba (✉)
LIFL – UMR CNRS 8022, Building M3, University of Lille 1,
59655 Villeneuve, d'Ascq Cedex, France
E-mail: djeraba@lifl.fr

N. Sebe
Faculty of Science, University of Amsterdam, Kruislaan 403,
1098 SJ, Amsterdam
E-mail: nicu@science.uva.nl

M. S. Lew
Leiden University Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
E-mail: mlew@liacs.nl

because typing simple keywords gives users instant access to billions of documents worldwide, including those composed of images and video. This is due in part to the fact that parsing text documents for keywords is simpler than content-based analysis of multimedia. Furthermore, textual information retrieval was pioneered at least a half century ago [4], providing the field with a deep and mature research platform. The important progress in text mining contributed also to making Internet search engines the successful tools we know today.

Multimedia information retrieval has advanced significantly, and it continues to be a rapidly growing field of research that is based on the background of research themes such as image, audio, and video analysis, speech recognition, scene detection, text extraction, and video tracking. More and more effort has been devoted to integrating these specific research solutions to truly multimodal systems. Some of the best examples are the multimedia data mining project of HP that includes multimedia information retrieval and personalization, and the multimedia mining project at IBM that deals with the associations between linguistic and audiovisual information.

Many new approaches resolve specific problems, and to clearly highlight the performances of each approach, it is important to specify evaluation recommendations with which these approaches can be tested and compared. Many possible criteria and adequate data have to be compiled and made available. Many projects, such as the TRECVID meetings, address this issue. Comparative methods have been rewarding in other areas such as speech recognition and textual information retrieval, so there is little doubt that these efforts will make significant contributions to the progress in multimedia information retrieval.

This special issue focuses on new architectures of multimedia information systems and describes the state of the art in this broad area of research. Are architectures for multimedia systems specific? What are their properties? What are the differences between several multimedia system architectures? The research results in this special issue hold potential for solving the technical problems of the field (Sect. 3). The basic idea behind the special issue articles is the system architecture of multimedia information retrieval systems. The research presented here should be a springboard for further development in this already exciting field.

1.1 Primitive/semantic features

Primitive features (e.g., object motion, shot intervals, color, texture, shape, pitch, etc.) are generally extracted automatically and computed efficiently. The most suitable applications are those where primitive features can be directly applied: crime prevention, including human face identification, searching in specific domain databases, etc. The main disadvantage of information retrieval based on primitive features is that there is little use for retrieval by semantic content. It is difficult to present queries that support semantic features. For example, it is not easy to find specific soccer/football players scoring a goal during a game held in spring or to find video sequences that contain traditional boats on the Nile River at the end of the day in video repositories based solely on primitive features.

Increases in storage capacity and reductions in cost make it possible to store massive amounts of multimedia information. The question is how to make this information useful in real-world applications. The important evolving and widespread adoption of cell phones and digital cameras is intensifying this phenomenon.

Past research addressed extensively primitive features, multimedia database models, algorithms to analyze media data, and algorithms to search for relevant or interesting data (e.g., query a video archive by identifying a melody or find pictures with similar textures) [9].

Semantic features involve varying degrees of semantics emphasized in multimedia information. They make multimedia content meaningful. In this context, we talk of semantic queries and we can distinguish objective features (including physical objects and semantic categories; things such as cars, mountains, rivers, people, flowers, and goals; actions such as walking, shooting, painting, and playing, etc.), outlines (accurate and descriptive information such as author, date, copyright, etc.), and subjective features (abstract concepts such as heroes in battle during the Second World War, religious adoration, innocence, entrepreneurial spirit, favorite children's movies, beauty, etc.). Semantic features permit queries such as "find video clips that contain the entrepreneurial spirit" or "find video clips preferred by young people."

Multimedia information retrieval based on semantic features is certainly a reasonable answer to the semantic drawbacks of information retrieval based on primitive features. Multimedia information retrieval based on semantic features supports robust semantic expression that highlights wide-ranging aspects of multimedia content. Semantic features are powerful enough to easily describe multimedia content at varying levels of complexity.

In many cases, multimedia retrieval efficiency, on the basis of semantic features, requires considerable effort on the part of the indexers because semantic features generally require manual annotation. The subjective features fall within the domain of manual interpretation and judgment and require more human intervention (e.g., application domain experts) than objective semantic features. Operational solutions are applicable to objective semantic features when the application domain is clearly framed. The advances in pattern recognition techniques contribute to this result. However, automatic operational solutions are rarely applicable to subjective semantic features and suffer from two important drawbacks. First, they are very time consuming. For example, annotating a geographic society image collection takes twenty minutes per image, and manual annotation for videos is certainly much more time intensive. Second, manual extraction of semantic features of multimedia content may result in different semantic features being linked to different indexers of multimedia content. For example, e-commerce sites maintain large repositories of fixed images to illustrate physical products. These repositories contain thousands of images that must be maintained at high resolution. Broadcasting corporations also deal with millions of hours of video footage repositories, which are hard to manually annotate. Automated assistance is clearly necessary.

Considering both primitive and semantic features is a realistic way to get around the drawbacks of current methods. How do we determine the relationships between primitive and semantic features? How do we index the content of multimedia information while at the same time limiting human intervention and increasing the automation of the indexing process? Automatic determination of the relationship between semantic and primitive features in an operational manner is an old problem and a great challenge.

Automatic extraction of text from speech accompanying video or recognition of text superimposed on images represents one pragmatic way to bridge the gap between semantic and primitive features, which is necessary for multimedia information retrieval. As stated previously, extraction of objective features such as cars, faces, oceans, rivers, etc. is possible in a limited application context, given recent advances in scene and object recognition and multimedia mining. In recent years, many researchers have pursued this direction of research.

A practical solution is to automatically extract primitive features (keyframes, shots, and other classical primitive features) from voluminous video bases (e.g., hundred of thousands of hours) and annotate them manually (semantic features) in a subset of video bases (100 h). The knowledge extracted automatically from the correlations between semantic and primitive features of the subset of video databases is a first step toward reusing video-footage bases for entire databases. It is therefore possible to generate an entirely automatic annotation for video. For example, we can use semantic features to annotate keyframes and TV companies now use this technology extensively. Current commercial products automatically create storyboards of thumbnail images, which users manually annotate. In the near future, we can expect further technological advances that will allow for direct search of video content with a much-reduced level of manual annotation.

We believe that this approach opens a challenging area of research that we call "semantic extraction" based on data-mining techniques. The key problem is predicting semantic features of multimedia content on the basis of the evidence of primitive features. This seems to be a typical probabilistic problem (e.g., Bayes classification). Data-mining approaches seem to be suitable for resolving such problems. An increasing number of multimedia and production film companies are using such approaches to organize their data for retrieval. The use of such approaches to index and classify multimedia information will permit large time and cost savings.

1.2 Multimedia usage mining

Other than the ever-challenging problem of automatically extracting semantic features from primitive features, we must resolve other problems, in connection with semantic feature extraction, to reveal whether we can transform multimedia information retrieval technology into operational solutions on real-world applications. One of the outstanding issues concerns multimedia usage mining [6].

Users create, exploit, and update multimedia information, and they are thus implicitly at the center of multimedia information. Hence, it may be interesting to consider users' behavior to understand their real requirements in order to extract semantically subjective features and, more generally, multimedia knowledge. Multimedia usage mining includes two aspects: the customization of user queries and the identification of best practices. The first aspect considers that each target application of multimedia information retrieval has its own range of specific needs. Methods that fail to address these needs are unlikely to perform well enough to convince users of the methods' usefulness. In the second aspect, we would like to identify the best practices in the research fields that could potentially benefit from multimedia information retrieval. These practices include frequent patterns of multimedia content retrieval and browsing. All these problems become much more complex when considering privacy issues and civil liberties that arise from using multimedia information in different application contexts.

Multimedia usage mining techniques can be roughly classified into two types of approaches. The first type addresses the analysis of user behaviors without considering temporal actions, i.e., the temporal relationships between retrieval and browse actions are not considered. This type may report on statistics of user behavior and on frequency counts of multimedia accesses. An example is the analysis of student usage of an educational multimedia system, as in [1]. We can define the notion of student profiles. Indeed, learning needs and expectations depend on the characteristics of the student type profile. User profiles include statistics such as the number of video viewing sessions, duration of video shows, number of video viewing sessions that last longer than 20 minuntes, average video session duration, average number of commands per minute during video viewing sessions, forward transitions, backward transitions, forward jumps, and jump ratio. As these numbers are based on the statistics collected on each type of student, we can analyze how the learning multimedia system can be improved.

The second type of approach addresses the analysis of user behavior by considering temporal relationships between retrieval and browse actions. It examines properties such as how user requests vary on a day-to-day basis and whether multimedia accesses exhibit any temporal properties. These techniques may benefit from system properties such as server proxy caches to analyze user behaviors. For example, the analysis may reveal that the users previewed the initial portion of a video to find out if they are interested in viewing the entire video. If the users like the initial portion, they continue watching; otherwise they stop viewing. This pattern suggests that caching the first several minutes of multimedia data should improve access performance. These approaches may also relate to behavior analysis of a single type of multimedia (e.g., video). For example, we could develop a framework that combines video content analysis and user log mining to generate video summaries. Thus, the objective is to develop a video browsing and summarization system that is based on previous viewers' browsing logs to help future viewers. We adopt the link analysis technique used in Web mining and propose the concept of shot rank [2], by which the importance of each video shot is measured, the user behavior is simulated with an interest-guided walk model, and the probability of a shot being visited is taken as an indication of the importance of that shot. The resulting shot rank is used to organize the presentation of video shots and generate video skims.

We believe that multimedia usage mining is still in its infancy and the proliferation of multimedia servers will certainly reinforce the need to automatically analyze user behaviors on multimedia content to ameliorate the quality of services. However, this approach requires suitable tools to save in real-time user browses and searches in log files. Hence, multimedia players and browsers need to be extended by this functionality—saving user browses and searches.

Analysis of user behaviors in large multimedia databases or broadcasts is an emerging problem. The growing importance of multimedia in everyday life (e.g., film production) automatically increases the importance of multimedia usage mining. To be able to cope with the abundance of available multimedia information, users of these videos need intelligent software systems that fully utilize the rich source of information hidden in user behaviors in large multimedia databases for retrieving and navigating through multimedia.

2 Ever greater challenges

We should deal with the great challenges of multimedia information retrieval; the currently unsolved and significant issues. The majority of these issues concern an inability to automatically obtain semantic features and user behaviors in multimedia content. The following issues are still unsolved:

– Scanning an archive of video broadcasts to find an interview with a particular individual, or a picture archive to find a photo of the person running for a high political office. Identifying where a particular photo was taken requires automatic capture of metadata regarding when and where the photo was taken. If the metadata are not captured when the photo is taken, then the problem becomes complicated by the fact that the other option is extensive image analysis with no insurance of correct responses.
– Finding video interviews of a particular person published on the Web. This problem might be solved by looking at the text associated with a streaming media file published on a Web page. However, it may be difficult to identify the text associated with a video clip if the Web pages are generated dynamically.
– Extracting user interest on video lectures of a particular video Web site. This problem is difficult to resolve because most commercial webcasting systems do not store user actions (play, forward, pause, etc.). Some experimental webcasting tools permit storage of actions in log files; however, these tools are still prototypes.
– Identifying a monument in the downtown of a tourist city. The idea is to point the cell phone camera at the monument and ask for the name of the monument. Solving this problem requires context as well as connecting to a database and a processing server. The obvious solution is to recognize the pattern of the monument using the captured image but this approach might return too many potential matches or take too much time. What the system should do is use the context of the situation (e.g., a map of the downtown and the localization of the tourist) to restrict the candidates to objects that might actually be at the place.
– Making the voluminous quantities of home video useful. On a daily basis, millions of people upload video captures on their computers. However, they do not have good tools to organize and store these data in such a form as to allow a user to say, "Show me the shot where my sons clean the car." As a consequence, developing fully automatic semantic extraction tools is still a great research challenge. Semiautomatic analysis tools that combine subset-annotated data coupled with primitive features of a whole dataset may index and organize data so they are easily accessible with less user interaction.

We believe the real challenge is to work on fundamental concepts (e.g., pattern recognition, frequent pattern extraction, multimedia usage mining, indexing and retrieval, parallel search, high-dimensional representation, combining partial results) toward overcoming domain brittleness so that operational approaches could potentially be built and deployed for normal people to use.

Finally, all the approaches presented focus on multimedia data, which invokes the basic problem of digital rights management. How can we progress in multimedia information retrieval without access to various collections of multimedia data to test and validate? We need comprehensive multimedia test sets that can be distributed along the lines of open source software. Without these test sets, which exist primarily at commercial companies, it will be very difficult to accurately assess the effectiveness of a new system or find its weaknesses.

3 Special issue articles

The articles of the special issue are based on the original submissions to the 5th ACM International Workshop on Multimedia Information Retrieval in conjunction with the

ACM International Conference on Multimedia 2003 (Berkeley, California, USA). The purpose of the workshop was to bring together researchers, developers, and practitioners from academia and industry to discuss challenges in retrieving multimedia information libraries, databases, and streaming-media sources. For this special issue we sought contributions from a wide range of theoretical and application areas.

When selecting the contributions for this issue, we aimed at providing a good balance of research areas and contributions to multimedia information retrieval. The 8 papers presented in the issue were selected from a total of 36 submissions that had been presented in the multimedia information retrieval workshop proceedings and which were in turn selected from a total of 92 original submissions to the workshop. Note that three other articles have been selected from the workshop but due to the size limitation of the special issue, they are going to be published in regular issues of the journal.

The selected articles fall into broad categories that reflect the variety of research directions in multimedia information retrieval. The purpose of this special issue is to present the current systems and architectures for multimedia information retrieval covering a broad range of issues.

– Wolfgang Muller, Martin Eisenhardt, and Andreas Henrich [10] propose a fast retrieval approach for high-dimensional feature vectors in peer- to-peer (P2P) networks using compact peer data summaries. Their research was motivated by the fact that most P2P systems are limited to queries based on an unique identifier or a small set of keywords and these are hardly applicable to a content-based retrieval application. To address this problem, the authors present two methods for peer data representation that permit to reduce the number of peers to be visited when processing a content-based image retrieval query.
– Silvia Pfeiffer, Conrad Parker, and Andre Pang [16] propose a distributed multimedia information retrieval architecture extending the World Wide Web to multimedia enabling existing Web infrastructure to provide seamless search and hyperlink capabilities for time-continuous Web resources with only minor extensions. Their article discusses the requirements for such an extension of the Web, contrasts the existing standards and technologies, and presents the developed solutions.
– Erdem Ural, Shrikanth Shri Narayanan, Hsuan-Huei Shih, Elaine Chew, and C.C. Jay Kuo [11] present a method that creates data resources for designing user-centric front-ends for query by humming systems. They also provide a statistical analysis for categorizing the collected data, focusing on inter-subject variability issues.
– Nicu Sebe, Ira Cohen, Fabio G. Gozman, Theo Gevers, and Thomas D. Huang [13] discuss the training of probabilistic classifiers with label and unlabeled data for human-computer interaction applications such as facial expression recognition, face detection, and skin detection. They provide an analysis that shows under what conditions unlabeled data can be used in learning to improve classification performance and investigate the implications of this analysis to Bayesian networks.
– Giang P. Nguyen and Marcel Worring [14] propose a content-based image retrieval system using relevance feedback based on saliency adaptation. They aim to dynamically update the user- and context-dependent definition of saliency based on relevance feedback. To that end, they propose an interaction framework for salient details from the perspective of the user.
– Ning-Han Liu, Yi-Hung Wu, and Arbee L.P. Chen [15] present an efficient kNN search approach in polyphonic music databases using a lower bounding mechanism. In their paper, they propose three polyphony representations with associated similarity measures and present a novel method to efficiently retrieve k music words that contain segments most similar to the user query based on the edit distance.
– Rufoei Zhang and Zhong fei Mark Zhang [16] propose an approach to make more effective and efficient the image retrieval process. They propose a new indexing method based on fuzzy logic to incorporate color, texture, and shape information into a region based approach to improving the retrieval effectiveness and robustness. They also introduce a new hierarchical indexing structure and a new retrieval algorithm to significantly improve the retrieval efficiency without sacrificing the retrieval effectiveness. Finally, the use relevance feedback to tailor the semantic retrieval to each user's individualized query preference through two new indexing algorithms.
– Yongqing Sun and Shinji Ozawa [17] present a hierarchical approach for region- based image retrieval based on the wavelet transform. First automated image segmentation is performed in the low-low frequency subband and the boundaries between segmented regions are deleted to improve the region-based image retrieval against the uncertainty of segmentation. Their approach show good tradeoff between retrieval effectiveness and efficiency as well as easy implementation for region-based image retrieval.
– Wei-Ta Chu, Wen-Huang Cheng, Jane Yung-jen Hsu, and Ja-Ling Wu [18] propose a semantic indexing and retrieval approach using hierarchical audio models. Their hierarchical approach models the statistical characteristics of audio events over a time series in order to accomplish semantic context detection. Two stages, audio event and semantic context modeling/testing, are devised to bridge the semantic gap between physical audio features and semantic concepts.

Acknowledgements We would like to take this opportunity to thank the program committee and the additional reviewers for their insightful and thorough review of the submissions to this special issue. We are also grateful to Klara Nahrstedt for her encouragement, valuable discussions, and patience during the long period of the special issue preparation. Many thanks to Larry Rowe and Forouzan Golshani Golshani for their help in organizing the MIR workshop. And of course, the special issue would not be possible without the great effort and prompt reactions of the authors. They really did a great job.

Here is the list of the program committee members of the special issue: Kiyo Aizawa (University of Tokyo, Japan), Aristeidis Diplaros(University of Amsterdam, The Netherlands), Laurent Amsalag (IRISA, France), Catherine Berrut (CLIPS, France), Liming Chen (LIRIS, France), Mohamed Daoudi (University of Tours, France), Alberto DelBimbo (University of Florence, Italy), Chitra Dorai (IBM, USA), Theo Gevers (University of Amsterdam, The Netherlands), Alan Hanjalic (TU Delft, The Netherlands), Thomas Huang (University of Illinois at Urbana Champaign, USA), Benoit Huet (Eurecom, France), Joemon Jose (University of Glasgow, UK), Brigitte Kerhervé (UQAM, Canada), Anil Kokaram (Trinity College, Japan), Rainer Lienhart (University of Augsburg, Germany), Frank Nack (CWI, The Netherlands), Wei-Ying Ma (Microsoft, China), Maude Manouvrier (University of 9, France), Nourreddine Mouaddib (LINA, France), Vincent Oria (New Jersey Institute of Technology, USA), Stefan Ruger (Imperial College London, UK), Marta Rukoz (Central University, Venezuela), Shin'ichi Satoh (National Institute of Informatics, Japan), Stan Sclaroff (Boston University, USA), Florence Sedes (IRIT, France), Hari Sundaram (Arizona State University, USA), Qi Tian (University of Texas at San Antonio, USA), Guangyou Xu (Tsinghua University, China), Hong Jiang Zhang (Microsoft, China), Xiang Sean Zhou (Siemens Research, USA).

References

1. Acharya, S., Smith, B., Parnes, P.: Characterizing user access to videos on the World Wide Web. In: Proceedings of Multimedia Computing and Networking, San Jose, CA (2000)
2. Yu, B., Ma, W.-Y., Nahrstedt, K., Zhang, H.-J.: Video Summarization Based on User Log Enhanced Link Analysis. In: Proceedings of MM'03, Berkeley, CA, vol 2 (2003)
3. Djeraba, C.: Guest editor's introduction: Content-based multimedia indexing and retrieval. IEEE Multimedia **9**(2), 18–22 (2002)
4. Lew, M.S.: Next generation web searches for visual content. IEEE Comput. 33(11) 46–53 (2000)
5. Luhn, H.P.: A statistical approach to mechanized encoding and searching of library information. IBM J. Res. Develop. **1**, 309–317 (1957)
6. Mongy, S., Bouali, F., Djeraba, C.: Video usage mining, multimedia dictionary. CRC Press, Boca Raton, FL (2005)
7. Rowe, L.A., Jain, R.: ACM SIGMM retreat report on future directions in multimedia research. Final report of ACM SIGMM (2004)
8. Sebe, N., Lew, M.S., Smeulders, A.: Video retrieval and summarization. Comput Vision Image Understanding **92**(2/3), 141–147 (2003)
9. Smeulders, A., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1349–1380 (2000)
10. Müller, W., Henrich, A.: Fast retrieval of high-dimensional feature vectors in P2P networks using compact peer data summaries. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
11. Pfeiffer, S., Parker, C., Pang, A.: The continuous media web: a distributed multimedia information retrieval architecture extending the World Wide Web. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
12. Unal, E., Narayan, S.S., Shih, H.H., Chew, E., Jay Kuo, C.C.: Creating data resources for designing user-centric front-ends for query by humming systems. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
13. Sebe, N., Cohen, I., Gozman, F.G., Gevers, T., Huang, T.S.: Learning probabilistic classifiers for human-computer interaction applications. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
14. Nguyen G., Worring, M.: Image retrieval using interactive saliency. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
15. Liu, N-H., Wu, Y-H., Chen, A.L.P.: Efficient kNN search in polyphonic music databases using a lower bounding mechanism. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
16. Zhang R., Zhang, Z.: Addressing CBIR efficiency, effectiveness, and retrieval subjectivity simultaneously. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
17. Sun Y., Ozawa, S.: HIRBIR: A hierarchical approach for region-based image retrieval. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)
18. Chu, W.-T., Cheng, W.-H., Hsu, Y.-J., Wu, J.-L.: Towards semantic indexing and retrieval using hierarchical audio models. Multimedia Syst. Special Issue on Systems and Architectures of Multimedia Information Retrieval (2005)

Author biographies

Chabane Djeraba has been a professor of computer science at the University of Science and Technology in Lille, France since 2003. He is a member of LIFL Laboratory, which is a joint research unit of both the French National Center for Scientific Research (CNRS) and the University of Science and Technology of Lille. At LIFL Laboratory he also heads a research team on mining, indexing, and recognition of multimedia and complex data. He was an associate professor of computer science at Nantes University from 1994 to 2003. He obtained a Ph.D. in computer science from Claude Bernard University of Lyon, France in 1993; a master's degree from Pierre Mendes France University, France in 1990; and an engineering degree from the National Institute of Computer Science (INI), Algiers, Algeria in 1989. He has organized major multimedia conferences and been cochair of workshops such as ACM Multimedia Information Retrieval and ACM Multimedia Data Mining. He has been the guest editor of special issues of top multimedia journals (e.g., IEEE Multimedia). During the last 15 years he has published 100 scientific publications in the areas of multimedia indexing and retrieval and mining, including one leading book on multimedia data mining, and he has participated on several program committees of major conferences and journals (e.g., International Journal of Multimedia Tools and Applications, Springer-Kluwer).

Nicu Sebe is an assistant professor in the Faculty of Science, University of Amsterdam, The Netherlands. He has more than 50 scientific publications in the areas of pattern recognition, multimedia information retrieval and human–computer interaction. He is the author of the books Robust Computer Vision–Theory nd Applications (Kluwer, April 2003) and Machine Learning in Computer Vision (Springer, May 2005). He has been a guest editor for top scientific journals including the ACM Multimedia Systems Journal, ACM Transactions on Multimedia Computing, Communications and Applications, Image and Vision Computing, and Computer Vision and Image Understanding Journal. He organized major multimedia conferences and workshops and was the co-chair of the IEEE Workshop on Human-Computer Interaction (2004 and 2005), ACM Workshop on Multimedia Information Retrieval (2003 & 2004), and of the International Conference on Image and Video Retrieval (CIVR 2003).

Michael Lew is codirector of the LIACS Media Lab and co-head of the Imaging/Multimedia Research Theme at the Leiden University Computer Science Department (LIACS). He has over 100 scientific publications in the areas of multimedia information retrieval, human–computer interaction, and biocomputing including a leading book on multimedia search and retrieval and has been a guest editor of numerous scientific journals including the ACM Multimedia Systems Journal, ACM Transactions on Multimedia Computing, Communication, and Applications, and the International Journal on Computer Vision and Image Understanding. His professional service includes founding the International Conference on Image and Video Retrieval and assisting in the organization of the major multimedia conferences such as ACM MM and IEEE ICME. He has been co-chair of the IEEE Workshop on Human–Computer Interaction and the ACM Workshop on Multimedia Information Retrieval. His group has received national and international coverage in scientific circles and newspapers and on television and has regularly been the recipient of leading grant-supported international research projects.