Multimedia Systems (2006) 12(1): 1–2
DOI 10.1007/s00530-006-0040-2

EDITORIAL

Arif Ghafoor · Zhongfei (Mark) Zhang · Michael S. Lew · Zhi-Hua Zhou

Guest Editors' introduction to the special issue: machine learning approaches to multimedia information retrieval

Published online: 26 July 2006
© Springer-Verlag 2006

With today's ubiquitous presence of multimedia information in almost all IT-based applications, it is imperative to address the indexing and retrieval issues for managing the explosive growth of the multimedia data for such applications. Research in the area of Multimedia Information Retrieval (MIR) aims at dealing with these issues and focuses on the development of effective and efficient indexing and retrieval techniques for the multimedia data. Here, effectiveness refers to how well the retrieved multimedia information meets the users' expectation, and the efficiency addresses how responsive the underlying retrieval mechanisms are as queries are posed by the users.

Due to its tremendous potential in a broad range of applications, research in MIR has received considerable attention in recent years by the multimedia community. Nevertheless, it has been observed that undoubtedly MIR research has been thwarted by two key challenges. The first challenge deals with the "semantic gap" as the majority of the existing methods in the literature generally use low-level features for retrieving multimedia information. However, as reported in the literature, such approaches do not provide satisfactory solution to the MIR problem due to the semantic gap that generally exists between the low-level features and the high-level semantic concepts embedded in the multimedia information. In essence, an accurate and comprehensive representation of high-level semantics is a daunting challenge faced by the MIR community. The second issue deals with the scalability aspect of the MIR problem. Currently, almost all the existing methods reported in the literature are shown to be only effective for clean data sets (e.g., the Corel data in image retrieval) and very small data sets (e.g., typically below 10,000 images in image retrieval). The issue of scalability has been explored by many researchers and it has demonstrated that achieving a logarithm search complexity is a nontrivial issue. Many existing methods have been demonstrated to be sensitive to the diversity of multimedia contents and quality of data. Clean data is generally preferred in reporting experimental results. Note that the scalability issue refers to two key aspects of MIR: the dimension of diversity associated with multimedia contents and the size of the multimedia data.

The aforementioned challenges have been steadily addressed by the MIR research community and several advances have been reported on this front. In particular, efforts using the machine learning approaches to MIR have begun to show noticeable promises for effectively overcoming these barriers.

Machine learning is a well-researched area with a wide range of applications, including MIR. Machine learning techniques provide powerful capabilities that consistently improve their performance from learning and experiences. Recent advances in machine learning have led to the development of systems that are capable of autonomous acquisition and integration of knowledge resulting in continuous self-improvement and increased performance in terms of effectiveness and efficiency. Undoubtedly, the philosophy behind the science of machine learning matches very well with the underpinning of developing effective MIR systems. In fact, many MIR problems are essentially machine learning problems under some specific MIR context. Therefore, effective solutions to MIR problems may very well be developed by tailoring the existing machine learning approaches. These approaches range from the classic learning paradigms, such as unsupervised learning,

A. Ghafoor (✉)
School of Electrical and Computer Engineering, Purdue University, Engineering Administration Building, Room 101 400, Centennial Mall Drive, West Lafayette, IN 47907-2016, USA
E-mail: ghafoor@ecn.purdue.edu

Z. (Mark) Zhang
Computer Science Department, Watson School, SUNY Binghamton, NY, USA
E-mail: zhongfei@cs.binghamton.edu

M. S. Lew
LIACS Media Laboratory, Leiden University, Leiden, The Netherlands
E-mail: mlew@liacs.nl

Z.-H. Zhou
National Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
E-mail: zhouzh@nju.edu.cn

supervised learning, and reinforcement learning, to more advanced learning paradigms that include active learning, semi-supervised learning, and multiple instances learning or learning with uncertainties.

However, it is worth mentioning that quite often if the same MIR problem happens to arise in different application scenarios, different methods from different machine learning paradigms may be required. For example, consider the relevance feedback problem in MIR. In this case, depending upon how much and how the relevance feedback data is provided by a user, the same problem may be mapped to the problem of either unsupervised learning, supervised learning, reinforcement learning, or semi-supervised learning.

On the other hand, given an MIR problem mapped to a specific machine learning problem, the abundant machine learning literature offers a wide spectrum of choices for developing effective and efficient solutions. For example, for a relevance feedback problem in MIR, if the user provides sufficient amount of such data, a rich toolbox in the area of supervised learning is available that can provide a wide variety of classification methods such as decision trees, neural networks all the way to the most recently developed techniques such as support vector machines. In addition, active learning may also be applied to develop more efficient learning strategies.

Consequently, with the rapid research advancement and development in both machine learning and MIR areas, the linkage between the two areas is strengthening. For the multimedia community, such a connection provides unprecedented opportunities to further promote and advance research endeavors in the MIR area. The primary focus of this special issue is to present recent advances in this direction.

In response to the Call for Papers for this special issue, we have received 24 submissions. After the first round of reviews, we have accepted six papers for the special issue. In addition, there are three papers that are required major revision with a second round of reviews. If accepted, these papers will be considered for publication in a regular issue.

The six accepted papers address a broad range of issues dealing with the application of machine learning to MIR. The first paper, titled "Support Vector Machine Active Learning for Music Retrieval" by Mandel, Poliner, and Ellis, describes an application of a SVM-based active learning technique to the retrieval problem for music data. The paper demonstrates that the proposed active learning requires half as many labeled examples to achieve the same performance as a standard scheme does. The evaluation is based on classifying 1,200 pop songs according to moods and styles selected from an online music guide and the performing artist.

The next article by Huang, Wu, and Lai, "Improved Adaboost-based Image Retrieval with Relevance Feedback via Paired Feature Learning", addresses the classic relevance feedback problem related to image retrieval. The proposed method uses an adaptive feature quantization scheme and exploits the feature dependencies through paired feature analysis. Subsequently, the scheme uses a novel approach that facilitates the selection of a few feature pairs with the best discrimination capabilities in the corresponding paired feature spaces.

In the third article, titled "A Probabilistic Semantic Model for Image Annotation and Multi-modal Image Retrieval" by Zhang, Li, Ma, and Zhang a probabilistic approach to the image retrieval problem has been proposed. The approach explicitly exploits the synergy among multiple modalities of information that typically coexists in many image retrieval applications. Specifically, through an EM-based training method developed in this paper, an implicit concept space is discovered that constitutes an explicit synergy between the imagery and text modalities. Based on this concept space, a multimodal image retrieval algorithm is proposed. The proposed model and the retrieval algorithm are evaluated to demonstrate the effectiveness and scalability using a noisy and diverse data set consisting of 17,000 images and 7,736 word vocabulary, that was generated through Web crawlers.

Luo, Jin, Hoi, and Lyu, in their paper: "Collaborative Image Retrieval via Regularized Metric Learning," revisit the relevance feedback problem in image retrieval by proposing a novel regularized metric learning approach. The learning is based on a similarity measure that uses the correlation between the image low-level features and the users provided relevance feedback data. The learning problem is mapped to a semi-definite programming problem that is efficiently solved using an existing software tool.

Video data poses unique challenges due to its inherent temporal dimension. Bashir, Khokhar, and Schonfeld's article, titled "View-Invariant Motion Trajectory-Based Activity Classification and Recognition", presents a novel affine-invariant classification and recognition system for indexing motion trajectory of objects in video databases. The authors present two affine-invariant representations for motion trajectories based on curvature scale space and centroid distance function. The proposed system has been implemented and a comprehensive performance evaluation based on known benchmark databases is presented.

In the last article, titled "Machine-Learning Based Classification of Speech and Music" by Khan and Al-Khatib, the authors present music-speech classification techniques based on several novel features. Subsequently, a fuzzy clustering approach is presented for selecting a viable set of features that enables a better classification yielding a degree of accuracy. Three different classification frameworks are presented that are evaluated through an extensive experimentation.

Overall, the collection of papers in this special issue provides a panoramic view spanning a variety of machine learning approaches for dealing with the challenges faced by the MIR research community, and gives an idea of the scope and directions of future research on this important and promising topic.