МЕХАТРОНИКА РОБОТОТЕХНИКА

МЕЖДУНАРОДНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ КОНГРЕСС

MECHATRONICS AND ROBOTICS

INTERNATIONAL SCIENTIFIC & TECHNICAL CONGRESS

MULTIMEDIA RETRIEVAL

Peter L. Stanchev[*&], Krassimira Ivanova[**]**

* Kettering University, Flint, MI, USA 48504, pstanche@kettering.edu

** Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria, kivanova@math.bas.bg

Easy and effective access to large information collections is of great importance in many aspects of everyday life. Enormous quantities of information in different forms and of different types (text, images, audio, video, etc.) are being produced and archived every day. For this reason, significant attention has been given to studying and developing techniques that support effective and efficient retrieval of multimedia data. Audio/video can be considered today as a primary means of communication, due to its richness in information content and its appeal. This implies that the development of techniques supporting the retrieval of audio/video documents is of primary importance for the general public as well as for professional users and represents an important asset of everyday life.

The process of image and video description consists of extracting the global image characteristics, recognizing the image-objects, and assigning semantics to these objects [6]. The image data can be treated as physical image representation and its meaning - as a logical image representation. The logical representation includes methods for describing the image and image-objects characteristics and the relationships among the image objects. Several visual descriptors exist for representing the physical content of an image, such as the MPEG-7 standard [5]. Historically, semantic retrieval was frequently based on computer vision. To reduce the semantic gap, the low-level content-based media features are frequently being converted to high-level concepts or terms.

A technique for improving the similarity search process of images and video in a Multimedia Content Management System is analyzed. The content based retrieval process integrates the search on different multimedia components, which are linked in XML structures. Depending on the specific characteristics of an image data set, some features can be more effective than others when performing similarity searches [2]. Based on this observation we propose a technique that predicts the effectiveness of the MPEG-7 image features that depends on a statistical analysis of the specific data sets in the Multimedia Content Management System. This technique is validated through extensive experiments with human subjects.

We analyze several aspects of the fine art databases [4]. We showed that MPEG-7 descriptors can be used, but they give different results when applyied to

different types of images. The use of the Color Structure descriptor only produces sufficiently efficient results in the query search. The new generation Semantic Web languages, such as RDF(S) and OWL will play a major role. The integration of semantic understanding of pictures with personalized delivery raises new questions. The query language for this type of system is not yet standartized. One of our investigations has been focused on analyzing the color distribution in art image paintings. Some of our conclusions follow.

- There is a great distribution of "warm" hues in the art pictures. In spite of expectations, the presence of "cold" hues in the group of pictures that were classified as landscapes was small. Only the Impressionism pictures show skewed distribution in this direction. The predominate presence of "warm" colors is due to painting of faces and bodies from one side and using the materials and varnish, which acquired yellowish from other side. Not without importance is the fact that "cold" hues as blue and green are non-durable under the influence of light;
- The difference in mean distribution of saturation among different movements is very small. We find a global trend of saturation distribution in the art image database;
- The analyses of mean distribution of luminance shows strengthen presence of dark colors in pictures from the 17-18 century (Baroque, Neoclassicism and 18th Century);
- The examination of the difference between distribution of luminance of landscapes, portraits, and subjects shows similarity of distribution for portraits and subjects and has significant difference between these two groups and landscapes in direction of higher presence of light colors.

Another big area for using multimedia retrieval is the field of medical systems. We discuss the problems which arise when working with medical images. As an illustration of a medical image processing tools, we discuss MR (Magnetic resonance) brain segmentation problem. Functional analysis of different medical systems is made. We emphasize on the fact that working with medical images is different from working with other kind of images. As an illustration two systems are presented [3]. The first system is MEDIMAGE, which is a multimedia database for Alzheimer's disease patients. The system design is motivated by the major need to manage and access multimedia information on the analysis of the brain data. The database links MR images to patient data in a way that permits the user to view and query medical information using alphanumeric and feature-based predicates. The visualization part permits the user to view or annotate the query results in various ways. These results support the wide variety of data types and presentation methods required by neuroradiologists. The database gives us the possibility for data mining and defining interesting findings. The second system is the Epilepsy system, which includes image data from MRI and SPECT scans, and EEG analysis results, and is used for patients with epilepsy. We determine topographic selectivity and diagnostic utility of brain

atrophy in probable Alzheimer's disease (AD) and correlations with demographic factors such as age, sex, and education.

Another important application, which has many similarities with information retrieval, is the one where information is not stored in an archive but flows to the users continuously in a single stream (or several streams). This happens, for example, with information delivered by news agencies or with broadcast TV programs [7], or in a surveillance system. In such cases, an enormous quantity of information is received by the users, who, however, are only interested in a very limited part of it. The process of selecting only the significant information is called information filtering. In general, this process can be applied to describe a variety of processes involving the delivery of information to people who need it. The complexity of the filtering process is linear with the number of streams, the number of filters, and the number of features used to represent the video data. The entire process must be performed in real-time so that we can arrive at a situation where most of the processing power can be dedicated to filtering. We present a novel approach for efficient video stream filtering that is based on the use of the MPEG-7 descriptors and exploits the properties of metric spaces in order to reduce the computational load of the filtering receiver [1]. The retrieval process is based on a simple schema: users specify their request needs (e.g. a set of keywords or a sample image) that are translated into a system query. The items in the archive are compared with the user's query in order to determine if they are relevant for the user's request. To process this type of query, it is necessary to determine a set of properties of the objects stored in the archive (usually called features) and a similarity measure to compare queries and archive objects. Video features can be described by using the MPEG-7 standard. In case the similarity measure is metric, many possible approaches to create indexes can be adopted. These indexes allow improving the efficiency of the retrieval process by comparing the query only with a limited number of objects in the archive. Our approach goes toward a solution of this problem by proposing a novel approach to Audio/Video filtering that makes use of simple additional information sent together with the video. This allows avoiding the comparison of the filter with video features for many non-matching videos or video components that cannot pass the filter. The proposed approach requires that the measure of the similarity between the filter and the video representative is metric and it is based on the use of the well known technique of pivots.

The specific challenges in the field are highlighted. Conclusion remarks about the future of the multimedia retrieval are drowning.

Bibliography

[1] Falchi F., Gennaro C., Savino P., Stanchev P., Efficient Video Stream Filtering, accepted for *ACM Multimedia,* 2007

[2] Stanchev P., Amato G., Falchi F., Gennaro C., Rabitti F., Savino P., "Selection of MPEG-7 Image Features for Improving Image Similarity Search on Specific Data

Sets", *7-th IASTED International Conference on Computer Graphics and Imaging, CGIM 2004*, Kauai, Hawaii, 2004, 395-400.
[3] Stanchev P., Fotouhi F., Siadat M-R., and Soltanian-Zadeh H., Multimedia Mining- "A High Way to Intelligent Multimedia Documents", *book chapter 7- Medical Multimedia and Multimodality Databases,* D. Djeraba (Eds) - Kluwer Accademic Publishers, 2002, 139-160.
[4] Stanchev P., Green Jr D., Dimitrov B., Semantic Notation and Retrieval in Art and Architecture Image Collections, *Journal of Digital Information Management, Vol. 3, No. 4, December 2005*, 218-221
[5] Stanchev P., Introduction to MPEG 7: Multimedia Content Description Interface, *International Conference on "Visualization, Imaging and Image Processing (VIIP 2004),* September 6-8, 2004, Marbella, Spain
[6] Stanchev P., Semantic Video and Image Retrieval - State of the Art and Challenges, ACMSE 2007, *The 45th ACM Southeast Conference*, Winston-Salem, North Carolina, USA, March 24, 2007, 512-513
[7] Zhong D., Kumar R., and Chang S.-F. Real-time personalized sports video filtering and summarization. In *Proceedings of the ninth ACM international conference on Multimedia*, pages 623– 625. ACM Press, 2001.

Author's Information

Peter Stanchev is a Professor of Computer Science at Kettering University, Flint, Michigan, USA and Professor and Department Chair at the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria. He is also affiliated with the Institute of Information Science and Technologies, Italian National Research Council, Pisa, Italy. He has published two books, more than one hundred and fifty papers in monographs, journal, and peer-reviewed conferences that have more than 250 citations. His research interests are in the field of Image Processing, Multimedia Systems, Database Systems, and Expert Systems. He lectures courses in the areas of: Database Systems, Information Retrieval, Data Mining, Human-computer Interaction, Computing and Algorithms, Web Technology, Computer Architecture, Design of Information Systems, Computer Operating Systems, Image Databases, Fuzzy Sets, Decision Support Systems, Multimedia Databases, Expert Systems, Image Processing, Computer Graphics and Game Design. He is a Co-Chair of the annual multimedia semantics workshops.

Krassimira Ivanova is a research fellow at the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria. Her research interests are in the field of Content Based Image Retrieval, Software Engineering and Theoretical Bases of Informatics. She is a technical editor of the International Journal "Information Theories and Applications" and member of the Program Committee of annual conference on Artificial Intelligence "Knowledge-Dialogue-Solution".