Proceedings The 45th ACM Southeast Conference

March 23-24, 2007, Winston-Salem, North Carolina

Darina Dicheva, V. Paúl Pauca, and Jennifer Burg, editors

Semantic Video and Image Retrieval — State of the Art and Challenges

Peter L. Stanchev
Kettering University
Flint, MI, USA 48504
1 810 762 7927
pstanche@kettering.edu

BRIEF SUMMARY

In this tutorial we will discuss the history, the state of art, and the future of semantic image and video retrieval. The process of image description consists of extracting the global image characteristics, recognizing the image-objects, and assigning semantics to these objects. 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 [2]. 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.

The MPEG-4 standard is presented threw the ANIMATION system [3] – a system for scene animation and content creation, retrieval, and display. The object recognition mechanism in the ANIMATION system includes three steps: (1) Low level animation images analysis - element recognition based on the Attribute Relational Graphs; (2) Object recognition - based on production rules with degree of recognition; (3) Image interpretation - based on the Dempster-Shafer theory of evidence.

The MPEG-7 descriptors can be classified as general visual descriptors and domain specific descriptors. The former include color, texture, shape, and motion features. The latter includes face recognition descriptors. Color is one of the most widely used image and video retrieval feature. The MPEG-7 standard includes five color descriptors, which represent different aspects of the color and include color distribution, spatial layout, and spatial structure of the color. The histogram descriptors capture the global distribution of colors. The dominant color descriptor represents the dominant colors used. The color layout descriptor captures the spatial distribution or layout of the colors in a compact representation. While the MPEG-7 standard accommodates different color spaces, most of the color descriptors are constrained to one or a limited number of color spaces for ensuring inter-operability. The image texture is one of the most important image characteristic in both human and computer image analysis and object recognition. Visual texture is a property of a region in an image. There are two texture descriptors in MPEG-7: a homogeneous texture descriptor and edge histogram descriptor. Both of these descriptors support search and retrieval based on content descriptions. MPEG-7 supports region and contour shape descriptors. Object shape features are very powerful when used in similarity retrieval. Although distance functions are not part of the standard, we will present the most used distance functions.

A technique for improving the similarity search process of images in a Multimedia Content Management System is analyzed [1]. 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. 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 will also illustrate several aspects of art databases [4]. We study which descriptors are best for art images. We use nine hundred art images from more than two hundred artists. Nineteen century paintings include Neoclassicism, Romantics, Realism, Impressionism, and Symbolism. Fifteen through seventeen century paintings represent Renaissance and Baroque. Middle ages paintings include Byzantium, Islam, and Gothic, while ancient art paintings include Middle East, Egypt, Greece, and Rome art. We will show that different color descriptors give substantially different output because the color is the most important feature in an art image.

Medical images are one of most significant image applications. As an illustration, two systems are presented [5]. The first system is MEDIMAGE, which is a multimedia database for Alzheimer's disease patients. It contains MR images, text, and voice data, and it is used to find correlations of brain atrophy in Alzheimer's patients with different demographic factors. The Epilepsy system

ACMSE 2007, March 23-24, 2007, Winston-Salem, North Carolina, USA

includes image data from MRI, SPECT, scans, and EEG and is used to analyze patients with epilepsy.

The main features of several well-known multimedia systems will be discussed.

Every day we are overwhelmed by information of many types: TV channels, news feeds, web sites, to name a few. Without an efficient and effective filtering support, much time and effort is required in finding the information that we really need in this highly dynamic information age. 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.

Conclusion remarks about the future of the multimedia systems will be drown.

REVERENCES

1. Amato G., Falchi F., Gennaro C., Rabitti F., Savino P , Stanchev P., "Improving Image Similarity Search Effectiveness in a Multimedia Content Management System", ***Tenth International Workshop on Multimedia Information Systems,*** August 25 - 27, 2004 College Park, MD, 139-146.
2. MPEG home page - http://www.chiariglione.org/mpeg/
3. Stanchev P., Dimitrov B., Rukov V., "Object recognition in the ANIMATION system", ***5th IASTED International Conference "Computer Graphics and Imaging",*** CGIM 2002, August 12-14, Kauai, Hawaii, USA, 2002, 27-31.
4. Stanchev P., Green Jr D., Dimitrov, Some Issues in the Art Image Database Systems, ***Journal of Digital Information Management***, Volume 4, Issue 4, December 2006, 227-232
5. Stanchev P., Fotouhi F., Siadat M-R., 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.

TUTORIAL GOALS

After the tutorial the participants:

1. Will be able to explain the main steps in the history, state of the art, and future of semantic image and video retrieval
2. Will familiarize themselves with the main descriptors in MREG 7
3. Will be able to build simple semantic image retrieval systems
4. Will know of examples of video filtering
5. Will be able to find the best descriptor for a particular image database
6. Will have knowledge about the specific features of medical and art databases

OUTLINE OF THE TUTORIAL

1. Introduction
2. History
3. Image Standards - MPEG-4
4. MPEG-7
   - 4.1. Multimedia Description Schemes
   - 4.2. MPEG7 - Visual descriptors
   - 4.3. Dominant Color descriptor
   - 4.4. Scalable Color descriptor
   - 4.5. Color Structure descriptor
   - 4.6. Homogenous Texture descriptor
   - 4.7. Texture Browsing descriptor
   - 4.8. Edge Histogram descriptor
   - 4.9. Region shape
   - 4.10. Contour Shape
   - 4.11. Evaluating the effectiveness of MPEG-7 descriptors
5. Art Systems
   - 5.1. High level color characteristics
6. Medical Systems
7. Examples of Image Database Systems
8. Examples of Video Retrieval Systems
9. Efficient Video Stream Filtering
10. Conclusions

INTENDED AUDIENCE

Computer science students, researchers, people willing to build and use multimedia systems. No special preliminarily knowledge is required.

BIOGRAPHY OF PRESENTER

Peter Stanchev is a Professor of Computer science, 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, Image Database Systems, 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, and Computer Graphics. He is the C Chair of the annual multimedia semantics workshops.