Video Retrieval with Multiple Image Search Strategies

Alexander G. Hauptmann
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 1448
alex@cs.cmu.edu

Michael G. Christel
CS Dept. and HCI Institute
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 7799
christel@cs.cmu.edu

Norman D. Papernick
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 USA
+1 412 268 7557
norm@cs.cmu.edu

INTRODUCTION

The Informedia digital video library has been on the forefront of automatic analysis of video streams, as well as pioneering interactive display and retrieval mechanisms for video data through various multimedia surrogates including titles, storyboards, and skims [1, 5]. This demonstration will highlight the use of Informedia processing and interfaces in the context of the TREC Video Retrieval Track, whose high-level goal is the investigation of content-based retrieval from digital video [4].

Multiple Image Search Systems for Interactive Use

In this demonstration we will show a newly developed aspect of Informedia data that allows a user to choose among multiple image retrieval systems to perform an image search over the video library. Among the image retrieval systems we will be demonstrating are:

- An implementation of the color-matching component of IBM's query by image content (QBIC) image retrieval system [2].
- Histogram-based systems utilizing the Munsell color space using different image partitioning resolutions (3x3, 5x5, etc.) [3].
- Systems using the RGB color space, in different partitioning versions, using color histograms as well as texture information.

Visualizing Key Image as the Search System Sees It

Since users have no a priori knowledge of the characteristics of the types of images in the collection, nor do they understand the space of color and texture representations for the individual image search methods, the system also allows the user to visualize the underlying representation of the search method. This is done by presenting the original query image together with an alternate image that results from the transformation or mapping of the particular search method. For example, when the QBIC system reduces the color space to 256 colors in 256 regions, the resulting transformation image shows a coarse granularity version of the original image, where certain color nuances and fine shape distinctions are lost. In contrast, a 3x3 RGB color histogram image search system divides the image into an even coarser 9-region set. In each region, the colors of the histogram will be considered randomly distributed for the purposes of the color match. As a result, this transformation gives a very different sense of what the image search system 'sees' as the target image to match. Users cannot only choose different image search systems, but they can also view the alternate representation of an image based on the transformation of the selected search method. This helps users get an intuitive feel for the different search systems and what it considered similar or important by each one.

THE TREC INTERACTIVE VIDEO RETRIEVAL TASK

The baseline data for this demonstration will consist of the 2001 TREC Video Retrieval collection consisting of 11 hours of video data. The data is further segmented into 80,000 I-frames over which the video search will be conducted. The I-frames are aggregated into 8000 shots, with each shot having a characteristic key frame as the visual thumbnail surrogate, i.e., summary representation.

This system is almost identical to the one used in the TREC interactive video retrieval task, which obtained a remarkable precision of 95% on general search items and 66% precision with 67% recall on known-item searches. Users will be able to take the same queries that were used for the TREC video retrieval evaluation and explore how they could find the relevant answers using the Informedia system. Our system will be demonstrated running on a 1GHz laptop computer, with users interacting with the system to retrieve video clips.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation under Cooperative Agreement No. IRI-9817496 and grant number IIS-012164. This work is also supported in part by the Advanced Research & Development Activity (ARDA) under contract number MDA908-00-C-0037.

REFERENCES

1. Carnegie Mellon University Informedia Digital Video Library, 2002, http://www.informedia.cs.cmu.edu.
2. Faloutsos, C., et al. Efficient and Effective Querying by Image Content. *Journal of Intelligent Information Systems 3*, 3/4 (1994), 231-262.
3. Mojsilovic, A., et al. Matching and Retrieval Based on the Vocabulary and Grammar of Color Patterns. *IEEE Trans. Image Processing, 9*, 1 (2000), 38-54.
4. NIST TREC Video Retrieval Track, 2001-2002, http://www-nlpir.nist.gov/projects/trecvid/.
5. Wactlar, H., et al. Lessons Learned from the Creation and Deployment of a Terabyte Digital Video Library. *IEEE Computer 32*, 2 (February 1999), 66-73.