MKLab Interactive Video Retrieval System

Stefanos Vrochidis, Paul King, Lambros Makris, Anastasia Moumtzidou, Vasileios Mezaris and Ioannis Kompatsiaris

Informatics and Telematics Institute / Centre for Research and Technology Hellas
6th Km Charilaou-Thermi Road, 57001 Thermi-Thessaloniki, Greece

+302310464160

{stefanos, king, lmak, moumtzid, bmezaris, ikom}@iti.gr

ABSTRACT

In this paper, the MKLab interactive video retrieval system is described.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – retrieval models, search process.

General Terms

Algorithms, Performance, Design, Experimentation

Keywords

Search engine, retrieval, visual, hybrid, video, MPEG-7, query

1.INTRODUCTION

The Search Engine implemented by MKLab is capable of handling video resources, integrating different search modules.

2.VIDEO RETRIEVAL SYSTEM

In general, the developed application is a hybrid interactive retrieval system, combining basic retrieval functionalities with a user-friendly interface supporting the submission of queries and the accumulation of relevant retrieval results.

The following basic retrieval modules are supported:

- Visual similarity search module
- Textual information processing module

The search system is built on web technologies (more specifically, php, JavaScript and a mySQL database) providing a GUI for performing retrieval tasks over the internet. Using this GUI, the user is capable of employing any of the supported retrieval functionalities. Retrieval results are presented ordered by rank in descending order with links to the temporally neighboring shots of each one. Storage of relevant shots uses a storage structure that mimics the functionality of the shopping cart found in electronic commerce sites.

3.RETRIEVAL MODULES DESCRIPTION

3.1Visual Similarity Search

Content based similarity search is realized using MPEG-7 visual descriptors capturing different aspects of human perception such as color and texture [1]. By concatenating these descriptors, a feature vector is formulated to compactly represent each image in the multidimensional space. An r-tree [2] structure is constructed off-line by using the feature vectors of all images. In the query phase, a feature vector is extracted from the example image and submitted to the index structure. Subsequently, the set of resulting images is ranked using custom distance metrics between their feature vectors.

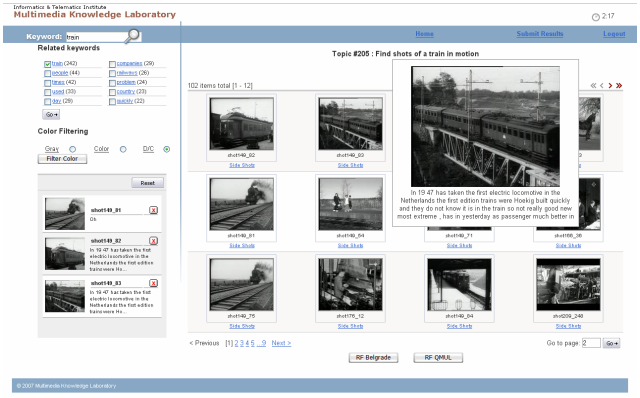

Figure 1. GUI of MKLab Video Retrieval System

3.2Textual Information Processing Module

Text query is based on audio annotations automatically transcribed from video. A controlled vocabulary paradigm is implemented rather than full-text. Indexing is implemented using thesauri, built from concepts identified in the LSCOM, the CalTech256, the U.S. Library of Congress Name Authorities, and the GEOnet Names Server. Thesaural relationships are automatically expanded by referencing the lexical databases of WordNet and Wikipedia as well as application of a clustering algorithm called the Levenshtein Similarity Metric [3].

Index and search terms are stemmed using the Porter algorithm and ranking is achieved using the TF-IDF weight algorithm (Term Frequency-Inverse Document Frequency). Facet analysis was used during domain analysis.

4.REFERENCES

[1] V. Mezaris, H. Doulaverakis, S. Herrmann, et.al. “Combining Textual and Visual Information Processing for Interactive Video Retrieval”, TRECVID 2004 Workshop, Gaithersburg MD, USA, November 2004.

[2] A. Gutmann, “R-trees: a dynamic index structure for spatial searching,”, ACM International Conference on Management and Data (SIGMOD’88), Siena, Italy, 1988.

[3] Lucia Specia and Enrico Motta, "Integrating Folksonomies with the Semantic Web", The Semantic Web: Research and Applications, pp. 624--639, 2007