K-Space Interactive Search

Peter Wilkins, Alan F. Smeaton, Noel E. O'Connor, Daragh Byrne
Centre for Digital Video Processing & Adaptive Information Cluster, Dublin City University
asmeaton@computing.dcu.ie

ABSTRACT

In this paper we will present the K-Space[1] Interactive Search system for content-based video information retrieval to be demonstrated in the VideOlympics. This system is an extension of the system we developed as part of our participation in TRECVID 2007 [1]. In TRECVID 2007 we created two interfaces, known as the 'Shot' based and 'Broadcast' based interfaces. Our VideOlympics submission takes these two interfaces and the lessons learned from our user experiments, to create a single user interface which attempts to leverage the best aspects of both.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems

General Terms

Design, Experimentation

Keywords

User Interface, Video Retrieval, TRECVID

1. SEARCH SYSTEM

For TRECVID 2007 we developed two user interfaces built on top of a common retrieval engine whose aim was to investigate the role of temporal context in interactive retrieval performance. To frame our hybrid interface, we describe these two interfaces.

Our research objective for TRECVID 2007 was to investigate the role of temporal context in retrieval performance. Whilst many previous studies have examined temporal context, we decided to take this examination to the extreme by constructing one interface with no temporal context, and a second interface which showed maximum temporal context, these were the 'Shot based' interface and the 'Broadcast based' interface (see Figure 1). Our motivation was that with the change in content for TRECVID 2007, videos may be more homogenous in content, therefore it may be of benefit to rank an entire broadcast as opposed to just shots.

[1]K-Space is a European Network of Excellence (NoE) in semantic inference for semi-automatic annotation and retrieval of multimedia content (http://www.k-space.eu)

Broadcast-based:
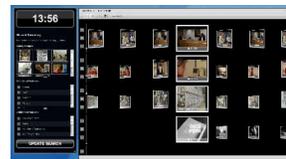
Shot-based:

Figure 1: User interface variations

The primary difference between these two interfaces is in the presentation of results. The 'Shot' based system displays shots in ranked based order, left to right, top to bottom. Adjacent shots displayed in this interface are not related. The 'Broadcast' based system however displays entire broadcasts to a user, rather than individual shots. Each horizontal line in the results represents a broadcast, with each made up of temporally sorted shots from that broadcast. When a broadcast is presented to a user it is centered on the highest ranked shot within the broadcast. Broadcasts are ranked through an accumulation function which examines the number of relevant shots in a broadcast, weighting shots which appear earlier in the rankings greater.

Our hybrid interface is the result of our analysis of our TRECVID 2007 participation. The interface leverages the recall power of the 'Shot' based interface, but for each shot the interface will also display a limited amount of temporal context which can be further explored. It uses the same retrieval engine, making use of ASR text, low-level visual features and semantic concept detectors developed from our participation in the High-Level Feature Extraction task.

Acknowledgments

2. REFERENCES

[1] P. Wilkins and et al. KSpace at TRECVid 2007. In *TRECVID 2007 – Text REtrieval Conference, TRECVID Workshop, Gaithersburg, Md., 5-6 November 2007,* 2007.