



University
of Glasgow

Urruty, T. and Hopfgartner, F. and Villa, R. and Gildea, N. and Jose, J.M.
(2008) *A cluster-based simulation of facet-based search*. In:
IEEE/ACM Joint Conference on Digital Libraries, 16-20 June 2008,
Pittsburgh, USA.

<http://eprints.gla.ac.uk/5747/>

Deposited on: 03 June 2009

A Cluster-based Simulation of Facet-based Search

Thierry Urruty, Frank Hopfgartner, Robert Villa, Nicholas Gildea, Joemon M. Jose
University of Glasgow
Glasgow, United Kingdom
{urruty, hopfgarf, villar, ngildea, jj}@dcs.gla.ac.uk

ABSTRACT

The recent increase of online video has challenged the research in the field of video information retrieval. Video search engines are becoming more and more interactive, helping the user to easily find what he or she is looking for. In this poster, we present a new approach of using an iterative clustering algorithm on text and visual features to simulate users creating new facets in a facet-based interface. Our experimental results prove the usefulness of such an approach.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Evaluation/ methodology

General Terms

Experimentation, Performance

1. INTRODUCTION

In our poster, we present a facet-based retrieval model which attempts to improve the usability and the effectiveness of a video retrieval system. The intuition behind our work is that facets may provide a good visualisation and structuring system for users engaged in complex searches. Instead of carrying out a search one query at a time, the users may execute, and view, multiple queries, enabling them to structure their retrieval process and to easily follow new retrieval directions. Most interactive video retrieval systems are evaluated in laboratory based user experiments. This methodology is often inadequate. An alternative way is to use an evaluation methodology in which interactive user actions are simulated on a test collection. Hopfgartner and Jose [1] employed a simulated evaluation methodology which simulated users interacting with state-of-the-art video retrieval systems. We adapt their simulation approaches in simulating users interacting with a facet-based video retrieval interface. We follow an iterative clustering algorithm approach based on the retrieved set of results. We aim to simulate users creating and interacting with new facets in the interface.

2. ITERATIVE CLUSTERING APPROACH

The main goal of our facet-based interface is to help the user to create a complex query with separated and structured views of different sub-queries. Our iterative clustering

approach mainly aims to simulate the user in their search tasks. First, we use textual and visual features to cluster retrieved results from an initial query. Then, we use these clusters to create more specific queries. Different strategies to create new queries have been tested and will be presented in the poster. Those queries are then used to propose automatically new sets of results in new facets. Finally, an iterative clustering process may be used on cluster-query results to refine the queries and consequently the retrieved results. The iterative process allows the display of the new result lists as new facets on the interface, or the launching of a new clustering call on each result list.

Our different experiments are based on the TRECVID 2006 collection using the set of 24 topics contained in the data collection. Each topic has several keywords and a judgement list of 60 to 775 relevant documents. In the poster, we will present various sets of experiments evaluating the precision of visual and text queries. They show that our approach using iterative clustering to simulate the actions of users creating facets can be useful and effective. Indeed, iterative clustering using visual features provides new relevant facets for the user interface, i.e. the facets have a better precision than initial query facets or present new relevant documents that were not included in the initial set of results. Additionally, we present some text experiments using query expansion of one key word and new queries with five representative key words. These experiments show that iterative clustering using text queries can also improve the precision of retrieved results. This is due to the diversity of keywords created and used in our approach which introduce different relevant results than the one retrieved by the initial text query.

The results of this cluster-based simulation approach have shown the probable benefits of a facet-based interface. In addition, they have also shown the benefits of clustering in suggesting new facets to the users during a retrieval session.

3. ACKNOWLEDGEMENT

This research was supported by the European Commission under the contracts FP6-027026-K-SPACE and FP6-027122- SALERO.

4. REFERENCES

- [1] F. Hopfgartner and J. Jose. Evaluating the Implicit Feedback Models for Adaptive Video Retrieval. In *ACM MIR '07 - 9th ACM SIGMM International Workshop on Multimedia Information Retrieval, Augsburg, Germany*, pages 323–332, 09 2007.