



[Hopfgartner, F.](#), [Hannah, D.](#) and [Jose, J.M.](#) (2008) An adaptive news video retrieval framework. In: Third International Conference on Semantic and Digital Media Technologies (SAMT 2008), Koblenz, Germany, 3-5 December 2008

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/5735/>

Deposited on: 12 April 2018

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

An Adaptive News Video Retrieval Framework

Frank Hopfgartner
University of Glasgow
Computing Science
Glasgow G12 8RZ, U.K.
hopfgarf@dcs.gla.ac.uk

David Hannah
University of Glasgow
Computing Science
Glasgow G12 8RZ, U.K.
davidh@dcs.gla.ac.uk

Joemon M. Jose
University of Glasgow
Computing Science
Glasgow G12 8RZ, U.K.
jj@dcs.gla.ac.uk

ABSTRACT

The increasing popularity of video sharing platforms such as YouTube and Google Video increases the need to further study how users can be assisted in their search for videos they are interested in. In this demo, we present a video retrieval system which guarantees the user easy and effective access to a large news video collection. This system can be used to further study interaction methodologies, aiming for a personalised video retrieval model which adapts retrieval results to the user's interests.

1. INTRODUCTION

With the improving capabilities of current hardware systems, there are ever growing possibilities to store and manipulate videos in a digital format, leading to a growing number of video archives. People build their own digital libraries from materials created through digital cameras and camcorders, and use systems such as YouTube or Google Video to place this material on the web. Even national broadcasters such as the BBC have extended their services in using YouTube to offer commissioned content and news clips.

Unfortunately, this data creation process is not matched by any comparable tools to organise and retrieve video information. There is a need to create new retrieval engines to assist the user in searching and finding video scenes he/she would like to see from many different video files. TRECVID [4], a set of workshops which provides common test collections in order to enable large scale evaluation of research approaches, plays an important role in the development of such engines. In the last few years, TRECVID concentrated on TV news as data corpus, a dataset which introduces various research questions including automatic segmentation of news videos into semantically related stories and retrieval of relevant results.

However, in order to study how users interact with a retrieval system, user-centred studies are needed, which are directly influenced by the video collection used in the study. Assum-

ing that users show a general interest in daily news, relying on news videos is reasonable. However, users are more likely to find what they want in an up-to-date news corpus rather than in older news. Relying on the TRECVID corpus, which has been recorded several years ago, can hence influence the result of a study.

In this demo paper, we present a news video retrieval system which automatically processes daily news broadcasts and enables the user to retrieve and to browse this collection. This system can be used to further study interactive video retrieval, aiming towards an adaptive retrieval model. In Section 2, we introduce the system and discuss its application in Section 3.

2. SYSTEM ARCHITECTURE

The video retrieval system is a web based application based on AJAX technology. AJAX takes away the burden of installing additional software on each client. As with most state-of-the-art video retrieval systems [5], it can be divided into four main components: Three components run on a web server (data collection, retrieval backend and user profiles) and one, the user interface, runs on the client side.

The current system automatically captures the daily BBC One O'Clock news. The bulletin has a running time of 30 minutes and is broadcasted every day from Monday to Friday on BBC One, the UK's main broadcasting station. The BBC enriches its television broadcast with Ceefax, a closed caption (teletext) signal which provides televisual subtitles for the deaf. Adopting the techniques introduced in [3], we use a colour histogram-based method to detect shot boundaries in our video files. Furthermore, we detect example keyframes by calculating the average colour histogram for each shot and extract the frames within the shot which are closest to the average. In an additional step, we combine the keyframes belonging to the same shot to form an animated presentation of the shot. Based on its textual, visual and audio features, we segmented the news videos into semantically related story segments, the unit of retrieval in our system. Moreover, we extract named entities from the transcript using GATE¹. Each story unit hence consists of a set of representative animated keyframes which have been extracted from each shot of the story, the transcript and the identified named entities. We use the open-source full-text search engine MG4J² to index and administer the collec-

¹<http://gate.ac.uk/>

²<http://mg4j.dsi.unimi.it/>

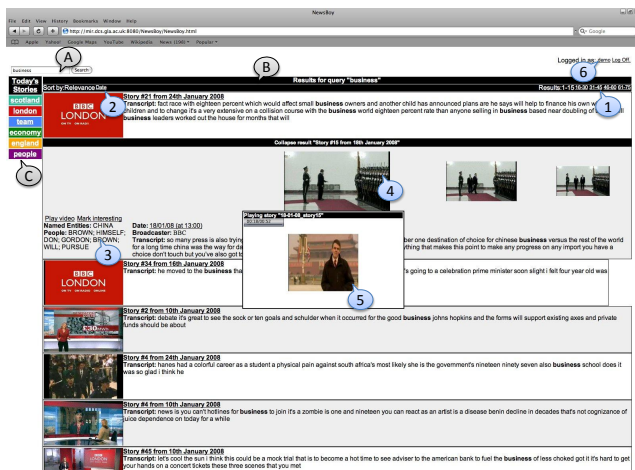


Figure 1: Video Retrieval Interface

tion. Further details about the system’s architecture have been published in [2]. In the remainder of this section, we will introduce the interface of the system.

Figure 1 shows a screenshot of the retrieval interface. It provides facilities to search within the video collection and to browse through it. The interface can be divided into three main panels: search panel (A), result panel (B) and a panel which contains content based on the users’ interests (C). In the search panel (A) users can formulate and carry out their searches by entering a query and clicking the button to start the retrieval. The result panel (B) lists retrieved results. As in most state-of-the-art video retrieval interfaces, results are presented by one keyframe and a shortened part of the text transcript. The panel displays a maximum of 15 results, more results can be displayed by clicking on the annotated page number (1). The results can be sorted with respect to their relevance to the query or chronologically by their broadcasting date (2). Moving the mouse over the keyframe of one of the results will pop up a tooltip with additional textual information. Users can get additional information about the result by clicking on either the text or the keyframe. This will expand the result and present additional information including the full text transcript, broadcasting date, time and channel and a list of extracted named entities such as people, locations and relative times (3). In the example screenshot, the second search result has been expanded. The shots forming the news story are represented by animated keyframes of each shot. Users can browse through these animations by clicking on the keyframe. This action will center the selected keyframe and surround it by its neighbored keyframes. The keyframes are displayed in a fish-eye view (4), meaning that the size of the keyframe grows larger the closer it is to the focused keyframe. In the expanded display, a user can also select to play a video or to mark it as interesting. Clicking on “play video” starts playing the story video in a new panel (5). The personalisation panel (C) guarantees direct access to the topics users’ are most interested in. The interests can be edited by clicking on the user name (6) on the top right side of the interface. In the next section, we introduce the research questions we studied using the system.

3. RESEARCH APPROACHES

The introduced video retrieval framework provides us a stable system which can be used to study various aspects in video retrieval. In [1], we automatically capture the user’s evolving interest in multiple aspects of news stories by unobtrusively profiling the user based on his interactions with the system. The aim of this study was to see whether implicit relevance feedback can be used to adapt retrieval results based on the user’s personal interestes. Therefore, we asked users who use the internet for their daily news gathering to include the news video retrieval system into this process for one month. The system automatically created user profiles based on the users’ interactions with the interface, which consist of weighted keywords that have been extracted from the results the user interacted with. The weighting for the keyword depended on the type of implicit feedback users could provide. Within this study, we took four features into account: Using the tooltip highlighting feature, playing a video, browsing through keyframes by clicking on them and expanding a result. The news aspects are identified by clustering the content of the profile based on the terms’ weighting. The study supported the hypothesis that implicit relevance feedback is useful for user profiling and adaptation of retrieval results.

In the demo, we will introduce this system and discuss the effect of different weighting functions on user profiling and content recommendation.

4. ACKNOWLEDGMENTS

This research was supported by the European Commission under the contracts FP6-027026-K-SPACE. It is the view of the authors but not necessarily the view of the community.

5. REFERENCES

- [1] F. Hopfgartner, D. Hannah, N. Gildea, and J. M. Jose. Capturing Multiple Interests in News Video Retrieval by Incorporating the Ostensive Model. In *PersDB’08 - Second International Workshop on Personalized Access, Profile Management, and Context Awareness in Databases, Auckland, New Zealand*, pages 48–55, 08 2008.
- [2] F. Hopfgartner and J. M. Jose. *Advances in Semantic Media Adaptation and Personalisation Vol. 2*, chapter Towards an Adaptive Video Retrieval System. CRC Press / Boca Raton, Florida, 2009. in press.
- [3] C. O’Toole, A. Smeaton, N. Murphy, and S. Marlow. Evaluation of Automatic Shot Boundary Detection on a Large Video Test Suite. In *Proceedings of Challenges in Image Retrieval*, Newcastle, UK, 02 1999.
- [4] A. F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and trecvid. In *MIR ’06: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, pages 321–330, New York, NY, USA, 2006. ACM Press.
- [5] C. G. M. Snoek, M. Worring, D. C. Koelma, and A. W. M. Smeulders. A Learned Lexicon-Driven Paradigm for Interactive Video Retrieval. *IEEE Transactions on Multimedia*, 9(2):280–292, February 2007.