Searching and Recommending Sports Content on Mobile Devices

David Scott[1], Cathal Gurrin[1,2], Dag Johansen[2], Håvard Johansen[2],

Dublin City University, Dublin, Ireland[1], Untiversity of Tromsø, Norway[2]
{david.scott,cgurrin}@computing.dcu.ie {haavardj,dag}@cs.uit.no

Abstract. As users rely more on mobile devices to access to video and web information we must adapt current technologies and develop new ones to support multimodal device access to digital archives. In this work we report on a prototype video retrieval system for TV sports content which utilizes sports summarization and personalization to deliver a multi-modal user experience.

1 Introduction

In recent years we have witnessed a revolution in how people locate and access video content. No longer are viewers restricted to accessing TV shows through a broadcaster schedule on a TV, now the viewer can access video content when they desire using DVRs or online WWW archives. In addition, not only do viewers access TV content using a computer, they now use mobile devices and the rate of employment of mobile video access has increased 52.2% from 2008 to 2009 [1]. This massive surge in usage is set to continue as more and more content is migrated to mobile devices. At present most video delivered to mobile devices is simply via a mobile front end to WWW archives such as YouTube, and there has been little effort to modify the video data specifically to support mobile access. Yet, although mobile devices have improved much recently in the way they handle user input (e.g. presence of qwerty keyboard), touch screens, etc., it is still more challenging for a user to access content using a mobile device and it remains necessary to keep user input to a minimum where possible. To support multimodal device access to a digital video archive, we have developed a video search and retrieval system (called DAVVI)) for sports video content which utilises automatic sports summarization and recommendation technologies to support both mobile and desktop device access.

2 Prototype System Description

The sports video retrieval prototype operates over an archive of a number of months of Norwegian and UK football content, and supports both WWW and mobile access, and is implemented using Silverlight. There are a number of innovative aspects of the system, including:

- Automatic event segmentation and tagging of events in sports matches.
- Real-time, automatic generation of result video documents (not lists of clips) as the unit of retrieval.
- Personalization and recommendation of sports content based on a stored user profile, primarily to support mobile access.
- Identification of a number of key event types, such as; goal, corner, free, yellow card, penalty.
- Multiple bit rate streaming to support multimodal device access.

The most innovative aspects of this demo include sports segmentation with keyframe extraction, the recommendation of content and the video documents result generation, each of which we will now discuss.

The event segmentation technique for sports content has received a lot of attention recently using purely visual processing techniques, for example our work [2] which uses an SVM to combine evidences from a number of key attributes that suggest an important event taking place, however the accuracy is still only in the region of 70%. By utilizing external sources of evidence we have been able to surmount this problem and achieve near 100% accuracy. However, this poses another research challenge, that of accurately identifying the optimal start and end of an event, which ideally should be defined by the specific user query. For example, one needs to define where a goal event begins and ends and not simply to return a specific length result. The external source of evidence we utilize was football match statistics trawled from the web, and this gives us an accurate timestamp of the actual event taking place (to the second). Once this is known, the event segmentation technique, based on [2], determined where the event started (build up to the event) and where the event ended (moved on to another separate event). Two segmentation approaches were employed, manual and automatic, and we are currently evaluating the effectiveness of both.

Manual: Each event type is segmented using an absolute time period judged from the event time, each event type can be given a different duration. Yellow cards for example were identified as short event, but goal events defined as thirty second events. The obvious disadvantage of this is non-optimal length of event segmentation.

Automatic: An automatic event segmentation technique optimises the length of the event segment, which can for example be shorter for mobile devices and longer for other devices. The starting point is defined by an increase in activity of crowd noise, fast visual motion and camera panning, which was also further reinforced by the parsed online match statistics. End points are determined as crowd excitement reduction, less camera movement and less player movement.

As mentioned, the unit of retrieval for this prototype is not individual video clips, but a generated video document that is comprised of a number of high ranked clips that best match a user's information need and profile. On the desktop device the result is a custom generated result video document, however the component clips are also retuned and can be manipulated by the user (e.g. reordered, deleted, etc.) to improve the resulting video document, which itself may be shared with buddies. However on the mobile device, taking into account the screen size limitation, the result of a query is only a custom generated result video document. For example, a result may contain all the goals scored by 'Ronaldo' after a corner kick in last season's football matches.

For recommendation purposes, content can be recommended on both mobile and desktop devices. The desktop prototype allows multi-user login, complex query formation, saving of playlists, recommendation of content to buddies and the creation of user profiles based on content access. To support mobile device access, the mobile version is bound to a single user, supports simple user query formation (e.g. free-text queries) and is mostly focused on utilizing the user profile to proactively seek out and recommend interesting content to the user (e.g. based on recent activity of a favorite player, or viewing habits of buddies, which is a form of collaborative filtering). The user profile is maintained by employing both explicit and implicit feedback, with explicit feedback being limited to the desktop interface (selection of favorite item) and implicit feedback (items fully watched) gathered on both mobile and desktop interfaces. Finally, hybrid filtering is employed to recommend content.

In Figure 1 we show screenshots of the desktop and mobile versions of the prototype DAVVI sports search engine. In the mobile interface you can see the text search box and a listing of the most recent video documents generated.

Figure 1. Prototype Interface to the Desktop and Mobile DAVVI (to be changed)

3 Future Work/Conclusions

In DAVVI we have developed a prototype sports event centered search engine which supports multimodal device access. For future work, an extensive user study is planned along with continuing to improve underlying algorithms and recommendation engines. Process migration to allow seamless video playback when swopping between different interaction devices is under development (e.g. changing devices from desktop to mobile would not affect.

References

1. Nielson Report, http://blog.nielsen.com/nielsenwire/online_mobile/americans-watching-more-tv-than-ever/

2. Sadlier, D & O'Connor, N, Event Detection in Field Sports Video using Audio-Visual Features and a Support Vector Machine. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 15, No. 10, , Oct 2005. (pp1225-1233)