SportsAnno: What Do You Think?

James Lanagan & Alan F. Smeaton
Centre for Digital Video Processing & Adaptive Information Cluster
Glasnevin
Dublin 9, Ireland
{jlanagan, alan.smeaton}@computing.dcu.ie

Abstract

The automatic summarisation of sports video is of growing importance with the increased availability of on-demand content. Consumers who are unable to view events live often have a desire to watch a summary which allows then to quickly come to terms with all that has happened during a sporting event. Sports forums show that it is not only summaries that are desirable but also the opportunity to share one's own point of view and discuss the opinions with a community of similar users. In this paper we give an overview of the ways in which annotations have been used to augment existing visual media. We present SportsAnno, a system developed to summarise World Cup 2006 matches and provide a means for open discussion of events within these matches.

1. Introduction

There is now a plethora of sports video and media on the Internet with sports channels offering live web casts of matches as well as recorded footage. Along with this video comes large numbers of reports written to capture all the major events within a match (e.g. Guardian Unlimited[1], BBC Sports[2], Sky Sports[3]). Since these written reports are essentially designed as a summary of the matches they describe, they may be used as a guide or key into the recorded video. They are as yet however, very disjoint and there appears to be no means available to a user for the simultaneous browsing of both written match reports and the associated video media. Beyond this, if a user wished to comment on the things said within a report, he/she must go to a third resource, a forum, to be able to write a reply. This loss of context and need to reference the original material requires a great deal of effort on the part of the user. More sites are beginning to see this problem and address this by allowing users to post comments at the bottom of articles published on the site. The style in which these comments are presented however does not appear to be of a discursive nature.

The advantage of the approach that is presented here is the continued augmentation of original data through the contribution of a community of users. Since video is always present, it is possible for a user, at any time, to see the arguments presented in writing first-hand and to couple it directly with the video. There is also the means to provide direct input into any discussions. Here we present SportsAnno, a video browsing system designed to allow users to read match reports taken whilst viewing the match video associated with the reports and to create in context comments used as the basis for discussion and searching between all the users of the system.

The immediate and easy access to both media types, coupled with the ability to leave comments within the text for other users leads to a more directed and communal style of annotation. The novelty of this system lies in it's ability to provide users with the chance to

[1] http://football.guardian.co.uk/
[2] http://news.bbc.co.uk/sports2/hi/football/world_cup_2006/default.stm
[3] http://home.skysports.com/worldcup

become up-to-date with any talking points and also to contribute easily to the on going discussion. As shown in Section 2, there is a great lack of systems which can currently give the user this opportunity.

This paper is structured as follows: In Section 2, background is given on current technologies and related research involving the summarisation of video and the annotation and linking of documents. Other systems designed for sports summarisation are also discussed. Section 3 gives an overview of the each of the components within the SportsAnno architecture which link together to create the full system. User studies and experiments are shown in Section 4 and Section 5 concludes with a presentation of future ideas and lessons learned during the creation and use of SportsAnno.

2. Related Work

Here we focus on the technologies and research areas which have been combined to create SportsAnno. It can be seen that while current and previous research has been extensive within each area, the overlap between the areas (involving the actual creation of a usable system comprised of elements from each area) is slight. This may be witnessed by the lack of comparable systems.

2.1 Video Summarisation

Sports video summarisation has been an active field of research for some time with researchers employing statistical algorithms both on the video directly as well as using natural language processing techniques on associated data in an attempt to index events within the video stream.

Summarisation of a video stream directly can be broken into two main approaches: multi-modal approaches which use a combination of audio/video/textual information to detect events of note in broadcast video (Cabasson & Divakaran, 2002)(Wang & Xu & Chang & Wan & Tian, 2005), and single-modality methods which use a single information source for event classification (Xie & Chang & Divakaran & Sun, 2002)(Zhang & Ellis, 2001). Both forms of classification have been used with success to analyse various genres of sports video ranging from baseball and soccer to snooker (Denman & Rea & Kokaram, 2003) and Gaelic football (Sadlier & O'Connor & Marlow & Murphy, 2003). While the single-modality approaches are able to yield much faster results, favouring them for real-time applications, the cost is a less accurate and incomplete classification of all possible events. It is for this reason that SportsAnno was implemented using a multi-modal approach so as to give the most complete initial classification as possible.

Direct analysis of the video stream however, whilst able to detect events based on audio/visual/textual onscreen data, does not provide the necessary level of information to narrow the semantic gap between low-level and high-level features present in video. Indeed, to do this it is necessary to look to associated streams of data. Two of the most widely available additional sources are website tickers and television closed captions.

Closed Captioning provides a textual transcript of all studio-audio spoken or commentary during the game/match broadcast. It is manually created and since it is a direct speech to text conversion, can consist of large amount of data which is of little relevance to the video itself. It can also be created live during the broadcast, sometimes resulting in errors in both synonym

recognition and spelling. It has however been shown to be of benefit in identify semantically different events when used in conjunction with audio/visual data (Assfalg *et al.*, 2003).

Web-tickers provide a far more directed and useful means of textual accompaniment to sports video data. These are sites that are updates minute-by-minute and provide brief descriptions of events as they happen. These descriptions can include keywords that are easily identified to aid semantic classification of events.

While systems have been developed to enable browsing of event-detected video highlights, there are no systems which allow for this video to become the focus of community discussion. As has been said earlier, nearly all broadcast sports are accompanied by newspaper reports, blog entries or personal email correspondence. When in a public form, this additional written material can be used to offer a richer interpretation of the broadcast video. Sport has always been a highly contentious topic and it seems that to ignore these valuable additional resources is to fail to utilise the analysed data to its full potential. These resources also provide within themselves a means for discussion and interaction between users. In the next section, we look at different forms of written media and the research which has been done into the annotation and augmentation of these document types.

2.2 Linkage and Annotation

With the advent of Web 2.0 and the technologies that contribute to it, the web-browsing paradigm of old has been shifted to give rise to a less author-centric internet. “Active browsing” refers to the process whereby a user not only browses the information already present but also creates new annotations. These annotations can be created for any number of purposes including to aid recollection, collation of ideas, creation of a group memory or to highlight particularly important sections within a document (Marshall, 1998).

Forums, wikis and blogs mean that is it now possible for users to change the linkage structure of web pages and provide arguments to whatever is written within the page, backed up by external evidence. This more commonly recognized form of annotation shows that users are becoming accustomed to this new paradigm, as evidenced for example, by a rate of growth of the blogosphere which is said to be doubling in size every 5-6 months. These are however specific types of webpage designed specifically to give the user the functionalities necessary to become part of the authoring process.

In order for this new approach to be extended and generalised, it is necessary to look at ways in which the advances within the field can benefit the user. Large amounts of research has concentrated on the digitisation of the annotation process in the context of study and research, work that has helped to show how best to bring the annotation styles used on paper to the digital medium (Schilit & Golovchinsky & Price, 1998), (Marshall & Bernheim Brush, 2004). Margilinia, underlining and highlighting are shown to be highly important indicators of the usefulness and value of specific sections within a document. It is therefore desirable that methods be introduced for equivalent annotations of digital documents.

Much work has already been done on creating tools that will allow for the establishment of new paths and trails through the web as well as the creation of additional content within a webpage without the specific consent of it’s author. Most intend to create a style of interaction akin to that envisioned by Vannevar Bush in his seminal paper “As We May

Think" (Bush, 1945). Tools such as iMarkup[4], Annozilla[5] and MADCOW (Bottoni et al., 2004) allow the user to create landmarks and form new paths and links. This means that new relationships may be formed between documents that were not originally envisioned by the author. Instead, these links grow out of particular user's needs and desire to collect information sources together.

The linking of two documents may be seen as a form of annotation since the link is created to imply that one document will provide additional information on the other. It is through this principle that the Google Page Rank algorithm gains its power (Brin S. & Page L. & Motwani R. & Winograd T., 1998). The creation of hubs and authorities is directly due to the linking of documents within the web (Kleinburg J.M., 1999). The extension of this principle to aid in automatic recognition of important social commentators is an area that deserves more research attention.

Linkage between documents can be used not only to help decipher which documents will be of most importance within a corpus but also which documents are to be judged most relevant in regards to any query issued by a user. This tenant holds true across several different genres of document:

1. E-mail: The most interesting documents can be seen as the documents that create the most feedback and response. Within an email corpus, the inherent thread structure also helps to weight the importance of a document. This approach does however lead to a different idea of importance to the traditional approach suiting it more to a serendipitous browsing session than the task of searching for a specific result.

2. Blogs: Much work has been done on the inter-linking structure of blogs so as to find the most influential bloggers. The linkage structure may be used to find communities, current topics of interest and people of influence (Chin & Chignell, 2006; Mishne & Glance, 2006).

Both of the above genres are recognised as important areas and now each have dedicated tracks within the Text Retrieval Conferences, TREC (Voorhees, 2005). Their purpose is to support research within the information retrieval community by providing the infrastructure necessary for large-scale evaluation of text retrieval methodologies.

The Enterprise Track in TREC (Craswell & de Vries & Soboroff, 2005) uses the Expert Finding task to help classify how knowledgeable a particular writer/creator is about certain topics of interest to the user. This same idea may be used to classify not how knowledgeable a writer is but instead how influential a writer is. In this context, influence is a measure of how much conversation and discourse is generated as a result of annotations/e-mail posted by a particular contributor. We consider a writer to be of importance if the messages/annotations they create in turn creates several responses. We intend to look at the automatic ranking of user/commentator importance in the context of SportsAnno, as discussed later in future work.

The Blog Track is a new track introduced in 2006 as a response to the growth of the Blogosphere as a whole. The aim of this track is to discover opinions and comments about a particular topic in response to a query. As of writing, the results from this year's participation were still to be made public. Weblogs can be seen a more direct link between the author and

[4] http://imarkup.com/download/plugin/server_plugin.asp
[5] http://annozilla.mozdev.org

his/her readership. As such, the comments made are of as great an importance as the original postings. There have been to date however, very few studies into the comments of weblogs themselves. These comments can provide valuable additional meta-data and help in the linking of users within the blogosphere. Many blogging sites do now have a "Trackback" feature that shows readers which other blogs have directly referenced their postings[6]. This feature enables blog conversations to be continued in different places but remain linked. Mishne and Glance (Mishne & Glance, 2006) have shown that if readers' comments are taken into account, a new relevance ranking can be achieved with "top-ranked results which are more discussion-oriented and attract more feedback from users". Again this is something we should look at more closely and perhaps incorporate into a ranking of most influential commentators within the SportsAnno project.

2.3 Systems

Systems for the summarisation and browsing of sports video do exist (Lui & Zhang, 2005). None of these systems, however, present a written source of complementary information with the summarized video. Indeed much of the work in this field is on the continued automatic detection of highlights, players and events of interest within the video.

Outside of the sports domain, a far greater amount of work has been done on the creation of complete video browsing and annotation systems. Work on the browsing of automatically annotated movies (Lehane & O'Connor & Smeaton & Lee, 2006) and segmented news video (Smeaton *et al.*, 2004) shows that sports could benefit from a similar approach. These systems allow the user a readily available access point into the annotated and segmented video, provide the final and crucial step in annotation and analysis cycle.

3. SportsAnno

The SportsAnno system was designed to give its users a comprehensive summary of all the action from the World Cup 2006 held in Germany. The aims of the system were:

- To allow users to become knowledgeable and form opinions about a sports event which they may not have seen live and be able to back up these opinions with visible evidence.
- To promote discussion about the sporting events and allow for the introduction of additional knowledge through this discussion.

Throughout the summer, all televised games were recorded and automatically marked up using event detection algorithms. At the same time, several newspaper websites were automatically scrapped to obtain the corresponding reports for each of the games. The aim of the system is to give users the opportunity to voice their own opinions about all the events in the competition, with all the evidence before them. The World Cup was chosen for its huge appeal and as stated above, since sports can be a very polarising, it was thought that this type of material would produce the most discussion. Another large advantage of using the World Cup is the enormous number of written reports that accompany each match, leading to many different viewpoints even within the official media. The reports were chosen to give a cross-section of this opinion.

[6] http://wordpress.org; http://www.sixapart.com/typepad

All sports reports in the media are in theory supposed to be objective in approach but this is never truly the case. Every sports report has an angle and the author, through their use of language, always portrays a certain bias. Nowhere is this bias more evident than in the field of sports. Sport has always been a highly contentious topic of conversation with each user having his/her own opinions about the events that take place. SportsAnno is designed as a system that allows users to comment both the events themselves and also the opinions of other users of the system.

It has been found that the loss of context whilst annotating can cause the focus of conversation to change (Brush *et al.*, 2002). Indeed, in-context commentary can allow for comments of a more specific and directed nature, as opposed to more general commenting on documents as a whole. In context means that comments are secured in place and neighboring the phrases/quotation to which they refer instead of being placed in a separate area. It was for this reason that we chose to allow in-context comments within SportsAnno.

In order to facilitate this interaction between the users, several technologies were used in the system. In the following section each of the components of the system is presented, along with the system architecture.

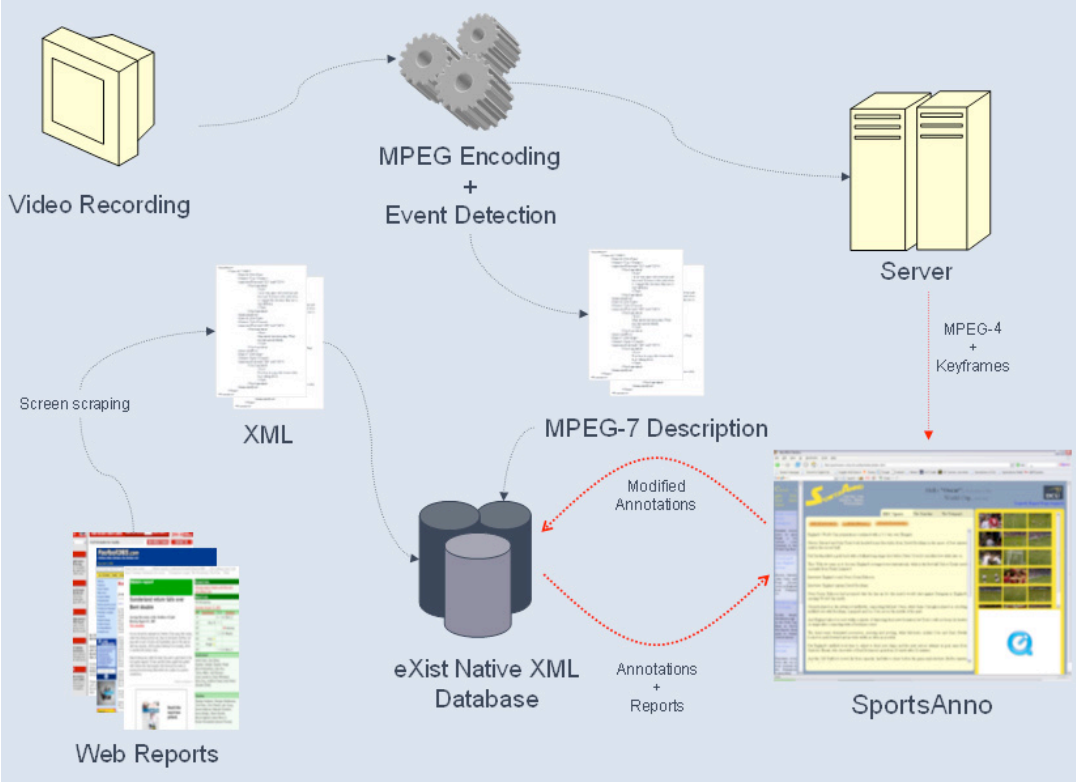

Figure 1. The SportsAnno Architecture

3.1 Architecture

Since SportsAnno brings together information from different media source and of different types, there is quite an extensive pipeline through which information must be sent before being presented to the user. Figure 1. illustrates this pipeline. Information comes from two main sources before being gathered into a single match record: video recorded from television and web reports taken from newspaper websites.

The initial video recording was made in MPEG-1 and later converted to MPEG-4, a step necessary to allow for streaming video playback through the Darwin Streaming Server. Once recorded, the video was processed so as to remove all none-game video. In this way the video begins just before the initial whistle is blown and ends just after the final whistle. Half-time is also removed so as to create a video which contains match footage only. In this way, no false events can be detected due to commercials or other footage. Each video was then approximately 90 minutes in duration, deviation being due to extra-time and penalty shoot-outs.

3.1.1 Video Segmentation

Once the video had been edited and cut to consist of just match footage, we ran a shot boundary detection algorithm. The shot boundaries detection algorithm used was the ***Cut_detect*** algorithm proposed by researchers within the CDVP (O'Toole C. & Smeaton A.F. & Murphy N. & Marlow S., 1999). ***Cut_detect*** is a shot-cut detection algorithm for MPEG-1 video files. The approach is based on the quantification of frame-to-frame dissimilarity, which is implemented via the generation of metrics relating to both histograms and statistical moments, for the colour components of each video image. Based on these descriptors, the algorithm invokes a threefold thresholding mechanism to quantify the significance of dissimilarity between frames, towards the detection of abrupt shot cuts in the video.

Since football video contains many hard cuts, the number of shots detected is very large while their duration can be very short. Each detected shot is assigned a confidence value based on how likely it is an event has occurred within the shot. Once events have been detected within the video, the shot boundaries are recalculated so as to provide shots that are of a useable length to a user. The minimum length of a shot was chosen to be 15 seconds. Initial shots are then amalgamated into new shot boundaries where all shots within the 15 second limit are concatenated to form a new shot as shown in Figure 2. If however the bounds of a shot containing an event overrun the 15 seconds limit, the amalgamated shot is increased so as to include all of the event shots. Key-frame extraction is also performed with key-frames chosen as the middle frame of a shot.

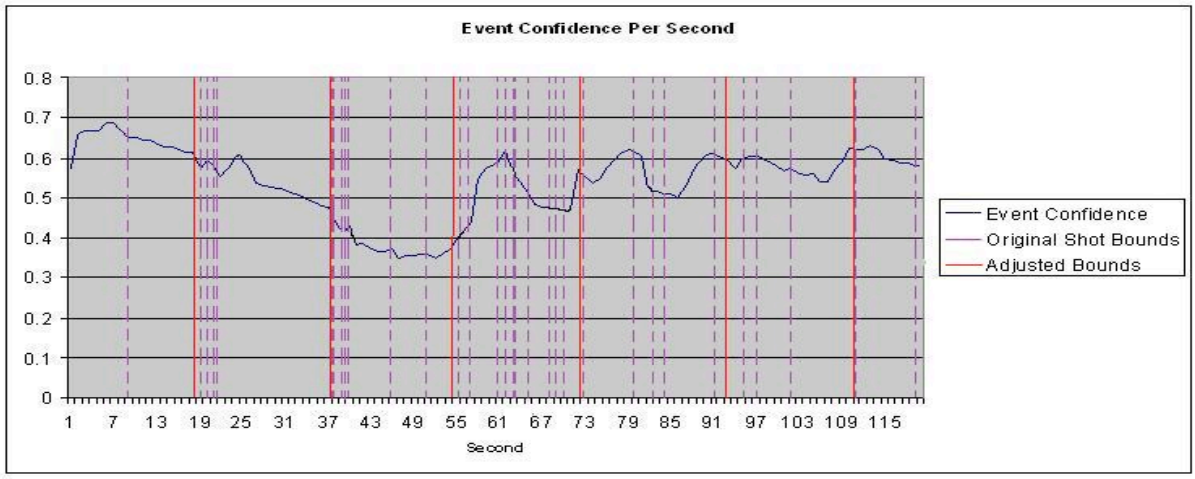

Figure 2. Adjusted Shot Bounds

Events were considered to occur when the event confidence value rises above a defined threshold and continue until this threshold is crossed again. This is shown in Figure 3. A

description of the manner in which events were detected is beyond the scope of this paper. All event detection was based on the work of Sadlier and O'Connor (Sadlier & O'Connor, 2004). The detection approach used was multi-modal and relied on both audio and visual information streams to determine confidence levels. 6 Support Vector Machine classifiers where used which detected the presence player close-ups, crowd shots, scoreboard changes, increased audio activity, playing field boundaries and increased visual activity.

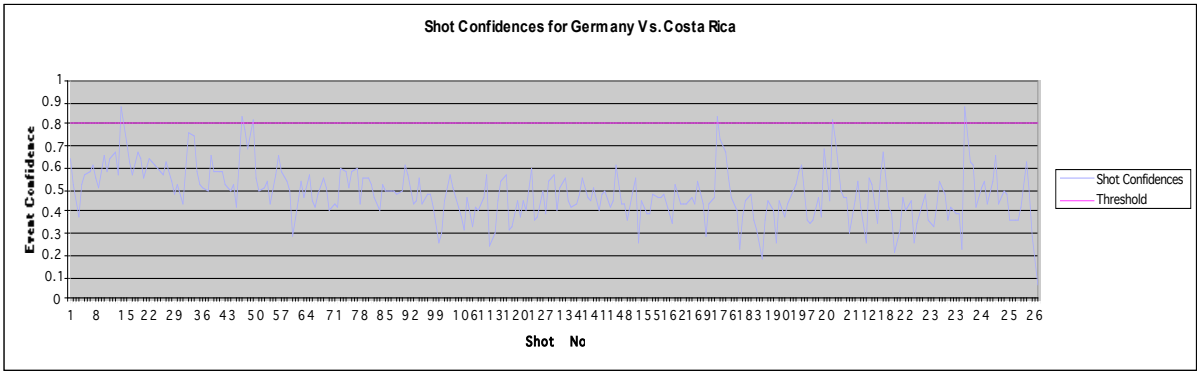

Figure 3. Event Detection using Threshold

After finalising the shot boundaries, an MPEG-7 file was created which contains all the shot information including duration, start point and confidence of an event occurring during each shot.

3.1.2 XML Document Storage

The second source of information required for each game is the match reports. Using a web parser, these reports were retrieved from the BBC Sport, Sky Sports and Guardian Unlimited websites. They were then transformed and stored as XML files. These three sites were chosen to give a cross-section of opinion. While the BBC and Guardian are less biased and brash in their coverage, Sky Sports was deliberately chosen as a site that would evoke more discussion due to its strong opinions.

Annotations were stored in separate files from the original report so as not to alter the original document. This was done so as to easily identify the insertion point of comments regardless of the number of comments already made. Annotations were stored as an XML document using a structure that was created specifically for this purpose. No standard method was followed but instead an XML structure was devised so as to maintain the thread architecture of the annotations. Each annotation has with in its record the name of the author, time of creation, quoted text and its content. It is possible to comment on any specific phrase within the original reports while replies to post are considered to be to the entire post as opposed to any specific part there of.

In order to organize the different files required for each game, a master file was created which links all the other files concerning a particular match together. This is the cross-reference file that contains the names of the reports, the annotation file and the MPEG-7 file of the match.

SportsAnno was built using an XML backbone so as to enable easy integration of existing standards whilst also providing easy extensibility. All data files required by the system are

stored within an eXist XML database[7], a freely available open-source project. The eXist database provide all the required functionality of a database for the storage and query of XML documents.

3.2 The User Interface

SportsAnno is a closed system requiring users to first register before being allowed access it. When first accessing the system, users are presented with a list of all games available to view. Attached to each game is the number of comments already made, the number of new comments since last logging in and a short description of the match, as shown in Figure 4. This description was chosen to be the subtitle from the BBC website and proved to be an adequate guide to the match.

Figure 4. The Navigation Frame for Match Selection

Once a game has been chosen, the user is presented with the full browsing interface allowing him/her to browse the reports and to see comments left by other users. This is shown in Figure 5. On the left of the screen the list of games is also available for easy navigation between matches. There are two major points of focus within the interface corresponding to the two complementary information sources: a collection of key-frames (representing the major events within the video) and the reports panel.

To the right are the key-frames, represent each of the shots within the video that have been marked as containing interesting events. Clicking on any shot will start playback from the beginning of the relevant shot. Each shot also has a small caption showing the time at which the shot starts. This is done so as to provide readers with an obvious correlation between anything written within the report and the video itself. Discrepancies arise between the time point within the video and the actual time displayed on the in-game clock. This is most often due to extra time played out at the end of each half or injuries during play. For this reason, a tilda is added to times after 45 minutes (half-time) to indicate approximation. Since times stated within the reports are never to the second but rather at a minute level, this slight inaccuracy was seen as no great inconvenience to the user.

Playback is shown through the Quicktime plug-in at the bottom right of the screen. It is possible to watch the entire match by clicking play on the player. Using the key-frames however will begin playback at the event chosen. The Darwin Streaming Server serves up the video in MPEG-4 format.

[7] http://exist.sourceforge.net/

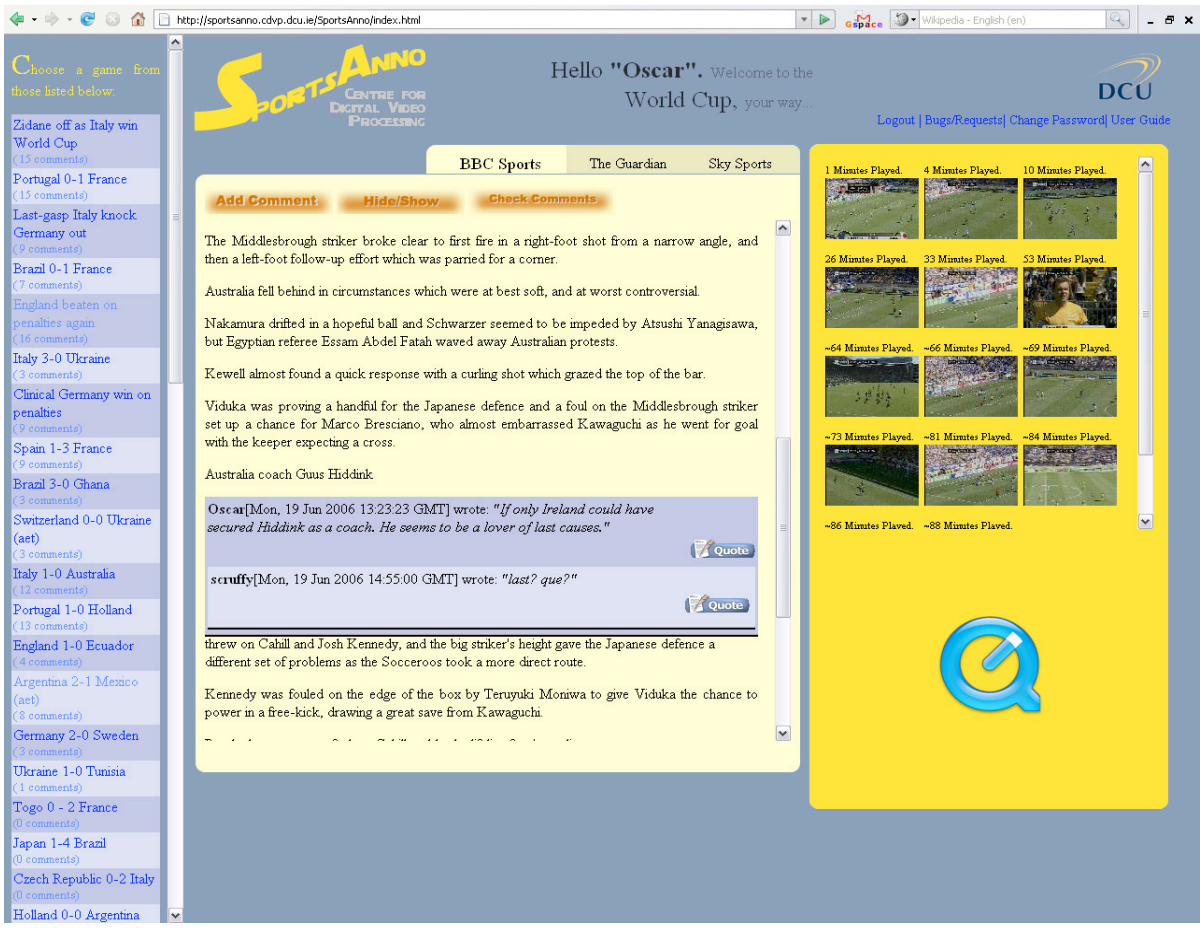

Figure 5. The SportsAnno Interface

Perhaps the most important element of the UI is the reports pane. Placed centrally within the display, it is here that users both read and annotate the newspaper reports. A tabbed panel is available, collecting all the reports into one place. Here the reports are shown with all comments made by users placed in context within the report. It is possible to hide these comments by clicking the button at the top of the reports pane so as to read the report more easily. By default however, all comments are shown.

Commenting and the threads of conversation these comments promote are the focus of SportsAnno. It was therefore of great importance to make the commenting facility as easy and intuitive as possible. In order to place a comment within a report, a user has simply to highlight a phrase within the report and click on the “Add Comment” button at the top of the reports pane. Commenting was restricted to phrases within a single paragraph (or a whole paragraph) so as to encourage discussion of specific points within the report. To reply to any comment posts, a user may click on the “Quote” button at the bottom of each post. This creates a thread attached to the comment. The alternating background colour for each post is used to signify the depth of the comment. If two replies are posted to the same parent, the same colour background is used. Also, a thick black ridge is used at the bottom of each comment depth as a visual aid to depth.

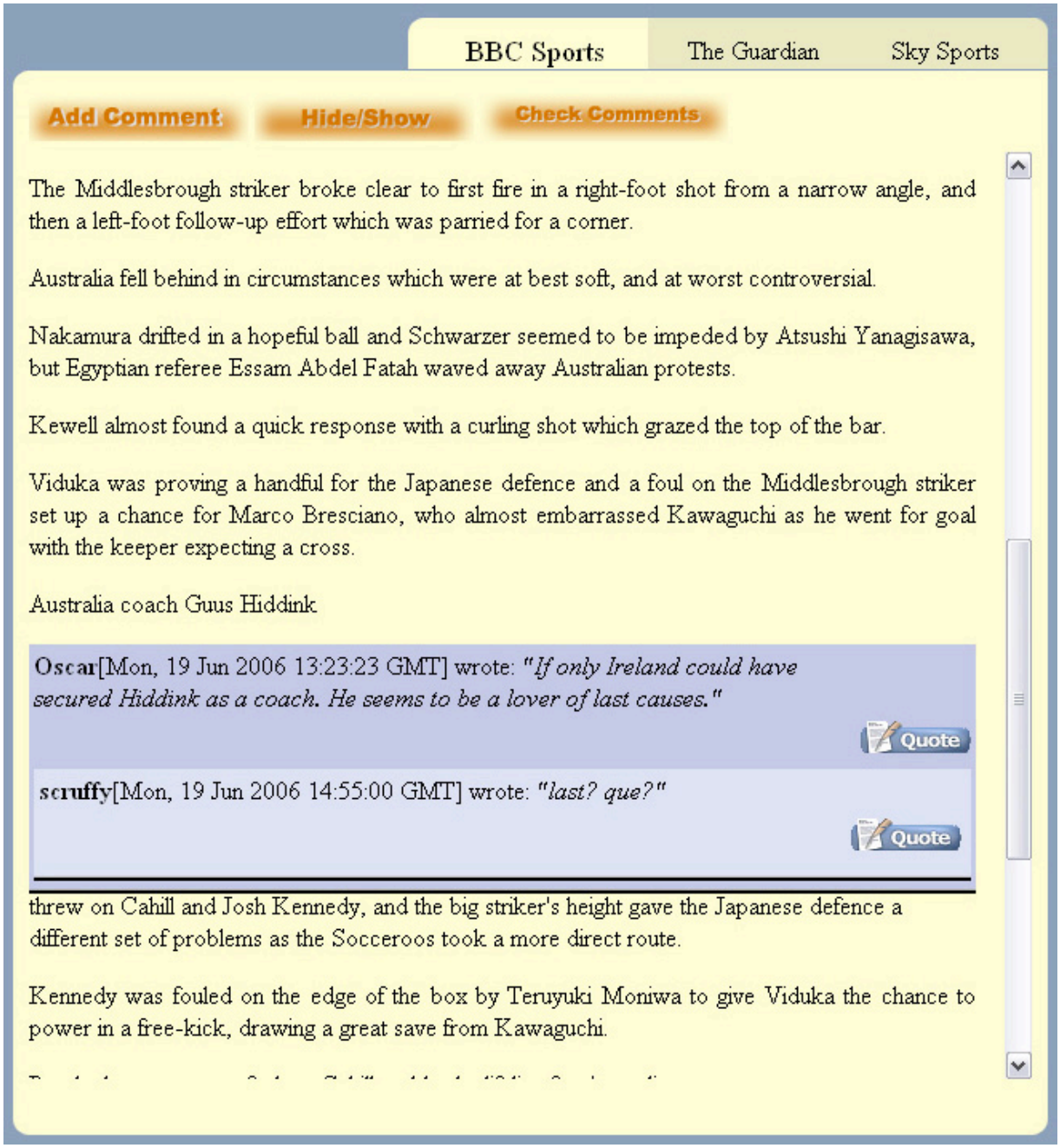

Figure 6. The Reports Panel with in-context annotations

4. Usage Study

70 people registered with the system, the dates of registration varying greatly from before the competition started to within the last week of the World Cup. All games were made available to all users and so even those who registered late could browse and comment on any match including those played before registering. 25 of the registered users where researchers within the group who have had had experience of annotation systems in the past. A further 12 came from associated research institutes who would again have had experience with annotation. The rest of the user community was made up from friends of registered users. Since the experiment was closed in nature, the user base was deliberately selected to be of people either directly known by the authors or know by a direct colleugue of the authors.

54 matches were recorded creating almost 83 hours of video data over the duration of the competition. This was accompanied by 162 newspaper reports. Not all matches were available for recording due to scheduling conflicts on television and one game (Serbia – Ivory Coast) was lost due to a recording error. The remaining matches were all fully indexed and event detection completed.

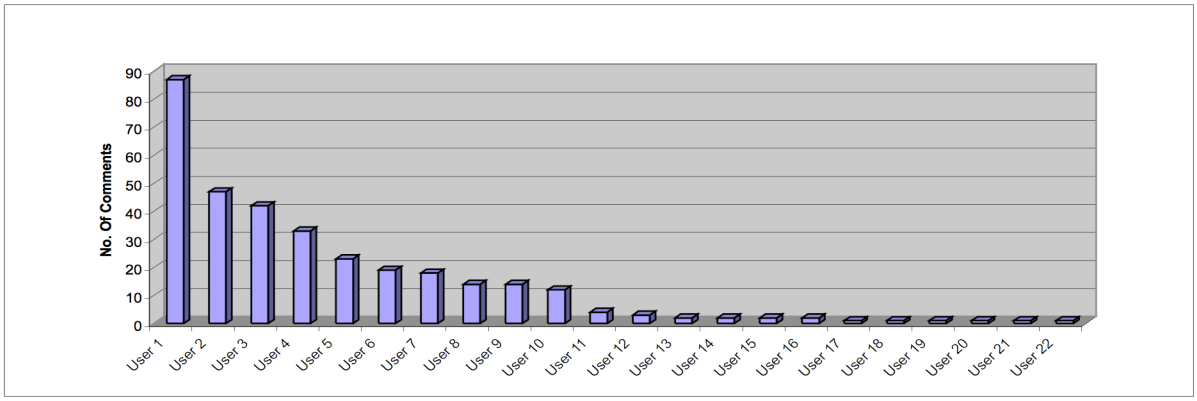

Figure 7. Comments per Active User

From the 70 people who registered, 24 made no further visit to the system. Of the remaining 46, 24 were active browsers viewing the comments left by others but not contributing themselves. 22 users made comments and took part in discussions about particular events within each match. The number of comments made by users varied greatly as seen in Figure 7. These comments are made up of both replies to previous posts and original postings. No differentiation is considered here. Figure 8. shows how many of these users commented on each match during the competition. The 22 users were not only those users who had had past experience of annotation systems.

It is clear that the first England game against Paraguay was particularly well commented. This is not a surprising result as the hype surrounding the England squad within the media of both the UK and Ireland generated lots of talking points. Only 6 of the 54 recorded games received no comments. Again, these games involved teams that would have little following within the registered audience, the only surprise perhaps being the Brazil – Japan fixture.

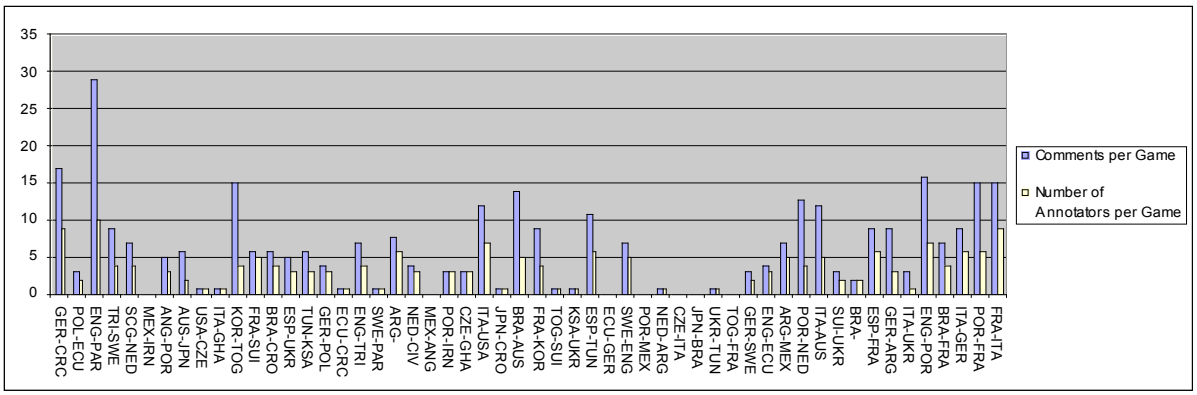

Figure 8. Comments and Annotators For Each Match

The ratio of commentators to comments shows that commenting is a useful way in which to generate discussion within a group. Within games with more than one user's comments, it can

be seen that it is not just new comments which are added but instead replies to the comments already left. Figure 9. shows the number of replies per game, broken down into original comments (i.e. comments which are not in reply to another comment) and replies. It can be seen that where original comments are attributed to more than one user, the number of replies versus original comments is high.

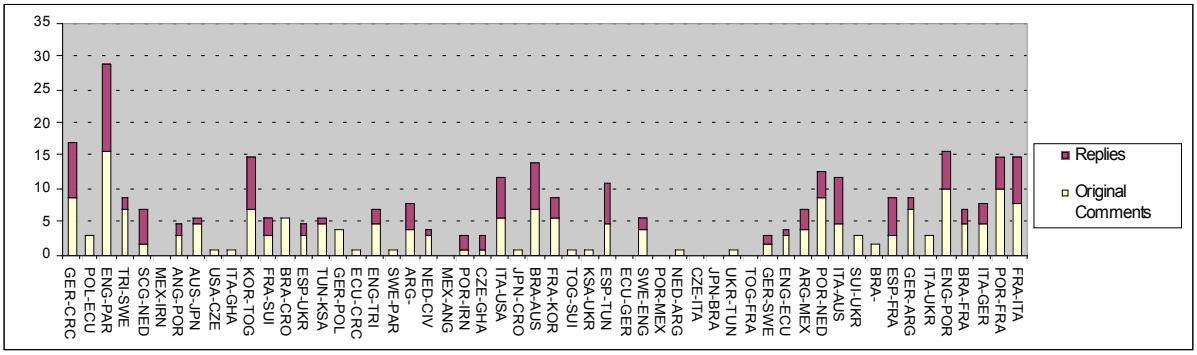

Figure 9. Comments per Match Divided Into Replies and Original Comments

One of the possible reasons for users not creating more replies to comments was the lack of a notification system which could notify users when a comment they had made was replied to. In this way, a user's attention would have been more readily drawn to the specific reply.

The time between first posting the match to the website and the last comments on a game being made was also recorded. Due to the type of data being presented, it is not altogether surprising that the number of comments made on a match fell dramatically 3 days after its first posting. Some games proved exceptional, mainly those involving teams that stayed in the competition for longer. Users did post comments on earlier performances involving teams such as Germany (the hosts) and France (a surprise to stay in the competition for so long) but in general, comments were of a more immediate and transient nature.

	Comments per Game	Original Posts	Replies	Comments After 3 days	Commentators Per Game	Thread Depth
Avg.	6.1	3.7593	2.35185	0.72222	3.25926	2.06
Std Dev.	5.84603	3.26754	2.9084	0.14973	2.49668	1.41976
Max.	29	16	13	6	10	6
Total	330	203	127	39		

Table 1. Comments and Conversation Threads

Table 1. shows the number of comments made and who was making these comments. These figures show that using a p-value of 0.05, the mean number of comments and commentators are significantly different from 0. Also, splitting the population into two disjoint samples reveals not statistically significant differnence in their mean values, again using a p-value of 0.05. The values are instead symptomatic of the group as a whole.

The maximum figures for comments in the first three columns all correspond to the England – Paraguay match. This was England's first game of the competition and so generated a lot of discussion. We can see that while there are more original comments than replies, there is on

average at least 2 replies to each post. The deviation is due to the existence of both highly commented games and those which received no real discussion. As mentioned before, replies are more prevalent for games where more than one person has created original comments.

The number of days after which the game was commented on was affected most strongly by the advent of weekends (during which very few comments were made) and the amount of time between the recording of the matches and when they were made available on the SportsAnno website. This time varied from same day to 2 days after the recording date. It is also thought that lack of a notification system prevented discussion from having an average lifespan of greater than 3 days, as mentioned earlier.

The presentation of video and associated reports seems to be a viable and useful manner in which to study the interaction of users. Since the reports are designed to contain all the events of interest to the user, in effect becoming a summary of the game, there is regularly a direct correlation between sentences within the reports and shots in the video. The presence of the video evidence is useful in providing support for subjective arguments made by a user. It also means that a user can bring him/her up to speed very quickly with discussions being had by other users without necessarily having seen the match live. Reviewing the video allows them to see exactly what comments related to the game itself mean. Without the video evidence, many users would not be able to comment with any sort of certainty on events within the game.

5. Lessons Learned

The analysis run on usage of SportsAnno throughout the World Cup 2006 were very useful in helping to establish the features of SportsAnno which were of greatest benefit to the user community. It also highlighted some features which would have helped in allowing that community to become both more active and those features which can add to the user's experience.

The lack of a notification system to alert a user as to when a reply to a post has been received seemed to be a particular weakness of the system. The use of such a feature would have aided in allowing conversations to stay active for a longer period of time. As stated above, the average lifespan of a match in terms of comments was 3 days, 90% of all conversation happening within this time. A notification system which alerts users upon login of both replies to their own messages and new discussions is something which will be investigated in future.

The possibility for information overload resulting from large numbers of comments being posted is something which is of concern with the present system. Since there is no way in which to select which annotations to view, or filter out those which are of little interest, the ability to keep conversations active for longer poses a problem. A method for prioritising user comments needs to be introduced, weighting the importance of user's comments on both a user specific and system wide level. The creation of friends lists which automatically flag the existence of new comments by friends/buddies of a particular user would help in reducing the number of irrelevant or uninteresting comments browsed in a short space of time. More interestingly, the system ranking of user importance is to be created. This is a list of top annotators. This list however, should not just be a list of users who have contributed the

greatest number of comments. More useful would be a list of users whose comments create the largest amount of discussion.

Direct linking of written reports and video as well as direct annotation of the video are features which we would like to include in future versions of the system. This would allow for easier navigation of the video and cut down on the locating of corresponding video events to points of interesting within the report. Linking of separate matches is also something which the authors would like to research. The opportunity to create comments which provide direct comparison between events within different games is a feature which we feel would create more discussion within the user community. It would also provide a means for serendipitous discovery of new information through user created annotation links.

SportsAnno has shown that the use of segmented sports video to augment conversation centred on associated written sports reports is of great benefit to the user. Through browsing and viewing of the video, users are able to become up-to-date with any major events. Through the viewing of others' comments, any user was also able to discover additional information which was not present in the original report. Examples of this are below:

- "… Joe Cole had done something similar not that long before this against Sweden, this was just that little bit better." *(Maxi Rodriguez's wonder goal against Mexico)*
- "This is actually the first shoot out that Argentina have lost in the World Cup, they have won their previous three." *(After Argentina lost on penalties against Germany)*

Whilst the second of these comments provides new and interesting information not previously included within the report, the first can only be assessed due to the fact that video of both goals is easily accessible.

6. Conclusion

In this paper we have presented SportsAnno, a system designed to promote user discussion about World Cup 2006 soccer. Through the use of both written and visual content, users are able to hold on-going discussions about the matches they view. This discussion has the potential to add new information to the original data and in such a way, grow the information source for each subsequent user. We see this as a valuable resource which could be adapted into such fields as medical research (where doctors, specialists and students alike would be able to comment on medical procedures), education (in areas such as Drama where there is a direct relation between recorded performances and the script) and sports in general. Results from the usage experiments carried out lead us to believe that it is also a viable approach for creating discussion about events which may not have originally been seen by all users. Instead, anyone with an interest in the material is able to contribute at any time. Analysis of the comments and annotations created by users may help in the creation of semantic tags for the video data. With the addition of filtration and ranking methods, it is envisage that large scale semantic assess to an increasing corpus of augmented content would be possible. The way in which SportsAnno brings together multiple modalities allows for this in an integrated fashion.

Acknowledgements

This work was supported by Science Foundation Ireland, under grant number 03/IN.3/I361. The authors would also like to thank Dr. David Sadlier for his contributions to the system architecture and background work.

References

Cabasson R. & Divakaran A. (2002) Automatic Extraction of Soccer Video Highlights Using a Conbination of Motion and Audio Features, In: *Symp. Electronic Imaging: Science and Technology: Storage for Retrieval for Media Databases,* Vol. 5021, January 2002, pp. 272 – 276.

Wang J. & Xu W. & Chang E.S. & Wan K. & Tian Q. (2005) Automatic Generation of Personalized Music Sports Video, In: *Proceedings of ACM International Conference on Multimedia, Singapore,* November 2005, (pp. 735 – 744).

Xie L. & Chang S. F. & Divakaran A. & Sun H. (2002) Analysis and Presentation of Soccer Video with Hidden Markov Models In: *Proceedings of IEEE. International Conference on Acoustics, Speech and Signal Processing,* (ICASSP) 2002.

Zhang D. & Ellis D. (2001) Detecting Sound Events In Basketball Video Archive *Technical Report, Dept. Of Electronic Engineering, Columbia University,* 2001.

Denman H. & Rea N. & Kokaram A. (2003) Content-based analysis for video from snooker broadcasts, *Journal of Computer Vision and Image Understanding, Special Issue on Video Retrieval and Summarization*, Vol. 92, pp. 141—306, November 2003.

Sadlier D. & O'Connor N. & Marlow S. & Murphy N. (2003) A Combined Audio-Visual Contribution to Event Detection in Field Sports Broadcast Video. Case Study: Gealic Football, In: *Proceedings of IEEE Symp. on Signal Processing and Information Technology, ISSPIT 2003,* Darmstadt, Germany.

Assfalg J. & Bertini M. & Colombo C. & Bimbo A. & Nunziati W. (2003) Semantic Annotation of Soccer Videos: Automatic Highlights Identification *Computer Vision and Image Understanding,* Volume 92, pp. 285 – 305, November 2003.

Bush V. (1945) *As We May Think* The Atlantic Monthly, 176(1), 101 – 108.

Bottoni P. et al. (2004) MADCOW: a Multimedia Digital Annotation System. In: Proceeding of the Working Conference on Advanced Visual Interfaces, ACM Press, (pp. 55 – 62).

Marshall C.C. (1998) Toward an Ecology of Hypertext Annotation. In: *Proceeding of UK Conference on Hypertext,* 1998 (pp 40 – 49)

Schilit B. & Golovchinsky G. & Price M. (1998) Beyond Paper: Supporting Active Reading with Free Form Digital Annotations. In: *Proceedings of ACM CHI'98* (Los Angeles, CA), ACM Press, (pp. 249 – 256)

Marshall C.C. & Bernheim Brush A.J. (2004) Exploring the Relationship Between Personal and Public Annotations. In: *Proceeding of The 2004 Joint ACM/IEEE Conference on Digital Libraries,* JCDL 2004, (pp. 349 – 357)

Brin S. & Page L. & Motwani R. & Winograd T. (1998) The PageRank Citation Ranking: Bringing Order to the Web. *Stanford Digital Library Technologies Project.* citeseer.ist.psu.edu/article/page98pagerank.html

Kleinberg J.M. (1999) Authoritative Sources in a Hyperlinked Environment, *Journal of the ACM*, 46(5), September 1999, pp. 604 –632.

Chin A. & Chignell M. (2006) A Social Hypertext Model for Finding Community in Blogs, In: *Proceedings of the seventeenth conference on Hypertext and hypermedia*, Hypertext 2006, (pp. 11 – 22).

Voorhees E.M. (2005) Overview of TREC 2005 *http://trec.nist.gov/pubs/trec14/papers/OVERVIEW14.pdf.*

Craswell N. & de Vries A.P. & Soboroff I. (2005) Overview of the TREC-2005 Enterprise Track *http://trec.nist.gov/pubs/trec14/papers/ENTERPRISE.OVERVIEW.pdf.*

Lui H.Y. & Zhang H. (2005) A Sports Video Browsing and Retrieval System Based On Multi-modal Analysis: SportsBR In: *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, Guangzhou, August 2005 (pp. 5077 – 5081).

Mishne G. & Glance N. (2006) Leave a Reply: An Analysis of Weblog Comments. In: *3rd Annual Workshop on the Weblogging Ecosystem, at WWW2006*.

Lehane B. & O'Connor N. & Smeaton A.F. & Lee H. (2006) A System For Event-Based Film Browsing, *3rd International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, TIDSE 2006, *Lecture Notes in Computer Science (LNCS)*, Vol. 4326, pp.334 – 345.

Smeaton A.F. & Gurrin C. & Lee H. & Mc Donald K. & Murphy N. & O'Connor N. & O'Sullivan D. & Smyth B. & Wilson D. (2004) The Físchlár-News-Stories System: Personalised Access to an Archive of TV News, In: *Proceedings RIAO 2004 - Coupling Approaches, Coupling Media and Coupling Languages for Information Retrieval*, Avignon France.

Bernheim Brush A.J. et al. (2002) Supporting Interaction Outside of Class: Anchored Discussion Vs. Discussion Boards, CSCL 2002 (pp. 425 – 434).

O'Toole C. & Smeaton A.F. & Murphy N. & Marlow S. (1999) Evaluation of Automatic Shot Boundary Detection on a Large Video Test Suite, *CIR'99 - The Challenge of Image Retrieval: 2nd UK Conference on Image Retrieval*, Newcastle, UK, February 1999.

Sadlier D. & O'Connor N. (2004) Event Detection in Field-Sports Video using Audio-Visual Features and a Support Vector Machine. In: *IEEE Transactions on Circuits and Systems for Video Technology*, 15(10), pp. 1225 – 1231.