TV News Story Segmentation, Personalisation and Recommendation

Alan F. Smeaton, Hyowon Lee, Noel E. O'Connor,
Seán Marlow and Noel Murphy

Centre for Digital Video Processing
Dublin City University, Glasnevin, Dublin 9, IRELAND
Alan.Smeaton@computing.dcu.ie

Abstract

Large volumes of information in video format are being created and made available from a number of application areas, including movies, broadcast TV, CCTV, education video materials, and so on. As this information is increasingly in digital format, this creates the opportunity and then the demand for content-based access to such material. One particular kind of video information that we are interested in is broadcast TV news and in this paper we report on our work on developing content-based access to broadcast TV news. Our work is carried out within the context of the Físchlár system, developed to allow content access to large volumes of digital video information. We report our work on Físchlár-News which provides text search based on closed caption information as well as our on-going work on segmenting TV News programmes and providing personalised intelligent access to TV news stories, on fixed as well as mobile platforms.

Introduction

The growth in volume of multimedia information, the ease with which it can be produced and distributed and the range of applications which are now using multimedia information is creating a demand for content-based access to this information. Digital video information is becoming commonplace through the development of DVD movies, broadcast digital TV and TiVo boxes, CCTV, and video on personal computers for gaming and educational applications. Besides the growth in volume of multimedia content, we can also observe an increasing and complex range of user needs scenarios where we require content-based access to such information. Our workplace and leisure activities are becoming such that when we search for information, and increasingly this is becoming searching for multimedia information, our search task can have a range of complexities. The old paradigm of ad hoc searching where a user's need is expressed as a query which is matched against a document or other collection of artifacts and a ranked list produced, and which is the dominant approach taken by web search engines, is only one type of search which actually misses many of the subtleties of helping users satisfy their information needs.

For example, sometimes when we search we want to find lots of information because our need is for a broad exhaustive search, while at other times we want a quick answer to a straightforward question and just a small amount of information is required. Sometimes we seek an answer to a specific question, sometimes our needs are more general. Sometimes we know what information we want, other times we're not sure, but we'll recognise it when we see it. Sometimes we're on desktop or home devices, sometimes we're mobile. Other dimensions to the search task exist, but all search tasks generally have several things in common, including the fact that we always have some "baggage" representing our prior knowledge, which we want to add to, as part of the task we are performing. By this is meant that we always know something about the topic of our search, either a lot, or very little, and the reason we are searching is to increase the amount we know about that topic. This feature of always having some "prior knowledge" is generally very complex to model and to capture in real-world applications and the approach of considering this prior knowledge as part of the search process and then judiciously providing "extra" information on top of what the user already knows, is difficult to achieve. We shall return to this point later.

In this paper we report on our work on developing an information retrieval system for one type of multimedia information (digital video), of one type of genre (broadcast TV news) and targeted at one type of user information need, namely an irregular viewer of TV news who wishes to be kept up-to-date with most developing news stories, but is not interested in all news.

The remainder of this paper is structured as follows. In the next section we briefly present a summary of our work in developing several flavours of the Físchlár system targeted at different users. We follow that with a description of how we feel access to broadcast TV news should be facilitated. In section 4 we describe 5 different video and audio analysis techniques which we have implemented and how the outputs of these can be combined together to yield an automatic segmentation of broadcast TV news into news stories. That is followed by a description of how segmented news stories can be used as the basis for a personalisation and recommender system which can ultimately be used to organise content for presentation to our user. A concluding section finishes the paper.

The Físchlár Digital Video Libraries

The Centre for Digital Video Processing at Dublin City University has developed the Físchlár system which allows indexing, browsing, searching and playback of digital video content on fixed and mobile platforms, to a real community of users on the Dublin campus. To date we have developed 4 working versions of Físchlár as follows:

Físchlár-TV

The Físchlár-TV system allows users to request recording of broadcast TV programmes and supports subsequent browsing/playback on a conventional web browser. Perceived as a web-based video recorder, the system has been operational for nearly 3 years at the time of writing and more than 1,800 registered users have been using web-based system to record and watch the TV programmes they like from both the University campus residences and from computer labs. Of all our Físchlár systems, Físchlár-TV has the largest number of real users who have been using the system for the longest time, and the application has become an important way for us to monitor usage and get feedback on different approaches to video browsing and searching. At any point in time there are about 100 recorded TV programmes available, with newly requested programmes being analysed and added to the collection and becoming available for browsing and playback on a daily basis. The oldest programmes automatically deleted to make room for newer material and to keep the current archive fresh.

Físchlár-News

The Físchlár-News system automatically records the 9 o'clock main evening news programme every day from Irish national channel RTÉ1 and thus has only TV news programmes in its collection. Currently about 2-years of recorded daily RTÉ1 news are available from Físchlár-News and this is made available to University staff and students, and is also conveniently accessible from any computer lab, library or residence from within the campus. The system is mainly used for study, teaching and research among journalism students and staff on campus.

Físchlár-Nursing

The Físchlár-Nursing system allows access to a closed set of 12 educational video programmes on nursing, and is used by the University's nursing students. In addition to the usual Físchlár keyframe browsing feature, a Table of Contents is provided for each video, allowing students to view a more semantic organisation of contents, an overview and in this way we provide access to different sections of the Físchlár-Nursing library.

Físchlár-TREC2002

The Físchlár-TREC2002 system allows access to a closed set of videos comprising advertising, educational, industrial and amateur films produced between the 1930's and the 1970's totaling 40 hours. This collection was defined by the National Institute of Standard and Technology (NIST), the organising body for the TREC (Text Retrieval Conference) video track (Smeaton and Over, 2002). The system allows users to perform shot retrieval and to formulate queries based on a set of simple semantic features of the video contents. This includes automatic determination of video shots as being indoor or outdoor, landscape or cityscape, containing faces, people or overlaid text, and where the audio is speech, music or a dialogue. This tailored version of the Físchlár system was developed for our participation in the interactive search task in the annual activity at the TREC Video Track in 2002 (http://www-nlpir.nist.gov/projects/t2002v/t2002v.html). More details on Físchlár-TREC2002 can be found in Browne *et al.* (2002).

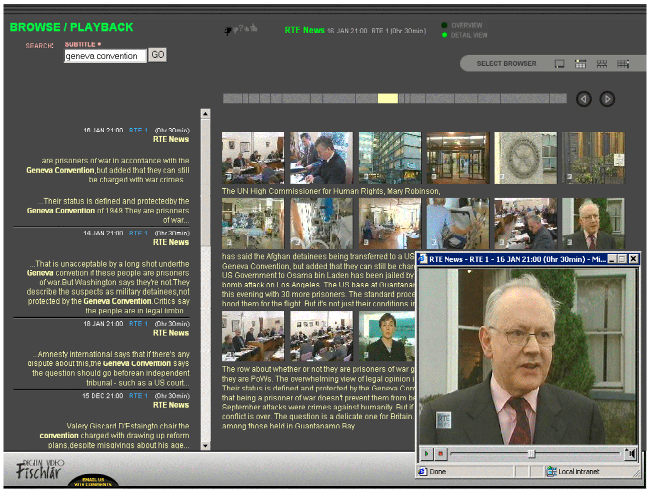
All the Físchlár Systems

All of the four Físchlár systems mentioned above are MPEG-7 compliant and all are designed and implemented with an internal XML architecture and all interfaces are generated from XML documents using XSL stylesheets. Físchlár-TV and Físchlár-News have both desktop and mobile handheld (Compaq iPAQ) interfaces while the others are desktop only. The development and deployment of these systems helps both to focus and direct our research work in that it gives us real goals for our basic research while providing an application for our more theoretical exploration into multimedia information management.

All Físchlár systems provide video browsing and playback and all provide some granularity of multimedia search. In Físchlár-TV the unit of retrieval is the TV program and users typically locate a TV program, perhaps browse its contents using the keyframe browser of automatic keyframe summary, and then play the entire program. In Físchlár-Nursing, users also want to play an either an entire program or large sections of a program and they do this by browsing the text descriptions and using these as jump-in points into the video. In Físchlár-TREC2002, the unit of retrieval is the shot, almost the smallest unit of video and the system was designed to retrieve shots as part of the TREC video retrieval task. Finally, in Físchlár-News we support searching based on a user's text query being matched against the captured closed captions from the TV broadcast. In the next section we shall explore whether this kind of ad hoc searching is adequate for searching TV news.

Interaction with Físchlár-News

The Físchlár-News system, as briefly introduced above, has a desktop interface (see Figure 1) in which presents a list of available daily news programmes on the left side of the screen, and the user selects one of them to browse and playback its content in more detail, on the right side of the screen. For interactive searching, a user can also type in text query terms and click the GO button to view an alternative list of video programmes that match the search terms. Having a large screen area for displaying multiple stages of interaction thus allows a text-based querying and a playback screen on top of it and has been proved useful; The reverse-chronologically ordered list of available news programmes and the keyframe-based browsers makes this easy and simple for the desktop user to interact with it.



16 JAN 21:00 RTE 1 (0hr 30min)
RTE News

...are prisoners of war.In accordance with the **Geneva Convention**,but added that they can still be charged with war crimes...

...Their status is defined and protectedby the **Geneva Convention** of 1949 They are prisoners of war...

14 JAN 21:00 RTE 1 (0hr 30min)
RTE News

...That is unacceptable by a long shot underthe **Geneva** convetion if these people are prisoners of war.But Washington says they're not.They describe the suspects as military detainees,not protected by the **Geneva Convention**.Critics say the people are in legal limbo...

18 JAN 21:00 RTE 1 (0hr 30min)
RTE News

...Amnesty International says that if there's any dispute about this,the **Geneva Convention** says the question should go beforean independent tribunal - such as a US court...

15 DEC 21:00 RTE 1 (0hr 30min)
RTE News

Valery Giscard D'Estaingto chair the **convention** charged with drawing up reform plans,despite misgivings about his age...

The UN High Commissioner for Human Rights, Mary Robinson,

has said the Afghan detainees being transferred to a US
Geneva Convention, but added that they can still be char
US Government to Osama bin Laden has been jailed by
bomb attack on Los Angeles. The US base at Guantanar
this evening with 30 more prisoners. The standard proce
hood them for the flight. But it's not just their conditions ir

The row about whether or not they are prisoners of war g
they are PoWs. The overwhelming view of legal opinion i
Their status is defined and protected by the Geneva Con
that being a prisoner of war doesn't prevent them from be
September attacks were crimes against humanity. But if
conflict is over. The question is a delicate one for Britain.
among those held in Guantanamo Bay

RTE News - RTE 1 - 16 JAN 21:00 (0hr 30min) - Mi...

Figure 1: Desktop interface to the Físchlár-News system

However, while this version of Físchlár-News is of good use to the casual ad hoc searcher who is seeking a particular clip of news form the archive, it does not satisfy many other user scenarios. In particular, it does not satisfy the most common scenario for users accessing TV news, namely an irregular viewer, who doesn't always manage to tune into the news when it is transmitted, who occasionally misses the news for some days, who has preferences for certain types of news stories and not for others, and for whom time is important, and not to be wasted. What this kind of user has as a significant, and measurable, characteristic, is that he/she has already seen a certain amount of news and this prior news viewed will have been logged. Thus, almost uniquely in an information retrieval scenario, we have a user for whom we can measure their context in terms of what they already know, and thus we can try to compute the "delta" on that, i.e. the part of the news that they don't already know.

Most news programs consist of a series of independent news stories, some of which are follow-ons and updates on previous news bulletins. In an albeit subjective analysis of news over a 28-day period during 2001, reported in (Lee and Smeaton, 2002) we found that a 30-minute broadcast typically has between 8 and 8.5 different news stories, of which 2.9 are updates on stories from the previous day, 1.5 are updates on stories from some day prior to that, and 3.9 of the stories are new, reflecting breaking news.

When the Físchlár-News user has been away for some time, or has missed the news, as happens, then what that user wants is an update on old or the news stories that are of interest and ideally the most recent broadcast on those stories. This is especially true of the user is accessing Físchlár-News from a mobile device where the scope for user-system interaction is much reduced because of smaller screen size, no keyboard, and a context whereby a user is on the move and unlikely to be able to give full concentration for long periods to the mobile device.

To address the various needs of the Físchlár-News user, either on a desktop or mobile device, we have developed a model for accessing Físchlár-News which pre-packages most of the content by automatically segmenting broadcast news into discrete stories, creating information links between these stories where appropriate and then personalising the set of stories not yet seen by the user on either their mobile or desktop platforms. This implies that the new functionalities (story segmentation, linking and personalisation) are required from the design end rather than from implementation end, in order to provide appropriate access to the system's news archive. The original outline and rationale of the mobile access to Físchlár-News system was presented at a workshop at SIGIR2002 in Tampere (Lee & Smeaton, 2002) and in this paper we give further details and an update on progress in the news story segmentation task..

Features Available for News Story Segmentation

The tasks of TV news story segmentation, linking and finally news story recommendation are ultimately based around detecting features from the TV video and measuring the similarity between segments of TV news based on these features. In our work we have a broad base of expertise in video and audio analysis and we can automatically process the video to extract features as described in the following sub-sections.

Spoken dialog indexing

In Físchlár-News we use closed captions automatically taken from the TV signal to provide a (sometimes dirty) transcript of what is spoken during the broadcasts and we are exploring the use of lexical chains for story bound segmentation (Stokes *et al*., 2002).

The dialogue for an entire TV News broadcast will contain topic shifts as the newsreader moves from one story onto the next and the subject matter changes. This will make the overall text less cohesive than a text which is all about the one topic. In our work we are using lexical chains in order to mark the boundaries in topic which indicate a news story boundary.

Lexical chaining is a linguistic technique that uses an auxiliary resource (in our case WordNet) to cluster words into sets of semantically related concepts e.g. {*motorbike, car, lorry, vehicle*}. In this work we determine lexical chains from the news transcript and then use the boundaries of these chains to indicate possible news story boundaries. In an evaluation of this approach on CNN transcripts (Stokes *et al.,* 2002) we have observed mixed results with lexical chains showing promising results and we are presently evaluating this on real closed caption materials, which will have more transcription errors and thus will make the task more difficult.

Speech vs. music discrimination

We have developed and tested a technique to automatically *detect speech vs. music* from video, something we used as part of our TREC2002 search task (Jarina *et al*. 2002). This is useful in determining where the news introduction (standard video introduction with music background) ends, and also when the news programme itself, is ended.

The algorithm for speech vs. music discrimination works on data directly taken from MPEG encoded bitstream thus avoiding the computationally difficult decoding-encoding process. The method is based on thresholding of features derived from the modulation envelope of the frequency-limited audio signal. The discriminator has been tested on more than 2 hours of audio data, which contain clean and noisy speech from several speakers and a variety of music content.

The inherent characteristics of different sounds such as speech and music are completely different; speech has a strong temporal and spectral variation, whereas musical sound has a smoother and longer-term variation as well as a high tonal content. As a result, the 4 characteristics found to be most relevant for discrimination between these two include the detection of peak occurrence and duration, rhythmic content and level of harmonic content.

A compressed audio signal in the form of an MPEG-1 Layer-II bitstream was analysed, the information of interest being the scalefactors. In summary, the scalefactors describe the energy of the audio signal grouped into 15 evenly spaced subband frequencies. Subband 1 consists of scalefactors describing the instantaneous energy of the audio signal from 0 to 700 Hz, each subsequent subband encompassing the next 700 Hz. By using this information as opposed to the decoded audio signal, a highly computationally and memory efficient approach can be taken to detection.

The bitstream was processed into shots and a sliding windowing method with an overlap of 33% was used to extract sections of the audio bitstream scalefactors from subband 2 to 7. For each audio frame, the duration of the largest peak and rate of peaks were computed, followed by a simple thresholding method. This process has been discussed in greater detail in (Jarina *et al*. 2001).

To detect the occurrence of rhythmic pulses, a long-term autocorrelation function was applied to the band-passed signal. Should a peak occur in the function, the magnitude of this peak would reflect the level of rhythm in the signal.

The final characteristic to be used in determining the nature of the audio is the harmonicity or level of harmonic content in the audio signal. Finally, the four sets of results were used to determine the certainty of the audio signal being speech or music.

Advertisement detection

We have a very reliable *advertisement detection* system (Sadlier *et al*. 2001) which is useful since the main evening news which we work on has a short commercial break somewhere in the middle of the broadcast. This is important to identify since we do not wish to include adverts in the news story content, and advertisement breaks indicate the end of a news story, and the beginning of the first one after the break.

Advertisement breaks may be isolated from actual programme material by the flags that most terrestrial and some satellite television companies signal during their broadcast: a series of ‘black’ video frames simultaneously accompanied by a decrease in the audio signal occurring before and after each individual advertisement spot. Our digital video system, Físchlár, captures television broadcasts and encodes the programmes according to the MPEG-1 Layer-II digital video standard. It was proposed that a ‘black’ and ‘silent’ classification of the individual frames of a captured television signal might be made as follows.

For analysis of the video: an examination of the DC Discrete Cosine Transform (DC-DCT) coefficients of a frame, which represent the weight of its zero-frequency content, with a view to establishing whether or not the frame is inherently dark enough to be labeled ‘black’. The DC-DCT coefficients of each Y-block of each frame were stripped from the video bitstream of the MPEG file and an average luminance DC-DCT coefficient was then calculated for each video frame of the sequence. The overall mean value of the average frame coefficients for the clip was then determined and the ‘black-frame’ threshold was then expressed as a percentage of this mean. Finally, each frame’s average DC-DCT coefficient value was compared to the threshold and if equal/less than, then the frame was labeled ‘black’.

For analysis of the audio, an inspection of the weight of the scalefactors of the signal’s (low) frequency subbands is done with a view to establishing whether or not a video frame’s accompanying audio signal power is minimal enough for it to be labeled ‘silent’. The scalefactors corresponding to the first 10 subbands of the encoded audio signal were stripped from the bitstream and an audio level for each video frame was determined by averaging the scalefactors corresponding to its associated audio signal. The overall mean value of the video frame audio levels for the entire clip was then determined and the 'silent-frame' threshold was then expressed as a percentage of this mean. Each video frame's representative audio level was compared to the threshold and if equal/less than, then the frame was labeled 'silent'.

Our ad-break detection method was tested against a corpus of 10 short television programme clips from 4 different channels [labeled (a), (b), (c) & (d)] coded in MPEG-1 Layer-II format. The recordings were meticulously chosen such that they exhibited significant content diversity with at least one complete ad-break somewhere in the middle. Results of ad-break detection are as Table 1.

Clip	Precision	Recall
(a) Chat Show	100	100
(a) News Broadcast	100	68.6
(a) Music Show	100	98.8
(b) News Broadcast	100	100
(b) Soap Opera	100	100
(b) Sports Show	100	89.3
(c) Youth Magazine	100	100
(c) Game Show	100	100
(d) Cartoon (1) (2)	100 100	100 100
(d) Comedy Quiz	100	86.9

Table 1: Results of the advertisement experiment by different clips

Anchorperson detection

We have developed a technique for determining *when the anchorperson is on-screen* at any point in time. This is based on clustering shots using colour, successively eliminating outlier shots and converging upon a set of anchorperson shots (O'Connor *et al.* 2001). We can then calculate image (keyframe) similarity between anchorperson shots and determine when the background image projected behind the anchorperson, changes, indicating a likely story change.

Anchorperson detection starts with a shot clustering algorithm based on the temporally constrained clustering approach (Rui *et al.,* 1999). The main difference between our approach and that of Rui et al is the choice of features used for each shot. We use a single feature extracted from each keyframe, rather than the multiple feature approach of Rui et al. We have found that this approach has worked well for our preliminary investigations but recognise that it will need to be extended in the future. The algorithm groups shots based on the similarity of their colour composition and the temporal distance between the shots. Each shot is represented by the colour signature vector of the keyframe associated with the shot. In this way, each shot is represented as a point in a multi-dimensional feature space. Possible distance measures in this space are the Euclidean distance measure and the co-sine distance measure, which is the one we have used.

The result of the shot clustering algorithm is a number of groups of shots that satisfies the spatial and temporal constraints. Experimentally results show us that anchorperson shots are usually (and often) in a single group, if it's a single anchorperson or in three groups, if they are two anchorpersons An anchorperson is a news presenter. Usually the anchorperson shots have a static background and the person who reads the news. This feature ensures high similarity values between subsequent anchorperson shots throughout the news programme. To detect the anchorperson shots we have set the temporal attraction function to a very long value, to avoid the disruption of the anchorperson shots. This together with a high threshold ensures that the anchorperson shots will be in the same group. We have a variety of cases to find the anchorperson shots:

1. the most common, where one anchorperson exists
2. two anchorpersons taking every second story
3. two anchorpersons, one main and one for sports stories or weather.

In order to decide which group is the anchorperson group, it is necessary to search through all the groups and attempt to discount them on the rules used. Any groups left after all the criteria have been applied are then considered anchorperson shots.

The anchorperson identification algorithm is a greedy algorithm that loops through all the groups of shots and selects the groups that satisfy the following conditions:

1. the range of shots of the group (the difference between the first and the last - in terms of shots) is higher than a predefined constant value (i.e. 10), because the anchorperson shots are widely spread across the news programme.
2. the group similarity mean is higher than a constant value. In the case of Cosine distance measure, for example 0.90. The group similarity mean is calculated as the mean of the shot similarities between adjacent pairs of shots in the same group as described in section 3.2.
3. the anchorperson shot length (in terms of frames) should be higher than a constant (e.g. 150). That means a possible anchorperson shot should be longer than 6 seconds (150 frames / 25 frames per second).
4. the number of shots within a group that satisfies the three above conditions.

Speaker segmentation and matching

The final technique we are developing as input into our story bound segmentation is the development of a technique for *speaker segmentation and matching* which will identify when different speakers are used, and will also help to identify when the anchorperson is speaking (Jarina *et al.* 2002).

At this point in time we are working on detecting and indexing the occurrence of speech as opposed to silence or music, and segmenting it in terms of different speakers (male/female, male/male, female/female). This is based on pitch tracking as well as other unique speaker qualities. From this the audio signal will be marked to show when a speaker starts and what the qualities of the voice is. Following that, we can cluster the various speakers in order to provide a tracking system which can detect the re-occurrence of speakers for what duration and in the order of which they speak. This signature of the dialogue during a news broadcast should provide input which will be of use to the story bound segmentation process.

Broadcasts into Stories, Making Users Happy

All of the five techniques mentioned in the previous section are working, individually, and our present efforts are to combine these together into an operational system for story segmentation, story linking and story recommendation, which has a mobile as well as a desktop interface. The combination of analysis is being explored through the use of Support Vector Machines (Burges, 1998) and preliminary work is showing promise of being able to effectively and efficiently combine these diverse analyses. The linking together of related news stories will be based on computing a similarity between potentially linked news stories, with some obvious temporal constraints build in. the similarity will be based on text dialogue, with proper name identification playing an important role.

Once news broadcasts have been segmented, and linkages between them identified, the final component is to arrange these on a per-user basis using a personalisation and recommender system. In Físchlár-TV we have been using the ClixSmart engine (Smyth and Cotter, 2000) to recommend TV programs for recording or for playback from those recorded and available in the Físchlár-TV library. We shall continue to use ClixSmart, in this case to recommend individual news stories. Once this is achieved then the fact that the user is on a mobile or a fixed line platform becomes just a matter of choosing many (fixed line) or fewer (mobile) stories to be recommended for the user.

Acknowledgements:

The support of the Informatics Directorate of Enterprise Ireland is gratefully acknowledged. We have benefited considerably from collaboration with Prof. Barry Smyth in UCD on the personalisation and recommender work reported in this paper.

References

Browne, P., Czirjek, C., Gurrin, C., Jarina, R., Lee, H., Marlow, S., Mc Donald, K., Murphy, N., O'Connor, N., Smeaton, A.F. and Ye, J. 2002. Dublin City University Video Track Experiments for TREC 2002. *Proceedings of the Text Retrieval Conference (TREC-2002),* Gaithersburg, Maryland, 19-22 November 2002.

Burges, C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121-167.

Jarina, R., Murphy, N., O'Connor, N. and Marlow, S. 2001. Speech-Music Discrimination from MPEG-1 Bitstream. In: *V.V. Kluev, N.E. Mastorakis (Ed.), Advances in Signal Processing, Robotics and Communications*, WSES Press, pp. 174-178.

Jarina, R., O'Connor, N., Marlow, S. and Murphy, N. 2002. Rhythm Detection for Speech-Music Discrimination in MPEG Compressed Domain. *Proceedings of 14th International Conference on Digital Signal Processing (DSP 2002),* Santorini, Greece, 1-3 July 2002.

Lee, H. and Smeaton, A.F. 2002. Searching the Físchlár-NEWS Archive on a Mobile Device. *Proceedings of the 25th International ACM Conference on Research and Development in Information Retrieval (SIGIR 2002), Workshop on Mobile Personal Information Retrieval*, Tampere, Finland, 11-15 August 2002.

O'Connor, N., Czirjek, C., Deasy, S., Marlow, S., Murphy, N. and Smeaton. A.F. 2001. News Story Segmentation in the Físchlár Video Indexing System. *Proceedings of the International Conference on Image Processing (ICIP 2001),* Thessaloniki, Greece, 7-10 October 2001.

Rui, Y., Huang, T.S. and Mehrotra, S. 1999. Constructing a table of contents for videos. *ACM Journal of Multimedia Systems*, vol. 7, pp. 359–368.

Sadlier, D., Marlow, S., O'Connor, N. and Murphy, N. 2001. Automatic TV Advertisement Detection from MPEG Bitstream. *Proceedings of the International Conference on Enterprise Information Systems, Workshop on Pattern Recognition in Information Systems (WPRIS 2001)*, Setubal, Portugal, 7-10 July 2001.

Smeaton, A.F. and Over, P. 2002. The TREC-2002 Video Track Report. *Proceedings of the Text Retrieval Conference (TREC-2002)*, Gaithersburg, Maryland, 19-22 November 2002.

Smyth, B. and Cotter, P. 2000. A personalized television listings service. *Communications of the* ACM, 43(8), 107-111.

Stokes, N., Carthy, J. and Smeaton A.F. 2002. Segmenting Broadcast News Streams using Lexical Chains. *STAIRS 2002 - STarting Artificial Intelligence Researchers Symposium*, Lyon, France, 22-23 July 2002.