

Searching the Físchlár-NEWS Archive on a Mobile Device

Hyowon Lee and Alan F. Smeaton
Centre for Digital Video Processing
Dublin City University
Glasnevin, Dublin 9, Ireland.
{hlee, asmeaton} @computing.dcu.ie

Abstract

The Físchlár-NEWS system provides web-based access to an archive of digitally recorded TV News broadcasts over several months, and has been operational for over a year. Users can browse keyframes, search teletext and have streamed video playback of segments of news broadcasts to their desktops. This paper reports on the development of mFíschlár-NEWS, a version of Físchlár-NEWS which operates on a mobile PDA over a wireless LAN connection. In the design and development of mFíschlár-NEWS we have realised that mobile access to a digital library of video materials is more than just the desktop system on a smaller screen, and the functionality and role that information retrieval techniques play in the mFíschlár-NEWS system are very different to what is present in the desktop system. The paper describes the design, interface, functionality and operational status of this mobile access to a video library.

1. The Físchlár NEWS Archive

The Físchlár-NEWS Archive is an online repository of daily evening TV news which is broadcast by the Irish National broadcast channel, RTE1. The system automatically records and processes news programmes along with their teletext (close-captioned subtitle of what is said during the programme) every evening and allows web-based access to the indexed programmes by both browsing and searching (querying).

Searching is an important access mechanism for any archive of considerable size, and this is also true in the Físchlár-NEWS Archive. All news programmes in the archive are accessible by browsing the reverse-chronologically ordered list of programmes, and thus the latest news programmes (such as yesterday evening's news) can be more easily viewed. However, immediate access to segments of the news which a user may want to see by allowing querying of the entire archive is a powerful alternative and complimentary access to browsing the archive. The Físchlár-NEWS Archive uses the teletext captured along with the video signal, for subsequent indexing and retrieval. Figure 1 shows a screen shot of searching through the Archive. A user types in query term(s) into the query box at the top-left side of the screen, and the result of the matching is displayed below, where matching news segments are grouped by



Figure 1: Físchlár NEWS Archive – the full desktop system interface

their entire news programmes and each of the retrieved segments within a news programme show their time of occurrence within that programme. When the user clicks on one of the matched segments, the programme will open on the right side of the screen in a content browser, with the matched point within the programme displayed. The browser shows keyframes and teletext together and the user can choose to further browse other parts of the programme or start playback at any point in the programme by simply clicking on a keyframe. The Físchlár-NEWS Archive is an adapted version of the Físchlár Digital Video System which records any TV broadcast programmes on users' requests and allows browsing/playback within the University's computer laboratories and campus residences from a desktop computer environment [O'Connor *et al.* 00].

The Físchlár system has been in operation on the University campus for almost 3 years and is accessed via a web browser on a desktop computer. At the time of writing we have over 1,700 registered users, of whom about half are "active" and regular users. Físchlár is used for research, for teaching and learning, and for entertainment. The Físchlár-NEWS system has been in operation for about a year and was developed specifically for students of journalism and media studies to provide them with easy access to an archive of broadcast TV news. Físchlár (which allows users to record any broadcast TV program) and Físchlár-NEWS (which records only the main evening news) are both very popular among the University community.

As an indicator of what it offers, the full Físchlár-NEWS Archive provides the following features for users:

1. Browsing of the list of daily news programmes in reverse-chronological order (by scrolling);
2. Text-based searching of all the news programmes by inputting word(s), with the resulting matched segments displayed by highlighting the matched word(s) with 3-4 lines of surrounding text providing the context of the matching;
3. Browsing individual news programmes with one of our various content browsers (keyframe-based browsers; for more detail see [Lee & Smeaton 02]);
4. Playback of news programmes by clicking on any of the keyframes in a content browser.

Our work in the development of Físchlár for a desktop environment has always been based on user issues and considerations – designing, building and deploying an operational Físchlár system whose interface functionality and layout has been designed with the user's goals, tasks and abilities in mind. This approach of leading the development of an interface to a video library system from the user's perspective, is continued in the development of a mobile handheld platform interface to Físchlár-NEWS. The most important set of issues which we will address in the development of Físchlár-NEWS are those of interface design – general interface concerns for mobile applications as covered in section 2 and interface design concerns for accessing our system in section 3. In section 4 we present the design and implementation of Físchlár-NEWS for a Compaq iPAQ, a PDA connected to the Físchlár database via a wireless LAN connection, and in section 5 we present our conclusions.

2. Interface Design Considerations for IR from Mobile Devices

User interface design for mobile devices has been discussed for many years now, with the consensus that a mobile interface should require a different interaction style from that of the GUI desktop interface, and that attempts to cram all the functionality of desktop system into a mobile device are a mistake [Thomas *et al.* 95] [Marcus *et al.* 98] [Kristoffersen & Ljungberg 99] [Longoria 01]. Though the current literature alerts to the fact that we do not have any established or known methodology on which to base an interface design for a mobile platform, rough design guidelines have been suggested based on experiences of individual researchers. These are:

- Determine what kind of information and what style of interaction a user will need/want to access via the mobile device, in order to have most benefit [Ruuska-Kalliokulju *et al.* 01] [Jordan *et al.* 01] [Longoria 01];
- Minimise user input – where applicable, provide simple user selections such as yes/no options, simple hyperlinking by tapping, etc. instead of asking the user to articulate query formulation or use visually demanding browsing that requires careful inspection of the screen;
- Filter out information so that only a small amount of the most important information can be quickly and readily accessed via the mobile device (e.g. use of automatic recommendation as provided in the Físchlár TV system [O’Sullivan *et al.* 02]);
- Proactively search and collect potentially useful pieces of information for a user and point these out, rather than trying to provide full coverage of all information via an elaborate searching/browsing interface;
- Use a layout that does not require a large space – converting where applicable, spatial information into a temporal format (as in RSVP [Bruijn & Spence 00]);

In terms of developing any system for a mobile device which is to support searching and information retrieval tasks, all these guidelines point to more pre-processing on the system’s side in order to determine which pieces of information a particular user will most likely to want to see. This encourages the development of systems which proactively recommend a particular piece of information (or pointers) to the user, and consequently demand less interaction on the user’s part. This aspect is even more important in the case of information retrieval from a video archive where browsing is such an important component of video access. What all this means is that in the development of search systems to be accessed from mobile platforms, the information retrieval functionality should be hidden as much from the user as is possible, and should form part of the information pre-processing. We shall see more of this in the next section.

3. Interface Design Concerns for Accessing the mFíschlár-NEWS Archive

In this paper we are concerned with developing a version of Físchlár-NEWS which is accessed from a mobile handheld platform. Our target user would use “mFíschlár-NEWS” from a mobile PDA such as a Compaq iPAQ, connected to the network with enough bandwidth to support streaming video. Before we delve into the way in which we have designed and built mFíschlár-NEWS, it is worth examining what are the characteristics of the TV news archive we are providing access to.

The RTE1 main evening news is a daily broadcast which summarises news developments over the previous 24 hours into a single 30-minute programme, with an advertisement break in the middle. It is typical of most TV news programmes in that it begins with a snapshot of the headlines followed by the signature tune and signature introduction and then a series of about 8 or 9 individual news stories before finishing up with a weather forecast. Each of the individual news stories is composed of a series of shots of an anchorperson with an illustrative graphic in the background (see the playback window in Figure 1 where the frame is taken from a story about a “Mother’s Appeal”), interspersed with outside footage. Within a single day’s news, some of the stories may be new, reflecting a new story which may unfold over time (first reporting of joyriders crashing into a police car) or may begin and end on that day (farmers protest at rising milk prices), some may be updates on a continuously developing story (Israeli siege of church in Bethlehem) or some may be updates on an old story from the past (killer sentenced to prison, reference to previous stories on murder, capture and trial).

As an indication of the structure of the RTE1 news, we took a period of 28 days during November 2001 and we analysed this to determine the characteristics of breaking news items during this period. Appendix A presents a graphical summary of how the stories in each broadcast are related to each other. Each column in that appendix represents a date, each row represents a unique news story and

shaded boxes across the row represent occurrences of that news story on that date. A series of shaded boxes across a row represents repeated occurrences of the same story across different days. There is of course an element of subjectivity in the way in which we have manually determined the relationship between stories and whether a news story on a given day is a new story or an update on an older one. Nevertheless we believe the characteristics we have identified are similar to those used in datasets from recent work on Topic Detection and Tracking [Allan *et al.* 98]

Our analysis of this period of news has identified 231 stories (about 8.25 per day) of which 115 are unique. On a given day an average of 2.89 of the stories presented on the RTÉ1 news are updates or developments on continuing stories from the previous day, 1.46 are updates on stories from somewhere prior to that, and 3.89 are new stories. The constantly appearing story across the 28 dates in Appendix A is about the Afghanistan conflict and some of the once-off stories are about things like the Government announcing a new health strategy and a US company cloning the first ever embryo.

The analysis of stories in the RTÉ1 news programme was important in helping us to design the functionality of a mobile platform access to this archive. The kind of work tasks which a mobile platform will naturally support are somewhat different to those tasks supported from a desktop environment. Because of screen size, absence of a keyboard and mouse, awkward data entry facilities, limited pointing ability and the fact that mobile users cannot be relied upon to have desk space around them, mobile user access to a video archive will generally support rather simple information seeking tasks. Sophisticated information seeking such as trying to find obscure TV News references to persons which may or may not even be in the archive, are best done from a desktop environment. Drawing up a simple usage scenario for a mobile user accessing a video archive may help [Carroll 00] in defining what a mobile device for accessing the Físchlár-NEWS Archive, is best suited for:

A user sitting in a bus (or waiting for a bus on a bus stop) brings out her PDA from the pocket to get recent updates of news stories since she last accessed the archive, either from her desk or PDA.

The above scenario, though rough and simple, implies a few important information retrieval techniques which are required as part of the system side processing:

- Video shot boundary detection and keyframe selection to allow keyframe browsing of a news story;
- Story identification and segmentation within a broadcast TV news programme;
- Linking of related stories in the news archive;
- Summarisation of a news story both visually (selection of keyframe) and as text (selection of teletext fragments);
- Identification of an individual user and tracking his/her access activity (through a user profile);
- Determination of which news stories on which date(s) would be most interesting or useful for a particular user at the moment of access;

While the kind of information retrieval techniques needed to support the above are sophisticated, on the other hand, the scenario implies quite a simple user interface, places little demand on the user, and has its main emphasis on good summarisation of each recommended story. From the interface perspective, it requires

- A listing of new and updated news stories in summarised format;
- An option for further, more detailed browsing of a story and for playback of that story;
- A temporal keyframe browser [Lee *et al.* 01];

As most work will have gone to the indexing-time processing of the video, and the resultant news listing will show only new or updated news stories, all a user should have to do is to tap on a title from one of the latest stories, and within a list of news items within a story, simply tap on an item to play or browse. This saves the user from elaborate and visually-demanding interaction from – which is mostly required on a desktop applications – and makes the interaction a more casual tapping action which does not require constant and uninterrupted attention, which is suitable for a mobile environment. This is a different philosophy from our desktop-based interface design for accessing Físchlár-NEWS as shown earlier in section 1, which allows full access to the news archive with a chronological listing of all news programs, keyword searching, and various keyframe browsers with elaborate interaction of the user's choice. The user accessing via desktop in an office will want not only recent updates to the TV news archive but may also want to study and research from within the news archive, and thus will benefit from such elaborate provision of features.

Figure 2 below illustrates the differences between the features provided on the desktop and on the mobile device for access to the Físchlár News Archive:

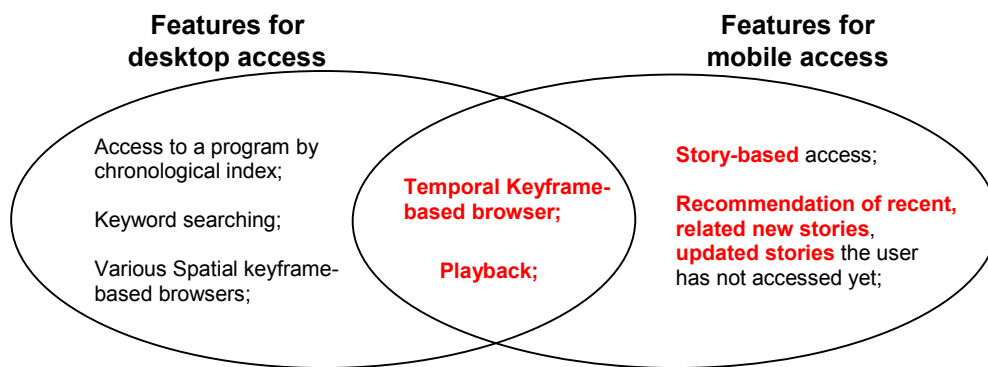


Figure 2: Features for desktop and mobile access to Físchlár News Archive

Note that Playback and a keyframe-based browser on the desktop environment have been determined as useful for a mobile environment, whereas the chronological index, keyword searching and various spatial browsers have been decided unsuitable for mobile environment.

4. Interface Design

As mentioned above, the user interface for system that has such functionality as mFíschlár-NEWS can be very simple, as most of the important filtering processes which involve information retrieval functionality have been gone through before the system is presented to the user. As a result, the presentation of each of the recommended news stories in a good summarised format becomes the main concern in the interface design. Although in the kinds of features provided in the interface are quite different from those of desktop interface, same overall look-and-feel was used to imply this being a part of Físchlár News Archive.

In designing the functionality of mFíschlár-NEWS we had considered something along the lines of that shown in Figure 3. Here a user has a choice between navigation through their own personalised listings of news stories from which they can browse and then play individual stories, and navigation through the generic daily listing of news stories, from which they can browse and then play stories. While this has a somewhat symmetric appeal, for reasons outlined earlier we decided to dispense with the generic access on the right hand side of Figure 3, and allow personalised access only.



Figure 3: Two possible ways of accessing the mFíschlár-NEWS

The mFíschlár-NEWS system is accessed from an iPAQ PDA running Windows CE and the Internet Explorer web browser. This means we can use the same underlying database as is used in the main Físchlár system which has cut down on development effort. The interface is illustrated in Figure 4 below:



Figure 4: Story list screen on iPAQ

The recommended stories are listed with each story composed of a single keyframe with a few lines of teletext summary. At the moment we choose the first keyframe from each story (normally consisting of the anchorperson with illustrative background image or graphic) and the first few lines of teletext from the story as a summary, but more elaborate summarisation may be investigated later. Presenting a thumbnail-size keyframe along with associated text has been known to be more effective for users compared to displaying only text, as studied in [Christel *et al.* 97]. When a user taps on the BROWSE link for each story, the content browser will be presented allowing the user browse the story's full set of keyframes quickly (see Figure 5 below). The browser allows previous/next keyframe flipping, as well as changing the two different levels of jumping: either shot-level keyframe jumping or story-level keyframe jumping. The keyframe browser has been designed to display one keyframe at a time temporally, rather than multiple keyframes spatially, thus making it more suitable for small screen for space-time trade-off [Bruijn & Spence 00]. This particular style of browser was developed as one of the many possible design options from the keyframe-based browser framework [Lee & Smeaton 02].

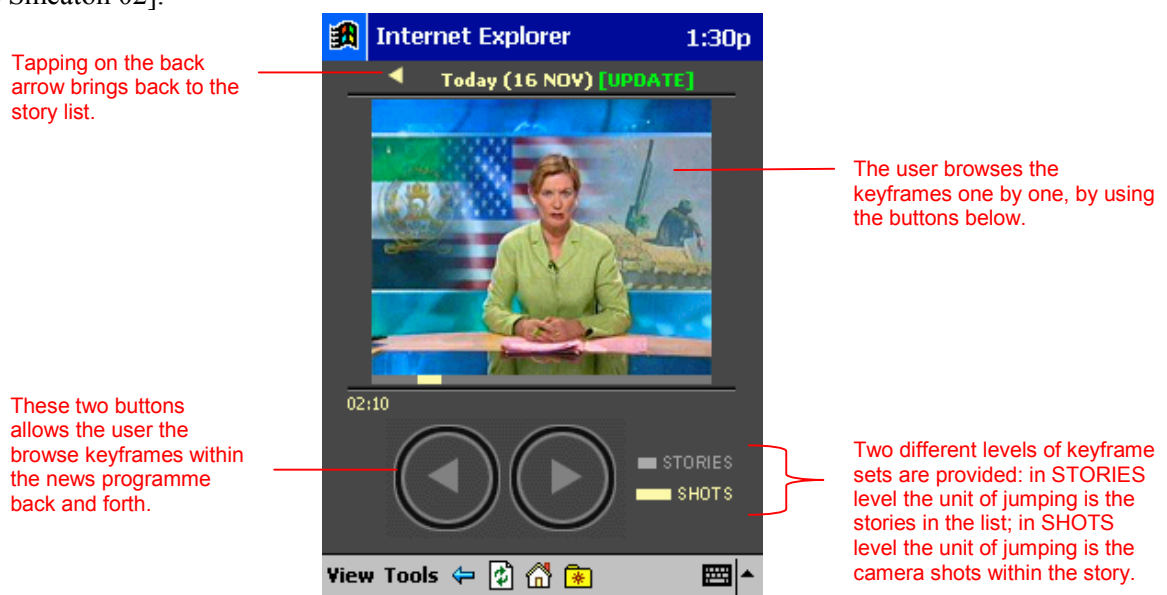


Figure 5: Browser screen on iPAQ

5. Conclusion and Status

One of the advantages for us in building the mFíschlár-NEWS system is that the underlying desktop system for broadcast TV news capture and streaming is already build and in place and we can re-use the capture, indexing and storage facilities from that. Físchlár has been in operation at the Dublin City University campus for 3 years and we have been operating Físchlár-News which captures TV news, for about 12 months. Our work on shot boundary determination and keyframe selection is already well-established and operational [Browne *et al.* 00]. As part of our work on access to TV news we are developing techniques for story bound segmentation via the visual aspects of news video [O'Connor *et al.* 01] and in collaborative work with others, via lexical chains derived from the teletext [Stokes *et al.* 02]. Work on automatically linking together related news stories will leverage our present work on linking in web searching [Gurrin & Smeaton 00] and our past work in constructing hypertext guided tours [Guinan & Smeaton 92] and will commence shortly.

Work on summarisation of an individual news story is more difficult since the duration of a news story is short anyway and news programmes are summaries of events in the first place. The Físchlár system uses the sophisticated ClixSmart personalisation engine from the Smart Media Institute at UCD [Smyth & Cotter 00] to recommend whole TV programmes for users. It is our intention to adapt this to allow this personalisation and recommender system to essentially rank segmented news stories for each user.

At the time of writing, the interface to the mFíschlár-NEWS system has been designed, built and tested. All of the backend components are already operational though we are still struggling to find a solution to allowing streaming of MPEG-1 video to an iPAQ, and we are presently putting the components together into an operational system.

More and more systems are starting to provide different means of access to their underlying information structures via different devices. Different devices require different user interfaces which should consider the particular environment of the device's usage. When designing such an interface for information retrieval on a mobile device, the constraints can be considerable compared to that on a desktop device, especially due to the currently available input/output mechanisms and the mobile environment which does not allow easy and efficient data input from users. The approach taken in our work reported here uses various intelligent information retrieval techniques to automatically extract the most relevant piece of information for the mobile user, thus lightening the potential burden of elaborate searching and browsing with the currently still awkward input/output mechanisms of mobile devices. We believe that this embedding of information retrieval functionality, rather than the facility of upfront searching, is a more appropriate use of information retrieval on mobile devices.

Acknowledgement: The support of the Informatics Research Initiative of Enterprise Ireland is gratefully acknowledged.

References

- Allan, J., Papka, R. and Lavrenko, V. (1998) **On-line new event detection and tracking**. *Proceedings of the 21st International ACM Conference on Research and Development in Information Retrieval (SIGIR 98)*, Melbourne, Australia.
- Browne, P., Smeaton, A., Murphy, N., O'Connor, N., Marlow, S. and Berrut, C. (2000) **Evaluating and combining digital video shot boundary detection algorithms**. *Proceedings of the Irish Machine Vision and Image Processing Conference (IMVIP 2000)*, Belfast, Northern Ireland.

Bruijn, O. and Spence, R. (2000) **Rapid Serial Visual Presentation: a space-time trade-off in information presentation.** *Proceedings of the Advanced Visual Interface (AVI 2000)*, Palermo, Italy.

Carroll, J. (2000) **Five reasons for scenario-based design.** *Interacting with Computers*, **13**, 43-60.

Christel, M., Winkler, D. and Taylor, C. (1997) **Improving access to a digital video library.** *Human-Computer Interaction: INTERACT97, 6th IFIP Conference on HCI*.

Guinan, C. and Smeaton, A. (1992) **Information retrieval from hypertext using dynamically planned guided tours.** *European Conference on Hypertext (ECHT 92)*, Association for Computing Machinery, New York, 122-130.

Gurrin, C. and Smeaton, A. (2000) **Dublin City University experiments in connectivity analysis for TREC-9.** *Proceedings of the 9th Text Retrieval Conference (TREC-9)*, Gaithersburg, MD.

Jones, M., Marsden, G., Mohd-Nasir, N. and Boone, K. (1999) **Improving Web interaction on small displays.** *Proceedings of 8th International WWW Conference*.

Jordan, P., Peacock, L., Chmielewski, D. and Jenson, S. (2001) **Disorganization and how to support it - reflections on the design of wireless information devices.** *3rd International Workshop on Human Computer Interaction with Mobile Devices (IHM-HCI 2001)*, Lille, France.

Kristoffersen, S. and Ljungberg, F. (1999) **"Making place" to make IT work: empirical explorations of HCI for mobile CSCW.** *Proceedings of the international ACM SIGGROUP Conference on Supporting Group Work*.

Lee, H., Smeaton, A., Murphy, N., O'Conner, N. and Marlow, S. (2001) **Físchlár on a PDA: Handheld User Interface Design to a Video Indexing, Browsing and Playback System.** *Proceedings of the International Conference on Universal Access in Human-Computer Interaction (UAHCI 2001)*, New Orleans, Louisiana.

Lee, H. and Smeaton, A. (2002) **Designing the User-Interface for the Físchlár Digital Video Library.** *Journal of Digital Information: Special Issue on Interactivity in Digital Libraries*, **2**(4)

Longoria, R. (2001) **Designing mobile applications: challenges, methodologies, and lessons learned.** *Proceedings of the Human-Computer Interaction International Conference (HCI International 2001)*.

Marcus, A., Ferrante, J., Kinnunen, T., Kuutti, K. and Sparre, E. (1998) **Baby faces: user-interface design for small displays.** *Proceedings of the International ACM Conference on Computer-Human Interaction (CHI '98)*.

O'Connor, N., Marlow, S., Murphy, N., Smeaton, A., Browne, P., Deasy, S., Lee, H. and Mc Donald, K. (2001) **Físchlár: An On-line System for Indexing and Browsing of Broadcast Television Content.** *International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*. Salt Lake City, UT.

O'Connor, N., Czirjek, C., Deasy, S., Marlow, S., Murphy, N. and Smeaton, A. (2001) **News story segmentation in the Físchlár video indexing system.** *Proceedings of the International Conference on Image Processing (ICIP 2001)*, Thessaloniki, Greece.

O'Sullivan, D., Wilson, D. and Smyth, B. (2002) **Improving case-based recommendation - a collaborative filtering approach.** *To appear in Proceedings of the 6th European Conference on Case-Based Reasoning*, Aberdeen, Scotland.

Ruuska-Kalliokulju, S., Schneider-Hufschmidt, M., Vaananen-Vainio-Mattila, K. and Von Niman, B. (2001) **Shaping the future of mobile devices - results of the CHI2000 Workshop on: "Future Mobile Device User Interfaces."** *3rd International Workshop on Human Computer Interaction with Mobile Devices*, IHM-HCI 2001, Lille, France.

Shneiderman, B. (1998) **Designing the user interface: strategies for effective human-computer interaction** (3rd ed). Addison Wesley Longman.

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. (2002) **Segmenting broadcast news streams using lexical chains.** *Starting Artificial Intelligence Researchers Symposium (STAIRS 2002)*, Lyon, France.

Thomas, P., Meech, J. and Macredie, R. (1995) **A framework for the development of information appliances,** *ACM Symposium on Applied Computing*.

Appendix A: Distribution of RTÉ1 News Stories during a 28-day period

01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27 28

