Físchlár on a PDA: Handheld User Interface Design to a Video Indexing, Browsing and Playback System

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

Centre for Digital Video Processing
Dublin City University, Glasnevin, Dublin 9, Ireland

The Físchlár digital video system is a web-based system for recording, analysis, browsing and playback of TV programmes which currently has about 350 users. Although the user interface to the system is designed for desktop PCs with a large screen and a mouse, we are developing versions to allow the use of mobile devices to access the system to record and browse the video content. In this paper, the design of a PDA user interface to video content browsing is considered. We use a design framework we have developed previously to be able to specify various video browsing interface styles thus making it possible to design for all potential users and their various environments. We can then apply this to the particulars of the PDA's small, touch-sensitive screen and the mobile environment where it will be used. The resultant video browsing interfaces have highly interactive interfaces yet are simple, which requires relatively less visual attention and focusing, and can be comfortably used in a mobile situation to browse the available video contents. To date we have developed and tested such interfaces on a Revo PDA, and are in the process of developing others.

1. Introduction

Físchlár is a web-based digital video system currently used by about 350 users within the campus environment in Dublin City University. The system allows the user to record broadcast TV programmes, and allows browsing and playback of the recorded programmes on a web browser. The user can easily browse 8 terrestrial TV channels' schedules for today and tomorrow, and by simply clicking on a programme sets the recording. The system then encodes the programme in MPEG-1 format when the broadcast time comes. The encoded programme is then subjected to automatic camera shot and scene boundary detection to extract representative keyframes, which are used for the user's browsing and playback interface. The interface allows the user to select one of the several browsing methods we have developed to see the keyframes and clicking on any of the keyframes will pop up a new window which starts streamed playback of the video from the clicked keyframe onwards. The system is capable of streaming to about 150 users concurrently. The system is a testbed for our technology development, wherein any implemented techniques such as various shot/scene boundary detection, integration with a programme recommender system (Smyth & Cotter, 2000), mobile application for video browsing and playback, and various user interface ideas are easily plugged in to the system and the outcomes are visibly demonstrated to our current user base. Users of the system are an important element of our work, as they provide new ideas from their own, real usage context. The Físchlár system is further described in O'Connor, et al.(2001).

In the recognition of the diversity of users' preferences and task contexts, we have developed a design framework for video browsing interfaces which allows us to come up with many different styles of browsing interface, and we have implemented 8 different browsers suitable for a desktop environment for the user to choose and use. As we are presently working on the use of mobile devices to access the Físchlár system's video contents, porting the system's browsing/playback feature to mobile devices has become an important concern for us. In this paper we apply the developed browsing interface design framework to a small, handheld PDA (Personal Digital Assistant) and demonstrate the resultant example browsing interfaces for the PDA. Section 2 briefly explains the rationale for developing such a design framework, and summarises the actual framework for keyframe-based browsing interfaces and how it can be used to design a specific browser. In section 3 we apply general interface design concerns for mobile devices to the design framework and demonstrate two example designs suitable for a PDA. Section 4 concludes with some future directions in video browsing interface design.

2. Design Framework for Video Browsing Interface

One of the main problems in designing a user interface for a novel system such as a digital video browser is the absence of usage information and lack of new ideas, making the conventional process of the initial design and subsequent iterative refinement through usage monitoring difficult. User interface design is a highly creative process (Shneiderman, 1998) requiring intuition and an artistic sense from the designer, and also some past design experience or mimicking from other systems' design (Sanderson, 1997) which should be understood as part of the design process. The designed interface usually goes through user testing and is iteratively refined. However, a single "optimised" interface developed this way cannot satisfy everybody, because people come to the system with different aptitudes, attitudes, preferences and task contexts. Furthermore, the current trend in technology shows diversification of devices for a single underlying system sharing the same data, such as email software accessed from an office desktop PC, a PDA or a mobile phone. This results in the need for designing different user interfaces for different devices suitable for different users. To address these problems, there have been efforts to streamline and turn the fuzzy, unpredictable and ill-defined interface design concept into a more structured and formalised process, exemplified by "design space analysis" (MacLean, Young, & Moran, 1989) and further adapted in various forms such as Stary (2000). In this approach, roughly the following steps are followed:

(i) analyse and identify important elements and alternatives in designing an interface, resulting in an exhaustive sets of possible design options, or design space;
(ii) consider the particular environment where the interface in concern is to be used, and
(iii) select a suitable set of options from the design space

This way, designing an interface becomes less of an intuitive, artistic task and more of a concrete, clear and simple decision-making process where the designer can come up with many different interfaces by selecting different combinations of options suitable for the target usage. A crucial part in this approach is the initial construction of the design space and the selection of the right set of options for the target usage. In designing video keyframe browsing interfaces for the Físchlár system, we have constructed a design space by identifying 3 important facets (or dimensions), and the possible options (or values) for each dimension.

2.1 Layeredness

The amount of keyframes extracted from a video programme crucially affects the user's browsing. A very large number of keyframes would provide detailed browsing, but become unsuitable for quick browsing. Provision of only a small subset of keyframes (for example, 10 to 30 keyframes extracted from one programme) could be a useful way of providing an overview of the programme content. The *Layeredness* dimension is concerned with the different detail or granularity of the keyframe set and the transition between different granularity, and has to be considered in designing a keyframe-based browsing interface. Some of the typical options for this dimension are:
<u>Single layer</u> - provides only a single set of keyframes, whether very detailed or selective;
<u>Multiple layer without navigational link</u> - provides more than one set of keyframes in a browser, thus the user can select the granularity s/he wants in the browsing;
<u>Multiple layer with navigational link</u> - provides more than one set of keyframes, and the user can jump between different sets of keyframes while maintaining the current point of browsing.

2.2 Temporal Orientation

Video is a time-based medium and the keyframes extracted from video are an ordered set of images by time. Thus an important concern is *what kind of time information*, if at all, should be provided to the user when browsing the keyframe set. Some of the typical options for this dimension are:
<u>No time information</u> - provides no explicit time information regarding each keyframe;
<u>Absolute time</u> - provides exact time information in numeric form (for example, timestamping each keyframe such as "15 minutes 30 seconds into the video");
<u>Relative time</u> - shows the time of the current browsing point in relation to the whole length of the video (for example, a timeline bar indicating the current viewing point).

2.3 Spatial vs. Temporal Presentation

There are two distinctive ways of presenting keyframes on the screen, and the designer has to decide which one should be adopted for the interface in concern:
<u>Spatial presentation</u> - displays many miniaturised keyframes side by side, allowing quick spatial browsing;

Underlined "Temporal presentation" - displays keyframes one by one.

2.4 Specifying a Browsing Interface

The three dimensions and their typical options introduced above form a design space where the designer selects one (or more than one) option from each dimension. Because each design option represents a distinctive design decision in a dimension, different combinations of options result in distinctive browsing interfaces. This makes it possible to design all conceivable and different browsing interfaces. For example, selecting the following options

Single layer from Layeredness dimension
Relative time from Temporal orientation dimension
Spatial presentation from Spatial vs. Temporal presentation dimension

results in a specific and distinctive browsing interface. What options to select should be based on the target usage context. The target users and the kind of devices they will be using roughly indicate which options are preferable and which are not. For example, on a very small, low-resolution screen such as a mobile phone, spatial presentation is not suitable because each keyframe content would be unrecognisably small if we try to further miniaturise keyframes and display many of them on one small screen.

In designing the browsing interfaces to the Físchlár system on a PDA, general guidelines and common sense can be used in this selection process. The following section considers these general ideas suggested for mobile device interface design, to be applied to the design framework to select suitable options, and two example interfaces emerged as a result.

3. Video Browser for PDA

User interface design for mobile devices is getting more attention nowadays, as more and more computational power and higher wireless bandwidth are becoming available for these devices. However, the usability and interface design of today's mobile devices has not been studied enough and most elements are simple replicas of desktop interface elements. The problem is that a good design for a desktop application with a large monitor and a pointing device might not be suitable for a small, handheld device such as PDA or mobile phone. Well-established graphical user interfaces with a direct manipulation style will enhance the usability greatly for a desktop environment where the user generally keeps on looking at the screen holding the mouse, ready for reacting to any visual feedback from the system (Kristoffersen & Ljungberg, 1999). This is far from the truth in the mobile environment (in a bus, on the street, etc.) where the user may be unable to keep on focusing on the screen and small visual details can easily be overlooked. However, a successful interface style on one environment can be completely unsuitable for another environment. While there should be more studies in developing a new interaction paradigm for mobile interfaces (Marcus, Ferrante, Kinnunen, Kuutti, & Sparre, 1998), in this paper we concentrated with designing interfaces based on the framework described in previous section, for a particular PDA called Revo from Psion. Revo was chosen within our project administrative resources and not necessarily ideal for developing fully-working versions. The interfaces designed were not connected to the Físchlár system via a high bandwidth wireless network to show live online information stored on the server, and the limited memory size (8 MB) and display rate of the device meant that a realistic evaluation involving user testing was not possible. However, for the purpose of demonstrating our design directions and ideas, this initial phase mock-up demonstration was sufficient.

The Revo has a landscape-oriented screen of 480 by 160 pixels with 16 shades of grey. The screen is touch-sensitive and is usually used with either a finger or a stylus that comes with the device. The first obvious design decision is that the low resolution and small screen of Revo makes spatial presentation unsuitable. Figure 1 shows one of our designed interfaces. On the right side of the screen, the list of recorded and processed TV programmes is displayed with a scroll bar, which the user can use with the right thumb while holding the device with the same hand. On selecting one of the programmes, the first keyframe extracted from that programme is displayed on the left side of the screen. Below the keyframe there are two buttons

Figure 1. Single layer / Relative time / Temporal presentation

(previous/next) for the user to flip through keyframes one by one, using the left thumb while also holding the device with the left hand. As the user flips through the keyframes by repeated tapping on the buttons, a timeline bar beside the buttons indicates relative time showing the current point of browsing in relation to the whole programme. Automatically flipping through keyframes (true temporal presentation) rather than using the buttons would be possible, but because it should not force the user to keep on concentrating on the screen, having the user controlling the keyframe flipping was considered better. Requiring a high degree of interaction (repeated tapping on the previous and next buttons) should be okay when the number of the interaction objects (buttons) is only two and these are always under the user's thumb. Note that it is possible to grab the device and interact with it with only one hand at either stage of interaction (either browsing to select a TV programme or flipping through a programme content), freeing the other hand for something else (holding a bag, opening a door, etc.). When both hands are available, the user can simply grab the device firmly with both hands, and use the right thumb for scrolling and selecting a TV programme, and the left thumb for flipping through the selected programmes keyframe content.

Another designed interface for Revo is in Figure 2 below. In this interface, browsing the keyframes of a single programme is considered, with multiple layers of keyframes available. With the two buttons on the right side (up/down buttons), the user can jump between 6 different layers, with the layer indicator beside the buttons showing currently selected layer. The top layer has very selectively-chosen 10 keyframes providing an overview of the whole programme; the bottom layer has the full camera shot-level amount of keyframes depending on the programme (usually 300-700 keyframes); other 3 middle layers provide in-between details of keyframes to browse. With the two buttons on the left side (previous/next buttons), the user can flip through the keyframes in the currently selected layer, and the current temporal position in the programme is indicated with the timeline bar above the buttons. The layers have navigational links between them, meaning that when the user jumps up or down a layer, the current point of browsing is maintained (i.e., when the user taps on the up/down button on the right, the timeline bar on the left will still be indicating the same position). Thus, the user can easily jump up a layer to move the current position more quickly to right or left, and when an interesting position is reached, jump down a layer to browse more detailed keyframes in that area. Note that all the widgets on the interface are on the sides of the screen and the keyframe displayed is in the centre. This browser is mainly meant to be used with both hands grabbing the device and continuously tapping buttons as if playing a pocket video game console, the user continuously moving up and down the layers while flipping through keyframes left and right in a highly interactive manner.

Figure 2. Multiple layer with navigational link / Absolute & Relative time / Temporal presentation

In both designed interfaces, the user has full control over the displayed information on the screen while the widgets to trigger any system feedback (i.e., buttons) are very obvious and always in easy reach. This style makes it acceptable and natural for the user to casually take attention away from the screen while stopping interacting with the system with thumbs, and a few seconds later focusing back on the screen, no more different from casually playing a pocket video game while sitting in a busy bus on the way home. Both interfaces were designed in such a way that the user need not pay careful visual attention or point at a small area in the middle of the screen, unlike the majority of desktop application interfaces (as well as the PDA interfaces which adopted the desktop interaction style) require. A rather simpler and blunt interaction, but still efficiently allows highly interactive and pleasant browsing of video content in different details. Although more sophisticated applications with more functionality would inevitably require finer-level pointing on the screen with the stylus, the overall styles used in the above example interfaces would be a good direction and guideline in designing other applications on a mobile device.

The design framework used in this design provided building blocks and the way to consider different elements in designing the interfaces, while general ideas on designing interfaces for the mobile environment provided selecting particular elements from the framework as well as dictating the overall interaction style of resultant designs. Revo provides as an option a small “tick” sound as aural feedback whenever a screen is touched, thus the user knowing that his/her action was sent to the system. However, mapping the “virtual buttons” on the above interfaces to physical buttons on the device would enhance the tactile feedback for the user (Myers, Lie, & Yang, 2000), which was not possible to implement with Revo for technical reasons. With the technical limitations of the Revo, informal discussion with a small number of test users has been conducted during and after the implementation mainly to get rid of minor usability problems (for example, the initial implementation of the first interface had smaller button sizes than the current version, which was picked up by a few test users), often with positive comments on the overall idea. As an alternative to using Revo, we are presently developing a mobile PDA interface to Físchlár for the Compaq iPAQ which has more RAM (32 MB as opposed to 8 MB), a more detailed screen (4,096 colours as opposed to 16 greys), a faster processor (206 MHz as opposed to 36 MHz) and has wireless LAN connectivity.

4. Conclusion

In this paper, designing a keyframe browsing interface to a PDA was considered with a specially constructed design framework for video browsing interfaces as a base. The commercial and research community are more and more aware of the importance of recognising people's individual differences and personal preferences, as evidenced in the idea of "personalisation" both as a live research area and appearing frequently in popular technology products and services. In the user interface design field, attempts to cater for the diversity makes it difficult to have a single user interface for a system which supports everybody's needs. Furthermore, the diversification of different devices for very different environments makes it impossible to stick to a single interface to support these different environments. Identifying all possible interface elements and specifying an interface from this list can be a good starting step for heading towards realising universal access which supports potentially all users and their circumstances. Already research is underway in the area of "unified interface design" (Akoumianakis, Savidis, & Stephanidis, 2000), which considers developing methodologies and tools for an intelligent system which can eventually automatically identify each individual user's preferences and needs at the time of use, and assemble suitable interface elements to provide this to the user dynamically, thus eventually being able to support everyone's needs on an individual basis.

Mere technological progress does not guarantee a wide acceptance of usage of that technology in the end product. Numerous failures in usability are found in small, handheld devices because the same interface paradigm for the so-far dominant desktop systems were used without further elaborate consideration. It is thus important to consider in depth the context of the use of the particular interface in concern.

References

Akoumianakis, D., Savidis, A., & Stephanidis, C. (2000). Encapsulating intelligent interactive behaviour in unified user interface artefacts. Interacting with Computers, 12, 383-408.

Kristoffersen, S., & Ljungberg, F. (1999). "Making place" to make IT work: empirical explorations of HCI for mobile CSCW. Proceedings of the ACM Conference on Supporting Group Work (GROUP '99), Phoenix, AZ, 276-285.

MacLean, A., Young, R., & Moran, T. (1989). Design rationale: the argument behind the artifact. Proceedings of the ACM Conference on Wings for the Mind (CHI '89), Austin, TX, 247-252.

Marcus, A., Ferrante, J., Kinnunen, T., Kuutti, K., & Sparre, E. (1998). Baby faces: user-interface design for small displays. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '98), Los Angeles, 96-97.

Myers, B., Lie, K., & Yang, B-C. (2000). Two-handed input using a PDA and a mouse. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2000), The Hague, Netherlands, 41-48.

O'Connor, N., Marlow, S., Murphy, N., Smeaton, A., Browne, P., Deasy, S., Lee, H., & Mc Donald, K. (2001). Físchlár: An on-line system for indexing and browsing of broadcast television content. To appear in the 26th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001), Salt Lake City, UT.

Sanderson, D. (1997). Technology design and mimicry. Proceedings of the ACM Symposium on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS '97*)*, Amsterdam, 311-313.

Shneiderman, B. (1998). Designing the user interface: Strategies for effective human-computer interaction (3rd ed.). Reading, MA: Addison Wesley Longman, p.89.

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

Stary, C. (2000). A structured contextual approach to design for all. Proceedings of the 6th ERCIM Workshop "User Interfaces for All", Florence, Italy, 83-97.