Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Displays 26 (2005) 79–96
DISPLAYS
www.elsevier.com/locate/displa

Modelling VCR-like video content navigation

Chris Crockford, Harry Agius*

School of Information Systems, Computing and Mathematics, Brunel University, St John's, Uxbridge, Middlesex UB8 3PH, UK

Received 4 January 2005; accepted 12 February 2005
Available online 13 March 2005

Abstract

Although various alternative control sets have been proposed for various purposes, VCR-like control sets remain the most commonly used for everyday video content navigation by many professional and home users, due to their predominance within popular commercial media players, such as Windows Media Player, QuickTime and RealPlayer. Consequently, this paper proposes a model of user interaction with VCR-like control sets, with the aims of better understanding the process and informing the development of future video content navigation control sets. The model is derived from the results of empirical experimentation. A digital video navigation system with a VCR-like control set was developed and subsequently used by a large sample of users ($n=200$), who were required to complete a number of goal-directed navigational tasks. Each user's navigational activity was tracked and recorded automatically by the system. Subsequent analysis revealed several patterns of interaction employed by users that were then used to construct the model.

Keywords: Digital video; Navigation; Browsing; Searching; VCR; Multimedia; Media player; Empirical study; Interaction

1. Introduction

A VCR-like control set consists of controls similar to those found on VCRs, such as play, fast forward, fast rewind, stop and pause. Although a number of alternative control sets for digital video have been proposed for various purposes [12,16,17,20,26], VCR-like control sets remain the most commonly used for everyday digital video content navigation by many professional and home users. Much of this can be attributed to their predominance within popular commercial media players, such as Windows Media Player, QuickTime and RealPlayer. A linear control set, it reflects the original requirement to transport analogue tape mechanically across playback heads. However, user familiarity with this control set, due to the success of VHS [19], has seen it adopted for the control of digital video too. Often, the control set is complemented by a visual timeline that shows the current playback position relative to the entire video stream. Despite the strong prevalence and popularity of VCR-like controls for navigating digital video, little research has attempted to better understand how users relate to it when attempting to navigate digital video, thus it becomes difficult to postulate what future video content navigation control sets should embody in order to better support the user.

Consequently, this paper proposes a model that represents user interaction with VCR-like control sets. The model is derived from the results of empirical experimentation. A digital video navigation system with a VCR-like control set was developed and subsequently used by a large sample of users ($n=200$), who were required to complete a number of goal-directed navigational tasks. Each user's navigational activity was tracked and recorded automatically by the system. Subsequent analysis revealed several patterns of interaction employed by users that were then used to construct the model. The remainder of this paper is structured as follows. Section 2 reviews relevant research literature. Section 3 presents the design and implementation of the digital video navigation experiment together with the results of the experiment; in particular, it describes the user searching and browsing techniques that were discovered. Section 4 presents the model that depicts user interaction with the VCR-like control set. Section 5 concludes the paper.

* Corresponding author. Tel.: +44 1895 265993; fax: +44 1895 251686.
E-mail address: harryagius@acm.org (H. Agius).

0141-9382/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.displa.2005.02.003

2. Existing research

The longstanding popularity of VCR-like controls for navigating video is something that has been demonstrated in the literature. For example, Branch et al. [9] discovered how frequently VCR-like controls were being used by users. Consequently, when developing their interface for a video-on-demand system, they decided to supplement, rather than replace, the controls with additional functionality, such as annotated bookmarks. Similarly, Drucker et al. [12] attempted to evaluate the use of VCR-like control sets within the home. In their experiments, the user controlled a digital PVR (personal video recorder) via a standard remote control while seated on a sofa. Three interfaces were evaluated in total: Tivo, Replay and a new interface that they developed known as Smartskip. Tivo and Replay's interfaces consisted of the standard VCR-like control set with additional skip forward buttons. Tivo provided fast forward scan speeds of 3×, 20×, and 60×, whereas Replay allowed users to jump forward through 30 s of media. Smartskip enhanced these two control sets by placing a scroll bar on the screen alongside a series of thumbnails which depicted the events occurring in the media in both forwards and backwards directions. This enabled the user to click on thumbnails to navigate to the requisite segment of video media. Drucker et al. found that, while their enhanced interface was not the most efficient in trials, it was deemed the most satisfactory by users. Pua et al. [23] provide a further example of VCR-like control enhancement in their VIDSEEK interface to the Vision digital video library. By presenting the user a timescale to undertake a keyword search within, they were able to reduce the information set returned. They then provide two further video interaction windows, one for key frames of interest and a separate RealPlayer window. The key frame window allowed users to navigate through the key frames to retrieve what they are looking for, whereupon the RealPlayer window could then be used to review the selected media. Van Houten et al. [30] proposed that the clustered nature of video specific information relevant to the user occurs in what they describe as 'patches'. In their system, they allow for the annotation of these 'patches of interest' by colouring in segments of a time line to represent the patches of interest. Such patches can then be jumped to and VCR-like controls used to navigate within those patches. Barletta et al. [7] proposed a framework for 'leafing' digital content. The framework was based on the same concept that users utilise when leafing through magazines, and was designed to encompass both the mental and manual processes that the user undertakes. The proposed framework was engineered to create a summarisation of semantic video excerpts, and VCR-like controls deployed as the ultimate navigation interface. In all of these cases, it is shown that users are not hostile to new control sets when some aspects remain familiar to them. However, different control sets often result in different viewing behaviour by users [19].

Other research has sought to provide alternative control sets. For example, Hurst and Stiegler [16] mooted that the VCR-like control set is limited to only being able to control media at an immediate temporal point, providing no ability to control media at remote temporal points. As a result, they proposed an alternative means for navigating forwards and backwards through a continuous media stream using scrollbars, where users zoom the scroll bar to adjust the granularity of navigation. Kamvar et al. [18] propose a similar approach for small-scale displays. They argue that VCR-style controls occupy too large a proportion of the screen on small displays and so justified the usage of the scroll bar and an associated magnifying function. For navigation of multiple video streams, Balachandran and Rangan [5] developed an interface that allowed users to view a matrix of different channel clips at differing times. This afforded the user the ability to look forwards and backwards with respect to time to see which video media were available. Using a directional arrow control set, the user could navigate and subsequently highlight the video window within the matrix that they wished to review, before issuing a command to select that specific media stream. Foote et al. [14] utilised a linear histogram depicting what they define as 'confidence scores' within the selected media as well as means of jumping to these areas of interest. Further control is presented through the deployment of scroll bars, to navigate through these areas of interest, together with a limited-functionality VCR-style control to provide a means of travelling at different speeds forwards and backwards through the media stream.

More advanced alternative video interfaces predominantly focus on filtering the video stream down to a manageable subset that contains only the most pertinent content. Hence, the user is offered a reduced, but more convenient, video stream to navigate, known as a *video abstract* [8,24]. Frequently, a control set more suitable to this reduced stream is also provided. User interaction with video abstracts can be seen as being subject to both media richness theory (MRT) and cognitive fit theory (CFT) to varying extents [15], depending on the particular type of video abstract. MRT is the ability of information to change a recipient's understanding within a given amount of time, such that 'rich' information can change a recipient's understanding more quickly than 'lean' information. CFT is the fit between the problem representation, the solution and the method involved. The greater the fit, the more effective the problem solving task will be. A key frame list in chronological order is the most basic type of video abstract. For example, British Pathé's [10] Web-based system uses a storyboarding technique where every 100th frame is displayed as a thumbnail. It does not, however, allow for the user to navigate directly to data that the first stage reduction has discarded. Similarly, the CueVideo project [26] provided a new style of interface for interacting with storyboarded video data, whereby