A Gesture-based Personal Archive Browser Prototype

Cathal Gurrin[1], ZhenXing Zhang[1,2,] Hyowon Lee[1] & Denise Carthy[2]

[1] Centre for Digital Video Processing, Dublin City University, IRELAND.
[2] Biomedical Diagnostics Inistiture, Dublin City University, IRELAND.

{cgurrin, hlee, zzhang}@computing.dcu.ie; denise.carthy@dcu.ie

Abstract. As personal digital archives of multimedia data become more ubiquitous, the challenge of supporting multimodal access to such archives becomes an important research topic. In this paper we present and positively evaluate a prototype gesture-based interface to a large personal media archive which operates on a living room TV using a Nintendo Wiimote for interaction.

1   Personal Media Archives in the Livingroom Environment

The increasing trend recently of people becoming content creators and not just consumers poses a challenge for organising and accessing the resulting personal archives of multimedia data. We have also noted a recent trend towards the integration of personal content management technologies into the enjoyment-oriented (lean-back) environment of the living-room, for example DVR functionality or WWW access on the TV. In this paper, we are concerned with the integration of personal content organisation facilities into the livingroom TV environment, which poses a number of challenges because it needs to be performed by non-expert users, with a remote control in a distractive (lean-back) environment, and not at a desktop computer with use of a keyboard and mouse (the typical lean-forward environment).

Previous research into managing personal media archives on a desktop device does not directly transfer to the living room environment, for reasons such as user interaction support, difficulty of querying and even device processor speed. It is our conjecture, however, that a living-room TV acts as a natural focal point for accessing personal media archives, and that taking into account the significant limitations and challenges of developing for such an environment is essential to successfully deploy multimedia content organisation technologies. Indeed initial work in the area by Lee et al. [1] suggests that simplicity of interaction is crucial for the livingroom environment, more crucial than in any other digital media domain and that ultimately, this simplicity of interaction determines the success or otherwise of any new applications. The challenge therefore is to marry the competing requirements of supporting complex digital multimedia archive organization technologies with the simplicity of interaction required when developing for the livingroom environment.

In this work we describe and demonstrate an interactive TV application, for the livingroom, which employs a gesture-based interface (using a Nintendo Wiimote) to manage a large archive of personal media gathered using a Microsoft Sensecam[2].

Studies on interactive TV interaction highlight the special characteristics of the livingroom environment and they show design implications and guidelines for a technology operating in such a context. Characteristics such as use of a remote control as input device, increased viewing distance and enjoyment oriented (not the complete focus of task-orientation) usage scenarios prevail. Based on these studies, previous research and our own experiences of developing information retrieval systems for lean-back devices [1], we have compiled a set of guidelines for developing interactive multimedia applications for the livingroom, or any lean-back, environment:

- *Represent complex digital multimedia objects visually.* Complex multimedia objects, such as photo collections, video archives or HDM archives need to be visually represented and easily manipulatable on screen.
- *Minimise user input where possible* and proactively recommend content or support information seeking via a small number of frequently used features.
- *Engage the user* with simple, low-overhead interaction methodologies, that are enjoyable to use, easy to learn and engaging in a distracting environment.

2. A prototype gesture based diary interface

Following the three guidelines, we developed a gesture-based browsing interface to an archive of Microsoft Sensecam [2] images operating on a living room (40 inch) TV, using a Nintendo Wiimote. Two weeks (about 50,000 photos or 3,500 per day) of Sensecam data, gathered by one wearer, was employed for this experiment, which we feel to be a good example of a challenging personal archive. The Sensecam images were tagged with date/time and location, as is standard for digital photos. How the three guidelines (above) impacted on the prototype is now illustrated:

- *Represent complex digital multimedia objects visually.* A HDM archive is an enormous repository of data and as such it needs to be summarised and visually easy to browse and interpret on any device. We used the event segmentation technique of Doherty et al. [3] to organize each day's images into a set of about thirty discrete events, which utilized visual processing of temporal image dissimilarity coupled with and analysis of Sensecam sensor data. A keyframe was automatically selected to represent each event based on an automatic analysis of its visual significance within that event.
- The prototype *minimized user input* by providing a diary-style calendar interface as the key access mechanism. A user could select next/previous days (a simple Wiimote gesture) and also select next/previous event (another simple gesture). Upon selecting an event, the event playback began which cycled through the images comprising that event at a fixed speed. The speed of this playback (from pause to fast-forward/fast-rewind) was user controlled by twisting the Wiimote as if one is twisting a dial or a knob.
- The prototype engages the user with *low overhead and low learning time* interaction methodologies that users found both easy and enjoyable to use.

A user evaluation was carried out of this prototype gesture based interface (Fig 1) with six novice users and the sensecam owner. All users received five minutes

training on both interfaces. The prototype was compared to a similar interface on a desktop device with mouse and keyboard (lean-forward) interaction. Six tasks (four known-item and two ad-hoc search) were allocated to each participant (on alternate interfaces) and ordered to as to avoid bias. The average time taken for known item search was 77 seconds for the gesture interface and 73 seconds for the desktop interface, which was similar. For the ad-hoc task however, users of the gesture interface found 50% more relevant images than the desktop interface under the same time constraints. For the Sensecam owner (with a good knowledge of the data), the gesture interface was significantly better at finding known items with little difference on the ad-hoc search. In a qualitative examination, users found the gesture interface to be more satisfactory, efficient, productive and easier to recover from error when compared to the desktop interface which was more comfortable and easier to learn.

In conclusion, we have presented a set of guidelines for developing interactive search/browsing systems in a livingroom environment and evaluated a prototype system adhering to these guidelines. We found that tailoring the underlying algorithms to suit the limitations of the target environment can result in an equivalent (or better) performing system than an equivalent desktop implementation. Since the integration of more personal content organisation technologies into the livingroom environment is likely, these guidelines and results are a valuable initial contribution.

Fig. 1. The prototype gesture-based interface showing playback from the seventh event of the 12th April 2009, which took place in early afternoon in Dublin, Ireland.

References

1. Lee, H., Ferguson, P., Gurrin, C., Smeaton, A.F., O'Connor, N. and Park, H.S. Balancing the Power of Multimedia Information Retrieval and Usability in Designing Interactive TV. In Proceedings of uxTV 2008, Mountain View, CA, 22-24 October 2008.
2. Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N. and Wood, K. "SenseCam: A Retrospective Memory Aid", In Proceedings of the 8th International Conference on Ubicomp (Sept 2006).
3. Doherty, A.R. and Smeaton, A.F. "Automatically Segmenting Lifelog Data into Events". WIAMIS 2008 - 9th International Workshop on Image Analysis for Multimedia Interactive Services, Klagenfurt, Austria, (May 2008).