

Hypermedia for Portable Video Players (PVP)

Mike Leggett, Shigeki Amitani
Creativity & Cognition Studios, Faculty of Information Technology,
University of Technology Sydney, Australia
{mleggett@uts.edu.au, shigeki@it.uts.edu.au}

Abstract

In this paper we propose the exploitation of high mobility portable battery operated Video Players (PVP) for the retrieval of video associated with the location in which it may be used. Reporting on an earlier interactive multimedia location-based prototype, we assess the possibilities for specific ontologies of a taxonomy of indexing procedure which avoids text-based retrieval methods, using instead the mnemonics of image association. We outline the proposed development of PVP firmware and a related user application enabling users to construct indexing procedures appropriate to their needs, using a metadesign approach.

Keywords: hypermedia, video, video player, authoring

1 Introduction

The proposal emerges from current interdisciplinary research into machine memory as a context for understanding its relation to human memory and methods for storing and retrieving movie files. It proposes an approach to indexing audio-visual media utilising an 'index-movie' file as the taxonomy of the indexing procedure, to which is linked related movie files. An interactive experimental prototype, PathScape, has provided initial evaluation of the concept using a real-world time-space representation as the basis for indexing. Further practice-based research approaches to user-defined storage and retrieval systems for the video iPod and other PVPs as advanced portable video systems, will be described.

The proposal is for the PVP user to interactively navigate the linkages between movie files, either as an exploration of a creative maze, or as a means of recalling a particular series of operations, directions, sequences explained in pictures and sound, but under the direct and immediate control of the PVP user. This feature will

enable complex data structures often represented visually - land surveys; mining topographies; design or biological sequences; architectural spaces; construction progress; cultural artifacts; etc - to be made accessible relationally rather than sequentially.

Whilst positioning of a pointer on a visible timeline provides instant access to a particular part of a movie in a conventional computer-based movie player, this is not an option in PVPs. However, the visibility of images during high speed spooling on a PVP, could assist locating entry points to a hyperlinked movies system utilising frame number metadata and mnemonics. An indexing approach of this kind implies special concerns in the design of such a system, for individual, specialised and public groupings and communities, for which metadesign approaches are being developed.

2. Navigation Principles

Interface design for multimedia databases has been the subject of investigation by earlier researchers for desk-based systems, though few have achieved avoiding the use of words or on-screen graphical devices to aid navigation. [1] [2] [3]. Experimental approaches by artists have included *Twelve of My Favourite Things*, effecting navigation using a touch screen over an image composite of three movies linked to other movies related by a colour selected on the screen.[4] In the late 1990s a website appeared documenting the Exeter Cathedral Vaulting: "There are two main routes into the material, Visual and Verbal. The Visual route is for those who are more at ease with images than text." The ceiling, built in the 14th Century, used the vaulting bosses as a mnemonic system related to the stories both sacred and profane, of an oral culture in the West Country of England of the time. The designers of the website echoed the memory system by using a plan of the vaulting and its bosses to access the database containing detailed photographs of each item together with several layers of metadata.[5] More recently the *Digital Songlines* project at the Australasian Centre for Interactive design in

Queensland uses graphical representations familiar in game engines, to map the GIS data relevant to 'country' and cultural artefacts, related to an indigenous community.[5]

The principle of this taxonomy does not seek to index video libraries or collections, nor provide machine-based 'importance sampling'. [7] The concept of *detail-on-demand* is a means of working with specific video material that avoids "...having to use a separate interface such as keyframes or a tree view". [8] As a means of navigation it has been explored by others [3, 9, 10, 11] based on earlier experiments with video and hypermedia theory [12].

The central novelty of an approach to *mnemonic movies* indexing is to enable an accelerated usage of movie based data or information. The movie being watched will provide the link to the related movie(s), without the need to return to scroll a text-based index menu at the root. It will enable PVP users to engage interactively with videos using links to move from one movie to another according to relational rather than sequential connections.

These approaches overlap with the Greek oracists and rhetoricians, who before the alphabet had been handed down, developed an elaborate form of artificial memory, described so fully in Yates' *Art of Memory*. Ars memoria, "...a series of *loci* or places. The commonest, though not the only type of mnemonic place system was the architectural type We have to think of the ancient orator as moving in imagination through his memory building whilst he is making his speech, drawing from the memorised places the images he has placed on them." [13] It could be claimed the first movies were a conceptual model made by the Greek rhetoricians, complete with wide shots, tracking shots, panning, tilts, close-ups and flashbacks. Played in the cinema of the mind's eye, the first 'classic film narrative' guided his oratory from theme to theme, detail to detail, by associating each element of the speech with the *loci* and the objects placed there and visible only to him.

2.1 Pathscape

Our familiarity with cinema and the reading of Cartesian spatial representation is exploited in the PathScape prototype system. It explores through demonstration, a means for augmenting human memory for the purposes of storing and retrieving movie files. The *detail-on-demand* principle employed however, has no overarching narrative, but a series of interactive option prompts. These access movie files in the system using a taxonomy based on fragmentary images, sounds, colours and shapes. The 'index-movie' file (I-MF) produces apparent motion in a central image for forward direction along an X-Y axis, perceived as a movement 'into' the cinematic space recorded, a landscape.



Figure 1: Screen images

The movement is controlled by gesture, using a mouse in the prototype (Figure 1 & 2) to 'move' towards point X accessing file I-MFX; by gesturing to the central image, movement ceases; gesturing to the bottom of the screen instantly loads I-MFY movie file, swinging the image through 180° to return along the path previously followed towards point Y.

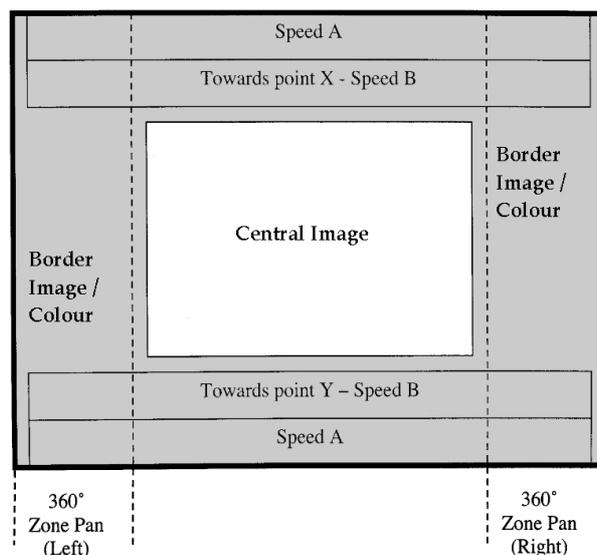


Figure 2: Screen area images and Cursor gesture outcomes

The taxonomy of the Path which the user traverses is ordered by *three* indexical devices. Two are located in the border area that surrounds a central image. The *first* level of indexing is within this border and seen at particular points as fragments of images, visible for short durations. These indicate a nodal junction which, when 'captured' by using gesture to halt movement in the central image, will enable with a click, the launch of a movie and associated sound from the database, replacing the central image movie of movement along the path.

Thus along the X-Y axis are the 1, 2, 3, 4, 5 etc interactive options, 'narrative branch nodes', which in effect are groups of movie keyframes representing a *loci* or location linked to an associated movie file. (Figure 3)

The *second* device uses changes in background colour in the border area and background sound to signify changes of zone. (In this prototype different colours represent different ecological zones). When a colour is visible in the border, gesturing to the left or right of the screen will launch the movie of a 360° panning movement of the landscape, (Figure 1 & 2) a movie representation of the zone through which the user is currently 'passing' - gesturing to the right will pan right, to the left will pan left : AA, BB, CC ... FF. (Figure 2 & 3) Within the pan will be 'found' further narrative branch nodes from where to launch movies set during the authoring process, associating each movie with the visible appearance of each locale.

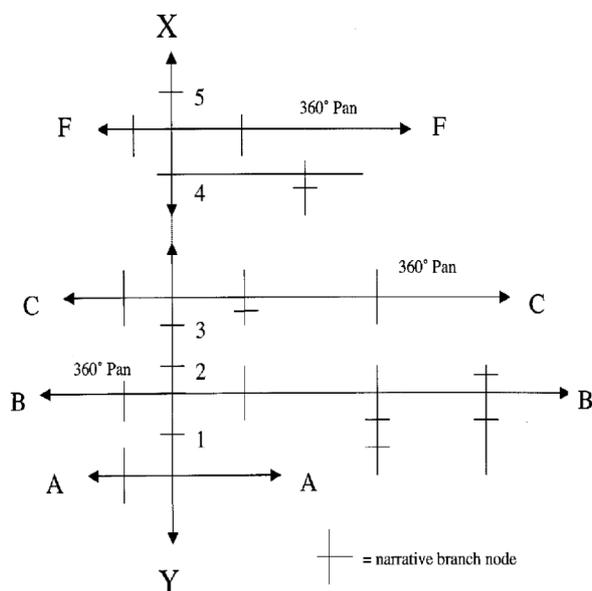


Figure 3: Schematic for accessing database

At the completion of a narrative, the *third* indexical device appears as a series of circle shapes that appear over the final frame of the movie. Blue, yellow and brown and green circles function as 'buttons' to linked topics, colour coded to symbolically represent a narrowing of the index path from the broad to the specific. [14, 15]

The encounter in this prototype enables the user to orientate within a given topography in a way not dissimilar to a regular route followed in the country or the city. Similarly, interaction with the surroundings reveals hidden evidence, concealed information and comment, delivered as stories, as samples of discrete

information enabling the interacting subject to put together knowledge of this place through information gathered. The interactive process is not through query structures addressed to a database, but as embodying gestures, using the relational terms, "more, same, less" within the interface of mnemonic cues to linked movie files. The experience is a procedure of constructing meaning through familiarity as part of a gathering process that adds to the individual's knowledge base accumulated during this and subsequent visits.

2.2 Prototype Outcomes

The prototype explored the means and the cinematic syntax of creating a multi-layered representation of the landscape, through time as well as space. As a multi-voiced 'interactive documentary' over which the visitor has agency to 'move', to be able to order the stories and the depth of detail which could be retrieved in the prototype, revealed four main areas of response:

- visitors who wholly embraced the visual and navigational experience together with the knowledge building process;
- visitors who wholly embraced the experience without much concern for the documentary and informational aspects;
- visitors for whom the knowledge acquired was unacceptable and without authority or specificity;
- visitors who resisted the responsibilities of interactive engagement.

The prototype demonstrated a wide range of responses from users but most acknowledged the novelty and applicability of the approach to a field of their interest. This indicated to us the need to develop an authoring tool that would enable individuals and groups to design their own system for linking their movies.

2.3 Video Acquisition

The prototype was completed in 2000 and since that time the video data stream has become more ubiquitous. Whether generated by a digital video handycam, mobile-phone, a web-based stream or download, optical media and broadcast television and video-on-demand databases, an ever increasing amount of digital media images and sounds need to be managed, whether for professional or recreational purposes. The PVP is an affordance for making use of the video data stream in a variety of ways in a range of ontological contexts.

2.4 Navigating the PVP

Codecs for video files and devices to handle them in creatively useful ways have developed exponentially. The Apple video iPod for instance, can store up to 3

hours of video playback and delivers high quality video using several codecs, 320 x 240 pixels at 30 frames per second with stereo audio. Interacting with the device is through gesture related to the navigational principles used in the Pathscape prototype (Figure 2) mapped to the device front panel (Click Wheel, Figure 4):

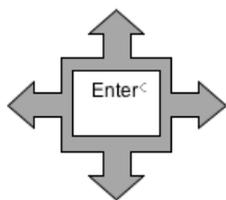


Figure 4: Click Wheel navigation controller on PVP

A simplified mapping, based upon the 'stories in a landscape' approach, will achieve similar outcomes (Figure 5):

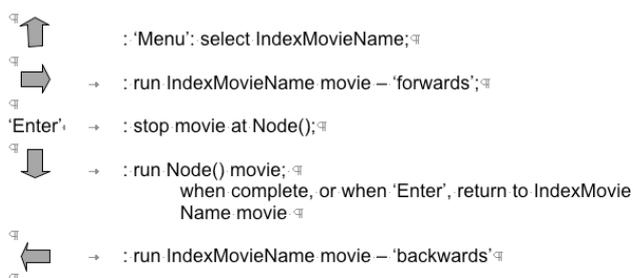


Figure 5: Click Wheel mapped functions

3 Metadesign and Authoring Principles

The use of consumer technology for productive as well as recreational purposes requires an adaptable design approach to the authoring process. Fischer and Giaccardi have shown that metadesign serves the interests primarily of the community of practice (CoP), the consumers, where the community of interest (CoI) are able to provide expert input to a complex design problem. Metadesign gathers potential from these convergences and becomes "...an emerging conceptual framework aimed at defining and creating social and technical infrastructures in which new forms of collaborative design can take place." [16] The metadesigner as CoI, in working with the CoE could advise in establishing a consistent (or even idiosyncratic) relationality for a specific collection of video files by advising on syntax, 'a connected order or system of things', [17] within an image-based indexing system.

In the context of using a modified consumer device to interactively produce outcomes based on relational rather than sequential ordering, it is important that the authoring principle of syntax to be applied in each design

and authoring process is determined. The authoring tool framework can then be applied to set the coordinates for the hyperlinking Node() governing the navigation options. Thus the design task can be seen to deal with mnemonic cues as much as the normally associated temporal aspects of 'editing' film or video, (though duration will be part of that decision-making process).

We propose two approaches to the design of the system. The first, On-Board Authoring is effected on the device itself and is capable of setting very basic relationships between the (suitably compressed) movie files uploaded to the PVP. The second, Off-Board Authoring, is more generic and involves an application external to the device on which the files and their relationships are established using drag and drop procedures before upload to the PVP.

As APIs for iPod are not publicised, we have developed a simulation to indicate how users of iPod or similar PVPs could author and navigate movies. The system was modelled with Java v.1.4.2 on Mac OS X.

3.1 On-Board Authoring

As the PVP has a limited interface, the authoring operations need to be simple and incorporated within the device's firmware. The prototype model has the following basic functions: (1) selecting; and (2) marking the related movies. The authoring operation is:

1. Select a file to use as the "IndexMovie". (Figure 6)

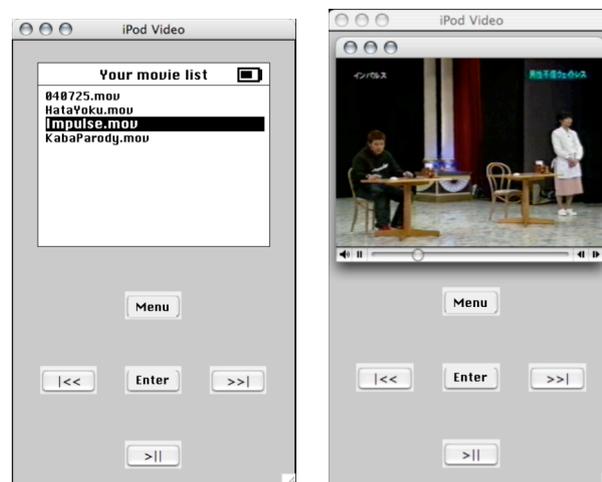


Figure 6: Choose movie. Figure 7: Play movie.

2. Play >|| (Figure 7)

3. Push >|| to Pause to stop the movie at the point a link is to be created.

4. Push "Menu" to see the movie list, and Select another file to link to;

5. Push 'Enter' to set the link, Node(), and return to Index movie.
6. Play >|| to continue
7. Repeat steps 3 - 6 to create additional movie links.

In this simulation the indexing information is stored as a simple text file recording movie file name and frame number from the IndexMovie for the PVP to reference during use. When IndexMovie is played a small arrowhead in the corner of the frame appears for two seconds to indicate where a linked movie can be played by pressing Enter. Otherwise the movie runs, (at fast speed if desired, in either direction, as is standard on PVPs), until the next required indicator is reached. The function of the indicator becomes redundant as the user becomes familiar with 'incidents', or specific images on the movie. Operating as mnemonics these enable the user to recall and so launch, the hyperlinked movie connected to Node() in the IndexMovie.

3.2 Out-Board Authoring

An out-board approach to authoring provides greater flexibility for linking, even to the extent of 'cascading' related movies without using one file as the key indexing file such as the Hyper-Hitchcock project have demonstrated. [8] The more recently demonstrated HyVal system uses authoring visualisation of video objects, metadata and the overall hypermedia document as parts of an Editor tool. Shot detection algorithms effect a semi-automatic function, giving it great potential for working quickly with large video file collections or through using search engine routines. [18]

The out-board authoring we propose for the PVP would employ a timeline similar to existing video editing applications, such as iMovie, as the receptor for linking the metadata associated with the linking options - a sprite dragged to position provides a pop-up window into which the linked movie thumbnail is dragged and dropped from the movie clip viewer. Following playback in the editing tool, adjustments and changes can be more easily effected than within the PVP itself.

4 Applications

Video acquired from many sources can be indexed using visual, non text-based protocols, determined by the individual, group or corporation, at a level of complexity appropriate to the ontological context or immediate application. Practical applications would be characterised through a need for dynamic non-linear navigation of movies, represent pedagogical issues for instance, or research data, media production study or methods, visualisation of spatial or temporal dimension etc. For example:

- as a user-centred product design / protocol analysis / software architecture analysis aid, the PVP becomes a mobile research tool;
- explaining the life-cycle of the frog, at various points in the tadpoles development, the PVP as personal teacher is able to show the detail of a specific moment in that development;
- the PVP as personal electronic tour guide enables the visitor to a place to determine, as with museum audio guides, at what point in a tour more detail is required;
- for the redevelopment of a city area the PVP becomes a planning tool, capable of integrating video-based data with the location in which the data was gathered, at which it is later referred;
- as the recreational device for which it was intended, the Singer Not the Song option will have the user command the iPod view behind the scenes of the recording session and concert footage.
- In the creative space of a classroom, the PVP as a teaching tool in the context of its well promoted use as an entertainment and recreational device will be promoted, in conjunction with an authoring tool, as a valuable learning system, engaging critical and creative assets amongst the student body.

5 Discussion

PVPs are 'hard-wired' devices with no facility at present for dynamic linking of the indexing movie(s) to external databases. Navigable media spaces of the kind described in which individual files can be accessed and / or updated from more centralised media resources and databases, become a 'soft-wired' installation possibility, using the appropriate protocols.

The user of the 'mnemonic movie' option on the PVP is also the designer. Design principles in each case will be approached according to the domain in which it will be employed. As a commercially marketable entity such as a music-based package, the design of the 'bundle of files' will reflect the 'culture of connections' of the target group. For a town planner, collecting data and compiling on-the-fly for examination by other stakeholders, the design approach will be different again. For the artist, hyperlinking will reflect a different set of issues to be explored by the interacting audience, as the mobility of the device enables the city or country environs to be used as the exhibition gallery.

6 Conclusion

PathScope, an experimental interactive prototype, provided initial opportunity to evaluate the concept of indexing audio-visual media utilising a real-world time-

space representation as the taxonomy of the indexing procedure. We propose a system for the PVP user to interactively navigate the linkages between movie files as a means of recalling a particular series of operations, directions, sequences explained in pictures and sound, but under the direct and immediate control of the video iPod or other PVP. The feature will enable complex data structures often represented visually - land surveys; mining topographies; design or biological sequences; architectural spaces; construction progress; cultural artifacts; etc - to be made accessible relationally rather than sequentially.

The contemporary burgeoning usage of the video data stream, whether generated by a digital video handycam, mobile-phone, a web-based stream or download, optical media and broadcast television and video-on-demand databases, determines an ever increasing amount of digital media images and sounds to be managed, whether for professional or recreational purposes.

We have proposed two practice-based research approaches to authoring suitably prepared digital video files, either on-board the PVP or off-board such that the hyperlinked prepared files are uploaded to the device for use 'in the field' of management and development professionals, or in the more familiar recreational ways for which the PVP is enjoyed.

References

- [1] Bolt, R. 'Put That There' Voice and Gesture at the Graphics Interface, (1980) *Computer Graphics* 4 (3) 262-270
- [2] Davenport, G. and e. al, Jerome B. Wiesner, 1915-1994: A Random Walk through the 20th Century. 1994. Accessed: 1.2.04 <http://ic.media.mit.edu/projects/JBW/>
- [3] Naimark, M. Place Runs Deep: Virtuality, Place and Indigenusness. in *Virtual Museums Symposium*. 1998. Salzburg, Austria: ARCH Foundation.
- [4] Hales, C., Portfolio Accessed 1.2.2006 from <http://www.smartlabcentre.com/4people/coreres/chaless.htm>.
- [5] Henry, A. and A. Hulbert, Exeter Cathedral keystones and Carvings. 1998. Accessed 1.9.04 from <http://hds.essex.ac.uk/exetercath/>
- [6] Leavy, B., Digital Songlines, Jones, J. Editor. 2004, Australasian Centre for Interaction Design, QUT: Brisbane.
- [7] Gatica-Perez, D. and M.-T. Sun. Linking Objects in Videos by Importance Sampling. in *ICME'02 IEEE International Conference on Multimedia and Expo*. 2002: IEEE.
- [8] Shipman, F., A. Girgensohn, and L. Wilcox. Hyper-Hitchcock: towards the Easy Authoring of interactive Video. in *Interact 2003*.
- [9] Tua, R. From Hyper-film to Hyper-web. in *Electronic Imaging and the Visual Arts: EVA 2002*. Florence.
- [10] Girgensohn, A., F. Shipman, and L. Wilcox. Hyper-Hitchcock: Authoring Interactive Videos and Generating Interactive Summaries. in *MM'03*. 2003. Berkeley, Ca.: ACM.
- [11] Girgensohn, A., et al. Designing Affordances for the Navigation of Detail-on-Demand Hypervideo. in *ACM Advanced Visual Interfaces*. 2004.
- [12] Tolva, J., MediaLoom: an Interactive Authoring Tool for Hypervideo. 1998, Georgia Tech: Atlanta. <http://www.mindspring.com/~jntolva/medialoom/>. Accessed 1.3.2006
- [13] Yates, F.A., The Art of Memory. (1992 ed) 1966: Pimlico, London.
- [14] Leggett, M. Losers and Finders: Indexing Audio-visual Digital Media. in *Creativity & Cognition Conference 2005*. Goldsmiths College London: ACM.
- [15] Leggett, M., Indexing Audio-visual Digital Media: the PathScape prototype, in *Scan*. 2005, Macquarie University: Macquarie University, Sydney. <http://scan.net.au/scan/journal/index.php>. Accessed 1.11.04
- [16] Fischer, G. and E. Giaccardi, *Meta-Design: a Framework for the Future of End-user development*, in *End User Development*, H. Lieberman, F. Paterno, and V. Wulf, Editors. 2004, Kluwer Academic Publishers: Dordrecht.
- [17] OED, Oxford English Dictionary. 2004.
- [18] Zhou, T. A Structured Document Model for Authoring Video-based Hypermedia. in *Proceedings of the 11th International Multimedia Modelling Conference (MMM'05)*. 2005. Deakin University, Melbourne: IEEE Computer Society