

PathScapes - Interface Options for Visual Indexing

Mike Leggett

University of Technology Sydney

E-mail: Mike.Leggett@uts.edu.au

legart@ozemail.com.au

ABSTRACT: A short paper outlining some issues concerning hypermedia, cinematic immersive states and navigation in the development of an interactive multimedia prototype, PathScape. The experimental model set out to critique Human Computer Interfaces that rely heavily on metaphors derived from the mechanical age and found more in common with classical pre-literacy interfaces that complemented the workings of human memory.

The modern computer is capable of circumventing the written word and demonstrating, on-demand, the spoken word, sound and picture. Though the technical provision for achieving this is rapidly occurring, an interface for indexing sounds and images which is not dependent on words, lies comparatively neglected in HCI advancement.

KEYWORDS: interface, indexing, interactive, cinema, hypermedia.

PAPER

Initialisation

'...writing is the technology of the inventory'. [1]

The traditional Human Computer Interface - HCI - relies heavily on metaphor derived from the mechanical age: the printed page, the desktop, the graph paper, the map, the film soundtrack dubbing sheet, etc. These two-dimensional spaces have acquired the bevel edge, two, three and four pixels broad - embossed frames, windows, work areas, palettes, icons, trash cans and so on. The baubles of the workaday computer world are seen in contrast to the ingenious contemporary games console which, in taking the cinema's Point-of-View (POV) shot, extends the subjects experience using 3D real-time rendering, mapping image to wireframe shapes and background perspectives. Whilst gamers crave action excitement and the rush of adrenalin, how can we begin to further develop this paradigm to become a tool for more complex thought and discourse? Are we moving towards HCIs which complement the workings of human memory and stimulate the individual users imagination? Or does the virtual desert war for couch potatoes become the norm?

As Derrick de Kerckhove has observed: "Literate people have a tendency to translate their sensory experience into words and their sensory responses

into verbal structures. This comes from the habit of translating strings of printed letters into sensory images in order to make sense of what they read." [2].

Random access in the computer as opposed to the book, is capable of circumventing the written word and representing, on-demand, the spoken word, sound and picture. Whilst the technical provision for delivering what amounts to little movies on the computer screen has advanced remarkably of late, (better compression of files, faster processors and networks, etc), the interface capable of becoming a design system for indexing sounds and images is yet to be adopted.

There is a tension here between the tools of the Enlightenment from where so many of our disciplinary patterns derive, and the tendencies within popular culture and post-modern interdisciplinary studies to, as it were, measure and compare less but experiment and juxtaposition more. Increasingly in the contemporary context of tools like the Macintosh lifestyle suite iLife, we can anticipate if not fewer words, then a lot more images to be digitally authored and then consigned to data media and the bottoms of drawers for want of a means of retrieving their intended, if not perceived meanings.

Traditionally the notion of visual indexing is based upon word interpretation - on-line picture libraries keywords are associated with location, subject, colour, date etc. - the Visual Art Data Service (VADS) image resources site is an example of this tradition as are many photographic archives and stock-shot libraries. Other websites have experimented with delivering knowledge via a combination of archive film footage, slides and text using intelligent interfaces. One of the earliest was developed by a research team in the Media Lab at MIT during the mid-90s, 'Jerome B. Wiesner, 1915-1994: A Random Walk through the 20th Century' [3] This monitors the users initial selection before reorganising subsequent options to cluster related topics, using a combination of image and words and re-shuffling their relative positioning on the screen. Each thumbnail image is able to operate as an iconographic link to play the archive footage. One of the team members, Glorianna Davenport, was to say in a later article that "keyword matching is a crude and unsatisfactory method for sampling the information content of complex sources..." [4]

The focus required to make an internet session or exploration of any interactive artefact productive, requires quite prodigious use of personal memory with the ability to store and retrieve whatever it is: text, images, sounds - data. Bookmarks, electronic as with real books, are prone to disappear and anyway, once the seeker is past a dozen or so, without an indexical notation how are they retrieved? Likewise, entering a new building or navigating around a new town or city, memory is tested. We may use maps and street names at first but eventually we seek to memorise the visual coordinates - the image of the flower-shop, the park with the fountain, the shape of a skyline, the pub on the corner. Eventually we have recorded the route in memory enabling us to play the movie later in order to arrive at a pre-determined destination.

The Greeks 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. [5] 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 [previously] placed on them."

The first movies it could be claimed, were a conceptual model made by the Greek rhetoricians, complete with wide shots, tracking shots, panning, tilts, close-ups and flashbacks, all played in the cinema of the mind's eye - 'classic film narrative'. The story, the diegesis of cinema, was equally inexorable from beginning, through middle to end. In this, the age of Deleuze and Guattari's rhizome, [6] linearity need not structure thought within the confines of logic and rhetoric. In the same way as the walk from home to the station may allow interventions of the everyday to structure the day itself, even enhanced by the imprecision of the visual cues that guide us during the walk, then too the invention or re-invention of a visual literacy based on the newer technologies, would enable us (with the happenstance of chance encounter), to employ indexing and classification appropriate to the task in hand and not subject to the reductive tendencies of the inventory and the catalogue, liable not to stimulate but to stifle imagination.

Experimentation

The interactive multimedia prototype of PathScape I developed with a small team of collaborators in association with the Australian Film Commission, has an interface and navigation system which gives access to knowledge through a connection with a specific place or location or series of locations represented with sequences of still digital photographs. It seeks to enable the navigator to

associate digital documents with a (fragmented) representation of contiguous cinematic space and thereby offer a means of retrieval based on visual memory.

This is achieved by gesture, (using a mouse in this prototype), to launch Quicktime movies from a database, or view them from RAM. Though the cursor performs a 'rollover', it is a gesture that achieves the desired outcome. This distinction highlights not the completing of a task but the responsiveness and background nature of the visible interface.

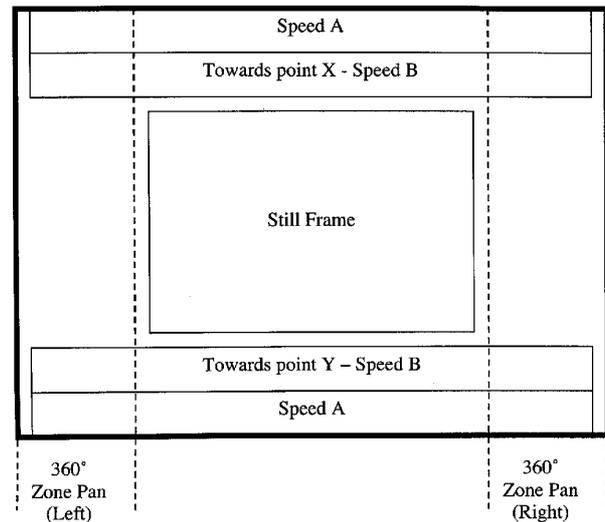


Figure 1: Screen Cursor Areas and Gesture Outcomes

A gesture with the cursor to the top of the screen launches the movie movement through the represented scene, perceived as still frames pixilated to produce apparent motion in a forward direction, through the landscape, as in a cinema POV tracking shot. By gesturing back to the centre of the screen, movement will cease. By continuing the gesture to the bottom of the screen, the image on the screen will be replaced by the view in the landscape visible 180° from the initial view - in other words 'behind' the POV of the initial image. By gesturing to top and then to bottom, the view through 180° can be instantly changed. By continuing the gesture to the bottom of the screen, apparent movement into the landscape will recommence, re-tracing as it were, the earlier steps. By gesturing further to edges of the screen, the apparent motion will speed up by a factor of two.

Thus in the prototype it becomes possible to traverse the full distance of 'the walk' through some three kilometres of bush, (X - Y), commencing at the low-water mark and ending in the rainforest, in 40 seconds at double speed (approximately 50kph 'real-time' Speed A in fig.1) and 80 seconds at the slower Speed B (25 kph). At any point the movement can be halted and a return made along 'the Path'.

The Path is ordered sequentially by two indexical devices. A border surrounds the central image that locates the user on the Path. Within this border are seen

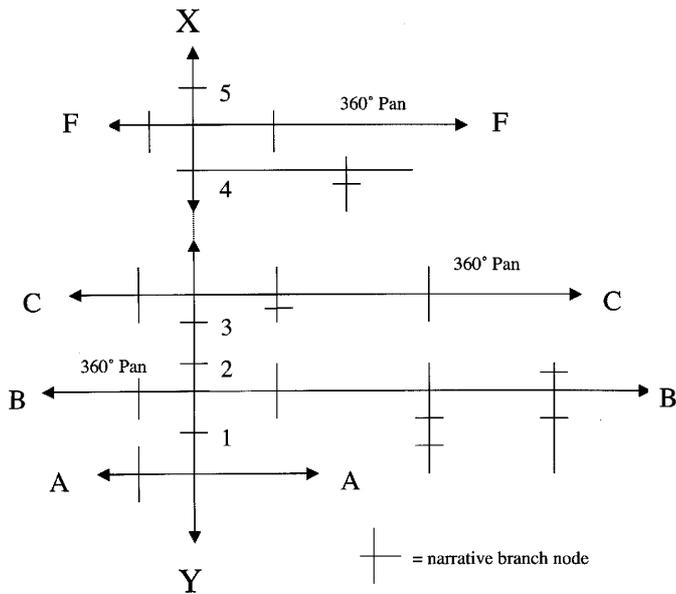


Figure 2: Schematic for accessing image/sound database

at various points along the path, fragments of images, visible for short durations. These indicate a nodal junction which, when 'captured' by halting all apparent forward movement, enable with a click the launch of a movie to replace the image and sound of the Path. Thus along an X-Y axis are the 1, 2, 3, 8, 9 etc options, or loci 'in' which are stored samples of discrete knowledge.

The second device is changes in background colour to the border and background sound, signifying changes of ecological zones. Along the X-Y axis are the A, B, C F etc axes. By gesturing to the left of the screen (or to the right) will launch a 360° panning movement, a movie representation of the ecological zone through which the user is currently 'passing' - to the right will pan right, to the left will pan left. Within the pan will be found further nodes to launch movies storing further narratives.

In the prototype these 30 movies range in duration from 20 seconds to 2 minutes, delivered in a range of presentational filmic styles exploring the possible approaches that could be taken to the notion of 'interactive documentary'. At the completion of a narrative a series of colour coded circles appears over the final frame. These are circles, abstract flat shapes, sitting in the same plane as the movie plane but in contrast to the illusionistic photo-image representation of spatially organized natural features.

Blue, yellow and green circles function as 'buttons' to linked topics colour coded to symbolically represent a broad sort under the descriptors Anecdotes, Historical Context, Commentary and Analysis, each option delivered as movies, slide shows, or audio with user

control of picture framing, extending and developing the background of what has gone before. A red circle option returns to the previous level and sequentially, back to the Path.

The soundscape, like the images, operate in the material realms of the symbolic and the representational. The encounter in this prototype enables the user to orientate within a given topography in a way not dissimilar to a walk in the country or the city and to interrogate the surroundings for hidden evidence, for concealed information delivered as stories.

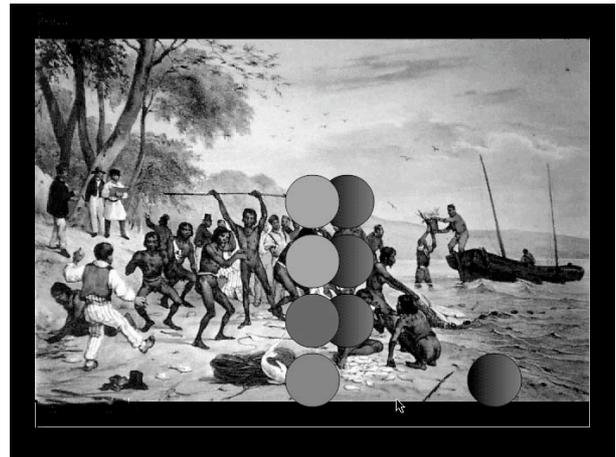


Figure 3: Screen grab within a narrative branch, with colour-coded circles.

The appeal is to the users knowledge and experience of moving through three-dimensional space, in the urban or rural setting and its remediation as an artificial topography. Encountering this range of immersive states engenders in the user a sense of a favoured space. In moving through the synthesis, the user will learn the pathway with the visual cues (loci) leading back to the combination of movie files that deliver the most pleasure or meaning. The user's cultural preference and temperament will determine this whilst (as an observer of other users) giving exposure to the cross-cultural preferences of others. There is a connection here with the Japanese concept of *ma* which connotes the complex network of relationships between people and objects. This notion of space-time continuum is distinct from the neutrality assumed of post-Renaissance space, which is personally and collectively colonized by the viewer.

The programming framework is designed to be dynamic and enable media assets, (in this prototype, movies, graphics and sound), as new discrete 'content' narratives to be added into the system and recognised each time the application engine is launched. Following demonstrations to several groups of the first prototype, another feature was added for those who had difficulty orientating and retrieving stories lodged in places represented on 'the walk'.

The grey/black circles on the screen that sit behind each of the coloured circles are the route through to the traditional text-based index, from where are launched the

encoding of the word literate technologies. The user in the prototype therefore has a choice - to navigate the index by using images and sounds, or by using words. This is not so much a cop-out from the original intention of making the project entirely image and sound-based but is regarded by us as the control factor in the second prototype, a useful resource tool for some potential users and also a useful inventory tool in managing the build of the overall content structure - writing is deeply embedded within multimedia developers too. Like many of the aspects of the interface design, the various devices and indexing systems could become options at application launch, easily switched on or off by the user.

PathScape can be delivered on disc (CD or DVD) or via the internet or broadband cable or conceivably, as it uses XML protocols, via a PDA or mobile phone. PathScape is a project progressing through several stages and adopting several iterative forms. The software framework is dynamic, rebuilding the database interface at each launch and thus can be regarded as a tool that others may find applicable for placing and retrieving multimedia assets.

Concurrence

In a recent edition of Nature Neuroscience, a study included a range of tests carried out on people who were highly ranked in the World Memory Championships. Whilst their brain capacity and structure was determined to be average it, it was found with functional magnetic resonance scanning (fMRI) that the regions associated with navigation and memory were more active than in a control group attempting the same memory tasks. The contestants confirmed they used the 'method of loci' strategy in which the objects to be remembered were placed along an imaginary pathway that could be retraced when recalling the items in order. "The longevity and success of the method of loci in particular may point to a natural human proclivity to use spatial context - and its instantiation in the right hippocampus - as one of the most effective means to learn and recall information" [7].

There exist several software tools such as ArcView related to topography, recorded time and place, widely used in the industries related to so-called environmental planning - water and land management, urban layout, national parks, mining and agriculture, etc. These are ingenious, specialised tool sets based on data derived from scientific method, measurement. Combined with GIS satellite data and a range of plug-ins that enable digital images, sound and text files to be attached to specific coordinates, this allows extensive profiles to be constructed and navigated in real-time.

The Humanities have adapted these tools - archaeologists and social scientists most notably. In

the west of Sydney, the NSW Migrant Heritage Centre has commissioned a website [8] using a specially developed Window/Linux-based application called TimeMap that links with a combination of text and map metaphors personal oral histories with localities around the City of Fairfield of Western Sydney.

My problem with such tools is the plethora of styles and codes that are incorporated, using maps, diagrams, graphical and typographic devices, each inflected with current tools and fashions in interface design. The user's encounter is like a visit to the aquarium, gazing through the glass at other peoples' lives, before moving onto the next container.

Whilst the Fairfield project takes an approach closely related to the archaeologist's inventory, making it possible to store and retrieve data about the past, but making the oral and written evidence distant and unproblematic, useful to and involving for only archaeologists. The PathScape project sets out to test the tide line between the practicalities of delirious immersion and the possibilities of indexing the cogent experience, as lived by the subject and lived by others.

In this sense it can become a tool, (related for instance to those used by amateur genealogists), for the accumulation of a library of personal audio and visual material. PathScape is also related to the notion of 'infovis', recently described by the mathematician Tamara Munzner as being "about tools that exploit the human visual system to help people explore or explain data. Interacting with a carefully designed visual representation of data can help us form mental models that let us perform specific tasks more effectively." [9]

Or indeed, in the case of John Tonkin's recent project Strange Weather, an installation consisting of a data base and a visualisation component employing various kinds of two and three-dimensional imaging to reveal patterns and underlying relationships in personal data. He describes it as "....a diagnostic tool that both elucidates and confounds." [10]

The elegant Exeter Cathedral website links treatise, catalogue and map to images using the hypertext book model: "There are two main routes into the material, Visual and Verbal. The Verbal route is for those who are more at ease with text than images." [11] There is an elegance and appropriateness in the visual part of the site that associates a contemporary on-line database design with a medieval equivalent - the vaulting and keystones in a 700 year old cathedral, the pathways and nodes storing 15th Century arcane and local knowledge using visual coding and systematic method.

Greg Ulmer's work around the term chorography is useful at this point. Based, again, on one of Plato's metaphors, the chora in the Timaeus, "...portrayed as the generative space that intervenes between being and becoming...." As Darren Tofts summarised: "Chorography is to hypermedia what the art of memory

was to the oral tradition. It sets the scene for imagining 'an electronic way of reading, writing and reasoning' that, while grounded in the concepts of memory and place, recognises that these concepts are subject to modification. The new world of immersion in information, as opposed to retrieval of information, is the frontier that chorography sets out to map." [1]

The prototype accesses representations of the natural world through a combination of gesture and iconic image fragment. In the outcome of a recent iterative version, GreenScape, the interactive encounter is in a public gallery-type space utilizing a screen and sound deployment which in conjunction with a control recognition system that goes beyond mouse-scale gesture will expand the meaning of a journey into the realm of the performance, as each user in turn is observed entering agentially the landscape.

As Nikos Papastergiadis has put it: "...I seek to grasp the sense of place that is created as art stimulates sensations and engages relations with other people." [12]

Bibliography

1. Tofts, Darren 'Memory Trade - a pre-history of cyberculture'(1997) Interface: Fine Arts Press, Sydney and G+B Arts International.
2. De Kerckhove, Derrick 'The Skin of Culture' (1995), Somerville, Toronto
3. Davenport, Glorianna et al : 'Jerome B. Wiesner, 1915-1994: A Random Walk through the 20th Century' (1994)
<http://ic.media.mit.edu/projects/JBW/> Access Jan 18 2003.
4. Davenport, Glorianna 'Indexes Are Out' Visions & Views, MIT Media Lab Fall 1996.
5. Yates, Frances A. 'The Art of Memory'(1992), Pimlico, London (1st pub 1966)
6. Deleuze, G & F Guattari 'A Thousand Plateaus: Capitalism and Schizophrenia'(1994), trans B Massumi, Minneapolis: University of Minnesota Press.
7. Maguire, Dr Eleanor, et al 'Routes to Remembering: the brains behind superior memory' (2002) in Nature Neuroscience V6 N1, (Jan 2003).
8. Archaeological Computing Laboratory (ACL), University of Sydney 'Peopling Fairfield City'(2001)
<http://www.acl.arts.usyd.edu.au/fairfield/>
Access Jan 18 2003
9. Leech, Angus 'Thinking with Vision - Tamara Munzner wants to make the world a better place with Infovis' (2002) Horizon° issue 6, Banff New Media Institute,
<http://www.horizonzero.ca/flashIndex.html>
Accessed Jan 18 2003
10. Tonkin, John 'Strange Weather' Experimental Art Foundation catalogue Adelaide 2003.
11. Henry, Avril and Hulbert, Anna 'Exeter Cathedral keystones and Carvings' (1998)
<http://www.exetercathedral.tell-com.com/> Access Jan 18 2003.
12. Papastergiadis, Nikos ' Creative Practice and Critical Thinking ' seminar paper University of Technology Sydney, 30.9.02.