

Meta-Design Approaches to Indexing Digital Media

Mike Leggett

**Faculty of Information Technology
University of Technology, Sydney
Mike.Leggett@uts.edu.au**

Abstract

The contemporary burgeoning usage of digital media – videos, audio and photographs – and media distribution through networks both electronic and physical, will be considered in the context of a convergence of these media with a contemporary and popular interest in personal and community history. I will outline some research that seeks to develop tools for storing and retrieving audio-visual digital media whilst accommodating the perceived needs of the ‘memory worker’, both amateur and professional, whether as an individual, or a closed or open group.

“Storage of artefacts is far easier than finding them again, as any dog will tell you” (Anon).

INTRODUCTION

The contemporary burgeoning usage of digital media – videos, audio and photographs – and their distribution through networks both electronic and physical, will be considered in the context of a convergence of these media with a contemporary and popular interest in personal and community history. I will outline the interdisciplinary practice-based research that seeks to develop tools for storing and retrieving audio-visual digital media whilst accommodating the perceived needs of the ‘memory worker’, both amateur and professional, whether as an individual, or a closed or open group.

CONTEXT: INTERDISCIPLINARY RESEARCH

MEMORY

Memory is a label for a diverse set of cognitive capacities by which humans and perhaps other animals retain information and reconstruct past experiences, usually for present purposes (Sutton, 2004).

Technologies, from notebooks to computers to language itself are good examples of what Andy Clark has described as:

...the pervasive tendency of human agents to actively structure their environments in ways that will reduce subsequent computational loads (Clark, 1997).

Interacting with external memory machines such as collections and libraries of knowledge located on computer servers around the globe is central to academic pursuit and increasingly, the education and edutainment of the population. The index has been central to retrieval of text-based data. Complex indexes have become subject to “classifying or arranging in classes, according to common characteristics or affinities” (OED).

DIGITAL MEDIA INDEXING

ICT manufacturers, following the indexing practices established by the industry have developed text / keyword-based multimedia file management applications (such as Extensis, Canto, etc) for aided-retrieval. Hewlett Packard Labs in Palo Alto developed a prototype application for non-expert users – Fotofile – that “blends human and automatic annotation methods.” The approach assumed that the ‘intuitive interface’ would be a text-based annotation system at worst and a thumbnail browsing system at best. In the final outcome it combined the two together with the addition of some ‘automatic’ (machine) features. A crude face recognition feature which offered users matched faces to confirm and name, used a ‘hyperbolic tree’ diagram visual device to link each face with its occurrence in other images. The paper provided no quantitative evaluation, though some compelling qualitative comments

Photography and home movies are activities that address deep human needs; the need for creative expression; the need to preserve memories, the need to build personal relationships with others. Digital photography and digital video can provide powerful and novel ways for people to express, preserve and connect. However, the new

technologies often raise new problems; the problems of multimedia organisation and retrieval... (Kuchinsky et al, 1999).

Browsing, as the means by which users match an image to memory or a perceived need, has itself been aided more recently by the work of Lim, Smith and Lu from Monash University, which like many similar systems captures data based on image shape, colour, texture and other content-based parameters – a kind of averaging of appearances. The outcome was i-Map, an interactive system for visualising and navigating a large scale image database...” that, by clustering magnified images onscreen (Content Based Image Retrieval), enabled the user to “explore areas which look more promising” before selecting an initial image for which the system would then seek further matches, before re-clustering (Lim, Smith, & Lu, 2004).

“Look more promising” is a query formulation used by many researchers in pursuit of non-specific information or data. Relational models of this kind were described by Ballard and Brown in the early 80s as turning away from representing models, to matching models from within a knowledge base. Thus proposition and inference became important aspects of interaction with the database. These approaches have become central to scientific, medical and surveillance sorting, storage and retrieval systems (D. Ballard & Brown, 1982).

A decade later Ballard used the term “personalised representations” (D. Ballard, 1991) to describe the means we use to facilitate everyday behaviour. Correctly identifying our toothbrush in a bathroom shared by the household is an example I suggest: some residents may use colour differentiation whilst others, distrustful of their colour memory, prefer placing their toothbrush in a part of the bathroom different to the others. Clark has described this as action-orientated representations “that simultaneously describe aspects of the world and prescribe possible actions, and are poised between pure control structures and passive representations of external reality” (Clark, 1997).

LOCI INTERFACES

The ancient Greek system of Ars Memoria was described by Frances Yates in the 60s, “a series of loci or places. The commonest, though not the only type of mnemonic place system was the architectural type” (Yates, 1966). Much of the work made using multimedia tools in the 90s developed this approach. Some researchers, whilst nominally connecting ‘place and knowledge’, take approaches that use the computer to link image symbols with specific narrative structures, replacing a book’s table of contents with the desktop e-book, for example the Xi-Hu landscape and culture and media model (Kiryama & Chen, 2000).

Other practice-based researchers have included:

- the visual artist, Chris Hales made Twelve of My Favourite Things, an interactive diaristic installation, accessed using a touch screen over an image composite of three Quicktime movies. Through interaction with ‘hot spots’ based on visible colour zones, movies narrating the world of some young children recorded talking about their favourite colours, places and people are accessed (Hales).
- ArcView, is related to topography, recording time and place, and is used widely in industries related to environmental planning, water and land management, urban layout, national parks, mining and agriculture, etc. These are specialised tool sets based on data derived from the scientific method of measurement using GIS satellite data combined with multimedia files. The City of Fairfield for instance, commissioned a system for relating the area’s history (Fairfield).
- the Digital Songlines project at the Australasian Centre for Interaction Design (ACID) in Queensland similarly uses the graphical representations familiar in game engines to map the GIS data relevant to ‘country’, relevant to an indigenous community and its traditional and contemporary cultural activity in the bush west of Brisbane (Leavy, 2004).
- Aspen Walk, an MIT project from the late 1970s, linked two video disc players with a computer program to enable a ‘walk’ around the Rocky Mountains

town, represented on a screen in front of the user, whereby the user/navigator could determine at each image of crossroads, to turn to the left, to right, or to carry straight on (Naimark, 1998).

- Swarm, by installation artist Alex Davies, uses a series of video projectors, camera and computer to create a data space that is constantly provisional, always in flux, your presence now absent, a previous presence now present. The space becomes charged with time as images of yourself and previous visitors are randomly stored and retrieved to be projected into the space of the gallery installation (Davies, 2003).
- Brenda Laurel moved the user into a virtual reality (VR) space with similar intent, where “Place motifs were embedded in the virtual environments as Placemarks – fragments of narrative ... We wanted to populate the environments with archetypal Critters with which human participants could merge. The narrative goal here was to give people character materials to play with” (Laurel & Strickland, 1994) (Author’s emphasis).

PRACTICE-BASED RESEARCH - PATHSCAPE

Play and exploration are central to Pathscape, a prototype system that emerged from a seed-funding project for the Australian Film Commission. The prototype has an interface and navigation system giving access to ‘narratives’ by their association with a specific place or location or series of locations. The links between movies have been set during the authoring process using Macromedia Director. The research objective for the prototype was to test the notion of ‘interactive documentary’ whereby multiple layers of information would be associated with a particular place and series of locations linked together (by a bush track through a landscape in this prototype). The authoring approach, having researched and gathered a wide variety of material for inclusion, sought to provide for the audience a cinematic experience of place, extended by an interactive process of knowledge seeking and meaning-making.

Each of the movies encountered in this interactive space employ, as part of the experiment, a range of genre approaches and / or narrative content that set out to

describe or make an association with, to tell a story about that section of the Path (in this prototype). By linking movie files from a database with the image of a place during the authoring process, we were seeking to examine the layers of meanings that could emerge and co-exist within the present, future, past of place.

The taxonomy of the database is represented with images of contiguous cinematic space. Individual photo images are pixilated to produce apparent motion – a movie – in a forward direction, perceived as a movement ‘into’ the space recorded of the bush track. The user controls this movement with gesture, (using a mouse in the prototype), to control the on-screen cursor.

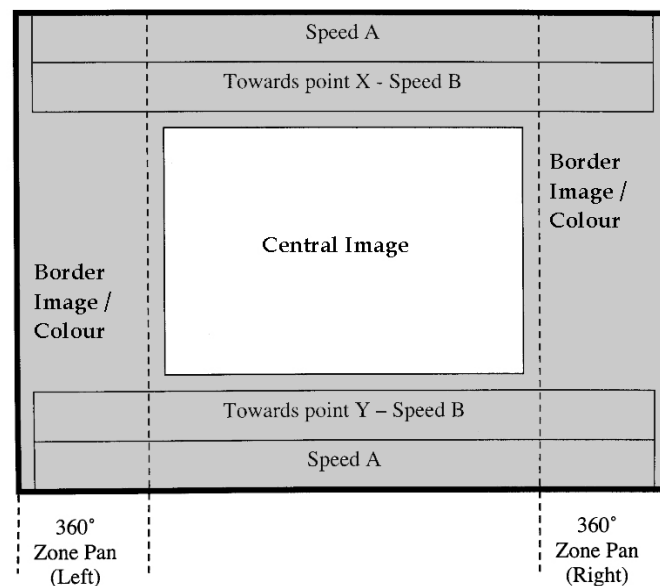


Figure 1 Screen Area Images and Cursor Gesture Outcomes

A gesture with the cursor to the top of the screen (Figure 1) launches the movie of movement through the landscape, as in a cinema Point-of-View (POV) tracking or dolly shot. By gesturing with the mouse to return the cursor 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, top or bottom, the motion 'into' the represented space will speed up by a factor of two.

Thus in the prototype it becomes possible to traverse the full distance of 'the walk' through the Bush, (X – Y in Figure 2), commencing at the low-water mark on the beach and ending in the rainforest three kilometres away. This takes about 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 taxonomy of the Path is ordered by three indexical devices. Two are located in the border area that surrounds the central image (Fig. 1). 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 a database, replacing the central image movie of movement along the bush path.

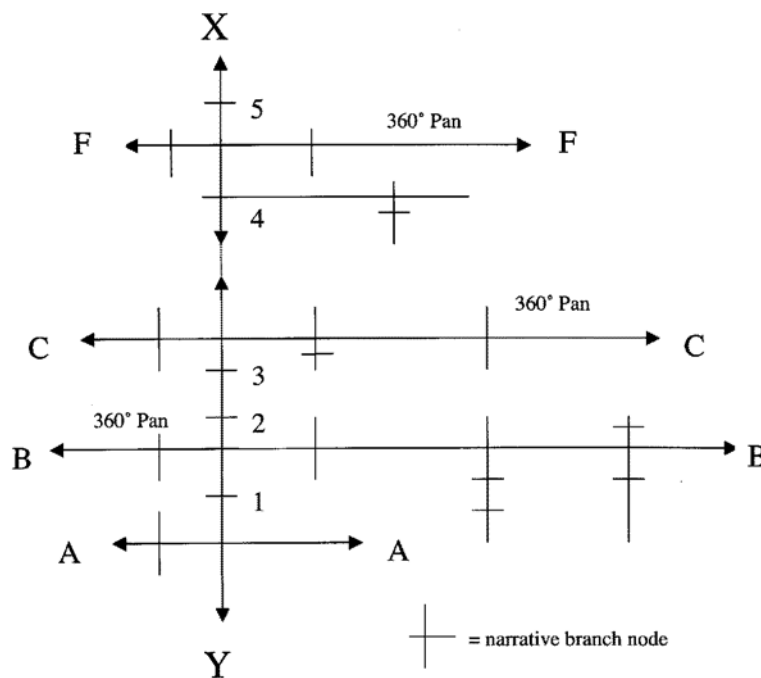


Figure 2 Schematic for accessing movie database

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

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, 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 in Fig. 2. Within the pan will be 'found' further narrative branch nodes from which to launch movies set during the authoring process, associating each movie with the visible appearance of each locale.

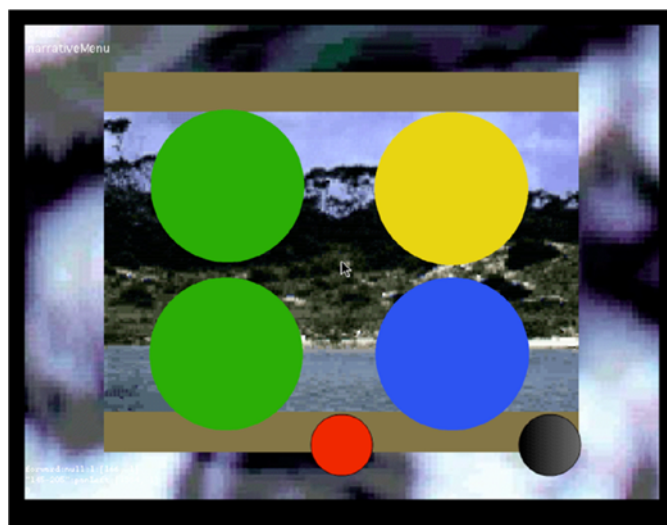


Figure 3: Screen grab: the end of a node movie, with colour-coded circles.

At the completion of a movie / narrative node, the third indexical device appears as a series of circle shapes 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 sort (in this prototype) under the descriptors: Anecdotes, Historical Context, Commentary and Analysis. Each option extends and develops the background of what has gone before, functioning as a taxonomy and narrowing the

index path to the specific, reducing from the broad. (At a final stage in the project development, following initial user responses and comment, a fourth indexical device was added whereby ‘shadow’ buttons behind the coloured circles were linked to text-based information such as: title, source, text transcription, content list, etc. This indexical convention was useful as a ‘comparative’ element, but lies outside the objectives of the current investigation.)

PROVISIONAL CONCLUSIONS

The encounter in this prototype is related to earlier and concurrent HCI research, and develops the concept of meaning derived from situated action (Suchman, 1987; Robertson, 2002). The emphasis here is not on goal orientated outcomes but on interactivity where action takes precedence and outcomes emerge. It privileges the heuristic as an essential component of knowledge building, reflection and critical perception of the process. The Pathscape 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. Interaction with the representation of the surroundings reveals hidden evidence, concealed information and comment, delivered as stories, as samples of discrete information, that enables the interacting subject to put together knowledge of the place expressed through movies of its individuals and communities. Less as query terms addressed to a database, more as embodying gestures, using the relational terms “more, same, less” within the interface. The experience constructs meaning as part of the gathering process, adding to the interacting subject’s knowledge base. Under these conditions, meaning emerges as a constantly shifting series of conclusions, the consequences of which flow on from the individual decision-making process about subsequent action. In the context of traditional cinematic experience, where reflexivity is rarely a component, the interactive experience of the audience, as explorers of a representation of that place, revealed four main areas of response:

- subjects who wholly embraced the immersive visual and navigational experience together with the knowledge building process;
- subjects who wholly embraced the experience without much concern for the documentary and informational aspects;

- subjects for whom the knowledge acquired was unacceptable and without authority or specificity;
- subjects who resisted the responsibilities of interactive engagement.

The prototype elicited a wide range of responses from users but most acknowledged the novelty of the interaction and potential 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 of linking the movies.

Accessing databases containing information expressed using discrete and specific video durations for the purposes of knowledge building is the subject of widespread research, as outlined in an earlier section. Using novel means that avoid the technology of text-based indexing requires careful and incremental movement in the design process. Accessing a database is reliant on levels of agreement between those storing data and those needing to retrieve it. The design of a 'visual' index is either determined by convention, or agreement between accessing parties, or according to other protocols, idiosyncratic or otherwise. It is not unusual for the individual to design a system which is highly productive for their needs only. It is more usual for specialised groups, (like the stone masons and cathedral builders of the 14th Century), to develop specialised and highly efficient systems, meaningful to their group but incomprehensible to others. The proximity of the authoring and user functions for situating action are key to the design process. Meta-design as an approach arising from such issues becomes of interest to the researcher where the tool needs to accommodate diverse systems of representation amongst potential users.

NEW STUDIES

The Pathscape prototype revealed the interactive principles that required further investigation, concentrating on the use of mnemonics as an indexing device. New studies will use practice-based research to investigate the development of mnemonic interfaces for storing and retrieving audio-visual digital media. These will encompass a

meta-design approach to aiding author/designer(s) to define the ontology and epistemology of personal and collective memory.

My research methodology will be practice-based, as an artist/designer of art systems. Acquiring evidence will be through data collection and data analysis. The models, artefacts, processes and events encountered during the course of exploration of the topic and the outcomes developed by this researcher will be recorded with camera, audio and notebook. The observations will be assessed through reflective analysis and recorded in the context of contemporary artefacts and systems developed by other artists and designers.

Primary Data will be gathered with a series of models and artefacts exploring the precept of a taxonomy based on visual mnemonics in digital media (video). New studies will pursue further paradigms utilising perceptive and cognitive facilities. Developing from the notion of recognising my toothbrush by its place in the bathroom rather than its colour, other space/time mnemonic representations will be investigated. For instance:

The Seasons or The Day (as time lapse) – temporal space; The Face – temporal, appearance birth to death; The Passing Crowd – ‘facial space’ and ‘the familiar face in the crowd’; The Plant – temporal, appearance germination to withering; The Performance – narrative order, gesture and incident; The Score/Schematic – a graphic space, and its sound; The Chest of Drawers containing a collection of objects associated with the media files to which they link; The Circle – clock face as a familiar indexical device; The Abstract – spatial/temporal hybrid; etc.

Towards the end of the research period, a final source of primary data will be the creation of a specific and developed artefact arising from the experiences and their analysis described above. The artefact will take the form of an installation illustrating the precept of a taxonomy based on visual mnemonics in digital media (video) and become an exhibition and/or performance.

Secondary Data will be gathered at the completion of each model through further reflection on the success or otherwise of the paradigm presented.

- Does the taxonomical device function for i) myself or, ii) others, as a method for retrieving a specific digital media file?
- Is there a limit to the size of the collection or quantity of files that can be addressed using each model?
- Or does it function better as an aleatoric memory aid?
- How might this paradigm be expanded, or related to another?
- Could there be qualities in each that might combine to form a new model?
- Is there value to maintaining a text-based component alongside the visual system?
- Can the precept be useful in developing and defining the broader topic of knowledge management, particularly in the context of a ‘design’ viewpoint?

META-DESIGN

Fischer and de Paula distinguish between two perspectives when considering knowledge management. The ‘commodity’ viewpoint is the classic corporate hierarchical set-piece that regards knowledge as an object, created by specialists, distributed to users, with access granted from the top downwards, enabling assignments to be set and with the system driving users to avoid making errors. The ‘design’ viewpoint on the other hand regards knowledge as something enacted, created by stakeholders, disseminated on-demand within a peer-to-peer network, enabling direct involvement with user-driven tasks that regard breakdowns as opportunities (de Paula & Fischer, 2003).

The design approach Fischer and de Paula describe will be familiar to those who have experienced the benefits of working collaboratively across different fields, particularly in the industrial-based creative arts, like cinema, print or television. This is common within many areas of the cool corporate sector, gathering synergies where they may - the new harvest. But the collaborative ventures familiar to many artists are less focused in their outcomes, expect little material reward, seeking primarily to enrich the culture of distinct communities. Fischer and Giaccardi describe this mechanism:

The inherently social and situated nature of knowing invites us to consider a meaningful social structure in which knowledge is enacted, created, and shared among stakeholders. Such a structure should represent the social and historical contexts in which they are capable of acting, participating, and making appropriate and informed decisions. ... Through practice, members of a socio-cultural community develop a shared understanding of what they do, how they do it, and how they are related to each other and to other communities and their practices (Fischer & Giaccardi, 2004).

In describing Communities of Practice (CoP) as stakeholders with the experience and sometimes the authority to work collaboratively, the Community of Interest (CoI) has the "...common concern or interest, to solve a particular complex design problem." Meta-design 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" (Fischer & Giaccardi, 2004).

Such a process, as proposed in this research, could assist the design of the representational system by its author(s), with different modes of taxonomy and mnemonics being suggested as part of the scenario, to provide ways of thinking about the representation of memory employing movie files. The author(s) may not necessarily be wishing this to be for exposure to an audience, (anymore than diaries are for a wider audience), so there are many questions about evidence and its remembering in this way, that may need to be approached.

Initially further development will be practice-based evaluation of different prototypes that address the following issues and conditions:

The audience – with interpretive action of the kind familiar to any user of the interactive encounter, whether computer game or network messaging:

- Taxonomy and indexing – visual and word
- Spatial orientation and navigation – presence
- Gesture and control – proximity devices

- Interaction, feedback, comment.

The author – with constructive action of the kind familiar to ‘an average’ computer user or video maker:

- Taxonomy and indexing – visual and word
- Design building and suitability of drag-and-drop media; thus awareness of sub-sets:
 - Media collection: movies, stills, sound, graphics (adaptability of ‘newsreels’)
 - Media composing: genres of address.

Database function / distributed systems – this ‘engine’ would be largely hidden, but configurable.

Generic, proprietary, prototype system for access via:

- disc or
- network.

Related inter-disciplinary research and literature:

- Mnemonic electronic devices – PDAs, mobile phones, etc
- Interactive sensor devices
- Theory and science of memory.

This is a notional structure for defining questions to be delivered to the author by the system within the context of meta-design. The central ‘problem’ for each author /designer will be situating meaning to flow from the process of ordering the author(s) audio-visual digital media?

As an expression of personal or communal culture, the resulting interactive artefact could in the words of Giaccardi, put:

... the tools rather than the object of design in the hands of the users. It creates interactive systems that do not define content and processes, but rather the conditions for the process of interaction. These objectives correspond to cultural shifts in the emerging aesthetics of interactive art (Fischer & Giaccardi, 2004).

CONCLUSION

We can anticipate a lot more images to be digitally authored and then consigned to the bottoms of drawers, for want of a means of retrieving their autobiographical or historical significance. The research project will investigate, using practice-based research, means by which a system of indexing based on representations of place can assist users in locating media representations of memory or evidence of history. Meta-design will be explored as a means of enabling users to identify an indexical system appropriate to their placing of media elements that represent the present, past or conceivably, the future.

References

- Ballard, D. (1991). Animate Vision. *Artificial Intelligence*, 48, 57-86.
- Ballard, D., & Brown, C. (1982). *Computer Vision*. New Jersey: Prentice-Hall.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, Mass: MIT Press.
- Davies, A. (2003). *Swarm*. Retrieved 1.7.04, from <http://schizophonia.com/frmindex.htm>
- de Paula, R., & Fischer, G. (2003). *Knowledge Management: why learning from the past is not enough*. In E. S. J. Davis, & A. Westerberg (Ed.), *Knowledge Management: Organizational and Technological Dimensions* (pp. 21-54). Heidelberg: Physica Verlag.
- Fairfield, C. O. *'Peopling Fairfield' website*. Retrieved 1.9.2004, from <http://acl.arts.usyd.edu.au/projects/consulting/fairfield/index.html>
- Fischer, G., & Giaccardi, E. (2004). Meta-Design: a Framework for the Future of End-user development. In H. Lieberman, F. Paterno & V. Wulf (Eds.), *End User Development*. Dordrecht: Kluwer Academic Publishers.
- Hales, C. *Portfolio*. Retrieved 1.9.04, from <http://www.smartlabcentre.com/4people/coreres/chaless.htm>

Kiriyama, T., & Chen, L. (2000). *The design of the Xi-Hu historical landscape and culture in media*. Paper presented at the Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques, New York City, New York, United States.

Kuchinsky, A., Pering, C., Creech, M. L., Freeze, D., Serra, B., & Gwizdka, J. (1999).

FotoFile: a consumer multimedia organization and retrieval system. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, Pittsburgh, Pennsylvania, United States.

Laurel, B., & Strickland, R. (1994). *Placeholder: landscape and narrative in virtual environments*. Paper presented at the Proceedings of the second ACM international conference on Multimedia, San Francisco, California, United States.

Leavy, B. (2004). *Digital Songlines*. Brisbane: Australasian Centre for Interactive Design, QUT.

Leggett, M. (2005). *Losers and Finders: Indexing Audio-visual Digital Media*. Paper presented at the Creativity & Cognition 2005, Goldsmiths College London.

Lim, S., Smith, R., & Lu, G. (2004). *i-Map: an interactive visualisation and navigation system of an image database for finding a sample image to initiate a visual query*. Melbourne: Monash University.

Naimark, M. (1998.). *Place Runs Deep: Virtuality, Place and Indigenouness*. Paper presented at the Virtual Museums Symposium, Salzburg, Austria.

Robertson, T. (2002). 'The Public Availability of Actions and Artefacts', in, *Computer Supported Cooperative Work*, vol. 11, Kluwer Academic Publishers, pp. 299-316.

Suchman, L. (1987). *Plans and situated actions: the probelm of human-machine communication*, Cambridge University Press, Cambridge, UK

Sutton, J. (2004). *Memory*. Retrieved 1.9.2004, from <http://plato.stanford.edu/archives/sum2004/entries/memory/>

Yates, F. A. (1966). *The Art of Memory*, (1992 ed) Pimlico, London.

ACKNOWLEDGEMENTS

To the reviewers of this paper, whose invaluable perceptions moved the central project up another notch of intelligibility. To Dr Shigeki Amitani and my other colleagues at the Creativity & Cognition Studios, UTS.

Biographical statement

Mike Leggett has been working across the institutions of art, education, cinema and television with media since the early-70s and has film and video work in archives and collections in Europe, Australia, North and South America. He has curated exhibitions of interactive multimedia for the Museum of Contemporary Art in Sydney, the Brisbane International Film Festival and Videotage Festival of Video Art, Hong Kong. He writes and lectures about media art, contributing to journals (Leonardo, Continuum) , is a regular correspondent for the Australian contemporary arts newspaper RealTime and is currently a PhD candidate in the Creativity & Cognition Studios, Faculty of Information Technology, University of Technology Sydney.