THE LUST TO EXPLORE SPACE:

THE ATTRACTIVENESS OF INTERACTIVE VIDEO WITHIN MULTIMEDIA APPLICATIONS

Victoria Burrill
vab@uk.ac.rl.inf

Dirk Fokken*
dirk@uk.ac.rl.inf

Gillie Evans
gle@uk.ac.rl.inf

Kaisa Väänänen
kaisa@igd.fhg.de

Informatics Department
Rutherford Appleton Laboratory
Chilton, Didcot
OXON OX11 0QX, UK

Computer Graphics Center (ZGDV)
Wilhelminenstr. 7
64283 Darmstadt
Germany

Keywords: *Multimedia/Hypermedia, Video, Audio, Human-Computer Interaction, Metaphors*

Abstract

In the computer domain, we understand the problems of interacting with text because we have been doing so for many years using word processors, hypertext and so on. The introduction of multimedia is relatively recent however; for the first time, we have the capability to really interact with other media such as video and audio. The issues associated with these new media in the computer domain are being explored, but there is still much to be learned.

This paper discusses psychological aspects of watching films in order to understand the cognitive process of how people perceive visual information and react to it. We consider the use of audio-visual media not only to communicate information *per se*, but also to communicate the way in which that information is organised and presented via user interface *metaphors*. The paper also considers more advanced methods of interacting with these dynamic media-based sequences. Of particular interest are the issues involved in interacting with the two distinct but complementary visual and/or audio components of video.

* At Rutherford Appleton Laboratory between September 1993 and March 1994.

1. Introduction

The human being is best able to perceive knowledge presented using a visual medium. We start to learn to see from the day we are born therefore visual media are very powerful communication vehicles which have the capability of being understood by almost everyone. But the understanding of films and audio-visual media has to be learned. One of the very first films to be screened publicly was *The Great Train Robbery* in 1903. This depicted a train rushing towards the audience. When this was first screened, many people in the audience paniced and ran out of the theatre thinking that they were about to be run over — they did not understand that what they saw was not real. Since then, directors have learned much more about how to use film as an exciting medium, and people have learned that film is an immersive effect, not a reality.

In the computer domain, we understand the problems of interacting with text because we have been doing so for many years. The introduction of multimedia is relatively recent however; for the first time, we have the capability of interacting with film. The issues associated with the use of this new medium in the computer domain are being explored, but there is still much to be learned. We know how to watch; film directors know how to make films. If we are to use multimedia applications successfully, we need to learn how to interact with films presented via a computer. Therefore we need to learn how to make interactive films which are suitable for use as part of multimedia applications. Such applications might include tourist information kiosks, teaching systems, interactive TV, technical information systems and so on. In order to do this we have to consider psychological aspects and understand the cognitive process of how people perceive visual information and react to it — which programs generate in the viewer the greatest lust for excitement and hence give the greatest pleasure. These issues are discussed in Section 2.

As part of our understanding as computer programmers, we need to consider the use of audio visual media not only to communicate the information *per se*, but also to communicate the way in which that information is organised and presented. A common method of doing this is the use of *metaphors*. Metaphors are useful because they allow specific complex structures to be expressed using a general, familiar term. The difficulty is that the metaphor chosen does not necessarily bear any relation to the information and information structure to which it has been likened, and yet it must enable the user to navigate and retrieve the underlying information. These issues are discussed in Section 3.

Many applications employ some form of metaphor (such as houses, libraries, index cards and guided tours) with which to organise their data; many use video and other dynamic media as part of the presentation of that data. However, very few applications use dynamic media with which to present their overlying metaphor because these media are relatively new and we do not (yet) have the technology to enable us to do so. Section 4 discusses examples of video-based metaphors.

Section 5 then considers more advanced methods of interacting with these dynamic media-based sequences. Of particular interest are the issues involved in interacting with the two distinct but complementary components of video — its visual component, its audio component — and both together.

2. The attractiveness of video

In this section we consider psychological aspects of the cognitive process of how people perceive visual information and react to it.

Are you able to walk past a TV without glancing at the screen? It is a primitive, human survival instinct to be aware of anything moving within the peripheral of our vision and hence to be ready for "fight or flight". Nowadays, life is very routine so we try to compensate for the lack of excitement of survival decisions by seeking-out other, equally powerful distractions; the more mundane our lives, the more we search for counteracting experiences **[CUB87]**. Many films and TV programs exploit this lust for excitement by presenting us with thrilling and highly entertaining programs which are preferred to more in-depth studies and documentaries by 80% of people **[BEE87]**. It could be argued that the modern equivalents of our lust for fight or flight are programs such as *Gladiators* and *The Crystal Maze*, and epitomised *in extremis* by futuristic films such as *Running Man*. (See notes.)

Film makers understand our primitive human instincts and have developed camera techniques to create thrilling shots which attract the viewer's eyes and stimulate their lust. Merely focusing a camera on the action as it happens (as in *Rope* (see notes) for example, or almost any home video!) may not be enough. Various camera techniques can be used; a zoom for example gives the feeling of observation, and draws the viewer's eye to some detail of the scene. The camera may also be used to portray different perspectives of the action; a worm's-eye view for example conveys the feeling of being small and defenceless, while a bird's-eye view conveys the feeling of superiority and being in control **[MCK91]**. So sophisticated have camera techniques become that they can influence our minds not just directly, but also indirectly on a more emotional level. Films can be paced in rhythms of highs of action and lows of inactivity in the same way that human attention is engaged or distracted. This mutual sympathy, or *Rhythmic Attraction*, between film and audience can create a very powerful environment, especially in cinemas where attention is focused solely on the dominating screen **[SAS88]**. (This effect can of course be experienced in the home, but is usually muted by the distractions of other lights in the room, ringing phones, making cups of tea and so on.)

Program and film makers understand how to attract people's attention and activate their lust for excitement during the presentation of the film. People have become spoiled by this constant flow of high-level attraction however, and quickly become intolerant to any presentation that fails to meet their expectations both during and after the presentation. If people are not attracted to the presentation because they are bored or irritated by it, or because they do not understand the content, then they start to fidget and become restless. In the cinema this may result in the audience walking out; at home it results in channel hopping in the search for something more attractive and exciting. It is equally important to maintain this feeling of viewer satisfaction <u>after</u> the presentation by giving some sense of episodic closure (even if this is immediately followed by "to be continued") **[BEE87]**. One method of sustaining people's attention and of providing the necessary satisfaction is to include some form of interactivity. Such interactivity may be indirect (such as displaying a film sequence and asking a question which requires the viewer to watch intently the sequence in order to be able to answer the question), or direct (such as video games). (It should be noted however that films and videos have not yet reached the same level of interactivity as video games although interactive films are being made **[BOD93]**; video games do not yet present the same level of realism as films because of the trade-offs between realism of graphics and speed of rendering.)

Keeping people continuously attracted and able to reach their level of satisfaction is very difficult because people bring different information backgrounds to their understanding of films and hence perceive the same presentation in different ways. These information backgrounds are dependent on many different factors including age, life experiences, mood, education, and indeed past experience in watching such presentations. There are three levels of film and TV literacy — the understanding of the language, conventions and culture of TV. The first level is simply reaction to the information in an uncritical manner; the second level is an ability to order the information presented and to comprehend effects such as flashbacks; the third level is the judgement of information where people allow themselves to be affected by the presentation but remain aware that it is not reality **[POP87]**. This type of classification may be applied not just to the audience, but also to the images themselves. Thus films can be created which speak to different people in different ways according to their past experiences **[WEI91]**. The wider the anticipated audience for a presentation, the simpler that presentation must remain so as to be understood by as many of these people as possible.

We have now examined various psychological aspects of people's understanding of conventional film and TV images. We now consider how this knowledge can be applied and exploited in the computer domain.

3. Metaphors to satisfy lust

This section defines metaphors in general and describes how they can be adapted and used in the computer user interface. Aspects of the design of a metaphor are also discussed in relation to the information the metaphor is required to represent. Finally, we consider how this knowledge of metaphors can be applied, together with our understanding of video, to video-based metaphors for a more effective and engaging user interface.

A *metaphor* is a phrase which describes one thing by use of another thing with which it can be compared. Metaphors are used in literature to describe something to the reader, comparing it to something else with which the reader is already familiar. For example, "the roses in her cheeks" where the reader obviously cannot see the girl, but the word "roses" conveys a suitable imagery **[LON90]**. Metaphors are also used in teaching to explain a new idea or concept by likening it to an old, familiar concept. For example, "the Return button on your computer keyboard does exactly the same thing as the return lever on your typewriter".

There are many different theories of metaphor **[SCH93]**. From amongst these theories, we identify four main historical theories. *Substitution metaphors* explain their contents by themselves, for example "Richard the Lionheart" **[HEL92] [BLA54]**; *comparison metaphors* are a special case of substitution, for example "Richard is like a lion"

[HEL92] [BLA54]. Both types describe the person as being brave. *Interaction metaphors* extend both their source and target to create a synergistic result **[BLA54] [BLA77]**. For example "Man is a wolf" conveys several meanings including "Man lives in packs", "Man is bloodthirsty" and "Man is subordinate to the pack leader". Finally, *structure-imaging metaphors* (*Struktur Abbildungstheorie*) are used to visualise structures between an unknown source domain and a known target domain **[HEL92] [CAR82]**. For example, "The computer file directory structure is like an inverted tree". Note that for the third and fourth types not all of the implications conveyed by the metaphor are necessarily true (man has its faults but does not actually behave like a wild animal), and that the metaphor used does not necessarily bear any relation to its target (trees have very little in common with computers). The third and fourth types are used extensively in computer user interfaces because of their wide applicability to complex domains.

Metaphors try to apply not only similarities between known and unknown objects of knowledge, but also make use of dissimilarities and omissions in metaphor relative to the new object of knowledge. Metaphors are open-ended in that they are not literal descriptions so must be incomplete, even indeterminate **[CAR85]**. Throughout our lives we use our experiences to build a framework or *mental model* of metaphors. With each new experience we re-evaluate our models and metaphors, modifying them if necessary, thus enhancing our experienced knowledge and understanding. If the new experience does not fit into our current model (the model is incomplete) we revise the metaphor, or possibly compose further metaphors to encompass the new experience. This sense of incompleteness and open-endedness initiates thought processes, stimulates interactivity and action **[CAR85]**.

Computer-based graphical applications exploit incompleteness and open-endedness to stimulate users into action. However, incompleteness and open-endedness alone do not stimulate interaction and action; it is also necessary for graphical applications to use metaphors in order to enable the user to visualise easily the underlying information and information structure. Designers of metaphors can be supported by Weidenmann's *Activation*, *Focus and Construction* and *Replacement* functions of image perception **[WEI91]**. These functions consider how people's past experiences and knowledge influence their understanding of images. This is discussed below.

The main purpose behind the use of metaphors is to impose a higher level organisation on an information structure. If we choose a metaphor that can be represented through an audio-visual medium, the underlying information structure can be visualised based upon the visual features of the metaphor. This audio-visual representation will enable the user to intuitively interact with the metaphor-based objects that act as "information containers" or access points to further information nodes.

In a hypermedia system the metaphor acts as an information organiser and visualiser of the underlying web of nodes and links. Hyperlinks may be visualised in the user interface via metaphors, and the information nodes may be "hidden" behind these objects. Furthermore, groupings of information nodes can be structured behind the metaphors and are accessed via clicks on the visual representations of those metaphors.

Perhaps the most well-known graphical metaphor for information management is the *desktop metaphor* where the user's files and applications are represented as documents on a desk which can be moved around and viewed according to the user's work at the time. A *pile* metaphor has been used to extend the desktop for casual organisation of information **[MAN92]**. Systems such as *VORTEXT* **[BUR89]** and *HyperBook* **[BEN87]** try to extend the metaphor of conventional books into their corresponding electronic versions. The *Domesday System* **[HOB86]** employs a metaphor whereby the user "walks" around an art gallery and views information by walking into pictures or through doorways.

The *ShareME* system (see notes and **[VAA93]**) uses a variety of metaphors to offer the user an intuitive method of organising and navigating in a hypermedia environment. (See figures 1 and 2.)

figure LibraryM4.ps

Figure 1: The ShareME library metaphor showing the library entrance with books on the shelves, and an open book

figure HouseM4.ps

Figure 2: The ShareME house metaphor showing the house entrance, and room in the house with information pictures on the walls

Having examined the use of metaphors in conventional information systems, we now consider their use within applications which include dynamic, time-dependent information — video and audio.

Conventional metaphors depict essentially static images (unless some form of animation is used). In contrast however, video-based metaphors depict moving images (assuming that the camera and/or objects it is filming are not stationary but move relative to each other). Thus objects can enter the image sequence, may exist for some time within that sequence, then may exit the sequence. In conventional metaphors all information can be represented simultaneously, whereas in video metaphors the information represented may have a finite lifetime over any one grouping level simply because of the inherent nature of the display medium. This finite lifetime of objects in the video sequence may affect the behaviour of the metaphors represented by that sequence. This is discussed in the next section.

Video may be represented on the screen in three dimensions: *x*, *y* and *time* (it is a flat, 2D image which moves in time). On conventional multimedia computer workstations, audio is usually represented in only one dimension: *time* (obviously audio has no visual component). The "missing" dimensions of audio can be added by use of an area on the screen in which users can indicate their selections. Some user interface issues associated with the mapping of non-visual information into an otherwise unrelated visual area are discussed in Section 5.

A video metaphor therefore is one which has a temporal component; a good video metaphor however is one which also meets our criteria of inciting lust in the viewer to explore further the information. This implies that a good video-based metaphor is highly structural, concrete, and customisable to the application requirements **[VAA94]**.

Examples of this are discussed in the next section.

4. The application of video metaphors

In this section we describe and discuss three video sequences. Although these sequences have not been designed specifically as metaphors, they embody features which are attractive and incite lust, and which present some form of metaphor to enable users to visualise information structures. The video sequences described are: *Diving Cars* (a TV advert for Nissan cars), *Ariston* (a TV advert for Ariston appliances), and *The Cube* (part of a music video for David Morales' song *Gimmie Luv* shown on ITV's *The Chart Show* in 1993). These sequences are then discussed in relation to the aspects of inciting lust and supporting metaphors presented in the previous sections. Note that we do not (necessarily) suggest that these videos be used in multimedia applications; they are described here purely to illustrate features that would seem to be important in the choice and creation of videos for metaphors.

Diving Cars

This sequence consists of two metaphor levels: two shots of three cars being dropped from a height into a swimming pool. The first shot shows, from a bird's eye perspective, a red car; the second shot shows, from ground level looking upwards, a blue car and a green car. In this second shot the fountaining water from the pool can also be seen as the cars fall into it. (See figures 3 and 4.)

figure colour.pool.half.grabbed.ps

Figure 3: The red diving car from the Nissan advert

figure colour.cars.half.ps

Figure 4: The blue and green diving cars from the Nissan advert

This video can be used to depict four access points to three information nodes at two information levels: one node (the red car) at the first level, two other nodes (the blue and green cars) at the second. The water in the second shot could also be used as an additional information node access point with which to indicate a first level access point at the same time as the second level access points. The two levels could be used to present different orderings of the primary and secondary information access points if the information structure of the controlling application requires this. (We have termed this *Information Switching*).

Ariston

The camera in this commercial is stationary; the set shows a somewhat futuristic room. On the left hand side of the room there is a staircase leading to a gallery along the back of the room. A washing machine (placed towards the back of the room on the left hand side) and a refrigerator (placed in the foreground on the right hand side of the room), are major objects with which the actors interact. A total of about a dozen actors enter the room from all sides, and mostly cross to the other side of the room before exiting. Actors also enter or leave the room by means of the gallery and staircase. During the sequence, the same actors may appear more than once, they do exactly the same actions each time, and groups of actors may appear in different combinations. Yet the actors still interact with each other by, for example, putting a champagne glass on top of the refrigerator before another actor takes the champagne away; or one actor puts a cake into the refrigerator, then another takes it out. (See figure 5.)

figure colour.ariston.half.ps

Figure 5: The Ariston advert

This overlay of actions of differing periodicity results in an endlessly shifting sequence in which actors enter the frame, do something, and then exit the frame. The fact that actors move across the frame and overlay their actions with those of other actors results in considerable complexity and confusion for the viewer. The entire back wall of the room is a mirror, which adds to the complexity. The total disorientating effect keeps the viewer’s eyes on the screen in an attempt to analyse the different action sequences. Continuity is also maintained by a monotonous voice repeating "Ariston: and on, and on...", the words being displayed as a running caption along the bottom of the frame.

This video can be used to depict many access points to many more information nodes at two information levels: the main appliances could be used to represent first level nodes, the various actors the second level nodes. Alternatively, combinations of appliances and actors can be used to represent different nodes. For example, the refrigerator and one person could represent a first level node, the refrigerator and two people a second level node, and so on.

The Cube

This is a music video incorporating various computer-generated buttons which float about the TV screen. The singers in the video "interact" with these buttons by, for example, using their finger to point to them. The buttons are superimposed on the video; thus sometimes the singers must interact with the reverse side of the buttons. (See figure 6.) The buttons take many forms, including a cursor control panel (with four direction buttons), a rotating cube (with various icons on its sides) and buttons superimposed on the three backing singers (simple push buttons).

figure colours.itv2.half.ps

Figure 6: The Cube from the music video

Of particular interest is the cube button depicted in the figure. This presents more than one access point to underlying information because as it rotates, different sides show a different icon. The cube moves across the screen in a parabolic path from bottom left to top right, and thus attracts the viewer’s attention, especially when the singer’s hand is seen moving towards it to push one of the icons.

One effect common to these sequences (and others identified during this work) is that of strangeness. Cars are not generally dropped into swimming pools, people cannot repeat precisely their actions many times over, and people do not generally push buttons from behind. Thus the sequences are very eye-catching. A second effect common to the sequences is that not all access points are presented to the viewer all of the time, but move into and out of shot. Thus the viewer is encouraged to stay alert, watching for the access point in which they are interested. The combination of these two effects — being eye-catching and exhibiting temporalness — is to incite lust and a desire in the user to interact with and explore further these strange objects.

The diving cars and Ariston videos are very short, but the music video is longer and could be used to present a more complex metaphor. This metaphor would be of a narrative form. The viewer has the choice of watching the sequence and learning the use of the buttons. Alternatively, the viewer can click on a button to explore further the story. This watching then doing is a form of inductive learning.

In the next section we consider more practical, user-oriented aspects of interacting with these objects.

5. Advanced Interaction

In the previous sections we discussed the use of video in the presentation of metaphors, and described some video sequences which encouraged the user's interaction. In this section we discuss more practical aspects of the user's interaction with video-based metaphors. We then extend our discussions into the realm of pure audio and describe how our ideas might equally well be applied to audio-based metaphors.

Throughout the preceding sections, techniques and methods have not been explicitly mentioned for the user's interaction with objects in video-based metaphors. From our experience with various computer-based applications however, we assume that most interactions will be via a mouse or possibly a touch-screen. Users will use the mouse to click on the (visual) object of interest in the video **[BUR94]**. The controlling application will then take some appropriate action such as following a hyperlink or displaying some other item of information. But the nature of video is that it usually consists of two distinct but complementary components of information — a visual component and an audio one — which can be presented separately, but are more likely to be presented together. It seems highly desirable therefore to allow the user to interact with both components (separately and/or individually), and some method must be defined to enable this distinction to be made. A simple solution might be to use the left mouse button to click on (visual) objects in the video, the right mouse button for objects in the audio. For example, clicking with the left button on actors or appliances in the Ariston commercial might result in some hyperlink to part of the information system describing specific aspects of a project; clicking with the right button might result in some hyperlink to information (audio and/or visual) about the commercial production company.

From this simple example however it is immediately apparent that there are several difficulties to overcome.

If the audio component of the video consists of two channels (background music and a voice-over for example), how can the user indicate to the underlying application the channel in which they are interested? It would be possible to use different areas of the video image to represent different channels. For example, the top 2cm strip of the image might be the access point for the background music, the bottom 2cm strip the access point for voice-overs. But such a scheme is not intuitive, and therefore it is difficult for the user to associate objects in an audio track with some arbitrary area on a computer screen.

What if the video depicts several people all talking together, out of shot, and above background music? The object’s audio component is "visible" as part of the action but its visuals are not involved so cannot intuitively be clicked on. Although this sort of situation is unlikely in most films, it has occurred *in extremis* in the film *Blue* (see notes). In these cases, we need some iconic form of the object to provide the visual access point which the video channel does not. Obviously the icons chosen should be clearly identifiable with the visual object they are replacing, and hence provide the user with an intuitive access point. In their simplest form, these icons could be simple areas in which the name or image of the object was displayed. In more complex forms the behaviour of the icons could have some significance in respect to the (implied) actions of their filmed but currently invisible counterparts. For example, the iconic form of a human character could be represented as a twig man most of the time, but animated in some way (such as waving its arms) when its associated object speaks. If the associated object exits from the left hand side of the frame and later in the scene reenters from the right hand side, the movement of the iconic form between points of exit and entry could be interpolated around the edge of the frame. Thus the mapping between visual objects and their iconic form when not visible is maintained. Although this solution has a sense of "cuteness", its visual and interactive attractiveness should not be dismissed.

The iconic form technique could be extended further and used to provide information access points throughout films such as *Blue*, or even presentations which truly do consist only of an audio component (radio broadcasts, for example). In these cases there may be a very deliberate design decision for the iconic form of objects to be very different from their actual form. A radio play involving aliens from outer space is of course presented using human actors; the iconic forms used to represent the aliens should not be humanoid, but rather should allow listeners to imagine whatever they wish.

As we have already seen, a good video metaphor is one which meets our criteria of inciting lust in the viewer to explore further the information. The objects depicted in such a video should be sufficiently eye-catching that it is obvious to the user that the objects can be clicked on. Therefore, it is unecessary for the controlling application to indicate in any way the existence of information access points by, for example, outlining them or using some form of icon.

Throughout this paper we have tacitly assumed that user’s interactions with objects will be via a mouse or touch screen. We have also assumed that applications which incorporate video-based metaphors will "display" the audio component of videos using a monophonic method. For the future however, when devices such as datagloves, holographic displays and 3D audio become more widely used, neither of these assumptions need necessarily be true. The applications developed using these techniques and devices would allow a great sense of realness, and the information structures thus visualised would incite lust in people to interact with them. The exploration of these information systems would be a very immersive, active and powerful learning experience.

Afterword

The video sequences described here, and the ideas presented on how to use them, are the preliminary results of on-going work that has yet to be completed. Such work would seem to be of great importance for hypermedia applications of the future, and provide additional understanding for other work into *Content-Based Retrieval (CBR)* from dynamic media. We have already seen that people are best able to perceive knowledge presented using a visual medium. Therefore the understanding of how people react to images will be valuable to developers of hypermedia applications and enable them to make their applications more attractive and effective. Thus people will be able to do more than merely watch and retrieve information from hypermedia applications; they will be able to interact with and experience information applications which have been designed to stimulate the user’s lust to explore further the information space.

References

[BEE87] U Beer *Bildung vom Bildschirm?* Medien zwischen Kultur und Kult Keck und Thissen (Eds.) Verlag Julius Klinkhardt, Bad Heilbrunn/Obb., 1987

[BEN87] ID Benest, G Morgan, MD Smithurst *A humanised interface to an electronic library* In Proc. Interact ’87, H-J Bullinger, B Shackel (Eds) Elsevier Science Publishers BV, 1987, pp.905-910

[BLA54] M Black *Die Metapher* in Anselm Haverkamp (hrsg.) Theorie der Metapher Wissenschaftliche Buchgesellschaft 1983

[BLA77] M Black *Mehr über die Metapher* in Anselm Haverkamp (hrsg.) Theorie der Metapher Wissenschaftliche Buchgesellschaft 1983

[BOD93] S Bode *Clocking the future* Sight and Sound, 3(8), 1993 p.10

[BUR89] VA Burrill, JA Ogden *VORTEXT: The Hard-backed Screen* Electronic Publishing — Origination, Dissemination and Design 2(3) October 1989, pp.143-156

[BUR94] VA Burrill, T Kirste, JM Weiss *Time-varying sensitive regions in dynamic multimedia objects: a pragmatic approach to content-based retrieval from video* to be published in Information and Software Technology Journal Special Issue on Multimedia, June 1994

[CAR82] JM Carroll, J Thomas *Metaphor and the cognitive representation of computing systems* IEEE Transaction on Systems, Man and Cybernetics September/October 1992

[CAR85] JM Carroll, RL Mack *Metaphor, computing systems and active learning* Int. J. Man-Machine Studies vol.22 pp.39-57 1985

[CUB87] F v Cube *Fernsehverhalten und Fernsehpädagogik aus der Sicht der Verhaltensbiologie und der Zeichentheorie* Medien zwischen Kultur und Kult Keck und

Thissen (Eds.) Verlag Julius Klinkhardt, Bad Heilbrunn/Obb., 1987

[FOK94] D Fokken *Research and development of interactive video metaphors for ShareME applications* Diplomarbeit thesis, ZGDV Darmstadt, March 1994

[HEL92] G Helm *Metaphern in der Informatik: Begriffe, Theorien und Prozesse* Arbeitspapiere der GMD Juni 1992

[HOB86] R Hobbs *The Domesday data* BCS Computer Bulletin 2(3) September 1986, pp.26-28

[LON90] *English Dictionary* Longman Group UK Ltd, new edition, 1990

[MAN92] R Mander, G Salomon, YY Wong *A 'Pile' Metaphor for Supporting Casual Organization of Information* Proc. SIGCHI'92, ACM, May 1992, pp. 627-634

[MCK91] J McKendree, J Mateer *Film Techniques Applied to the Design and Use of Interfaces* Proc. of 24th Annual Hawaii International Conference on Systems Sciences, Hawaii, 1991 pp.32-41

[POP87] K G Pöppel *Über drei Bedingungen, gut zu sehen und sehen zu lernen* Medien zwischen Kultur und Kult Keck und Thissen (Eds.) Verlag Julius Klinkhardt, Bad Heilbrunn/Obb., 1987

[SAS88] J Schmidt-Sasse *"Macht Filme wie komplexe Werbespots!"* Ansichten einer künftigen Medienwissenschaft, Berlin, Sigma Bohn 1988

[SCH93] J Schmidt *Entwurf und Implementierung multimedialer Benutzeroberflächen mit verschiedenen Interaktionsmetaphern für ShareME* Diplomarbeit, Technische Hochschule Darmstadt, Fachbereich Informatik, 1993

[VAA93] K Väänänen *Interfaces to hypermedia: Communicating the structure and interaction possibilities to the users* Computer and Graphics 17(3), 1993

[VAA94] K Väänänen, J Schmidt *User interfaces for hypermedia: How to find good metaphors* To be published in Proc. CHI94, April 1994

[WEI91] B Weidenmann *Lernen mit Bildmedien, Mit den Augen Lernen I* Beltz Verlag, Weinheim und Basel, 1991

Acknowledgements

This paper is one of the results of a research collaboration between the *MOVie* project in the Informatics Department of Rutherford Appleton Laboratory, Didcot, UK and the *ShareME* project at The Computer Graphics Center (ZGDV) Darmstadt, Germany (see notes). Specifically, it is a distillation of part of Dirk Fokken's Diplomarbeit thesis concerning psychological aspects of Content-Based Retrieval from video **[FOK94]**.

The authors would like to thank their colleagues at RAL and ZGDV for their support, comments and understanding during the writing of this paper, and thank Professors Hopgood and Encarnacao for supporting such an interesting and fruitful Fellowship exchange.

Notes

Blue: A 1993 film by Derek Jarman which consists of a sound-track of people discussing such subjects as treatment for AIDS, the situation in Sarajevo, and meditation on the colour blue. Throughout, the cinema screen is an unchanging plain blue colour, portraying, amongst other things, the director’s loss of sight as a result of his illness.

Crystal Maze: A British Channel 4 program in which teams of contestants perform a series of mental and physical challenges in order to collect crystals which finally enable them to win money. Each challenge has a set time limit. If the contestant has not exited the game area at the expiration of that limit, they are trapped. Unless a crystal is forfeited to free them, they are also out of the game (and money).

Gladiators: A British ITV program from LWT in which contestants compete in games such as tug-of-war against body-builder "Gladiators" to win points and hence the championship.

MOVie: The MOVie project (Mapping Objects on Video by interactive editing) is a stand-alone application which allows a human editor to delineate a video sequence in a cinematic context. MOVie provides multimedia application authors with a means whereby a visual area in a video image may be mapped to the contextual object which it represents, thus enabling *Content-Based Retrieval* (CBR).

Rope: A 1948 *Who dunnit* film by Sidney Bernstein and Alfred Hitchcock in which all the action takes place in one room, the camera remaining stationary throughout.

Running Man: A 1987 film directed by Paul Michael Glaser about a game show of the future in which criminals and other "enemies" of the state have to cross an area of bombed-out Los Angeles whilst battling high-tech "Stalkers" sent to kill them. The prize is that the "competitors" stay alive long enough to reach their goal.

ShareME: The ShareME system (**Share**d **M**ultimedia **E**nvironments) is an interactive authoring tool for the construction of multimedia information systems based on user interface metaphors. ShareME offers a suite of pre-designed and programmed metaphors, from which the author chooses one which suits the application’s information. The end-user then interacts visually and intuitively with the multimedia environment through this metaphor-based user interface.