Digital Creativity
2002, Vol. 13, No. 2, pp. 71–82

1462-6268/02/1302-0071$16.00

Interacting with pictures: film, narrative and interaction

Stephen Boyd Davis

Middlesex University, UK

s.boyd-davis@mdx.ac.uk

Abstract

The article explores the spatial practices of film in order to better understand the design of digital interactive media. The criteria used in the design and selection of the film image are discussed and a novel view of film as a pragmatic and highly economical form of picture-making is proposed. The difficulties are highlighted of simply transferring such imagery to an interactive context, but it is argued that the same guiding principles can be applied in the newer medium. It is suggested that the demands of visual interaction are leading to the development of new pictorial modes, but that much work remains to be done. In particular it is proposed that the mature expressivity of traditional film is a goal to which designers of pictorial interactive media should aspire.

Keywords: computer games, film making, film space, interactive media

1. Introduction

As the technical difficulties of delivering filmic imagery on interactive devices are overcome, the problems raised by combining such imagery with interaction become more obvious. Though film, whether conceived in terms of classical Hollywood or other cinema, has proved a remarkably powerful medium of expression, how helpful is it to adopt filmic practices for interactive media? By analysing film's methods we might be able to identify practices which we can borrow and reapply in the newer medium; suggestions in this regard have been made by Persson (1998), Boyd Davis and Athoussaki (1999) and Clarke and Mitchell (2000), among others. But more fundamentally, by determining the criteria which have given us film's pictorial 'language' over the past hundred years, we can increase our understanding of the guiding principles of depictive media and perhaps reapply those principles to newer forms.

Rather than proposing specifically cinematic lines of development for interactive media, the argument presented here tends to liberate rather than constrain: we cannot easily foresee those visual practices which subsequently become defining characteristics of a medium or genre and we should experiment freely, as film pioneers have done.

2. Picturing

It is assumed in what follows that there are two broad approaches to depiction which can be regarded as extremes of a range. One aspires to efface itself, to seem transparent to what is depicted; it has been conceived in terms of the

Albertian window (Elkins 1994 45–62), or as operating without a code (Barthes 1977 17). The other thrusts its pictorial nature before the user's attention: examples abound from the paintings of the Cubists to the films of Greenaway. It is assumed here that this polarity will continue in digital media and that an important goal of many digital interactive artefacts will be to seem transparent to the subject matter they present. There is an old controversy over the relationship of depiction to vision which concerns the artificiality or otherwise of perspective (for example Panofsky 1925/1991, Edgerton 1976, Ivins 1938/1975, Crary 1990), but what matters here is not the truth-status of one technique or another, as whether the viewer in a particular culture is or is not strongly aware of the mediation. Bolter and Grusin (1999) seem to regard digital media as outside history in this regard. Though they accept the dichotomy between transparent and 'remediating' media (those which work principally by referencing other media) they seem to believe that digital media will never acquire the transparency of their precursors. There seems little evidence for this ahistorical view: already computer games, virtual environments and many other forms of computer graphics aspire to suppress as far as possible any sense of their own mediation. There is surely more evidence for the view of Murray (1997 26): "If digital art reaches the same level of expressiveness as these older media [print and film], we will no longer concern ourselves with how we are receiving the information". Nevertheless, while some depictions aspire to be taken as unencoded representations and others have an overtly presentational quality, it is important not to be simplistic about this distinction, and avoid the position taken for example by Allen (1995 82 passim) and others that film spectators forget that they are watching a film. Clearly this is hardly ever the case: film viewers are well able to think two apparently incompatible things at once, experiencing both a present reality and a constructed artefact, and as Armes points out, this is a source of their aesthetic pleasure (Armes 1994 48). Wollheim (1987) and Podro (1998) have shown how this is the case for painting too, emphasising the interplay between awareness of the depiction and the depicted.

What does it take for a pictorial technology to seem transparent? It is not simply a matter of making the depiction more like the optical stimulus of an actual scene. Filmmaking, even when it aims to evoke a sense of effortless access to what is depicted, employs pictorial strategies which have little basis in natural vision. Often, the viewer mistakes demonstrable artificiality for naturalness, and it is important to ask why. Techniques which are taken to be natural now may have been perceived as artificial when they were new. This is clearly relevant to the design of any new media, because if artificial techniques can be used to construct an apparently natural sense of seeing in an existing medium, it follows that novel techniques can be employed to solve the problems particular to new technologies, and that, if they are rightly used, these new techniques too will come to seem natural: users will come to see 'through' them.

3. Depiction in film

Many writers, conceiving the spatial representations of film in terms of realism (usually without defining clearly what this might mean), puzzle over what it is that the viewpoint represents. At times Bazin (1967 46) considers realism as having an unproblematic relation to the scene: we know what scenes look like and film should look the same. The depictive strategies by which film evokes a sense of seamless vision have proved so effective that even experienced film theorists have tended to write as though the camera's viewpoint were analogous to a situated eye (Silverman 1983, Aumont 1989), an 'invisible witness' (Bordwell 1985 54). Even Currie, dismissing the various attempts to posit

the camera as an eye present in the diegetic space, resorts to considering the viewpoint as imagined vision (Currie 1995 xv).

Lothe remarks that "film displays space superbly" (Lothe 2000 52) but this is an odd claim. Shots may display space more or less adequately, but the many extra-photographic innovations in the relation between shots made over the years of film's history suggest considerable difficulties with the depiction of space. The representation of the overall space is constructed, not reflected: in Gombrich's terms when referring to painting, this is 'making' rather than 'matching' (Gombrich 1977 248). While the content within each shot generally has all the informational redundancy associated with photography, between shots it is highly pragmatic. The space of film is essentially pictorial. Unlike a virtual environment, where the viewer make look almost anywhere, the successive images are the only visual evidence the viewer has for constructing the diegetic space, even though, notoriously, the parts of the scene may never have existed together in any actual pre-pictorial space. The film-maker offers fragmentary evidence, organised with a view to affording certain assumptions and interpretations, and the film-viewer (partly on the basis of shared conventions) duly makes those interpretations.

designers of digital interactive media have a great deal to learn from [film's] informational economy and its suppression of every extra-diegetic device

4. Information and expressivity: the optimal view

How are the shots chosen which build the larger diegetic space? Despite his claims for various kinds of realism, Bazin also advances arguments which are essentially about expressivity. For the continuous deep focus shot he argues that "it brings the spectator into a relation with the image closer to that which he enjoys with reality" (1967 35) and that such shots produce "both a more active mental attitude on the part of the spectator and a more positive contribution on his part to the action in progress" (p35–6), which advances the argument from matching scenes to affording experiences for the viewer—essentially a pragmatic rationale. Incidentally Eisenstein used the same argument for the practice of montage against which Bazin was protesting (Eisenstein 1949/1977 34).

At one level, editing can be considered as simply the omission of the uninteresting, rather than as meaningful construction. To avoid wasting film when photographing an event of only intermittent interest, early cinematographers would stop the camera until something else interesting occurred (Hepworth 1897 127). Subsequently the guiding principle of film's spatial practice became narrative drama, but it has continued to be guided by a ruthless informational economy.

Generally film-makers (for example, Callaghan 1973, Harrington 1973, Englander and Petzold 1976, Reisz and Millar 1982) have—as one might expect—a more pragmatic approach to their task than film theorists. Except in relation to point-of-view shots, where the audience seems to look through the eyes of a character, the question of who or what is doing the looking hardly ever arises. The film-maker's task is conceived instead as a pragmatic one of showing, not of imitating any aspect of vision.

Digital Creativity, Vol. 13, No. 2

Narrative and drama are crucial determinants—in Bordwell's words (1985 50): "Hollywood cinema subordinates space to narrative causality". Bann notes how viewpoint and organisation in painting combine together to offer the optimum visibility of the various elements and the optimum comprehensibility of the situation (Bann 1987 88) and film is organised according to the same pictorial principles. Whereas in the space of a single picture this might or might not imply some kind of 'realism', in film considered over time it certainly does not—at least not in any straightforward sense.

Film technique is used to achieve both affective and informational expressivity at one and the same time: a camera-angle can be used to provide a certain kind of information about the action and at the same time to set up a relation between the viewer and the scene, such as looking up at a dominant character. Affect and information are intimately bound together, and the informational aspect—such as what is shown when—is as important to the viewer's engagement as any overtly affective content. The optimal view is, crudely, that shot which gives in context the greatest information, and which therefore is the shot which the viewer most wants to see. Indeed, following Carroll (1996 125–138) we could say needs rather than wants, given our compulsion to discern what others are thinking and intend to do.

For a given situation in the narrative, no single viewpoint is likely to fulfil the requirement: hence the close-up, for example. Furnham (1999 55) suggests that the rationale for the close-up of an individual character is that it is equivalent to a component of live theatre missing from film, namely stage presence (a criterion of affective expressivity), but a more prosaic explanation is that close-ups simply provide better information about the actor's expressions. Film history is instructive here: Brownlow documents the difficulties which early audiences experienced with close-up views of faces (Brownlow 1968 98) and recounts how Gance was ordered by an executive of his film company in 1913 not to use them (Brownlow 1968 524). There is a difference between a technique being unproblematically realistic and its coming to seem natural through a process of acculturation.

The information value of a shot is contextual: a shot may offer little information seen alone, but in a sequence may provide what Hochberg (1987 607) calls the answer to a visual question. In a typical classical fiction film scene, when Vivien Leigh falls down stairs in *Gone with the wind* (Fleming et al. 1939), six camera positions are used in eleven seconds, each providing the optimal view when seen in context. Quite apart from its affective qualities it can be regarded as the most informationally expressive articulation of the event and the reactions of the participants. Patently there is no position which could be adopted by an actual observer situated in the staircase scene which would yield these views. Editing is here not the omission of the irrelevant but the provision of the psychologically necessary. Again it must be noted that editing too was considered problematic at first (Musser 1991 393–4), especially when using straight cuts (Bottomore 1990 105) or where changes of scale were involved (Tsivian 1990 251).

Making the camera viewpoint behave as much as possible like an eye does not give the strongest sense of natural access to the depicted world. In a 30-minute Steadicam[1] sequence comprising all of Act Four of *La Traviata* (Griffi 2000) the opening view roams in close-up across details of the characters hands, medicine bottles and bedclothes—in conventional filmmaking this deprecated technique is nicknamed hosepiping (Callaghan 1973 70). One might have thought it would present a convincing imitation of natural vision, but though it may arguably resemble the input to vision, it notably fails to seem like looking. Later in the sequence, Violetta points to a drawer where her money is kept. The conventional approach would be to

show her pointing and then to show the drawer itself, but in continuous camerawork the view instead pans as rapidly as the film-maker dares across the room from the pointing hand to the drawer. If the next view required is again of Violetta, the camera must pan back again. In a technical sense there is a very substantial quantity of information—or at least of data—and arguably it is what a real eye would observe, yet it provides little information of value to the narrative, nor, ironically, of much value in depicting the coherence of the space. Similar difficulties arise in *The lady in the lake* (Montgomery 1946) a film which really does position the film-viewer in the diegetic space, "an experiment that failed" (Walker 2000). The protagonist through whose eyes we watch is only seen twice, once in the prologue and once when he passes in front of a mirror. Again the camera must pan where information about another part of the scene is required in order for the story to be comprehensible and engaging. Similar problems of information economy also arise in virtual environments, including those of games, discussed below.

Of course films do not consist just of optimal views. The single most important qualification to the concept is its deliberate denial—preventing viewers from seeing that which they crave. Murray documents the nineteenth-century literary antecedents of filmic technique in Brontë, Dickens and Tolstoy adopted by DW Griffith (Murray 1997 29). Griffiths' contribution, most notably in the dramatic cut-away, was to see that the inherent naturalism of photography within shots could be combined with an arbitrary, pragmatic approach to space between shots. This practice is in itself entirely dependent on the optimal view. If the viewer did not anticipate the provision of that optimal view, the drama of its denial would not work.

If it were not for the lessons of film, one might have thought that a virtual eye, offering the closest possible match of the image in the representation to the scene before the eye of a situated observer, would be ideal when users must believe that they are interactive observers of a world. This of course is the assumption that virtual reality systems are built on. Instead, we crave the optimal view. If two people are seen to meet—Lara Croft and Werner for example—we want (need) at each moment to have the optimal view of them speaking and reacting to one another. Likewise when an avatar in a shared virtual environment has a pictorial or modelled face, the user whose avatar it is needs at vital moments to see that face, not be compelled for ever to see an over-the-shoulder or through-the-eyes shot. On pointing at a distant object towards which the avatar should move, the user needs to see the avatar moving towards the object and towards the camera, as we are accustomed to see the protagonist in film. In a project with young children using *ActiveWorlds* (® Activeworlds.com Inc.) by Bailey and Moar (2000) children were enabled to apply scanned photographs of their own faces to their avatars. They wanted to be able to see not only the faces of their friends but their own as well. An ad hoc solution sufficient to satisfy the children was to apply the face to the back of the avatar's head as well as to the front so that it was always visible in the over-the-shoulder view. The depicted world had to be altered to compensate for the inadequacies of the viewing system.

5. The evolution to maturity of the spatial in film

Gance remarked of filmic innovation that "what now appears the simplest of things may once have seemed the most incredible of inventions" (Brownlow 1968 528) and the hundred year history of film is filled with examples of both sudden and subtle refinements to existing spatial practice. Limitations have been turned to strengths: the fact that the camera is not an eye allows it to do things the eye cannot do, just as the fact that a painting is not a replica of natural

vision allows many possibilities to be exploited. However, there has not been a simple incrementing of techniques: many which have been invented have subsequently been expunged. The rationale for this purging is as important to interactive media as the techniques which have survived.

Though space is treated cavalierly by the film-maker in the interests of authorial narrative, the classical fiction film aspires in general to seem like natural vision. Part of the evidence for this lies in the historical elimination from the genre of spatial practices which excessively draw attention to the fact of representation. These include the use of split screen, frontal views, orthogonal camera movements, superimposition and symmetry. For example, whereas in some genres it is acceptable to solve the problem of showing both detail and context by displaying more than one image at once, in the fiction film it is not. This is one of the principal spatial differentiations between genres. Polyptychal approaches survive, indeed flourish, in some kinds of factual television, where the agenda is a quite different one from that of fictional narrative, and of course in the standard windowed graphical user interface. But in the classical film, only temporal, not spatial, juxtaposition of separate views is generally permitted. Significantly, spatial juxtaposition is extensively used by Greenaway as part of his project to bring depiction to the fore and avoid the naturalism of the classical Hollywood tradition. Whereas Ridley Scott (another art-school trained English director of the same generation) uses his strong sense of pictorial space as a means to conventional ends, Greenaway exploits his to objectify the screen image—making the viewer media-aware—in a way which will always be of only marginal interest to most cinema-goers, precisely because it deliberately prevents the psychological immersion which is the essence of mainstream film-making. Greenaway has referred (1997 9) to the "tyranny of the camera", but it is already the case that film has liberated itself from that 'tyranny' while at the same time managing to be taken as 'natural', a remarkable double success (though not in Greenaway's terms).

6. From film to interactive media

The interactive digital medium is not a single genre, but a technology, or set of technologies, supporting multiple (albeit emergent) genres. In this sense it is like television, or like publishing, more than it is like film, since in film one genre—the classical fiction film—dominates. The spatial practices adopted by multimedia genres are diverging and will continue to do so: elsewhere I have categorised seven different spatial usages in interactive media, and no doubt more could easily be discerned (Boyd Davis and Jones 2002). For some of these, as in some genres of traditional media, the sense of largely unmediated visual access to a world will be essential.

Three themes emerge from the discussion so far. One is that the spatiality of film is pragmatic, and this offers inspiration and hope to the designer of new media. It is clear that even where the illusion of natural vision is intended, simple mimesis of vision is not necessarily (perhaps is never) the answer, since even those films which aspire to seem highly naturalistic are really informed by a deeply pragmatical approach. In the history of cinema, innovations are tried and if they 'work', they are retained. We can say that film imitates vision, but only in the sense that it aspires to evoke visual experience, real or imagined, not to imitate optics. The designer of interactive media is therefore free to experiment, untrammelled by any requirement to copy some putative model of natural vision.

Secondly, in film there is a close fit of spatial practice to objectives: this is evidence of film's maturity. What film-making attempts to convey and the effects that it attempts to exercise on the audience are well served by its

spatial forms. This contrasts with pictorial interactive media which are still struggling to find spatial forms appropriate to their objectives, partly because in some cases those objectives are unclear. However interactive media also suffer from the characteristic problem of any young medium that there is a lack of shared knowledge of a body of conventions between makers and users of each multimedia genre. We have seen how important this process of mutual education has been for film.

The third lesson from film is that the specific spatial practices of narrative film can provide only part of the visual vocabulary of interactive media, for the simple reason that the spatiality of film is driven by narrative, narrative which is authored in a way quite unlike the user's interactions with interactive media. This is a problem largely overlooked by writers such as Murray (1997) who seem to believe that the facility for the user to freely explore the diegetic space, to look anywhere and see anything they choose, can be achieved with no loss to the narrative. On the contrary, it is clear that the diegetic space is a fundamentally pictorial construct, and to enter it is in many ways to destroy the narrative. Computer games designers are still struggling with this problem, which they generally solve by splicing together as smoothly as possible two distinct modes, one comprising short chunks of fixed narrative ('cut scenes') and the other the interactive game play. Even so, most games hardly qualify as interactive narratives, for the simple reason (as Cameron pointed out in 1995) that there is insufficient difference between the action and its representation. Where the apparently transparent presentation of the fiction film is made truly transparent, narrative itself tends to collapse.

7. Constraints on the spatiality of interactive media

There is currently a very particular constraint on the spatiality of digital interactive media, arising directly from the technologies which it employs. This is that the user must be provided with visual objects to interact with by pointing, normally by directing a cursor using a mouse, trackpad or similar device. This requires the presence of an object on the screen at the time when the user wishes to make the interaction. It is thus quite unlike an object in the space of film which has earlier been seen by the viewer and which is assumed to be still present even when not on screen. Though Poole (2000 73) lumps together joysticks, joypads, mice and keyboards as 'curiously alienating devices' it is visual pointer-based interaction which dramatically constrains spatiality. Multimodal interaction, allowing the user to interact using a mode such as speech (Cassell et al 1999, Oviatt and Cohen

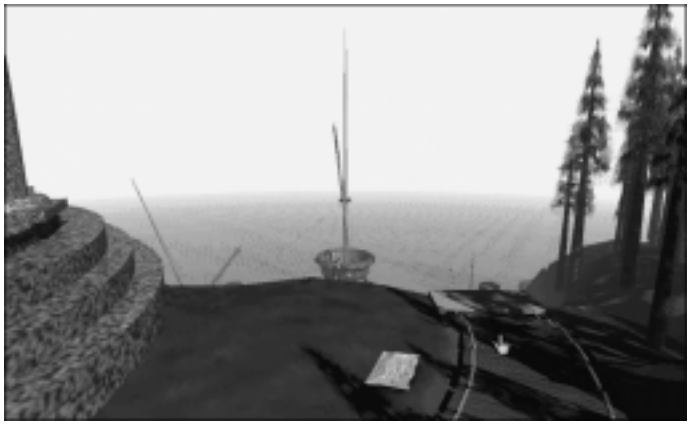

Figure 1. Myst. Digital adventure game on CD-ROM designed by Rand and Robin Miller, 1993.

Digital Creativity, Vol. 13, No. 2

2000), could radically affect this spatial characteristic since users would then be able to address elements which they could not see. Though Grasso, Ebert and Finin (1998) enumerate several differences between graphical and multimodal user interfaces, they omit this basic fact, that graphical interfaces are bound to make all available objects visible. This is a fundamental problem, especially since, as already noted, the classical film excludes the use of multi-part displays, allowing only a single view at a time to occupy the screen.

Having considered film as essentially pictorial, we can evaluate interactive media within the same framework. In some approaches every aspect of depiction is under the authors' control, while in another very little is—perhaps too little.

In *Myst* (Miller and Miller 1993) not only is every view decided by the authors, but so is the relationship between 'shots': there are severe limits on how users may alter their location. These limits are to a certain extent made to seem legitimate by the design of the environment. What compensates for the lack of freedom is that many effective pictorial devices from film are available to the designers. They can control, if necessary on a shot-by-shot basis as a cinematographer does, every aspect of focus, depth of field, framing, angle of view, and so forth. This has direct informational benefits: if the user must see a certain item (such as the discarded paper on the ground in the illustration, Figure 1), then as with film, the designer can ensure that it will be suitably placed in the view. Every option available to the user is similarly embedded in the diegesis.

Like film, *Myst* is pictorial: the user never has access to the world depicted, other than through the depictions devised by the authors. At the opposite extreme lie various virtual environments, such as *ActiveWorlds* already mentioned. In these, users can look where they wish in the space and the resulting view is an entirely automatic outcome of the virtual viewing device's processing of the model. The geometry and other attributes of this world exist independently of any particular view: a given depiction at any moment is the automatic outcome of the application of generalised viewing parameters to the world-data, so there are none of the opportunities to engineer a specific pictorial outcome in order to fulfil particular objectives which are so important in authored picture making. The pictorial characteristics are uniform no matter what is viewed and whatever the context. Whereas every aspect of a shot in film is designed to offer exactly pertinent information, not just in terms of the relationship to adjoining shots but also pictorially within the frame, the context-insensitive 'cinematography' of virtual worlds is informationally inexpressive. Work has begun on these problems, for example to enable the events in virtual environments to prompt their own appropriate cinematography (He, Cohen and Salesin 1996)—essentially a means of ensuring that the optimal view is presented at all times. Efforts are also being made by the games industry to make viewpoint dependent both on the action and on some notion of dramatic relevance, for example in the *Resident evil* series (Capcom/Eidos 1999).

Another way of offering interaction with pictorial media is to devise a hybrid display. In a limited sense *ActiveWorlds* is an example of this, if one considers the controls and other screen furniture around the main window to be part of a single picture. However such programs show a strong demarcation between the diegetic and extra-diegetic elements. Closer forms of integration are exploited by other products which in their different ways attempt to solve the problem of making visible everything with which the user may interact while at the same time not compromising too grossly the sense of natural seeing.

One approach is simply to aggregate in the display the different objects with which the user can interact. Standard graphical user interfaces adopt this approach. Generally there is little articulation of the relationship between the

Figure 2. 'Contact'. Prototype interactive documentary. Designed and produced by Dulce Maltez, Brett Bennett and Mattia Cova, 1997. Lansdown Centre for Electronic Arts.

parts. On opening a package such as *Photoshop*, a bundle of rectangular objects is displayed, but there is little or no articulation of the relationship between them: these hardly count as pictorial.

Other amalgams of pictorial segments go beyond mere accretion of parts and approach the status of pictures in their own right, pictures whose spatiality is designed in response to the special demands of interaction. In Contact (Maltez, Bennett and Cova 1997), a prototype interactive documentary comprising a continuous narrative with associated sub-narratives, the story segments are pre-authored: the only role for interaction is to choose when to experience the sub-narratives, and to advance or step back in any narrative (Figure 2). However, it is significant for this discussion in two ways. Being a documentary, the provision of multiple views in a single display is not problematic as it is for fiction. The other aspect of interest here is the form of the interaction: all visible control devices have been eliminated. While standard graphical user interfaces are often considered examples of direct manipulation (Shneiderman 1992 202–205), the form of interaction offered by *Contact* is often more direct in that it involves manipulation of the objects of interest rather than with control devices which in turn affect objects. For example, to control a digitised video a user would normally drag a slider, while in *Contact* the user drags the pointer to right or left on the image itself. At present the use of a slider and buttons is familiar, while the direct control of video by interacting with its image is not. It is easy to imagine a situation in which the reverse is true: the provision of external controls may come to seem obscure and alienating and the direct style natural and intuitive. Only familiarity, the sharing of a convention by makers and users which has been so productive in the development of film, could make this happen. The simplified presentation and direct style of interaction taken together give some of the sense of unmediated access to content which I noted in relation to film, and emphasise the need to conceive of visual and interaction design as aspects of a single set of objectives. Another project which uses simple spatial juxtaposition is *Lovebytes* (Alozie et al. 2001). As the illustration (Figure 3) shows, the screen is almost filled by the main action, but branching narratives are offered when relevant in sub-displays: they too are cinematic and entirely pictorial. By choosing them, users alter the outcome of the drama as well as the view they get of that outcome. Again, the elimination of screen-furniture, together with the intuitive simplicity of the interaction, seem to hold some promise of making even an aggregate display feel natural. What is offered to the user is another segment of the diegesis, not an extra-diegetic button or control.

Where user investigation leads to the discovery of new scenes, the naturalism of this

Figure 3. 'Lovebytes'. Prototype interactive fiction film. Designed and produced by Pierre Alozie, Sotiris Haritos, Rupen Shah and Astrid Pawlowitzki, 2001. Lansdown Centre for Electronic Arts.

experience seems to compensate for the evidently representational qualities of the objects and scenes encountered. Again there is an analogy with the acceptability of editing in film which though unlike natural vision is nevertheless accepted as in some deeper sense natural and has through custom become almost invisible to the film viewer. Whether such displays could ever be taken as 'realistic' is open to doubt, but it is certain that the close binding between the pictorial strategy and the interaction makes them seem more natural than they otherwise would.

Further interesting examples of spatial practice are offered by games. Strategy games employ a pictorial hybrid where visual objects in the periphery support functions which the scrolling world itself cannot. These extra-diegetic components are anchored to the frame of the view, not to the world. In some sense this is no more than is offered by a word-processor with its fixed tools and menus around scrolling documents, or by the virtual worlds browsers already described. However, in games there is an attempt to disguise the controls as part of the diegesis. In *Railroad tycoon II* (PopTop Software 1998) a map-like aerial view of the terrain accompanies the main axonometric '3-D' scene (Figure 4). In other words, the designers have chosen to offer two optimal views on screen at once, a solution already noted as taboo in classical film. Not only do the buttons have modelled shading to impart some sense of concreteness but they are attached to the surround using structures reminiscent of Victorian engineering in an attempt to justify their intrusion into the overall display. When a dialog-box is displayed, the extraneous pressure gauges and dials operate and a valve emits steam! Perhaps most interestingly in pictorial terms, the

Figure 4. Railroad Tycoon II. Digital strategy game on CD-ROM. Created by PopTop Software, 1998. Published by Gathering of Developers, Inc. ©Gathering of Developers, Inc 1998. Reproduced with permission.

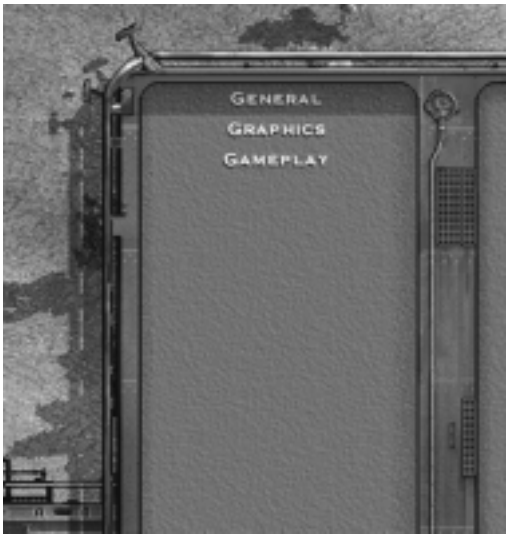

Figure 5. 'Railroad tycoon II' (detail). Digital strategy game on CD-ROM. Created by PopTop Software, 1998. Published by Gathering of Developers Inc. ©Gathering of Developers Inc 1998. Reproduced with permission.

diegetic shadows of the trees in the landscape are indistinguishable from the shadows of the extra-diegetic controls (Figure 5).

Such hybrid spaces are reminiscent of early Renaissance pictures, and for good reason. A work such as Crivelli's *Annunciation with Saint Emidius* (National Gallery, London) can itself be seen as a solution to competing demands: there is a coherent three-dimensional space in which the action takes place, but iconic objects—a marrow(?) and an apple—are placed at the front of the space. They are present in a symbolic role, but to our post-Renaissance eyes sit strangely in the context of an otherwise naturalistic setting. This is picture-making at the awkward cusp of two traditions, and could be said to show similar qualities of immaturity to those of many current computer displays.

If pictorial strategies are best conceived as pragmatic, based on the objectives of the artefact, in many ways games are a perfect illustration: at every turn it is clear that the designers are concerned with what seems to 'work' and that the various realisms and spatial principles are selected and balanced to that end. At present such artefacts seem to lack the maturity of expressiveness that I attributed to narrative film. However, it is also possible that the impression of incongruity is itself a temporary phenomenon arising from the relative unfamiliarity of these spatial configurations and that in future they will come to be seen as transparent and natural. Certainly these hybrid spaces do not have the failing that their purpose is unclear: they support rather well the functions which it seems they should. In that sense their expressivity is high.

8. The future spaces of digital interactive media

The pragmatic view of depiction suggests that any mode of picturing which satisfies users' information needs (as defined above) can come to seem natural.

Three developments seem of particular importance for the future. One is the development of context-sensitive 'cinematography' in which the games industry seems likely to lead the way. Here there is a great deal of work to be done not just in devising ways to automatically generate appropriate viewpoints on the action but also of controlling the pictorial parameters of each view, and, even more difficult, enabling intervention by some overall dramatic and narrative 'intelligence' which would select viewpoints and generate pictorial images based on an overarching narrative intention even while the user is free to navigate and operate in the world.

A second decisive influence on future spatiality will be the success or failure of multimodal interaction. To be able to address elements, including characters, in an interactive mode which does not require them to be on screen at the time would radically alter the spatiality of the medium. This could be combined with intelligent automated cinematography so that user interventions are responded to in a way which comes to seem natural and intuitive.

In the short term the most promising line of inquiry seems to be to continue to

develop hybrid pictorial modes in which informational economy and relevance are the guiding principles. Even when a film is opulent or poetic, no spatial or pictorial device in the fiction film is used gratuitously: designers of digital interactive media have a great deal to learn from its informational economy and its suppression of every extra-diegetic device.

However, there is no 'short cut' to maturity of spatial expressivity, since this maturity lies in the relationship between representations and those who use them, not in the representation alone. The pragmatic, almost accidental, approach by which spatial innovation is achieved means that no precise predictions can usefully be made. The fact that many innovations in film space were makeshift inventions which have subsequently been adopted into mainstream practice—and that the success of so many of film's spatial techniques could not possibly have been predicted—gives pause for thought in relation to new media. It emphasises the need to rethink inherited pictorial practices to suit the demands made on the artefact. It suggests that substantial open-ended experiment is called for and that the spatialities of new media may be quite other than we currently imagine them. This includes adapting concepts of visual naturalism to suit the technology and its uses.

Note

[1] Steadicam is the tradename of a device which helps to stabilise hand-held cameras. The *Traviata* scene was claimed by Channel Four Television as the longest unedited sequence to date.

References

Allen, R. (1995) *Projecting illusion: film spectatorship and the impression of reality*. Cambridge University Press, UK.

Armes, R. (1994) *Action and image: dramatic structure in cinema*. Manchester University Press, UK.

Aumont, J. (1989) The point of view (trans. Denner, A.). *Quarterly Review of Film and Video* 11 1–22.

Bailey, F. and Moar, M. (2000) Children's creation of shared 3D worlds. *Proceedings of Digital content creation conference.* Bradford, April 2000 [no page numbers].

Bann, S. (1987) Art. In Cohn-Sherbok, D. and Irwin, M. (eds.) *Exploring reality*. Allen and Unwin, Boston, pp. 83–108.

Barthes, R. (1977) *Image—music—text* (trans. Heath, S.). Fontana, London.

Bazin, A. (1967) *What is cinema?* Volume 1 (trans. Gray, H.) (Originally published in Editions du Cerf, Paris: Qu'est-ce que le Cinéma? in 4 vols. 1958–1965.) University of California Press, Berkeley.

Bolter, D. J. and Grusin, R. (1999) *Remediation: understanding new media*. MIT Press, Cambridge MA.

Bordwell, D. (1985) Space in the classical film. In Bordwell, D., Staiger, J. and Thompson, K. (eds.) *The classical Hollywood cinema*. Routledge, London, pp. 50–59.

Bottomore, S. (1990) Shots in the dark: the real history of film editing. In Elsaesser, T. and Barker, A. (ed.) *Early cinema: space, frame, narrative*. BFI Publishing, London, pp. 104–113.

Boyd Davis, S. and Athoussaki, H. (1999) VRML: a designer's view. In Vince, J. and Earnshaw, R. (eds.) *Virtual worlds on the Internet*. Proceedings of Virtual environments conference, Bradford, 15–16 April 1997, IEEE Computer Society, pp. 35–51.

Boyd Davis, S. and Jones, H. (2002) Screen space: depiction and the space of interactive media. In Jorge, J.A., Correia, N.M., Jones, H. and Kamegai, M. B. (eds.) *Multimedia 2001*. Proceedings of Eurographics workshop on multimedia, UMIST, Manchester, 8–9 September, 2001. Springer, Vienna.

Brownlow, K. (1968) *The parade's gone by*. University of California Press, Berkeley.

Callaghan, B. (1973) *The Thames and Hudson manual of film-making*. Thames and Hudson, London.

Cameron, A. (1995) Dissimulations: illusions of interactivity. *Millenium Film Journal* 28 (Spring 1995) online: http://mfj-online.org/

Carroll, N. (1996) *Theorising the moving image*. CUP, Cambridge, UK.

Clarke, A. and Mitchell, G. (2000) Screen play: film and the future of interactive entertainment. *Proceedings of Digital content creation conference*. Bradford, April 2000 [no page numbers].

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H. and Yan, H. (1999) Embodiment in conversational interfaces: Rea. *Proceeding of the CHI 99 conference on Human factors in computing systems: the CHI is the limit*. May 15–20, 1999, Pittsburgh, PA USA. ACM, New York, pp. 520–527.

Crary, J. (1990) *Techniques of the observer: on vision and modernity in the nineteenth century*. MIT Press, Cambridge MA.

Currie, G. (1995) *Image and mind: film, philosophy and cognitive science*. Cambridge University Press, Cambridge UK.

Elkins, J. (1994) *The poetics of perspective*. Cornell University Press, Ithaca NY.

Edgerton, S. Y. Jr. (1976) *The Renaissance rediscovery of linear perspective*. Icon/Harper and Row, New York.

Englander, A. A. and Petzold, P. (1976) *Filming for television*. Focal Press, London.

Eisenstein, S. (1949/1977) *Film form: essays in film theory*. (Ed. and trans. Leyda, J.). Harcourt Brace, Orlando, Florida.

Furnham, D. (1999) *Documentary practice*. Unpublished PhD thesis. Middlesex University, UK, February 1999.

Gombrich, E. H. (1977) *Art and illusion. a study in the psychology of pictorial representation*. 5th edn. Phaidon, London.

Grasso, M. A., Ebert, David, S. and Finin, T. W. (1998) The integrality of speech in multimodal interfaces. *ACM Transactions on Computer-Human Interaction* 5(4) 303–325.

Greenaway, P. (1997) *Flying over water/Volar damunt l'aigua*. Fundació Joan Miró, Spain / Merrell Holberton, London.

Harrington, J. (1973) *The rhetoric of film*. Holt, Rinehart and Winston, New York.

He, L-w.,Cohen, M. F. and Salesin, D. H. (1996) The virtual cinematographer: a paradigm for automatic real-time camera control and directing' *Proceedings of the 23rd International conference on computer graphics and interactive techniques*. August 4–9, 1996, New Orleans, LA, USA, pp. 217–224.

Hepworth, C. (1897) *Animated photography: the ABC of the cinematograph*. The Amateur Photographer's Library. no. 14. 1890.

Hochberg, J. (1987) Perception of motion pictures. In Gregory, R. L. (ed.) *The Oxford companion to the mind*. Oxford University Press, Oxford, pp. 604–8.

Ivins, W. M. (1938/1975) *On the rationalization of sight* (originally published by Metropolitan Museum of Art 1938). Da Capo Press, New York.

Lothe, J. (2000) *Narrative in fiction and film*. Oxford University Press, Oxford.

Musser, C. (1991) *Before the nickelodeon—Edwin S Porter and the Edison Manufacturing Company*. University of California Press, Berkeley CA.

Murray, J. H. (1997) *Hamlet on the holodeck: the future of narrative in cyberspace*. The Free Press/ Simon and Schuster, New York.

Oviatt, S. and Cohen, P. (2000) Perceptual user interfaces: multimodal interfaces that process what comes naturally. *Communications of the ACM* 43(3) 45–53.

Panofsky, E. (1925/1991) *Perspective as symbolic form*. Zone Books, New York.

Persson, P. (1998) A comparative study of digital and cinematic space with special focus on navigational issues. *Proceedings of Ninth European conference on cognitive ergonomics*. University of Limerick, Ireland, August 1998, pp. 67–72.

Podro, M. (1998) *Depiction*. Yale University, New Haven, CO.

Poole, S. (2000) *Trigger happy: the inner life of videogames*. Fourth Estate, London.

Reisz, K. and Millar, G. (1982) *The technique of film editing*. 2nd edn. Focal Press, London.

Shneiderman, B. (1992) *Designing the user interface—strategies for effective human-computer interaction*. 2nd edn. Addison Wesley Reading, MA..

Silverman, K. (1983) *The subject of semiotics*. Oxford University Press, New York.

Tsivian, Y. (1990) Some historical footnotes to the Kuleshov experiment. In Elsaesser, T. and Barker, A. (ed.) *Early cinema: space, frame, narrative*. BFI Publishing, London, pp247–255.

Walker, J. (ed.) (2000) *Halliwell's film and video guide 2000*. 15th edn. HarperCollins Publishers.

Wollheim, R. (1987) *Painting as an art: the AW Mellon lectures in the Fine Arts 1984*. Princeton University Press, Princeton NJ.

Digital and film works

Capcom/Eidos (1999) *Resident evil 3: nemesis*. Digital game for Sony PlayStation. Created by Capcom Co. Ltd. Published by Eidos Interactive Limited, UK.

Alozie, P., Haritos, S., Shah, R. and Pawlowitzki, A. (2001) *Lovebytes*. Prototype interactive fiction film. Unpublished postgraduate project at Lansdown Centre for Electronic Arts at Middlesex University, UK.

Fleming, V. with Cukor, G. and Wood, S. (1939) *Gone with the wind*. Selznick International / Metro Goldwyn Meyer, US, 220 minutes Technicolor.

Griffi, G. P. (2000) *La Traviata* (cinematography by Vittorio Storario; Steadicam operator Garrett Brown). RADA Film in association with Channel 4, RAI and TV France, broadcast on Channel 4 Television, 2 and 3 June 2000.

Maltez, D., Bennett, B. and Cova, M. (1997) *Contact*. Prototype interactive documentary. Unpublished MA Project at Lansdown Centre for Electronic Arts at Middlesex University, UK.

Montgomery, R. (1946) *The lady in the lake*. MGM, US, 103 minutes, b/w.

Miller, Rand and Robin (1993) *Myst*. CD-ROM produced by Cyan. Published by Broderbund. © Cyan, Inc. Myst ® Cyan, Inc.

PopTop Software (1998) *Railroad tycoon II*. Digital strategy game on CD-ROM. Created by PopTop Software. Published by Gathering of Developers, Inc.