Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Minimal Graphics

I. Herman, D.J. Duke

Information Systems (INS)

INS-R9803 March 31, 1999

Report INS-R9803
ISSN 1386-3681

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Minimal Graphics

I. Herman

Centre for Mathematics and Computer Sciences (CWI)
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
email: I.Herman@cwi.nl

D.J. Duke

Department of Computer Science, University of York
Heslington, York, YO1 5DD, United Kingdom
email: duke@cs.york.ac.uk

ABSTRACT

The problem of producing a photorealistic rendering of a graphical model continues to be the focus of considerable research effort in the computer graphics community. However, photorealism is not the only possible criteria for judging the value of an image. In this paper we step back from the physically–based model that underlies many of the existing approaches to rendering, and instead consider the rendering problem from a more fundamental view: how is graphical information processed by the user? Using differences in artistic traditions as our initial motivation, we identify the need for an approach to rendering that is based fundamentally on cognitive theory. Existing work on non–photorealistic rendering has started to take steps that address this need, but using a model of cognitive information processing we identify a significant research problem: the quest for a minimal rendering process.

1991 Computing Reviews Classification System: H.1.2, H.5.0, I.3.3, I.3.6, I.3.7, I.3.m, J.4, J.5
Keywords and Phrases: computer graphics, non–photorealistic rendering, cognition
Note: At CWI, the work has been carried out under the INS3.1 project, “Information Visualisation”. The online version of this paper contains the reproductions in colour.

1. Introduction

This is a typical programmatic paper: it probably raises more problems than it solves. It describes the authors’ long term research vision in an area which, in their view, should gain a lot of importance in future. The direct inspiration for this line of work came when one of the authors (IH) visited an exhibition on Japanese prints in Amsterdam but, indirectly, the visit the authors made together some years ago in the old imperial city of Kyoto had a great influence, too; the reader will soon understand why.

Consider the paintings in Figure 1 and Figure 2 on the next page. The first is a reproduction of a painting of Johannes Vermeer, one of the outstanding Dutch painters of the 17^{th} century. His “The Little Street” is a typical example of Dutch, but also of classical European painting. The second is a detail taken from a landscape painting by the Japanese artist Senzui Byobu, dated to the 11^{th} century*. While the paintings both convey a similar subject matter (a building within its context), the artistic techniques that they use are quite different, as will be discussed shortly. If we consider these paintings as forms of communication, then a natural concern is how the different styles of representation affect that communication. It is not the purpose of this paper to explore such issues, which properly lie in the area of art theory and visual perception. Rather, we make a simple observation: computer graphics, which today plays a major role in visual communication within the information society, has concentrated much on one particular approach to rendering information, so called photorealistic rendering.

* The painting and the original JPEG reproduction are the property of the Kyoto National Museum, Japan. The web site of the Museum (http://www.kyohaku.go.jp/) contains other examples of traditional Chinese and Japanese paintings.

Vermeer tried to represent *reality* on the canvas, with all the intricate effects of lights, of shadows, of reflections, etc. Such minute details as the texture of the brick walls or the garment of people are also represented with great care, although they are hardly noticeable to the naked eye*. This attempt for realism has been one of the main characteristics of European painting up to the beginning of the 20^{th} century. Some artists, like Dürer or Leonardo da Vinci, and indeed Vermeer himself, too, conducted life experiments to understand the propagation of light, human vision, the nature of shadow, etc. In doing so, they became precursors of an early form of experimental mathematics; for example, modern projective geometry, or the rule of perspective mappings, grew out of these experimentation.

Figure 1. Vermeer, "The Little Street"

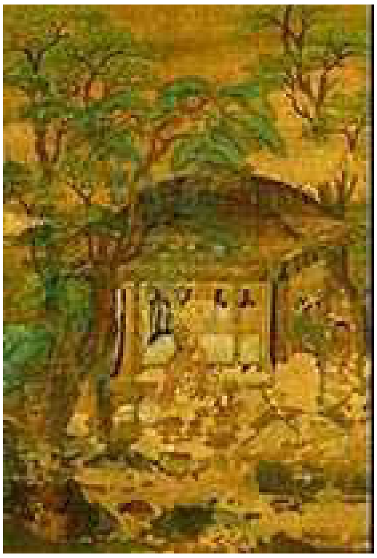

Figure 2. Senzui Byobu, "Lanscape with Figures" (detail)

This European approach to art is in sharp contrast with the art of China and Japan. The contrast between the two paintings is striking. Clearly, Senzui Byobu, as most traditional Chinese and Japanese painters, did not try to *reproduce* nature. He did not know about the mathematics of perspective views. The picture conveys an *impression* of the landscape; only parts of the contours, of the main lines of objects (of the hills, the trees, etc.) are represented. The whole of the picture is remarkably void of details. Nevertheless, the "message", the "information content" is there, and the undeniable aesthetic beauty of this painting is just as appealing as the one of Vermeer's.

Why is this interesting for information scientists? Traditional computer graphics, as it developed in the past 15 to 20 years, may be considered as a direct continuation of traditional European painting, at least up until the end of the 19^{th} century: the goal is to reproduce nature through images generated by computer graphics (making use, by the way, of the different projections originally developed by some of those artists!). The ideal is "photorealism", or its generalisation into concepts of "virtual reality" or "virtual humans". It is not the goal of this paper to criticise these lines of research, which are stimulating, exciting, and full of extremely difficult and challenging research problems. However, one should not forget an essential issue. A significant goal of computer graphics is *to help the human observer to understand information through pictorial means*, as part of human–computer interaction. In some cases (e.g., a virtual walk–through of a building) photorealism has a clear role, but one should realise that this is not always necessarily the case. The example of Chinese/Japanese painting shows that conveying information about one's environment can also be achieved *without* a strive for photorealism, judiciously choosing instead a level of graphical information which is enough to communicate the intended message. In addition, this can be done without losing the expressiveness and the aesthetic beauty of the image.

* Obviously, the JPEG reproduction of the picture does not give back all the details well. The original painting is the property of the Rijksmuseum, in Amsterdam, The Netherlands; high quality printed catalogues for Vermeer's art are also available, if the reader wishes to see all these details.

Figure 3. Jane Avril (Henri de Toulouse–Lautrec, 1896)

Underlying Chinese and Japanese art is an aesthetic of visual simplicity. The objective of this paper is to set out an alternative model for rendering that incorporates this aesthetic. For a lack of a better name, the term “minimal graphics” will be used throughout this paper for the following research goal: based on some model of information (which may be either a traditional geometric model of a full scene or something different) one should *produce images which strive for a minimum level of complexity for a task*, which should be as simple as possible, but which should convey the intended amount of information to a human observer. Furthermore, (although this is even more difficult to describe in algorithmic terms) the generated images should be “pleasing” to the human eye, should be therefore readily accepted by humans as a means of communication. The contrast between the European and the Far Eastern schools in painting is a perfect illustration of the differences between photorealistic computer graphics and this new approach.

One could also characterise the goal of the research in minimal graphics in more “artistic”, albeit much less precise form: is it possible to reproduce the artistic style of Far Eastern painting on a computer?*

2. Motivations

Although research could be motivated by a sheer intellectual challenge or aesthetic requirements, computer graphics has always been driven by practical needs, too. Hence the question: why develop minimal graphics at all? Why is this of any practical interest?

Schuman et al.[23] have made an interesting assessment on the usage of sketchy figures in CAD systems. They show that architects, when talking to their clients in the early phase of a development project, prefer to use sketches rather than photorealistic images. Sketches have an affective quality that encourages interaction, as they convey a sense of only partial commitment to a design. In contrast, a photorealistic image suggests immutability. This example is, we believe, an illustration of a general principle — that minimal images may be better suited for interaction than their photorealistic counterparts.

Another important area, where minimal graphics may become useful, is applications with new interaction methods. It is now widely recognised that the current methods of human–computer interaction will have to undergo radical changes in the coming years through the introduction of new kind of input and output devices; such changes are

* Note that, in this text, the example of Far Eastern painting is used as a contrast; one could also have referred to some schools of modern European art (see Figure 3), or to cartoon and caricature drawings. Beyond issues of personal taste, a reason why Chinese/Japanese paintings might be appropriate to direct the developers’ thoughts is their ancient traditions, accumulated throughout the centuries, the experience and philosophical background of this art, which may help in developing new methods.

necessary to achieve a greater acceptance of computing by society. A typical example is provided by the haptic devices. Although these devices (e.g., the Phantom haptic device) are still expensive and clumsy, it is only a matter of a few years when they will become easily accessible. Computer users may then "feel" the contours of objects on their fingertips, so to say, which will be of an enormous advantage for, e.g., visually impaired users. With its much–reduced level of complexity, variants of minimal graphics might be more adapted to rendering haptic information than algorithms derived from photorealistic approaches. As we will discuss shortly, the foundations of minimal graphics in cognitive theories mean that it can be more readily adapted to human needs.

A further example of a practical problem that may justify a minimal approach to graphics is the challenges raised by the wide–spread usage of the Internet. An image, generated through a photorealistic image generation process, is typically very complex, with a high probability that adjacent pixel values will be different. Such images do not compress well, because practically all image compression make use of image coherence. Images generated through minimal graphics, which do not necessarily reflect physical laws, may have a much higher coherence, which means that they will compress much more efficiently. Similar challenges are created by the usage of PDA–s, of various devices for ubiquitous computing, etc.

Finally, it has to be emphasised that minimal graphics should not be regarded as a "competitor" to photorealistic rendering. On the contrary, the results anticipated in this paper will contribute to the line of research which has recently become known as perceptually based rendering, e.g., [16].

3. Non–photorealistic rendering: is it the same?

One of the interesting development in computer graphics in the last years is non–photorealistic rendering (NPR). The concerns described in the previous sections are very similar to those which led to the development to this field. These techniques imitate non–photographic techniques, such as painting or pen–and–ink, to create images and illustrations. The various methods may differ greatly in visual appearance, and they usually rely on some artistic technique or style. An underlying assumption in NPR is that artistic techniques developed by human artists have intrinsic merit based on the evolutionary nature of art; in this sense, the goals sound very similar to those we have developed for minimal graphics so far.

Most of the work in NPR are of a rather "post–processing" nature, insofar as they use techniques to modify images (whether scanned or synthetically generated) to achieve painting–like effects[5,18,19,21] (see also the overview of Lansdown and Schofield[17]). Winkenbach and Salesin[28] also describe a modified graphics output pipeline to produce pen–and–ink as well as a number of general principles (drawn from the literature, e.g., [9]) on the usage of different brushes and strokes. Hsu et al.[11,12], or Strothotte et al.[25] also give an overview on the different stroke techniques which exist in the literature. Some of these techniques rely on physical models (e.g., how ink is absorbed by paper), whereas others represent more heuristic approaches. An interesting alternative is the imitation of watercolour strokes, as described in Curtis et al.[5].

Results in NPR have been significant in the past years, leading to impressive results. However, the common characteristics of virtually all NPR systems is that the final effects are achieved with a very strong participation of the end–user and that they are based on a traditional modeller to extract the image to be rendered. The main question which still remains is *how to automatically extract the minimal amount of information necessary for a particular task*? To quote one of the papers on pen–and–ink drawing[28] on what they call "indication problem":

> Indication is one of the most notoriously difficult techniques for the pen–and–ink student to master. It requires putting just enough detail in just the right places, and also fading the detail out into the unornamented parts of the surface in a subtle and unobtrusive way. Clearly, a purely automated method for artistically placing indication is a challenging research project.

Using this terminology, the goal of "minimal graphics" may also be described as to develop an *automatic indication technique*, which has to be combined with NPR. In this sense, minimal graphics is the logical continuation of, and complementary to non–photorealistic rendering.

4. Fundamentals of Minimal Graphics

The extraction of a minimalist information from a model seems to be, at first glance, some sort of geometric task. One would try to extract and use, for example, geodesic or other characteristic curves (see, for example, the ink drawings of Elber[8], or even the early example of Sasada[22]), use some sort of silhouette detection algorithm, or special forms of dithering. However, these approaches do not "simplify" the images, and the result may lack the "symbolic", abstract nature that minimal graphics is seeking to achieve. Other techniques could complement this approach, for example the use of some sort of smooth (not necessary convex) hull of the 3D objects, which could be used for the final image. Wavelet–like encoding could be envisaged; multiresolution methods in modelling objects might help in extracting the "sweep" of a curve or a surface[24]. In the longer term, a minimal approach to graphics will require a review of the scene modelling techniques in use.

However, adapting existing techniques from graphics is not sufficient. When trying to formulate the issues raised by minimal graphics, one soon realises that "abstract" or "minimal" are not concepts that can be described in purely algorithmic terms. Rather, "minimal" means that the image is rich enough so that *the human mind would complete the information through the cognition process.* A clear(er) idea of the cognition process is therefore necessary in order to decide what "minimal" really is in a specific context. In there lies, in our view, the greatest challenge in minimal graphics.

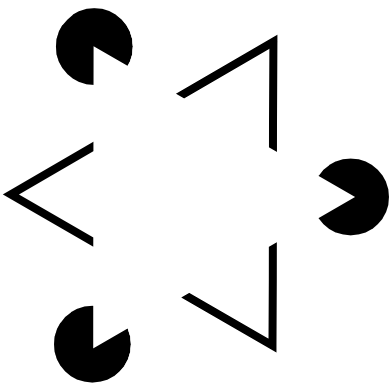

Figure 4. The Kanisza triangle

Figure 5. The Duomo of Milan

A number of optical illusions have been described which exemplify how human cognition is capable of "completing" an image. Figure 4, for example, shows the so–called Kanisza triangles: the three wedges in the black circles create an illusionary white triangle. There are numerous such illusions in, e.g., a recent book of J. Ninio[20].* To take another example, consider the image of the Duomo of Milan (Figure 5).The façade of the building has a very complicated edge, consisting of a complex pattern of stone carving. Nevertheless, the human mind clearly perceives a triangular façade, by "smoothing" the edges in the image. Looking at a cloud in the sky, the contours of a fractal image: these are all examples of the same effect. The effect can also be experienced in the temporal domain: a well–known example is Johansson's dancing figures[13]†. Generalising from these examples, it seems that the human cognitive process is somehow able to fill in some "emptiness" (the "triangle" in the middle of the Kanisza figure, the empty space at the edge of the Duomo). This duality between "empty" vs. "full" seems to play

* J. Ninio also emphasises the extremely important role played by our cultural heritage in perceiving various "illusions". Similarly, the basic principles underlying photorealistic graphics, like perspective view, should by no means be considered blindly as inherent to human cognition in all of its details; much of it may be determined by our own education and cultural and educational background… (see also [10]).

† People with lights on their joints are filmed moving in the dark. If the film is viewed frame by frame one sees nonsense — just static dots. Seeing the film at the right speed, one clearly sees people moving around.

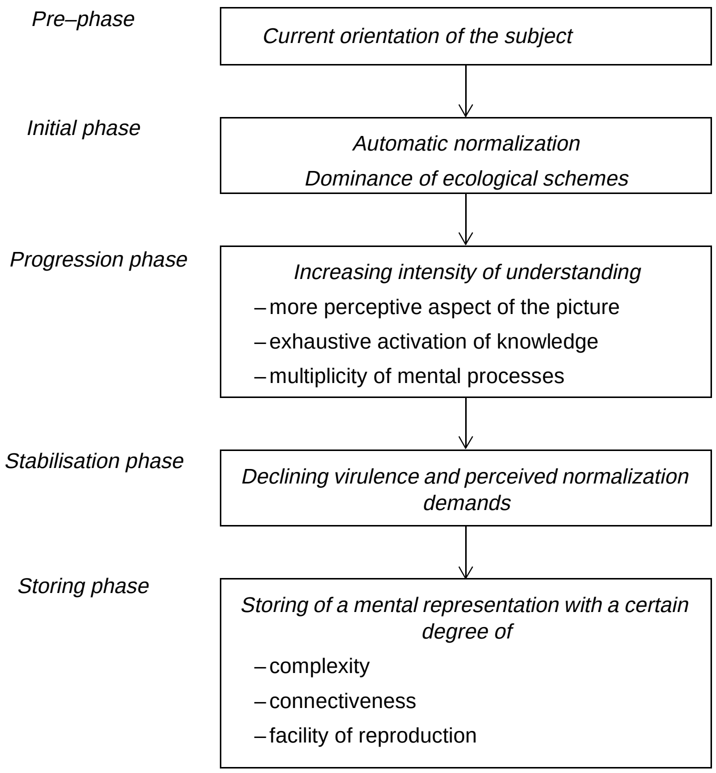

Figure 6. Phases of picture understanding in Weidenmann's model[27] (see also Strothotte[25])

an essential role in the way humans perceive their environment. This is why sketch images can show the Duomo of Milan as a simple triangle with some additional ornaments: minimal graphics should be able to generate similar sketchy images automatically. While a geometric model will be a necessary input to this process, it may not be sufficient; one of the research questions posed by minimal graphics is what additional information or data might be necessary or helpful within the rendering process. Although there is no simple, complete model that accounts for all aspects of visual illusion in terms of cognitive processes, there are theories that explain significant aspects of the problem at particular levels of operational detail, for example from neurological properties of the pre–cognitive phase, through to cognitive effects grounded in the interplay between top–down and bottom–up processing. It is not the goal of minimal graphics to produce new cognitive theories, but rather to draw on the existing knowledge of how such processes contribute to our understanding and interpretation of images.

It is obviously not possible to give a complete overview here for all the various theories which describe human perception and cognition. We can, however, point to a number of cognitive, and "process oriented" approaches that have already found use in the context of rendering. Strothotte[25], for example, gives a description of Weidenmann's scheme for picture understanding[27] (see also Figure 6), based on the concept of mental models, i.e., the particular representation of the context for every person. One important conclusion is that "…the mental models of novices and experts of the same event, for instance, are very different from one another"[25], which seems to indicate that, in order to be effective, a minimal graphics system should be configurable to the viewer. ICS[1,2,7] (see also Figure 7) is another model for cognition, which has been used successfully in clinical domains, as well as for theoretical work. A separate section (see Section 5 below) contains an example of the kind of analysis ICS allows us to do in order to characterise the various cognitive mechanisms related to minimal graphics.

We believe that there is an important analogy between the foundations of photorealistic graphics and those of minimal graphics. The principles of photorealistic graphics rest on an approximation of physical reality that is elegantly captured by Kajiya's well known rendering equation[14]. In contrast, minimal graphics does not (necessarily) seek to reproduce physical aspects of reality, and requires instead a model of cognitive information processing. Kajiya's equation is not in itself an algorithm for rendering images, but rather provides the theoretical foundation for families of approaches (ray tracing, radiosity) that implement particular aspects of photorealism. Similarly, we do not expect or require that the cognitive theory underpinning minimal graphics will provide an explicit approach to rendering. Rather, we believe that such a theory will provide the basis for defining a number of new rendering techniques that achieve a minimal approach.

Research in minimal graphics has an interdisciplinary bonus. While photorealistic graphics has drawn on theories developed within physics, it has not (to the best of our knowledge) promoted new developments within physics. In contrast, the fact that computer graphics is able to systematically generate dynamic representations also provides a unique opportunity to feed back into the development of new cognition theory. Such research will combine the growing awareness of the importance of dynamic aspects of visual perception with novel techniques for computing and displaying graphical representations.

5. Minimal Graphics and ICS

This section is an example of how a particular model of cognition can help us to formulate and understand some aspects of the minimal graphics problem. It exemplifies the type of research which has to be pursued.
A key point in the idea of minimal graphics is that in a reasonable number of situations, a “minimal” image is as good as, if not better than, one produced by “photorealistic” modelling. Why should this be? Implicit in the argument is a notion that somehow a human observer will be able to extract particular information from a given image. What minimal graphics requires is that some notion of information content, and the process by which information is extracted, be addressed explicitly in the theory on which a rendering technique is founded.

It is not the purpose of this paper to review the state of cognitive theory. We can, however, say that theories of cognition fall into two broad groups. Microtheory is concerned with the explanation of phenomenon within some restricted scope; for example, theories of vision fall into this group. Macrotheory, on the other hand, attempts to provide a framework in which the operation of different micro–theories can be situated and organised. In this section we demonstrate how one particular approach to macro–theory, namely Interacting Cognitive Susbsystems (ICS), might provide the kind of foundation that we are seeking for minimal graphics. ICS has already been used in the context of HCI and computer graphics to explore the usability of gestural interaction[6] and multimodal techniques[7].

The ICS model consists of a cognitive architecture and a collection of principles that govern and constrain the operation of that architecture for information processing. The architecture consists of nine distinct cognitive subsystems, four of which play a role in human understanding of images. Two independent and qualitatively different paths within the system are involved in the transport and processing of visually derived information. These are illustrated in Figure 7.

The first path begins with the vis–obj transformation that maps visual characteristics, such as texture, colour, hue, shading, etc., into a representation that concerns shapes and position within space. This can be seen as the transformation that is of primary interest within image analysis, i.e., extracting objects from a raw image. Next in the pipeline comes obj–prop, which maps spatial objects and information into propositional information, i.e., knowledge about what is in the image. If the results of vis–obj could be understood in terms like “there is a square–ish shape oriented parallel to the ground, with a triangular shape on top of it”, then the product of obj–prop might be “there is a house over there”. The final step which we will concern ourselves with here is the prop–implic transformation, that produces a higher–level representation encompassing affect and emotion, for example, perhaps a feeling of security of familiarity if the building that has been recognised is the viewer’s home.

However, the path from vis to implic via obj and prop is not the only route taken by visually derived information. The visual system also contains a transformation directly into implicational space, vis–implic. This operates in parallel with the path to implic via obj and prop. Certain aspects of the visual field give rise to implicational responses: sharp objects or shapes may connote threats or hazards; softer, rounded shapes may convey a sense of safety or harmlessness. Facial expressions in particular are quite rich in features that have implicational meaning.

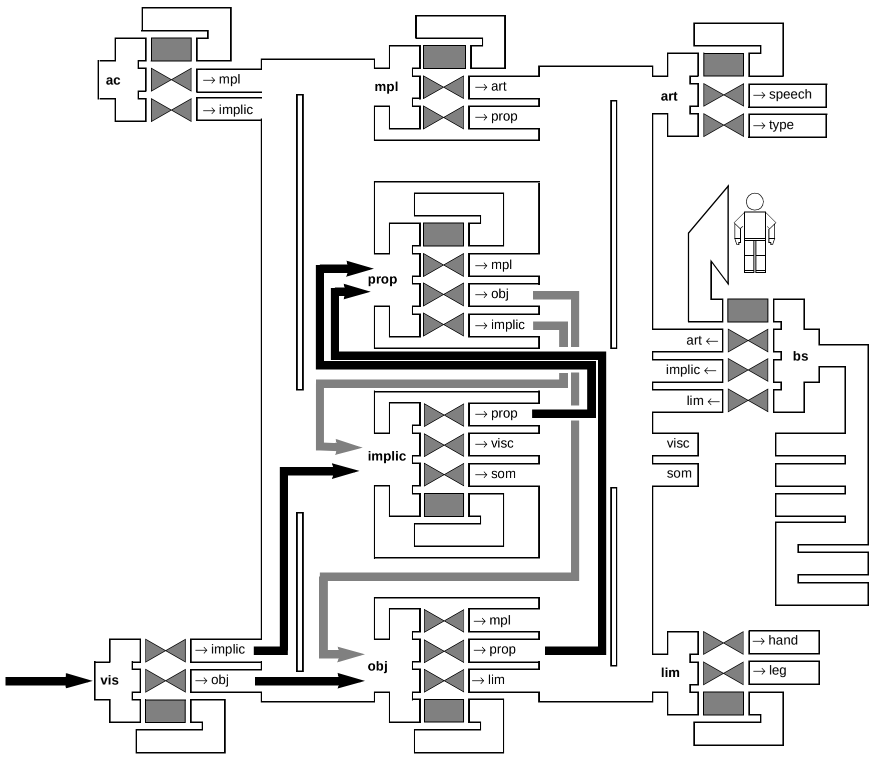

Figure 7. The ICS model, and flows supporting visual information processing.

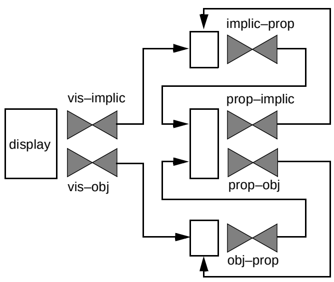

Figure 8. ICS resources for visual processing

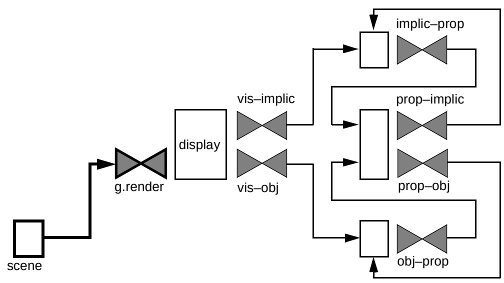

Figure 9. Geometric rendering in context

In addition to the two 'bottom–up' paths that take visual information to implicational meaning, there are also the two 'top–down' transformations, implic–prop and prop–obj, that can interact to construct an object–level representation from meaning or understanding. In Figure 7, the flows through these transformations are shown in light grey. Through a process called blending, top–down models can re–enforce or interfere with models of the visual or object–level scene being constructed at the object or propositional system. A classic example here is the Necker cube, which has two object–level interpretations [2].

We now illustrate the role that a model such as ICS might play in the development of minimal graphics by working backwards from the cognitive architecture to yield some insight into the structure and functionality of a minimal renderer. If we elide the details of the representation in Figure 7 that are irrelevant for visual information processing, and simplify some of the elements that remain, we are left with the model shown in Figure 8. This is a representation of the human image analysis process. Computer graphics is about image synthesis, producing an image which we hope that the user will interpret in a particular way; it is thus an inverse to the analysis process. We can start by taking the simplest approaches to graphics, in which we render a purely geometric model. This can be visualised as the process shown in Figure 9.

The flow of information between transformation processes is governed by a number of principles. All processes are operating in parallel, and as more than one process can produce a given kind of output representation, each system is actually receiving multiple streams of input data. For example, the obj system receives representations from the vis–obj process corresponding to perceptual input, as well as representations from prop–obj, representing top–down "mental imagery". While a transformation may at times operate on a stream of representations derived from a single source (e.g., the mind focusing on what is being viewed), more generally the different streams of data arriving at a subsystem can be combined (blended) to produce a composite representation. A good example of this is where an existing mental model of what a user expects to see is integrated with a consistent model of an image that the user is actively viewing. Like Kajiya's equation, the model described is an approximation of reality. In practice, the result of a transformation may for example depend on the strength to which different input–output mappings have been learned over time, or on qualitative properties of the data themselves, for example stability and jitter. However, as the space of inputs to transformations is richer than we can describe at the moment, we can simply assume that there is some property of the product of incoming data representations that determines the result, so the transformation can be considered a function. What is of more interest is that the inverse mapping, e.g., from object to visual space, is a relation: more than one visual representation may be generated from a given object–level model. So for any given outcome, i.e., an object–level model, there is a set of visual models that could generate it. For example, a photograph of a house and a painting of the house may have significant visual differences, yet still be understood as denoting the same structure. Minimalist graphics is about mapping geometric models into parts of visual space that are not utilised by photorealistic approaches. There are further consequences of thinking about the nature of inverse transformations, which we will come to later in this section.

In the model of Figure 9, all information processed by the human visual system is derived from the geometric model. This includes any implicationally–derived data. For example, cartoonists have for a long time used the "sharp/threat, round/friend" paradigm in drawing their characters: those with whom the viewer should empathise are often drawing with exaggerated, almost child–like, curves, while "villains" are typically give sharp features, particularly on the face (e.g., eyes and mouth). Essentially, information about intended *affect* is being hard–wired into the geometry of the models or drawings. However, recent work in human animation, particularly on facial expressions, has moved towards separating out the basic geometry from the problem of configuring the geometry to capture expressions of mood, emotion etc. This can be viewed as a first step towards separating, on the rendering side, the models that generate those aspects of the display that are handled by the object and the implicational levels of the human side. This separation is illustrated in Figure 10In practice, the process operating on the two sources of data (scene and affect) is probably the same; the two transformation symbols should be understood as two concerns, rather than necessarily independent processes.

Existing approaches to non–photorealistic rendering can be described in terms of this model. For example, the work of Strothotte et al. [25] on sketch rendering is controlling the drawing of lines to give an image that has certain affordances, for example, it has a softer, more "pliable" look to it than a comparatively "hard" image produced by standard illumination and shading models. Other approaches to non–photorealistic rendering, described in Section 3, concentrate also on modifying drawing attributes or shading models, and fit into this framework.

On the human side, implicational and object level data is being extracted from an image. To synthesise such an image, it is probably necessary to inter–relate attributes which influence affect with the geometric structure to which that affect should be related. In the approaches mentioned above, the separation of affective and structural information is made *a priori*, coded into the rendering algorithm. However, it is not just the object–level representation on the human side that we are constructing. Suppose that we want to communicate certain information to the user via an image, in other words some form of propositional understanding about the model from which the image

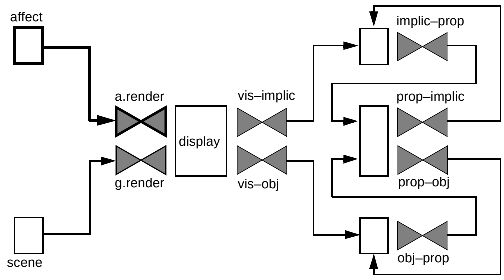

Figure 10. Rendering of affect

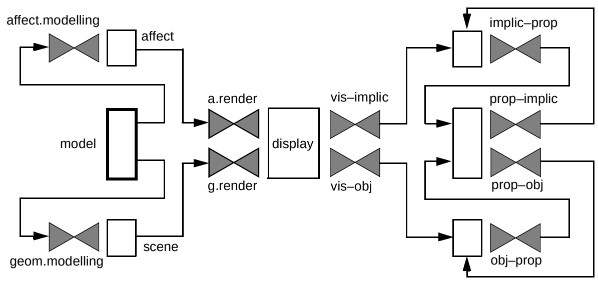

Figure 11. Trading of affect and structure

. was generated. The propositional system obtains data from both the object and implicational systems, so one issue that should be addressed in carrying out non–photorealistic rendering (or any rendering, for that matter) is the role of *both* of these channels in the process of understanding the image. On the side of synthesis, the problem can be restated as: "Given a model about which we want to convey information to the user, how can we control the use of rendering techniques to find an appropriate or effective way of conveying the necessary information?"* A model of this process is shown in Figure 11. Information in the model is used to determine how to trade–off and integrate structural information from that which may give rise to particular implicational responses, and/or to determine which of a possible range of images should be generated.

One aspect of the right–hand (human) side that is still missing from the left hand side figures (the rendering side) is the bottom–up links from (in our case) the scene and affective model to the original model. On the human side, the process of interpreting visual information involves interaction between the cognitive processes and levels of representation concerned. The act of interpreting an image may involve generating bottom–up plausible interpretations and modifying or refining these to accommodate the data being produced from higher–level subsystems. In principle, the dual process of *generating* a suitable image could involve iteration or convergence towards a particular model; a system might for example attempt to combine a possible structural and affective model and compare the product against the model that it is attempting to render. Such a cyclic approach to image generation would complete, at least conceptually, the symmetry of the roles.

In the case of human processing of visual information, the interplay between top–down and bottom–up processing is important for two reasons. The first is in resolving ambiguities or uncertainties in mental representations. The second is in utilising the experience and knowledge of the user in forming an interpretation of an image. For example, a technical diagram may have graphical elements or components that only have meaning in particular domains. Someone looking at such a diagram without knowledge of that domain is not going to form the same propositional representations of the image than someone with that knowledge. Similarly, as the image generated by a renderer moves from the "photorealistic" end of our hypothetical ordering, towards the non–photorealistic end, we would expect that there will come a point where we start to rely on a user's top–down processing ability

* Given this analysis, minimal graphics could also be referred to as "affective graphics"…

to extract detail from the image, or “fill in the blanks”. A simple illustration of similar “Gestalt” principles in cognitive psychology involves showing a viewer an apparently random collection of dots. However, some of the dots actually form the rendering of a dog. When the user becomes aware of this, it is possible to “see” the dog in the image[2]. The dots could be said to be a non–photorealistic rendering of a dog, but presumably, at this point, the image is of little use as a tool for information exchange. So some care is needed in deciding just how much detail can be elided from an image if the interpretation of the picture is to match that which is intended, at least to some required level of certainty.

In this section we have tried to show how non–photorealistic rendering and minimal graphics could be related to the cognitive processes and resources that humans deploy in interpreting images. By viewing the structure of a renderer in a way that is dual to these cognitive resources, it seems to be possible to understand what existing approaches to non–photorealistic models are aiming to do, and where there is room for significant new work. The analysis also begins to reveal the necessity of basing minimal graphics on cognitive theory, and how the role of such theory becomes analogous to that of physical models for photorealistic rendering.

6. Conclusions

The contribution of this paper is simple to state: we have argued that the conventional view of rendering, so elegantly captured in Kajiya’s equation, is but one part of a much broader enterprise of graphics–based communication in which there is a need to consider fundamentally different approaches to the rendering problem. To make progress on this enterprise, it is necessary to understand, and in some cases discover, cognitive theories that explain how graphical information is processed and understood by humans. What is not simple to state, of course, is how this new view of the rendering problem can be addressed. However, by working with a model of human information processing, we have at least been able to show where some existing approaches to non–photorealistic rendering fit into this problem, and consequently to highlight directions for further work. While we have focused on comparisons with non–photorealistic rendering, there are, of course, other lines of work which deal with the issues of generating effective presentations, these include work on visual communication[26] and , from an AI perspective, on presentation planning[3]. The contribution of these areas to minimal graphics is still to be explored.

We started the paper by showing the inspiration Chinese and Japanese painting may have on minimal graphics; let us close with another aspect of Far Eastern art which refers to the topics discussed so far. The duality of “emptiness” and real content is a central principle of traditional Chinese and Japanese painting, and of Buddhist and Taoist philosophy in general[4]. The traditions behind the Chinese/Japanese painting schools may become very helpful in understanding some relevant aspects of human cognition, too: much like in other areas, Far Eastern philosophy may have accumulated experiences which Western science cannot fully explain yet. Referring to the Taoist traditions of Japanese art, Kakuzo Okakura, one of the first Japanese scholars attempting to present Japanese art to the Western public, writes in his famous “Book of Tea”[15]:

> [Laotse] claimed that only in vacuum lay the truly essential. The reality of a room, for instance, was to be found in the vacant space enclosed by the roof and the walls, not in the roof and walls themselves. The usefulness of a water pitcher dwelt in the emptiness where water might be put, not in the form of the pitcher or the material of which it was made. Vacuum is all potent because all containing. In vacuum alone motion becomes possible. One who could make of himself a vacuum into which others might freely enter would become master of all situations. The whole can always dominate the part.
>
> These Taoists' ideas have greatly influenced all our theories of action, even to those of fencing and wrestling. Jiu–jitsu, the Japanese art of self–defence, owes its name to a passage in the Tao–teking. In jiu–jitsu one seeks to draw out and exhaust the enemy’s strength by non–resistance, vacuum, while conserving one’s own strength for victory in the final struggle. In art, the importance of the same principle is illustrated by the value of suggestion. In leaving something unsaid the beholder is given a chance to complete the idea and thus a great masterpiece irresistibly rivets your attention until you seem to become actually a part of it. A vacuum is there for you to enter and fill up the full measure of your aesthetic emotion.

This quote seems to be a perfect example on how Far Eastern art, tradition, and philosophy may help in understanding some issues of cognitive problems related to minimal graphics.

Acknowledgements

Obviously, the ideas and concerns described in this paper have been discussed with a number of different colleagues and friends, and they have all, in some ways, contributed to the ideas described here. We should mention the names of Phil Barnard (MRC–CBU, Cambridge, UK), Jon May (University of Sheffield, UK), David Duce (RAL, UK), Patrick Olivier (University of York, UK), and Pere Brunet (UPC, Spain).

References

[1] P.J. Barnard and J. May: "Cognitive Modelling for User Requirements". In: *Computers, Communication and Usability: Design Issues, Research and Methods for Integrated Services*, North Holland Series in Telecommunication, Elsevier, Amsterdam, 1993.

[2] P.J. Barnard and J. May, "Interactions with Advanced Graphical Interfaces and the Deployment of Latent Human Knowledge". In: *Interactive Systems: Design, Specification, and Verification, Proc. of the First Eurographics DSVIS Workshop*, Springer–Verlag, Wien, 1995.

[3] M. Bordegoni, G. Faconti, S. Feiner, M.T. Maybury, T. Rist, S. Ruiggieri, P. Trahanias, M. Wilson: "A Standard Reference Model for Intelligent Multimedia Presentation". *Computer Standards and Interfaces*, **18**(6–7) 1997.

[4] F. Cheng: *Vide et plein — Le langage Pictural Chinois*. Points Essais, Éditions du Seuil, Paris, 1991.

[5] C.J. Curtis, S.E. Anderson, J.E. Seims, K.W. Fleischer, and D.H. Salesin: "Computer Generated Watercolor". In: *Computer & Graphics, SIGGRAPH'97 Proceedings Issue*, 1997.

[6] D.J. Duke: "Reasoning About Gestural Interaction". *Computer Graphics Forum, Eurographics'95 Proceedings Issue*, **14**(3), 1995.

[7] D.J. Duke, P.J. Barnard, D.A. Duce, and J. May: "Syndetic Modelling". In: *Human Computer Interaction*, **13**(4), 1999.

[8] G. Elber: "Line Art Illustrations of Parametric and Implicit Forms". In: *IEEE Transactions on Visualization and Computer Graphics*, **4**(1), 1998.

[9] A. L. Guptill. *Rendering in Pen and Ink*. Watson–Guptill Publications, New York, 1976.

[10] P.A. Heelan: *Space–Perception and the Philosophy of Science*. University of California Press, Berkeley, 1983.

[11] S.C. Hsu and I.H.H. Lee: "Drawing and animation using skeletal strokes". In: *Computer & Graphics, SIGGRAPH'94 Proceedings Issue*, 1994.

[12] S.C. Hsu, I.H.H. Lee and N.E. Wiseman: "Skeletal strokes". In: *UIST'93 Proceedings of the ACM SIGGRAPH and SIGCHI Symposium on User Interface Software and Technology*, 1993.

[13] G. Johansson: "Visual perception of biological motion and a model for its analysis". In*: Perception and Psychophysics*, **14**, 1973.

[14] J.T. Kajiya: "The Rendering Equation", In: *Computer & Graphics, SIGGRAPH'86 Proceedings Issue*, 1986.

[15] Kakuzo Okakura: *The Book of Tea*. Original printing: Fox, Duffield & Company, New York, 1906. (The book had numerous translations and reprints since and it is also available on–line at the following URL: http://www.takase.com/KakuzoOkakura/Chapter00.htm.)

[16] J. Prikryl and W. Purgathofer: “State of the Art in Perceptually Driven Radiosity”. *Eurographics’98 STAR Reports*, A. A. de Sousa and F.R.A. Hopgood (eds), Eurographics Publications, Aire–la–Ville, 1998.

[17] J. Lansdown and S. Schofield: “Expressive rendering: a review of nonphotorealistic techniques”. In: *IEEE Computer Graphics and Applications*, **15**(5), 1995.

[18] P. Litwinowicz: “Processing Images and Video for an Impressionistic Effect”. In: *Computer & Graphics, SIGGRAPH'97 Proceedings Issue*, 1997.

[19] B.M. Meier: “Painterly Rendering of Animation”. In: *Computer & Graphics, SIGGRAPH'96 Proceedings Issue*, 1996.

[20] J. Ninio: *La Science des Illusions*. Odile Jacob, Paris, 1998.

[21] M. Salisbury, M.T. Wong, J.F. Hughes, and D.H. Salesin: “Orientable textures for Image–based Pen–and–Ink Illustration”. In: *Computer & Graphics, SIGGRAPH'97 Proceedings Issue*, 1997.

[22] T.T. Sasada: “Drawing natural scenery by computer graphics”. In: *Computer–Aided Design*, **19**(4), 1987.

[23] J. Schumann, T. Strothotte, A. Raab, and S. Laser: “Assessing the effect of nonphotorealistic rendered images in CAD”, In: *Proceedings of the CHI'96 Conference*, ACM Press, 1996.

[24] E.J. Stollnitz, T.D. DeRose, and D.H. Salesin: *Wavelets for Computer Graphics*. Morgan Kaufmann Publishers, Inc., San Francisco, 1996.

[25] C. Strothotte and T. Strothotte: *Seeing between Pixels*. Springer–Verlag, Berlin, Heidelberg, 1997.

[26] E.R. Tufte: *Envisioning Information*, Graphics Press, Cheshire, Conn., 1990.

[27] B. Weidenmann: *Psychische Prozesse beim Verstehen von Bildern*. Verlag Hans Huber, Bern, 1986.

[28] G. Winkenbach and D.H. Salesin: “Computer Generated Pen–and–Ink Illustration”. In: *Computer & Graphics, SIGGRAPH'94 Proceedings Issue*, 1994.