IMAGE SYNTHESIS:

OPTICAL IDENTITY OR PICTORIAL COMMUNICATION?

Michael I. Mills

The Interactive Telecommunications Program
New York University, New York, New York 10003

ABSTRACT

This paper reviews some ideas from E.H. Gombrich (an art historian) and Julian Hochberg (a perception psychologist) on artists' attempts to simulate reality in painting. These ideas suggest that workers in image synthesis enlarge the scope of their mission. Rather than trying to get computers to make images which are indistinguishable from direct perception, they should pay more attention to the use of computer graphics for "pictorial communication"--where departures from objective realism can sometimes lead to more effective displays and where the mental process of image interpretation should itself be a legitimate focus of research and experimentation.

INTRODUCTION

"Our goal in realistic image synthesis is to generate an image that evokes from the visual perception system a response indistinguishable from that evoked by the actual environment." (Hall and Greenberg, 1983, p. 10) This quote, taken from a recent article in IEEE Computer Graphics and Applications, conveys how many workers in the field of image synthesis interpret their mandate. At first glance, this seems like a sensible research objective. It provides a straightforward criterion for progress. We judge the success of this synthesized image by comparing our perception of it to our direct perception of the identical scene in the real-world.

Now it is true that researchers in image synthesis are constantly finding new ways to mathematically model the physical behaviour of light as it is propagated throughout the environment. And one day these models may enable observers to experience experience in computer "surrogates" many of the important patterns of stimulation available in the visual world--including perhaps those patterns dependent on binocular vision (Schmandt, 1983) and on real-time changes in the observer's station point (note 1). But I believe that short of actually building physical replicas, there will always be fundamental differences between looking at images and looking directly at the real thing. (c.f. Gibson's discussion of "Pictures as substitutes for visual realities" in Reed and Jones, 1982).

An important question one can raise then is this: Is there a better--or at least more sophisticated--way of conceiving of the task of depiction in computer graphics than as a contest with reality, a quest to "fool" the perceptual system via imitation of the real world? The communication of visual truths in an image may have less to do with the manufacturing of perfect copies of retinal images than with the skillful manipulation of evocative forms; less to do with transcribing reality than with suggesting it pictorially. And to learn more about the underlying principles of this communicative process--in which the brain imposes interpretations on patterns of light that reach it--we must turn to the domains of art history and perceptual psychology, rather than computer graphics.

CHANGING STANDARDS OF REALISM

If there is one book I would urge computer scientists in the synthetic imaging area to read, it would not be a treatise on technology or the physics of light, but E.H. Gombrich's masterpiece, Art and Illusion (Gombrich, 1960). I think that graphics scientists would find this book particularly valuable because they would see many similarities between artists' "experiments" in using paint to simulate the effects of light and their own quest to simulate the visible world in computer displays. Obviously, I cannot do justice to all the insights in Gombrich's wonderful book. But I would like to highlight several examples of particular relevance to the problem of making realistic images by computer.

In a chapter entitled "From Light to Paint", Gombrich tells of a controversy surrounding "Wivenhoe Park"--a marvelous landscape by the nineteenth-century English artist John Constable in which we can see "...the play of sunlight on green pastures, the gentle ripples on the lake with its swans, and the beautiful landscape that encloses it all." (P. 33) Now we would never mistake this oil painting for a photograph; but we easily accept it as a reasonably faithful record of what Constable might actually have seen when he stood on that spot to paint the picture. It is therefore difficult for us to grasp that this painting was considered by nineteenth-century eyes as an avant-garde, if not revolutionary, "experiment" in depiction.

Constable was widely rebuked for trying out a greater range of tonal contrasts--especially the use of brighter shades of green--in reconciling the local colour of grass with the tonal gradients needed to suggest depth, than had heretofore been attempted in landscapes of the era. At first, there was great resistance to his use of lighter shades. Legend has it that Constble's patron, Sir George Beaumont, chastised him for not making the grass in the foreground "the requisite mellow brown of an old violin." Constable's retort, so Gombrich tells us, was to "take a real violin and put it on the grass before him to show his friend the difference between fresh green grass and the warm tones demanded by convention." (Gombrich, 1960, p. 44)

Over time, Constable's transgression became the rule, not the exception, and paved the way for brighter and more "realistic" landscapes. But we should not come away from this story thinking that Constable's victory lay in his ability to perform as a kind of human camera--a copying machine capable of matching the "true" colour of grass to the corresponding green on his pallette. In fact, as Gombrich points out, his genius was rather more subtle. The colour of the grass in Constable's painting is still closer to the brown violin than to the green of fresh grass. Constable's gift was not that he could copy nature better than his predecessors, but rather his intuition that the human visual system could be stretched in its response to the way pigment can suggest relationships in light: that by expanding the contrast between light and dark tones, the artist could force the perception of new "visual truths".

MORE DETAIL, MORE REALISM?

A first lesson from Gombrich, then, is that judgements of realistic portrayal are not so much a process of comparing images with reality, but rather that our criteria for what we accept as visual reality shifts with the discovery of novel techniques for suggesting the effects of light. A second lesson, perhaps even more important to our present purpose, concerns the assumption by many workers in digital imaging that, simply by cramming more and more details (adding more pixels), we automatically achieve more realistic and "useful" images of things and people.

Interestingly, Gombrich tells us that the "more detail, more realism" hypothesis was also shared at first by the old masters, who learned through trial and error that adding more and more detail did not necessarily lead to better paintings.

Rembrandt, for example, in his early portrait work, would struggle to render the "microstructure" of small segments of the visual array corresponding to, say, shiny gold braid on his subject's garment. In his later work, however, he learned that a few well-placed brushstrokes enabled the observer to achieve an even superior experience of gold braid. The secret was that when viewed at the appropriate distance, the small segment representing the gold braid fell on the periphery of the eye and the viewer's mind would "fill in the gaps" so to speak; i.e. the broader brushstrokes could convey the immediacy of the visual system's response to smooth shiny surfaces and glitter better than the more detailed, laborious rendering.

A modern-day counterpart to the old masters' experiments on the relationship between "redundancy removal" and "realism" is the by now well-known computer-generated "block portraits"--for example, a grid of say 14 x 18 squares, each of which can assume any of sixteen grey levels. What is interesting about these reduced-information block portraits, of course, is that viewed close-up, they are merely arrangements of light and dark squares. Seen from a distance of several feet, however, the individual squares seem to fuse magically into a clearly recognizable picture of a human face. Moreover, if the viewer shakes his head or someone jiggles the picture, there seems to be more apparent detail in what is actually quite a coarse-grained image. The intriguing question for students of perception is why stepping back from such images, or blurring them intentionally--actions which are the opposite of what we usually do when we want to get a better look at something--should improve the identification process. No one as yet has a complete answer; but as Rembrandt realized in his rendering of gold braid, the secret must lie not just in the physical qualities of the image, but in accounting for what we will call, after Gombrich, the "beholder's share" in picture perception. (In the case of block portraits, "blurring" through head movement or backward lococmotion serves to "filter out" the high frequency spatial components -- the sharp block edges -- allowing the critical low frequency "portrait" information to get through. C.F. Harmon, 1974).

OPTICAL IDENTITY VERSUS OPTICAL SIMULATION: INSIGHTS FROM HOCHBERG

It would appear then that Rembrandt beat Bell Laboratories by several centuries in exploring mehods of redundancy removal in picture making! But surely Rembrandt did not paint this way just to save on bandwidth. Somehow he knew (although we cannot be certain what was going on in his mind) that a few brush strokes could conjure up an impression of smooth, shiny metal better than a more detailed rendering. Can we say more about why he may have been right? Fortunately, we can--thanks to Julian Hochberg, one of the few perceptual theorists who has probed the complex links between art, pictures and the workings of the human visual system: what we have been calling, following Gombrich, the beholder's share.

In an important article called "Some of the things that paintings are", Hochberg hypothesizes that the techniques used by Rembrandt are breakthroughs in depiction, in the way they exploit fundamental differences between central and peripheral vision. Hochberg reminds us that it is only the center of the retina--the foveal region--that picks up full colour and resolves fine detail. As the distance from the fovea increases, the eye's ability to resolve detail falls off dramatically. Peripheral vision responds mostly to abrupt changes in luminance, indicating edges, large surfaces, and movement. Consider what would occur if, when viewing a Rembrandt portrait, our gaze happens to land on a spot away from the middle region of the painting--i.e., away from the portion of the picture encompassing the face and hands. Even though our foveae can resolve great detail in these outer regions, there is no fine detail to be had. Hochberg notes two important consequences of this fact. Firstly, our gaze will drift back to the middle region of the picture--the face and hands--which are the points of interest, where fine detail does exist. According to Hochberg, this "forced focusing" has obvious advantages from a compositional point of view.

A second, more subtle consequence concerns the use of "simultaneous contrast" effects to overcome the limits of paint to represent apparent brightness: highlights, reflections and so on. Simultaneous contrast effects occur when we look at two neighboring areas of a scene, one of which is dark and the other light. Our brains do not merely register the objective difference in brightness between the two adjacent patches, but actually enhance the perceived brightness contrast. This works as follows. The neural receptors in the eye are linked so that the more a particular retinal region responds to light, the more it inhibits the response of adjacent receptors. In a given scene, a light patch surrounded by a dark region will be seen as subjectively brighter than a patch of the same tonality surrounded by a light region. The receptors receiving the central image are being less inhibited by neighboring receptors in the former case, with the net effect of raising the apparent brightness contrast. (Painters discovered early on that by surrounding a grey patch with a black background, they could make the grey patch appear brighter than it would have looked on a white background.) Moreover, as Hochberg points out, simultaneous contrast effects are even stronger in the periphery of the eye than in focal vision. Although Rembrandt may not have been able to articulate his awareness of this phenomenon, he seems to have used it to good effect. Hence, as we mentioned earlier in conjunction with Rembrandt's attempt to render gold braid, he learned that by placing large brushstrokes of extreme lights and darks just outside the area just outside the area of focal interest (outside of foveal vision), he was able to enhance the apparent brightness of the highlights and shiny surfaces in the picture.

We can now better understand Gombrich's claim that Rembrandt's rendering of gold braid in his later portraits have an "immediacy" and "glitter" not conveyed by its more detailed counterpart--but with an important corollary. The enhancement of brightness only works if the braid lies outside the focal area of the painting. Once we fix our eyes on the gold braid, inspecting it closely, the global quality--the illusion of glitter--is destroyed as our eyes resolve the individual brushstrokes.

To Hochberg, these facts are extremely important if we are to grasp the special, dual character of pictures: i.e., pictures are themselves perceptibly flat objects, yet they can be seen as "surrogates" for other objects, usually three-dimensional. In normal vision, when we search a scene for information, we do resolve fine detail when we move our gaze onto the periphery. When we shift from the central to the outer regions of the Rembrandt painting, however, a mental "switch" occurs: instead of fine detail which we interpret in terms of the scene--paint depicting a face--we notice the rough-hewn, individual brushstrokes, dobs of paint; not an object. This fact, according to Hochberg, serves to remind us (perhaps unconsciously) that we are indeed looking at a painting--itself a flat object covered with pigment--not a real scene or a replica. This is the artist's way of telling us that he is not engaging in a contest with reality: not trying to copy, in a one-to-one mapping, the light emitted from the scene. Rather, he is using paint to "simulate" certain important effects of light on our visual system--a process of pictorial communication better thought of as optical simulation than as optical identity or equivalence.

The difference between the self-conscious use of paint to achieve optical simulation, as opposed to optical identity, can most clearly be seen in the experiments of the Impressionists. Their avowed goal, like that of modern-day workers in image synthesis, was "...to use optical science to produce paintings that would provide the same impression to perceptual experience as does the light in a (usually outdoors) scene." (Hochberg, 1979, p. 33) No one, however, would ever mistake Monet's impressionist painting "Rouen Cathedral" for a photograph of the real scene. Unlike Rembrandt's portraits which contained pockets of fine detail in focal regions, Monet's painting provides no resting place for our eyes--no islands of fine detail. The entire canvas is built out of rough-hewn swatches of colour. How could Monet do this and still claim to be following the goals of Impressionism?

Hochberg, again, provides a penetrating analysis which brings us full circle to the computer-generated block portraits mentioned earlier:

> In an impressionist painting, when it is viewed from a normal distance, there are no places at which the fovea can pick up fine detail. To peripheral vision, on the other hand, the Impressionist painting looks veridical, as it does when viewed with deliberately out of focus ("abstracted") gaze and when viewed from a distance that is considerably greater than normal. As Harmon and Julesz (1973) showed, the perceptibility of the object represented by a patchwork picture, like those most characteristic of Impressionist painters, is increased when the perceptibility of the small details or "high spatial frequencies" (provided by the edges of the brushstrokes) is reduced, as in peripheral viewing or viewing from increased distance. At a normal viewing distance, the scene dissolves

into patches wherever the fovea leaves them, and they are reclaimed by peripheral or parafoveal vision. (Hochberg, 1979, p. 35)

For Hochberg, then, one might say that Monet's intent was to convey, or more precisely, to simulate the first fleeting impression--the dazzle and freshness of a momentary glance at a sunlit catherdral; an evoking of the events in peripheral vision, undetailed, with gross volumes and colours. Monet's goal was not to fool us into believing that we are looking at a real scene, but to show us how pictures can be used to capture novel "truths" of visual experience.

CONCLUSION

At the outset we saw that the goal of image synthesis research was to use computers to make images of objects and people which would be optically identical to direct perception of the visible world. Most of the effort to date has been spent developing mathematical models of the physics of light, the geometrical modeling of solid objects, and building the hardware to generate scenes based on these formal descriptions. I hope that my brief review of pictorial realism in art, based on the writings of Gombrich and Hochberg, has convinced you that image synthesis and research might profit not only by considering the physics of light and solid geometry, but by learning more about the "beholder's share" in picture perception. Let me summarize the lessons we have gleaned thus far from these authors.

First, we saw that judgements of realism in picture-making are always context-bound; what is considered a realistic image may depend more on comparisons between one picture and another, according to a particular period's criteria for realistic picture-making, than on any fixed or absolute standards of reality based on looking at the visible world. Second, we saw that adding more and more detail, either with paint or pixels, does not necessarily lead to more realistic looking pictures--if one takes into account the limits of the display medium and the functioning of the visual system (i.e., the differences between central, parafoveal, and foveal vision, simultaneous contrast effects, etc.). Finally, we saw that by acknowledging and experimenting with the interactions between the visual system and the limits of the medium of depiction, artists redefined their basic mission. Rather than trying to create objects (paintings) which would be optically identical to the real scene, they strove instead for optical simulation through painting: a self-conscious process of pictorial communication which acknowledges the dual character of pictures (e.g., perceptibly flat objects treated so as to depict a different set of objects. And where the departures from realism--the techniques of optical simulation themselves--become a focus of interest and investigation in their own right.

We should not conclude from these arguments that computer graphics researchers should abandon the quest for realism via sophisticated formal models of light propagation and solid modelling. Nor are we saying that computer imaging should imitate the techniques of Rembrandt and the Impressionists. We are claiming, however, that learning more about the "beholder's share" in pictorial communication, including developing an awareness of how computer imagery fits in with the history and psychology of picture making, could have payoffs in such nitty-gritty matters as developing techniques for redundancy removal in picture encoding, and might tell us more about what styles of pictorial rendering might be acceptable in specific contexts, in different task environments. (C.F. Hearty, 1983).

ACKNOWLEDGEMENTS

Many of the ideas in this paper stem from a larger report entitled "Simulating Reality in Computer Images: Insights from Art History and Cognitive Science" written by the author and funded by the Canadian Department of Communication (contract 36100-3-0057). I would like to thank Dr. Paul J. Hearty of the Canadian Communications Research Centre for his valuable guidance in helping to formulate these ideas.

FOOTNOTES

Note 1. "Viewpoint dependent imaging" in Discursions, a videodisc by the Architecture Machine Group, MIT.

REFERENCES

Gibson, J.J. Pictures as substitutes for visual realities. In E.Reed and R. Jones (eds), Reasons for Realism: Selected essays of James. J. Gibson. Lawrence Erlbaum, 1982.

Gombrich, E.H. Art and illusion. Princeton University Press, 1960.

Hall, R. and Greenberg, D.A. A testbed for realistic images synthesis. IEEE Computer graphics and its applications, vol. 3, no. 8, 1983.

Harmon, L. The recognition of faces. Image, object and illusion: Readings from the Scientific American, 1974.

Hearty, P.J. Behavioural research on Telidon graphics. Paper presented at the workshop on Telidon and the graphic arts, Science and Technology Centre, Vancouver, British Columbia, 1983.

Hochberg, J. Some of the things that paintings are. In C.Nodine and D.Fisher (Eds.), Perception and pictorial representation, Praeger, 1979.

Schmandt, C. Spatial input/output correspondence in a stereoscopic computer graphic workstation. Computer graphics, 17(3), 1983, 253-261.