Perception and Data Visualization: The Foundations of Experimental Semiotics

Colin Ware
Faculty of Computer Science
University of New Brunswick
Box 4400, Frederiction NB
CANADA, E3B 5A3
cware@unb.ca

Abstract

Data Visualization is a new discipline involving the application of computer graphics to the understanding of information. It is successful because it can take advantage of the remarkable pattern finding capability of the human visual system. Visualization techniques are applied in medicine, software engineering, stock market analysis and many areas of science. But is it a science or a design discipline? If it is a science, how should research be conducted? In this paper it is argued that visualization can be productively regarded as an area of applied perception research, building on new advances in our understanding of how people see. The intellectual basis for such a discipline is outlined and illustrated with three example relating to color, object perception and stereoscopic space perception respectively.

Keywords: data visualization, semiotics, human perception.

Introduction

Why should we be interested in visualization? Because the human visual system is a pattern seeker of enormous power and subtlety. The eye and the visual cortex of the brain form a massively parallel processor that provides the highest bandwidth channel into human cognitive centers. At higher levels of processing, perception and cognition are closely interrelated which is the reason why the words "understanding" and "seeing" are synonymous. However, the visual system has its own rules. We can easily see patterns presented in certain ways but if they are presented in other ways they become invisible. Thus for example, the word DATA shown in Figure 1 is much more visible in the bottom version shown below than in the one at the top. This, despite the fact that an identical amount of the letters is visible in each case and in the lower figure there is more irrelevant "noise" than in the upper figure. The rule that applies here, apparently, is that when the missing pieces are interpreted as foreground objects, continuity between the background letter fragments is easier to infer. The more general point is that when data is presented in certain ways the patterns can be readily perceived. We can think of a "grammar" of perception and this grammar of perception can be translated directly into a rules for displaying information. If we can understand this perceptual grammar then we can present our data in such a way that the important and informative patterns stand out. If we disobey the rules our data will be incomprehensible or misleading.

Note: Until recently, the term "visualization" meant "to construct a visual image in the mind (Shorter Oxford English Dictionary). However, in recent years it has come to mean something more like *a graphical representation of data or concepts* and it is this newer sense that the word is used here.

Figure 1. Adapted from Nakayama et al. 1989

There are four basic stages in the process of data visualization together with a number of feedback loops. These are illustrated in Figure 2:
They consist of.

- the collection and storage of data itself,
- pre-processing, such as filtering or statistical averaging designed to transform the data into something we can understand,
- the display hardware, and the graphics algorithms that produce an image on the screen.
- the human perceptual and cognitive system (the perceiver).

The longest feedback loop involves gathering data itself. A data seeker, such as a scientist, or a stock market analyst, may choose to gather more data to follow up on an interesting lead. Another loop controls the computational pre-processing that takes place prior to visualization. The analyst may feel that if the data is subjected to a certain transformation prior to visualization it can be persuaded to give up its meaning. Finally the visualization process itself may be highly interactive. For example in 3D data visualization the scientist may fly to a different viewpoint to better understand the emerging structures. Alternatively, a mouse may be used to interactively select parameter ranges that are most interesting. Both the physical environment and the social environment are involved in the data gathering loop; the physical environment is a source of data, while the social environment determines in subtle and complex ways what is collected and how it is interpreted.

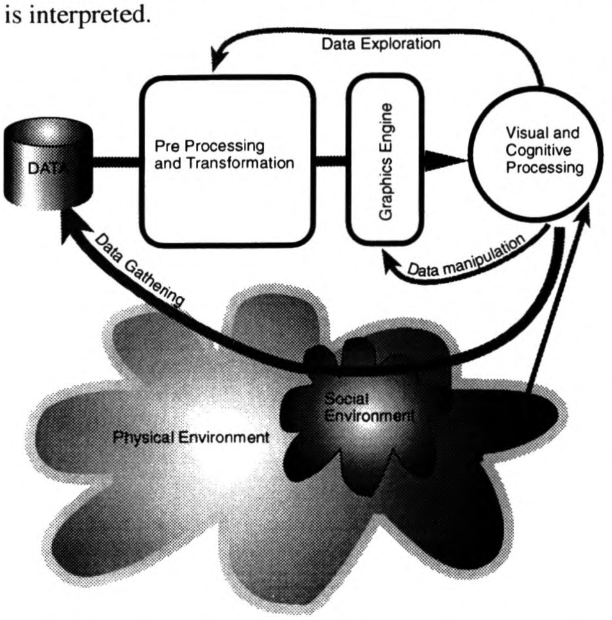

Figure 2. A schematic diagram of the visualization process.

A science of visualization can be based mainly on data and on human perception. Algorithms are important, but only in so far as they aid the perception of data. The computer can be treated, with some reservations, as a universal tool for producing interactive graphics. The critical questions are how best to transform the data into something that people can understand, for optimal decision making. However, before plunging into the detailed analysis of human perception and how it applies in practice, we must establish the intellectual basis for the endeavor. The purpose of this discussion is to stake out a theoretical framework wherein claims about visualizations being "visually efficient" or "natural" can be pinned down in the form of testable predictions.

The study of symbols and how they convey meaning is called semiotics, a discipline originated by the French philosopher and linguist, Ferdinand de Saussure (reprinted 1959). This field of study has been mostly dominated by philosophers and those who construct arguments based on example, rather than formal experiment.

The most profound threat to the notion that there can be a *science* of visualization originates with Saussure. He defined a principle of arbitrariness as applying to the relationship between the symbol and the thing that is signified. Saussure was also a founding member of a group of structuralist philosophers and scientists, who although they disagreed with one another on many fundamental issues, were unified in their general insistence that truth is relative to its social context. Meaning in one culture may be nonsense in another. A trash can as a visual symbol for deletion is meaningful only to those who know how trash cans are used. Thinkers such as Claude Levi-Strauss, Barthes and Lacan have condemned the cultural imperialism and intellectual arrogance implicit in applying our intellects to characterizing other cultures as "primitive", and as a result, they developed the theory that all meaning is relative to the culture. Since it seems entirely reasonable to consider visualizations as communications, this argument strikes at the root of the idea that there can be a natural science of visualization. Here is philosopher, Nelson Goodman, statement of the case.

"Realistic representation, in brief, depends not upon imitation or illusion or information but upon inculcation. Almost any picture may represent almost anything; that is, given picture and object there is usually a system of representation, a plan of correlation, under which the picture represents the object" (p.38). Goodman, Languages of Art.

However, there is a counter argument that can be launched based on the tools of science. There are a number of lines of scientific evidence that suggest that the basis of perception is common to all humans and indeed to many higher animals, and in addition that simple line drawings and diagrams are understood through these basic mechanisms. Here is a small part of the evidence.

Hochberg and Brooks (1962) raised their daughter nearly to the age of two years in a house with no pictures. She was never read to from a picture book and there were no pictures on the walls in the house. Although her parents were not able to completely remove pictures from her environment on trips out of the house, they were careful never to indicate a picture and tell the child that it was a representation of something. Thus she had no social input telling her that pictures had any kind of meaning. When the child was finally tested she had a reasonably large vocabulary, and she was asked to identify objects in line drawings and black and white photographs. She was almost always correct in her answers despite the lack of instruction in the interpretation of pictures.

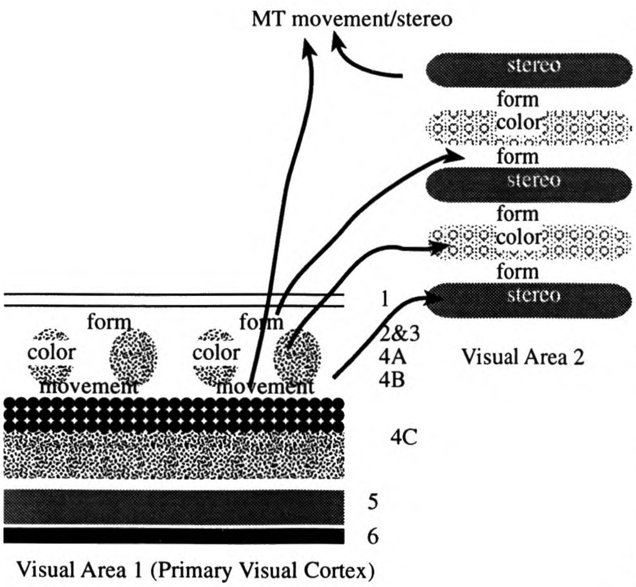

Figure 3.14 Architecture of primary visual areas. Adapted from Livingston and Hubel (1988)

Probably the best evidence for there being a hard-wired grammar of perception comes a convergence of evidence from two disciplines: visual psychophysics and neurophysiology. Visual psychophysics is an experimental discipline involving carefully controlled experiments on humans where subjects observe and respond to patterns of light that vary in precisely controlled ways. In these experiments, processing capabilities of the human visual system are either directly measured or inferred. An example of an early result is trichromacy theory. This theory predicted that there are exactly three different types of color receptor in the eye long before physiologists actually confirmed their presence.

Modern neurophysiology, has made considerable progress in identifying functional processes and pathways in the brain. Figure 2 gives a single example, an overview of the architecture of the primary visual areas of the Macaque monkey (Livingson and Hubel, 1988). These are the areas where input from the retina is first processed. Various sub-systems respond to different features of the input such as form (size, orientation) color (red-green, yellow-blue and black-white opponent channels), stereoscopic depth, and local motion. There are two points to be made, for the present argument. The first point is that the fact there is an architecture. This strongly suggests specialized processes that, if understood, may result in efficient ways of presenting information. The second point is that there is general agreement between the set of features found by in the Livingston and Hubel model and the set of features inferred from psychophysical measurement of humans. This suggests that a similar set of features is primitive in human vision (See Triesman and Gormican (1988) for example). Thus there is much psychophysical evidence to support the idea that simple form, color and motion are primitives in human perceptual processing, as they appear to be in animal processing.

Finally, theories of object perception, suggest that a critical part of the processing involves extracting the contours or boundaries, segmenting the object from its background. The mechanisms of contour perception will be strongly stimulated by simple outline line drawings. This provides an explanation for why simple outline drawings are readily recognized as representing three dimensional objects, even be people who have never seen pictures before (Kennedy, 1974).
This is not to say that culturally derived design rules should not play a role in the design of data visualization systems. Clearly culture has played a huge role in the

graphical design conventions that we used. The following anecdote illustrates the point. An Oriental student in my laboratory was working on an application to visualize changes in computer software. She chose to represent deleted entities with the color green and new entities with red. It was suggested to her that red is normally used for a warning, while green symbolizes renewal, and perhaps the reverse coding would be more appropriate. She protested that, green symbolizes death in China while red symbolizes luck and good fortune. Thus the use of color codes to indicate meaning is highly culture specific.

Therefore, the design of data visualization techniques can draw from social science disciplines, as well as the hard science approach of vision research. However, the rules for study in the social sciences, are very different than those that apply in the hard science of vision research. Most social scientist believe all knowledge to be relative to social context. At best the techniques of social science are based on careful observation that result in "thick description", a kind of translation of knowledge from one culture to another (Geerz, 1973). The methods of historians and anthropologists are well suited to understanding computer interfaces placed in a rapidly evololving social context but the hard science approach can yield design rules that transcend culture and time. For a more complete analysis of the relative roles of social science and physical science approaches see Ware (1993).

The remainder of this paper is devoted to three examples intended to illustrate how the science of vision research can be directly applied to solving problems in data visualization.

Example 1: Color sequences for Univariate and Bivariate maps

The first example applied our understanding of the mechanisms of color perception to the creation of color sequences for pseudo-coloring scientific data. The opponent process theory of color vision states that the input from the cone receptors of the retina is transformed in to three independent color "channels". Two are chromatic, a red-green channel and a yellow-blue channel. The third is the achromatic black-white (Luminance) channel. (The channels are shown schematically in Figure 3). The different properties of the color channels have profound implications for the use of color for data display.

Both chromatic channels have considerably lower spatial resolution than the luminance channel and this is especially so at high spatial frequencies. The implications of this is that purely chromatic differences are not suitable for displaying kind of fine detail.

It is impossible or at least very difficult to see stereoscopic depth in stereo pairs which differ only in terms of the chromatic channels (Gregory, 1975). Stereo space perception is based primarily on information from the luminance channel. In addition the shape-from-shading information that we get from the interaction of light with oriented surfaces is not understood if the shading is transformed from a luminance gradient to a purely chromatic gradient. When this is done the impression of surface shape is much reduced. In general, perception of shape and form appears to be processed mainly through the luminance channel (Gregory, 1975).

On the other hand, chromatic coding is very good for displaying the type of an object. We can easily sort objects into different categories according to color providing that not too many colors are used. Thus in order to perceive the *shape* of cells on a slide, it is important that there be a luminance difference at the boundary. However if the cells are differentiate in color (perhaps because of a staining technique) it easy to classify them rapidly on the basis of color alone.

In pseudo-coloring a map there are two important tasks that must be considered: Perceiving features or forms in the map is one, and reading data values using a map key is the other. The most common coding scheme used by physicists is a color sequence that approximates the physical spectrum. It is important to notice that this is not a perceptual sequence according to opponent process theory This can be demonstrated by the following test. Give a person the colors red, green, yellow and blue and ask them to place the colors in order, and the result will be varied. Give the person a series of gray paint chips and they will happily comply with either a dark to light ordering of a light to dark ordering. However, *parts* of the spectrum are perceptually ordered. For example the sequence between red and yellow, or between yellow and blue. This is exactly what the opponent color theory predicts. the spectrum approximation is poor both in terms of allowing the scientist to perceive form in the data (unless there are considerable luminance changes in the sequence). However, it is very good if subjects are required to read values back using a key Simultaneous contrast effects can drastically distort perceived color values and this can lead to errors in reading pseudocolored maps. Simultaneous contrast effects occur independently in the opponent channels. Precisely because the spectrum sequence is not monotonically ordered with respect to any opponent channel, contrast effects tend to cancel and readings are more accurate (Ware, 1988).

As a general rule it is important that if shapes and patterns are to be perceived in the pesudo-colored data then it is important that there be a strong, monotonically varying luminance gradient in whatever sequence of colors that is chosen.

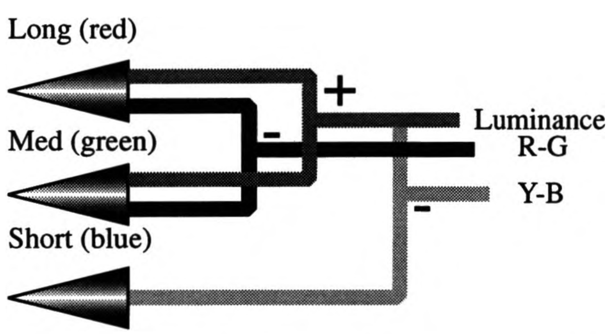

Figure 3. The opponent processing model of color vision. The cone signals are differenced to produce the red-green and yellow-blue channels and summed to produce the luminance channel.

Example 2: Applying rules from form perception

Figure 4 provides a somewhat simplified overview of a neural network model of shape perception developed by Hummel and Biederman (1992). The important point here is not whether this model is correct or not in all of it's details, rather it is to illustrate how a such a model can be used directly to provide guidelines for information display.

To give a very brief summary. The model is essentially hierarchical, Visual information is first decomposed into edges, then into component axes, oriented blobs, and vertices. At the next layer three dimensional primitives such as cones, cylinders and boxes (called geons) are identified. Next the structure is formed that specifies how the geon components interconnect; e.g. the arm cylinder is attached near the top of the torso cylinder. Finally, object recognition is achieved. Although not represented in this model, the geons must also have surface attributes such as color and texture.

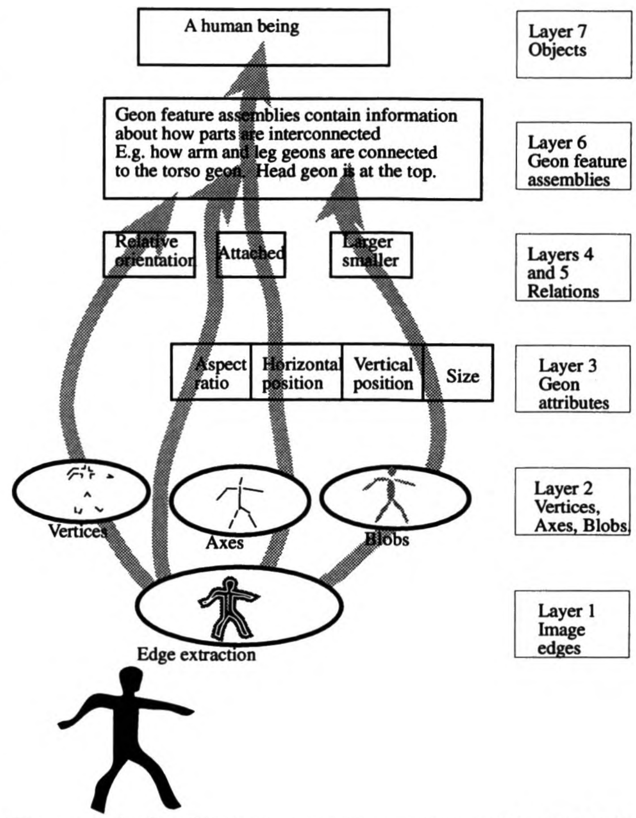

Figure 4 A simplified view of Hummel and Biederman's neural network model of form perception.

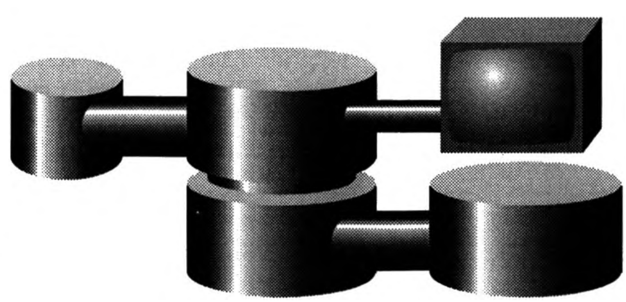

Figure 5. A geon diagram

Example 3: Creating a usable stereo display

We can apply this model directly to data visualization. If cylinders and cones are indeed visual primitives then we can construct diagrams out of geon-like primitives, and such diagrams should be easy to interpret. This concept is illustrated in Figure 5. In order to represent the architecture of a system, geons should be used to represent the components. The system architecture can be represented by the skeletal structure linking the geons. The result might be called a geon diagram. The size of the components becomes a natural metaphor for their relative importance (or perhaps complexity. The strength of the connections between them is given by the neck-like linking structures. In Figure 3 it is clear that there is a large central structure that has two roughly equal sub-components - the two cylinders one on top of the other. Attached to this are three other system components with varying strengths of connection.

Note that everything about Hummel and Biederman's theory could be wrong and the point still valid. Cylinders may have no special role as geons, we may have no special mechanisms to extract connections between sub-components (although this seems inconceivable). Nonetheless, if there is some general process used by all human to understand objects, we can use information about this process to construct effective displays.

Silicon Graphics has for a number of years, made it easy for people to construct stereoscopic views of data - by providing software support and by making the hardware (both graphics boards and monitors) stereo capable. However, despite this, and despite the additional fact that many people regard stereo viewing as synonymous with 3D (this view is badly misguided) there are very few applications that are actually built to use stereoscopic viewing.

The vision research literature on this topic provides some of reasons why it is difficult to provide an effective stereoscopic display using computer graphics. Here is a brief introduction.

Stereoscopic vision refers to the ability of humans and other animals to extract spatial distance information from the fact that the images receive by the two eyes are slightly different. The differences between pairs of features in the two images are called disparities. If the disparity between the two parts of an image becomes too great then *diplopia* occurs. Diplopia is the appearance of the doubling of part of a stereo image when the visual system fails to fuse the images. In the worst case diplopia can occur with remarkably little depth. At the fovea the maximum disparity before fusion breaks down is only one tenth of a degree. When viewing a monitor at normal distances this corresponds to only a few centimeters in front or behind the screen. Depth judgments can still be made outside the fusion area, however, these are less accurate. In real-world situations, factors such as depth of focus and motion minimize the problem of diplopia, but there are severe in stereographic computer displays.

A second problem with monitor based stereoscopic displays is conflict between vergence and focus. When we fixate objects at different depths, two things happen: the convergence of the eyes changes (called vergence) and the focal length of the lenses in the eyes adjusts (accommodation) to bring the objects into focus (Patterson, and Martin, 1992. The vergence and the focus mechanisms are coupled in the human visual system. If one eye is covered the vergence and the focus of the *covered* eye changes as the uncovered eye focuses on objects at different distances. The problem occurs in a monitor based stereoscopic display because all objects lie in the same focal plane regardless of their apparent depth. However, accurate disparity and vergence information may fool the brain into perceiving different depths. Thus screen based stereo displays provide vergence and disparity information *but no focus information.* There is some evidence that the failure to correctly present focus information may cause a form of eye strain presumably because of the coupling described above (Noro, 1993). Certainly it contributes to double image problems.

In view of the above observations, how may we reduce the problems associated with the decoupling of focus and vergence in stereo displays? One solution is to reduce screen disparities. Veron et al. (1990) produces a guideline a that screen based stereo displays should be placed 2.3 meters from the viewer to give fusable images out to infinity. This assumed maximum eye separation of 6.9 cm and that virtual objects are always behind the screen.

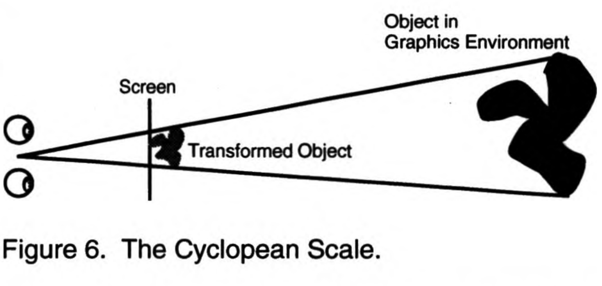

Figure 6. The Cyclopean Scale.

Another, more general solution is the cyclopean scale algorithm illustrated in Figure. ?? In this algorithm the scene is scaled about a mid point between the observer's two eyes until it lies just behind the monitor screen. Following this the eye separation is adjusted to select an optimal range of disparities (Ware et al, 1998). This algorithm has two beneficial effects. The first is that it increases the eye separation *relative* to distant images. A distant large object, such as a mountain will have no useful disparity under normal viewing conditions. However, after a cyclopean scale the eye separation relative to an object will be increased. The second benefit is that vergence-focus conflicts are reduced because objects are brought closer to the screen and hence into the area where the eye can be expected to focus.

Conclusion

The purpose of this brief essay has been to promote the idea that a science of information display can be viewed as a kind of applied area of vision research. This idea is already implicit in much research that appears in the human factors and data visualization literature. However, it is the author's contention that the work that has been done to date only scratches the surface. There is a very large amount of research already available, in perception, in psychophysics and in visual neuroscience that directly applies to vision research. However, the vision research literature is large and employs an arcane vocabulary. It is daunting to the non-expert. In some instances the problem is merely one of knowledge transfer from on discipline to another. In others, new research is needed to bridge the gap between the theoretically focused experiments of the vision scientist and the concerns of the information display designer. Because information displays are far more complex than the displays used by vision researchers, it may be necessary to carry out experiments to confirm that the research results do indeed generalize to the display situation. Another kind of research may be needed to develop new display algorithms based on a knowledge of human perception. These must then be experimentally verified to determine that they are indeed effective.

References

Geertz, C. (1973) The Interpretation of Cultures, Basic Books, New York.

Goodman, N.(1968) Language of Art. Bobbs Merrill, New York.

Gregory, R.L, (1975) "Vision with Isoluminanct Color Contrast: 1 A projection technique and observations". Perception, 6(1) 113-119.

Hochberg, J.E. and Brooks, V. (1962) Pictorial recognition as an unlearned ability. American Journal of Psychology. 75, 624-628.

Hummel, J.E. and Biederman, I., (1992) Dynamic binding in a neural network for shape recognition. Psychological Review, 99(3) 480-517.

Kennedy, J.M. (1973) A psychology of picture perception. Jossey Bass.

Livingston, M.S. and Hubel, D. (1988) Segregation of form, color, movement, and depth: anatomy and physiology, and perception. Science 240, 740-9.

Nakayama, K., Shimono, S. and Silverman, G.H. (1989) Stereoscopic Depth: Its relation to image segmentation, grouping and the recognition of occluding objects. Perception, 18, 55-68.

Noro, K. (1993) Industrial applications of virtual reality and possible health problems. Japanese J. Ergonomica, 29, 126-129.

Patterson, R. and Martin, W.L. (1992) Human Stereopsis, Human Factors, 34(6) 669-692.

Pearson, D., Hanna, E. and Martinez, K. Computer Generated Cartoons. In Images and Understanding, Barlow, Blakemore and Weston Smith, Eds. Cambridge University Press.

Saussure, F. de (1959) Course in General Linguistics. Reprinted by Fontana/Collins.

Triesman, Anne, and Gormican, S. (1988) Feature analysis in early vision: Evidence from search asymmetries. Psychological Review, 95(1) 15-48.

Ware, C. (1988) Color Sequences for Univariate Maps: Theory, Experiments, and Principles. IEEE Computer Graphics and Applicatrions, Sept, 41-49.

Ware, C. Gobrecht, C. and Paton, M.A. (1988) Dynamic adjustment of stereo display parameters, IEEE Transactions of Systems, Man and Cybernetics. Part A: 28 (1) 56-65.

Ware, C. (1993) The foundations of experimental semiotics: A theory of sensory and conventional representation. Journal of Visual Languages and Computing. 4, 91-100.

Veron, H. Southard, D.A. Leger, J.R. and Conway, J.L. (1990) Stereoscopic displays of terrain database visualization, in Proc. SPIE Stereoscopic Display Applications, Santa Clara, pp. 124-135.