

Visualization of Three-Dimensional Form: A discussion of theoretical models of internal representation

Abstract

Current research in perception, cognition, and human-computer interaction is explored as a source of theories explaining the process by which three-dimensional forms are visualized. A number of theoretical models are summarized and compared in an attempt to uncover variables important to future research conducted by the EDG profession. The appropriateness of dynamic versus static imaging techniques to develop a student's visualization abilities becomes a vehicle for applying some of these theoretical models and proposing specific display techniques.

Wiebe, E. N. (1993). Visualization of three-dimensional form: A discussion of theoretical models of internal representation. Engineering Design Graphics Journal, 57, (1).

Introduction

Animation of computer graphics has been a topic of considerable interest among computer users. With the ever-improving price/performance ratio of computer equipment, the dynamic display of graphic information using traditional animation techniques has become affordable to a broad range of computer users. Educators involved in teaching engineering and technical graphics realize the potential of these powerful display tools for assisting students in developing their visualization skills (McCouston, 1991; Wiley, 1990a).

An important skill taught to engineering students is the interpretation of multiview drawings. The mental synthesis of orthographic views into a three-dimensional form can be a difficult technique to learn. One approach to teaching students how to perform this synthesis would be to start with a 3-D object, choose a stationary viewpoint, and rotate the object in small increments between the standard orthographic views. This demonstration can not only be done with a real object, but also with a dynamic sequence of images on a computer. If we generalize this process, we can see how using computer animation techniques can be used to help develop visualization techniques to describe how objects undergo all types of changes over time: in location and orientation and also in shape and size (Zsombor-Murray, 1990).

In addition to generating dynamic displays, computers are just as capable of producing static images much like the ones that have historically been used in teaching engineering and technical graphics. The question arises: are dynamic or static image presentation techniques more appropriate for teaching various visualization skills? A question such as this certainly does not have an obvious and straightforward answer. The visualization of three-dimensional form does, though, strike directly at issues of considerable current interest to the Engineering Design Graphics (EDG) profession. Authors such as Wiley, Miller, and Bertoline have advocated developing curriculum and a methodology of future research based partly on existing visualization research in psychology and education (Miller & Bertoline, 1991; Wiley, 1990b).

Experimental psychology in the areas of cognition and perception have a rich tradition in exploring issues related to the visualization of three-dimensional form. In addition, more recent research in this area has been directed specifically at the design of computer displays. Perceptual and cognitive principles are critical for understanding Human-Computer Interaction (HCI) and, of most interest here, the use of the computer to assist in the visualization of three-dimensional form (Allen, 1982; Haber & Wilkenson, 1982). For these reasons, this paper explores research in perception and cognition for potentially useful insights into understanding what techniques of visual display on the computer might show the most promise in teaching visualization skills to students. The central interest is a better understanding of the mechanisms involved with the mental representation of three-dimensional form as it is understood by researchers today.

Research work across disciplines is an inherently difficult task to achieve. Differing end goals, research methods and semantics all interfere with the

melding of useful information. The EDG profession is not only attempting to make use of psychological research, but many distinct areas within this discipline. Eliot (1987) uses the concept of models as a way of transcending some of these boundaries. In his definition, theoretical models are descriptions or analogies that help us to visualize something that cannot be observed. Models may be substructures of theories, or they may stand on their own. Most importantly, they draw attention to variables of importance for future research. There is a synergy in working with theoretical models in our minds to help understand the representation of physical forms in the minds of students.

Theories of Internal Representations

Imagery

One does not have to look very far to find professionals in science and technical fields who use mental imagery of three-dimensional form to assist them in solving problems. It might be a medical researcher exploring the structure of a protein molecule or a mechanical engineer working on a mechanism design for a car. Though psychologists, philosophers and others have attempted to answer questions pertaining to mental imagery over the years, a major resurgence of interest took place among cognitive psychologists in the early seventies. Instead of looking at the nature of abstract, logical thought, Shepard & Metzler (1971) looked at a more analogical issue: how do we mentally image and manipulate a set of connected cubes.

As the result of a series of experiments that he and his colleagues performed, Shepard (1978) came to a number of broad conclusions. First, though it may be unwise to think of the brain process underlying a mental image as some sort of picture, there is a close, analogous relationship between the mental image and a perceived object. Secondly, subjects are able to formulate a mental image of an object even if it is not physically perceived. Furthermore, subjects use very similar mental processes on the image in both the perceptual and imaginal conditions.

These results confirm what graphics educators have always known; that perceived objects and mental images of objects influence the way we understand 3-D form. If, indeed, there is a close analogous relationship between the perceived object and the mental image formed of it, consideration must be given to the way the object is presented to students if a sound mental image is going to be formed. This becomes particularly critical if it is not the object itself, but a computer generated image of one that is being presented.

Retinal Patterns

For those mental images formed by perceiving an object directly or as a graphic representation, the visual information must be received through the eyes and form a pattern on the retina. Not only is a single, static pattern useful in discerning form, but changing patterns (such as those produced by a computer animation) can create its own unique information. These changing patterns of retinal stimulation — even when the cues are viewed monocularly and showing

no perspective convergence — can produce the perception of three-dimensional form (Wallach & O'Connell, 1953).

Gibson (1966) also concluded that these patterns, or optic array, contain enough information to fully comprehend three-dimensional form in the environment around us. An important point he makes is that visual sampling which forms a retinal image should not be thought of as discrete, but as an overlapping mosaic. The resulting perceived three-dimensional form is not so much a product of what is in each discrete image, but comparisons made between them.

Researchers such as Julesz (1975) agree that considerable processing can be done on these patterns with purely automatic perceptual mechanisms. At the same time there is a distinct limit to how much pure perception can contribute to form recognition. Julesz contends that the purely perceptual pattern detection can only extract the simplest features of a form from an image and that for full comprehension of a form, higher level processing involving conscious thought is necessary.

Marr (1982) is one of many contemporary researchers who have worked on a computational model of perception. He considers that, from a historical perspective, Gibson came the closest to the level of computational theory. Gibson asked the critically important question of how does one perceive constancy in the world around us on the basis of continually changing sensations. We internally work with a representation of an object that constitutes its "real shape", not one based on a single retinal pattern.

It is at this point that Marr diverges from Gibson. Marr agreed that we are actively searching for physical invariants in patterns that add to our understanding of a three-dimensional form. Where he disagrees is in the belief that this problem is exactly and precisely an information processing problem and that the detection of these invariants is vastly more complicated than Gibson stated. Rather, these mental processes can be viewed as a symbol-manipulating system. The computation it supports is the construction of useful descriptions of the visible environment.

Approaching the interpretation of retinal images with many different models, researchers have developed a much clearer understanding of how visual information is processed. Disagreements do, however, exist on a number of fundamental issues. For example, how are sequences of images on the retina combined and how is this information processed?

The Mind as an Information Processor

The analogy of the human mind as an information processor affords some unique approaches to understanding the generation of internal images or representations. Whereas Shepard emphasized the functional aspects of imagery, Kosslyn (1980) developed a model of imagery more structural in nature. The structural mechanisms involved in internal imagery are similar to those used in perception and involve information being stored in specialized locations. Furthermore, the information stores are acted upon by different processes that manipulate and modify the information. Kosslyn took this model to its logical extreme by developing computer programs that simulated this model.

Another approach to this information processing model is to look at the "bandwidth" of the mind to receive and process information (Baecker, 1987). Perception, as an active process, needs resources to process visual information. When these critical resources are limited, our ability to make use of visual information degrades. Though there is a point of diminishing return, if more resources are made available, performance will be enhanced. If no more resources are available, another way to increase performance would be to improve the quality of information being received so not as much processing has to be done. Yet another way to improve performance is to be able to rely on "previous experience" to assist in the processing of the information. The efficiency in which this information can be retrieved from long-term memory will influence how much use it is in processing information currently being received.

Paivio (1983) looked specifically at the role that long-term memory played in imagery. He maintained that long-term memory contained two different but interconnected symbolic systems for coding and representing information. One, imagery, is a memory code for processing spatially synchronous information, whereas a verbal coding system processes in a sequential fashion. This dual-coding model helps to remind us of the uniqueness of generating and manipulating internal representations of spatial forms over working with verbal or written information.

Looking again at the computational model, Marr (1982) postulates that the retinal patterns generated through vision go through a number of stages of processing in order to build up a final internal representation of an object. At the early stages, a *primal sketch* is formed through algorithms that make explicit important information within a retinal image. This information is made up of such primitives as edge segments, boundaries, curvilinear organization and the like. Later stages of processing are *2 1/2-D sketch* and finally *3-D model representation*. As in most of the previous models mentioned, this model takes into account both static and dynamic imagery. Dynamic imagery, in particular, has led to interesting models of how the final three-dimensional form is perceived from a stream of visual stimulation.

Transformation of Stimulation into Form

The models outlined so far hold that patterns resulting from retinal stimulation are transformed into internal representations of the object being viewed. It is a big jump to now try to outline more specifically how the transformation actually takes place. A seemingly profitable approach of a number of researchers has been to try to understand this transformation not for a static object, but one in motion. The question can then be asked, what type of geometric or mathematical model best describes the changes in the retinal image resulting from the object in motion?

Much of our thinking about the description of objects and space continues to be influenced by Newton's and Kant's view of three-dimensional Euclidian space (Eliot, 1987). It is a conception that gives rise to the notion that we can use an extrinsic measurement system (usually linear and two-dimensional) to directly measure the retinal image. A number of researchers have pointed out that when

applied to the real-world case of geometry in motion, this Newtonian approach leads to many artificial complications (Johansson, 1975; Lappin, 1986). If we are to believe that our internal representation of a three-dimensional form is independent of any particular orientation, a geometric model that preserves the perceptual invariance of an object as it undergoes motion will be the most profitable approach. Given the temporal nature of an object undergoing motion, Eliot urges us to consider non-Euclidian models that capture the spirit of Einstein's contention that timeless space does not exist.

Johansson (1975) advocates a complete break from the Euclidian model by proposing a nonmetric geometry based exclusively on relations rather than particular measurements. This geometry, *projective geometry*, is based on the concept that certain relations remain invariant under perspective transformation. Like Gibson, he believes that the perceptual system abstracts these invariances in the changing retinal image and constructs percepts of rigid objects moving in three-dimensional space. Using a model titled *perceptual vector analysis*, Johansson has performed a number of experiments exploring phenomenon surrounding this concept. Though the model proposes a non-Euclidian method for processing the visual information, the resulting perception (from the internal representation) is as we would expect: constant Euclidian shapes in rigid motion in a three-dimensional world.

The computational model of Ullman (1979) also directly addresses the issue of objects in motion but from the standpoint of symbol processing rather than the spontaneous abstraction of three-dimensional form from changing retinal patterns. The first step (not elaborated on here) is the matching of different images representing the same physical object by the correspondence process. Once this is done, one of two distinct interpretation processes is used to transform the images into the representation of three-dimensional form.

The first interpretation process, *Structure from Motion* (SfM), states that we have the capability to infer 3-D structure from a changing image when each static projection by itself contains no useful 3-D information. The fundamental problem underlying this interpretation process is the ambiguity that arises from the lack of a one-to-one correspondence between a given 3-D structure in motion and a given 2-D (retinal) image transformation. The application of *reflective constraints* — uniform use of knowledge that reflects general properties of objects — to 2-D image transformations reduces the range of possible interpretations. The primary constraint proposed is the *rigidity assumption* : that the image transformations should be interpreted with a bias towards understanding it as a rigid object.

Building on the rigidity assumption, Ullman derives his *structure from motion theorem*: "Given three distinct orthographic views of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined." (Ullman, 1979, p. 148) Again, we can see how the thinking of researchers such as Ullman so closely parallels the experiences and work of members of the EDG profession. He goes on to highlight some readily identifiable points. Namely that 3-D structure can be derived from as few as three distinct orthographic views and of the importance that the four identifiable

points are non-coplanar. This theorem forms the minimal nucleus on which the interpretation scheme can operate.

Ullman resolves the fact that the projection of the environment to the eye is perspective and not parallel by proposing that perception of objects is based on a polar-parallel system. Under this system, local features are resolved through rigid interpretations applied to a nuclei of elements using an approximation of orthographic projection. The polar component of the projection system is then applied globally to the object to fully resolve the 3-D form of the object.

The second interpretation scheme proposed by Ullman, *Motion from Structure* (MfS), recognizes that often 3-D structure is perceived from single static frames. Under this scheme, known structure can be used to derive motion in space from a series of static frames. It follows that when static 3-D perception is present, both SfM and MfS operate simultaneously. It is not perfectly clear, however, to what extent the static and dynamic interpretations interact.

This last section demonstrates that when attempting to explain the internal processes that transform 2-D retinal images into an internal representation of a 3-D form, researchers in perceptual and cognitive psychology are relying on tools very familiar to those in the EDG profession. The computational model proposed by Ullman was a good example of this connection. There are a number of elements the models of Lappin, Gibson, Johansson, Ullman and others have in common. For example, they all find the exploration of geometry in motion and the resulting dynamic retinal patterns a valuable approach to trying to understand how we perceive three-dimensional form. Also, just as important, is all of the proposed geometric systems attempt to achieve a resulting internal representation independent of any particular viewpoint of the object. This internal representation is not a collection of individual snapshots of an object in motion but, rather, a singular, canonical form.

Dynamic versus static imaging

Arguments for the use of dynamic imaging for enhancing visualization abilities can be formulated from a number of models outlined in the previous sections. Very directly, Gibson argues that we normally perceive the world dynamically and that this dynamism removes ambiguity in our perception of three dimensional objects (Kaiser & Proffitt, 1989). McCuiston (1991) argues that enhanced spatial schema skills should be measurable through mental rotation tests. His studies show that computer-assisted instructional material containing dynamic imaging improved performance on mental rotation tests relative to similar instructional material containing only static images.

Empirical studies by Barfield, Lim, & Rosenberg (1990) have shown that dynamic image sequences created by the rotation of the subject's viewpoint about a perspective display improved the accuracy of some aspects of the spatial relationship between two bodies. Braunstein (1986) would contend the most important component of Barfield's enhanced display was the ability to rotate the scene rather than the perspective projection. Braunstein found that convergence, the cue that separates perspective from parallel projection was not as important as dynamic rotation in understanding 3-D form. The earlier stated model of

Ullman (1979) would also support this contention. His theorem of structure from motion states both that a parallel projection is perfectly adequate for resolving 3-D form but that at least three distinct orthographic views are needed (easily generated by dynamic rotation).

Many of the previously outlined models make use of comparisons of multiple retinal images to extract information used in the construction of an internal representation of the 3-D object. Where is the raw retinal information stored as these comparisons are made? One explanation of the information processing model holds that as a subject views a sequence of dynamic images, the unstructured visual information is received in a visual buffer (Norman, 1982). This is called a buffer in part because what has been seen is retained for some length of time. This memory trace allows us to compare the current frame of the animation to the ones seen in the near past, giving context to the animation. Though the visual buffer allows for context to some past visual information received, the time span is very short. The information in this buffer decays exponentially: by 1/3 every 100-150 msec.

Some of the information in the animation may be retained by being passed on to short-term or long-term memory. There is, however, a severe limit to how many "chunks" of unrelated material can be held in short term memory (Haber & Wilkenson, 1982). These chunks in turn have a short life-span and will only be passed on to long-term memory if the information can be organized. Because of the nature of long term memory, the information stored there is filtered and incomplete and can be hard to recall relative to short-term memory and the visual buffer.

This model of how human memory functions argues against flexibility in making mental comparisons between what is currently being displayed in an animation and what was displayed at some time beyond the near-past. In our normal everyday lives we have the capability of extracting the necessary information from a constant real-time stream of images to navigate the environment successfully. When teaching our students, however, we are working with an impoverished visual information source — the computer display — and asking them to interpret novel forms at a level of detail with which they are not accustomed to. It is in this environment that the bandwidth of visual information processing may be surpassed.

Techniques for organizing information

On one hand dynamic imaging techniques seem to support those models that use comparisons of multiple retinal images to develop an internal representation of 3-D form. Yet, information processing models seem to suggest that there may be a bandwidth limitation as to how many frames of a dynamic image sequence we are able to effectively process. Employing techniques of organizing and enhancing the information presented on the computer display may overcome some of these barriers (Haber & Wilkenson, 1982; Morse, 1979).

One potential solution would be to present all or some of the frames of a dynamic sequence on the computer display at the same time. Using selective attention, the viewer could make comparisons between different portions of a

single computer display rather than multiple unique displays. Here the potential is to generate retinal patterns that allow comparisons to be made on discrete portions of the image without the temporal pressure of the display changing on the viewer as it would in a dynamic sequence. Paivio's (1983) model indicates that visual information is encoded in a parallel — or synchronous — manner versus serial encoding of verbal information. This would indicate that performance in processing visual information, such as that depicting a 3-D object, would be less sensitive to an increase in information in a single display than would verbal information. This gives support to the potential of simultaneously presenting a larger temporal slice than a single frame of an animation at one time.

Displaying multiple frames of a dynamic sequence simultaneously increases the complexity of the display considerably. If a large number of frames are displayed simultaneously, there is a need to organize the information in ways that the maximum amount of spatial information is derived from it as efficiently as possible. When looking at the interpretation of a 3-D object, Sanford, Barfield, & Foley (1987) and Yuille & Steiger (1982) both found that when figure complexity increased, the object was evaluated on a feature-by-feature basis. This feature-based analysis is akin to symbol manipulation in the computational model of perception (Marr, 1982; Ullman, 1979) or Paivio's (1983) contention that an image is a visual nested hierarchy based, in part, on levels of complexity. These features can be seen as components of the overall retinal pattern. If comparison of these components contributes significantly to the generation of the internal canonical representation, then techniques that enhance the viewers ability to locate and organize these components of the overall retinal pattern should increase visualization performance.

Numerous researchers have proposed employing spatial organization techniques in order to relieve the perceptual and cognitive load on the viewer. In some cases the suggested spatial organization is based on an understanding of cognitive structure (Haber & Wilkenson, 1982), whereas in other cases historic principles of aesthetics and graphic design are employed (Morse, 1979; Tufte, 1990).

One recommended approach — referred to as *Small Multiples* by Tufte — is to organize graphic information based on a Cartesian grid or matrix, much as in a spreadsheet. This differs from a spreadsheet because instead of holding only a single numerical value, each node of the grid contains a unique graphic image. Though the graphic information contained within a node of this grid may not be highly structured, the organization of the overall image on the screen is. This arrangement allows for scanning large quantities of information with little effort. Since all of the nodes are displayed as a static image, the viewer can scan at will, conjecturing and testing visual relationships at a pace suitable for learning.

One of the real powers of a static display of a 2-D matrix is in the holistic nature of the information displayed. Not only does each node of the matrix represent a self-contained image, but the combination of all of the nodes creates a whole new image. If each node is thought of as an element in a pattern, then the subtle shifts between each node can set up a larger, overall perceivable pattern.

This technique has been used in the encoding of multivariate information derived from satellite images (Pickett & Grinstein, 1988). The same techniques can be applied here with multiple views of a 3-D object.

One of the distinct disadvantages of the Small Multiples technique is the loss of spatial context of the individual frames. Each frame has been placed in a uniform spatial order based on the defined axes of the Cartesian grid rather than the actual location of the object in the original reference frame. It may be that preservation of spatial context is an important factor in making visual comparisons between frames. An alternate technique would be to allow the frames of the dynamic sequence to remain on the screen rather than overdraw them. This can be particularly effective if the object underwent a spatial displacement, leaving a trail of previous locations and orientations on the screen (Pollack, 1974).

For both of the above mentioned organizational schemes, various rendering techniques could be used to assist in the coding of the visual information. Color has been found to be a very powerful tool for coding information on the computer screen (Salomon, 1990; Truckenbrod, 1981). One way that color can be used is to highlight critical information. Since the goal is to identify the changes between a sequence of patterns, those elements of the pattern that do change can be highlighted using color coding. Another rendering tool to use is transparency. Though transparency does not have the same degrees of freedom that color does, it can be used in conjunction with other rendering techniques to depict a single variable such as time. Transparency used in conjunction with the trail technique can give the effect of fading over time.

The use of photorealism in rendering 3-D objects has been explored extensively as a method of increasing visual understanding of the object. Among the most common techniques explored are shading surfaces of the object to simulate a light source and hidden line removal. Research has not, however, come to a definitive conclusion as to whether these techniques improve the viewer's judgement of spatial relationships or performance on fundamental cognitive tasks such as mental rotation (Barfield, et al., 1990; McWhorter, et al., 1990; Sanford, et al., 1987).

Specific Examples

There are a number of ways that static and dynamic display techniques can work together. One way of remedying the problem of bandwidth limitations in viewing an animated sequence of images is to record each of the images in a computer file. This history file provides a quick reference to past images seen on the screen. Used like a video recorder, the animated sequence could be played backwards or forwards at any speed and stopped for inspection at any point. Unlike the information stored in short or long-term memory, this historical record on the computer would consist of an unfiltered, complete record of past images seen. Though only one image would be seen at a time, the viewer would have a high degree of control over the speed and sequence of images. This should allow for flexibility in making visual comparisons with holistic information. One of the disadvantages to this technique is the tremendous

amount of computer disk space that can be taken up with these history files. This problem is exacerbated when the resolution and number of supported colors is increased to make more realistic images.

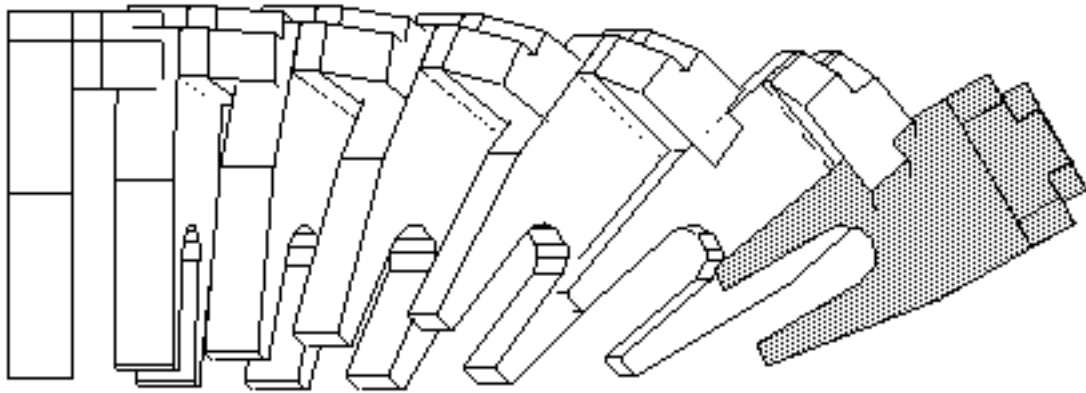


Figure 1.

A modification to the history file technique is leaving all or some of the frames of an animation on the screen. The trail of images on the screen provides for a tracing of the transformation through which the object has gone (Fig. 1). Each of the individual images in the trail has a specific time associated with it. That is, the image is part of a sequence with some images being drawn before and/or after it. This time dimension, in turn, can be represented by a degree of opacity. The most current image could be completely opaque with the older images being progressively more transparent, giving the effect of the images fading over time. An alternative would be to use hidden line removal and to overlap the images in the sequence they are displayed. The most current image could be further highlighted by using a color contrasting with the rest of the images. This technique could be used while a dynamic image sequence is being played or to create static images seen by themselves. In many ways this technique mimics the way the visual buffer is thought to work. The difference is that the temporal control of the visual information on the computer screen can be returned to the viewer. For instance, the viewer can control the number of frames to remain on the screen, the speed of the decay, and time spacing between frames.

Though the tracer technique by and large avoids the computer storage problems of the history file, it still has weaknesses. By its very nature, the images in each frame reside on the screen unchanged in their location from where they would be in the animation. This is needed because the change in relative locations and orientations between frames is the primary information being conveyed by the overall image. The visual structure of the overall image is dependent exclusively on the dynamics of the object being observed. If the object undergoes very little displacement across the computer screen, the images may become bunched to a degree that features of individual frames become illegible.

One solution to the cluttering is to impose an overt ordering to the individual frames of the image sequence. This can be done by organizing the frames into a one or two-dimensional grid or matrix using the Small Multiples concept. Though the absolute location of the object to some global or screen coordinate system is lost, the absolute coordinate system can be used to represent specific variables and to order the visual information in a way that is easily scanned. Students have had many years of training in reading information on a Cartesian grid. This skill can now be put to work at interpreting spatial information and improving visualization skills.

For example, with a 1-D matrix, the axis can represent the time dimension creating a history timeline. Each node of the grid can represent the change in the spatial relationship between two objects. With a 2-D matrix, each axis could represent rotation of the object about orthogonally opposed axes (Fig. 2). A matrix such as this could incorporate a traditional multiview drawing by highlighting the nodes in the matrix that represent standard orthographic views in a different color.

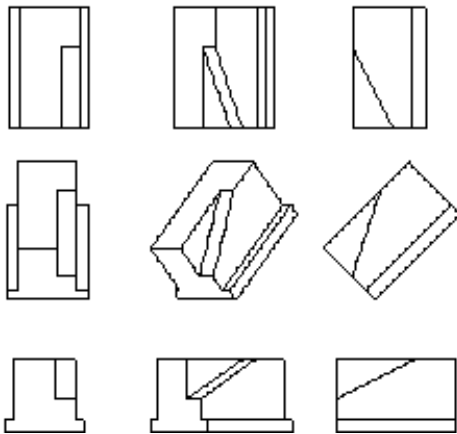


Figure 2.

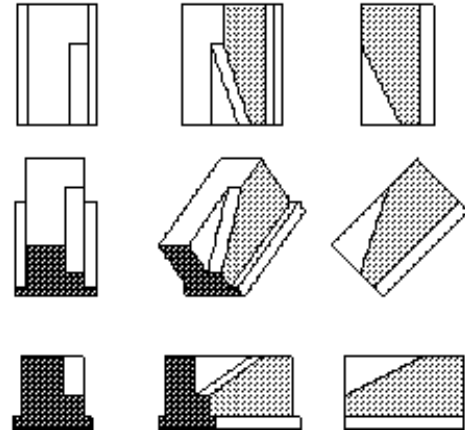


Figure 3.

Another possibility for the use of color coding in the matrices is in highlighting particular edges or faces of the object and tracing their foreshortening as they undergo rotations relative to the viewer (Fig. 3). Length and angle coding, like color coding, are powerful perceptual cues (Morse, 1979). Though length and angle are poor ways of representing an absolute value, they are quite useful at showing relative change. A student can scan a matrix of an object undergoing rotation and perceive the change in length and angle of the same edge highlighted in all of the nodes.

Conclusion

Research in perceptual and cognitive psychology provides a rich source of information for reshaping the EDG curriculum. The models outlined show considerable overlap in interest with the EDG profession and provide mechanisms for furthering our understanding of how 3-D visualization skills develop in students. In addition, these models can be of assistance in evaluating

how computer technology could most effectively be integrated into the classroom.

Central to the models presented is the examination of the retinal images created by the viewing of real or virtual 3-D forms. The perception and subsequent processing of these images form an important component of our understanding of the world around us. This internal processing is facilitated by mental representations we form of the 3-D object being viewed. These representations, in turn, are intimately linked to larger mental structures that organize the visual information.

As mentioned earlier, perception of images are not only processed in an existing mental structure, the process of perception also influences the structure itself. Our goal as educators is to influence the formation of these structures so that the student can most effectively understand and mentally manipulate 3-D form. Hochberg (1964) believes that forms are perceived via their prototypical elements such as edges and corners. These prototypical elements — whether they are edges and corners or other features — may also be the building blocks of the internal representations. For us to perceive an object in motion, photorealistic images are not as important as images that contain the characteristic elements that make up its canonical form. That is, the canonical form becomes the internal structure into which we gather the information from successive images and ultimately understand the nature of an object's three-dimensional form.

A number of specific display prototypes were presented that attempted to embody some of the paradigms outlined in the models. In addition, research on information display techniques was used to propose ways in which visual information might be organized, highlighting those elements that may be critical to the formation of internal representations of 3-D objects. Only further experimentation with these and other prototypical displays, both dynamic and static, will reveal which ones support the best learning environment.

Bibliography

- Allen, R. B. (1982). Cognitive factors in human interaction with computers. In A. Badre & B. Schneiderman (Eds.), Directions in human/computer interaction (pp. 1-26). Norwood, NJ: Ablex Publishing.
- Baecker, R. M. (1987). Cognition and information processing theories. In R. M. Baecker & W. A. S. Buxton (Eds.), Readings in human-computer interaction: a multidisciplinary approach (pp. 299-307). Los Altos, CA: Morgan Kaufmann.
- Barfield, W., Lim, R., & Rosenberg, C. (1990). Visual enhancements and geometric field of view as factors in the design of a three-dimensional perspective display. In Proceeding of the Human Factors Society - 34th Annual Meeting, (pp. 1470-1473).
- Braunstein, M. L. (1986). Perception of rotation in depth: the psychophysical evidence. In N. I. Badler & J. K. Tsotsos (Eds.), Motion: representation and perception (pp. 242-247). Elsevier Science.
- Eliot, J. (1987). Models of psychological space: Psychometric, developmental, and experimental approaches. NY: Springer-Verlag.

- Gibson, J. J. (1966). The senses considered as a perceptual system. NY: Houghton Mifflin.
- Haber, R. N., & Wilkenson, L. (1982). Perceptual components of computer displays. IEEE CG&A, 2(3), 23-35.
- Hochberg, J. E. (1964). Perception. Englewood Cliffs, NJ: Prentice-Hall.
- Johansson, G. (1975). Visual motion perception. Scientific American, 232(6), 76-88.
- Julesz, B. (1975). Experiments in the visual perception of texture. Scientific American, 232(4), 34-43.
- Kaiser, M. K., & Proffitt, D. R. (1989). Perceptual issues in scientific visualization. In SPIE (Eds.), Three-dimensional visualization and display technologies (pp. 205-211).
- Kosslyn, S. M. (1980). Image and mind. Cambridge: Harvard University Press.
- Lappin, J. S. (1986). The representation and perception of geometric structure in moving visual patterns. In N. I. Badler & J. K. Tsotsos (Eds.), Motion: representation and perception (pp. 79-92). Elsevier Science.
- Marr, D. (1982). Vision. San Francisco: W. H. Freeman.
- McCuiiston, P. J. (1991). Static vs. dynamic visuals in computer-assisted instruction. EDG Journal, 55(2).
- McWhorter, S. W., & et al (1990). Evaluation of 3-D display techniques for engineering design visualization. In Proceedings of the ASEE Engineering Design Graphics Division mid-year meeting, (pp. 121-129). Tempe, AZ:
- Miller, C. L., & Bertoline, G. R. (1991). Spatial visualization research and theories: Their importance in the development of an engineering and technical design graphics curriculum model. Engineering Design Graphics Journal, 55(3), 5-14.
- Morse, A. (1979). Some principles for the effective display of data. (ACM) Computer Graphics, 13(2), 94-101.
- Norman, D. (1982). Learning and memory. San Francisco: W. H. Freeman.
- Paivio, A. (1983). The empirical case for dual coding. In J. C. Yuille (Eds.), Imagery, memory, and cognition (pp. 307-332). Hillsdale, NJ: Erlbaum.
- Pickett, R. M., & Grinstein, G. G. (1988). Iconographic displays for visualizing multidimensional data. In Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, (pp. 514-519).
- Pollack, I. (1974). Detection of changes in spatial position: IV. Multiple display fields, display aiding and interference. Human Factors, 16(2), 93-116.
- Salomon, G. (1990). New uses for color. In B. Laurel (Eds.), The art of human-computer interface design (pp. 269-278). Reading, MA: Addison-Wesley Publishing.
- Sanford, J., Barfield, W., & Foley, J. (1987). Empirical studies of interactive computer graphics: perceptual and cognitive issues. In Proceedings of the Human Factors Society - 31st Annual Meeting, (pp. 519-523).
- Shepard, R. N. (1978). The mental image. American Psychologist, 33(2), 125-137.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171(Feb.), 701-703.

Truckenbrod, J. R. (1981). Effective use of color in computer graphics. (ACM) Computer Graphics, 15(3), 83-90.

Tufte, E. R. (1990). Envisioning information. Cheshire, CT: Graphic Press.

Ullman, S. (1979). The interpretation of visual motion. Cambridge, MA: MIT Press.

Wallach, H., & O'Connell, D. H. (1953). The kinetic depth effect. J. of Experimental Psych., 45(4), 205-217.

Wiley, S. E. (1990a). Computer graphics and the development of visual perception in engineering graphics curricula. EDG Journal, 54(2), 39-43.

Wiley, S. E. (1990b). A hierarchy of visual learning. EDG Journal, 54(3).

Yuille, J. C., & Steiger, J. H. (1982). Nonholistic processing in mental rotation: some suggestive evidence. Perception and Psychophysics, 31, 201-209.

Zsombor-Murray, P. J. (1990). 2-D and 3-D CAD: complements to visualization. EDG Journal, 54(3), 17-29.