PERSONALIZED PERSPECTIVES IN 3-D ASSEMBLY

By

Lawrence Scarritt Stead

S.B. Massachusetts Institute of Technology
1975

Submitted in Partial Fulfillment
of the Requirements for the
Degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1978

Signature of Author______________________________
Department of Architecture
June 19, 1978

Certified by______________________________
Nicholas Negroponte
Associate Professor of Computer Graphics
Thesis Supervisor

Accepted by______________________________
Professor N. John Habraken
Departmental Committee for Graduate Students

The work reported herein was sponsored by the Office of Naval Research, Contract Number N00014-75-C-0460.

DISCLAIMER NOTICE

The accompanying media item for this thesis is available in the MIT Libraries or Institute Archives.

Thank you.

Personalized Perspectives in 3-D Assembly

Table of Contents

Personalized Perspectives in 3-D Assembly

By

Lawrence Scarritt Stead

submitted to the Department of Architecture on June 19, 1978 in partial fulfillment of the requirements for the degree of Master of Science.

Abstract

This thesis aims to establish some basic tools for investigating people's ability to interpret various types of renderings of 3-D solids, and how a rendition aids or inhibts visualization and "visual problem solving". Image generation is performed by a powerful computer graphics facility, versatile enough to permit several different display schemes to be implemented and evaluated.

A general purpose system was created which supports flexible generation of both single frames and animated sequences. Objects are given colors and located in space by a simple "script". This script can then be used to make images in each 3-D representation style for comparison.

The system's capability is used to create animated 3-D sequences showing the assembly of a multi block puzzle into a cohesive figure. It is a demonstration aimed at raising visualization issues pertaining to 3-D assembly tasks. A color videotape of the assembly using the implemented display techniques is planned.

Thesis Supervisor: ______________________
Nicholas Negroponte

Title: Associate Professor of Computer Graphics

CHAPTER ONE

Personalized Perspectives

1.1 The Maintenance Problem

It is a sign of our times that many of the goals and tasks of contemporary society are so complex that the digital computer, in its various guises, has become an essential partner in any effort to sucessfully manage these challenges. We believe that yet another field in which computers will soon be intimately involved is repair and maintenance of sophisticated equipment.

State of the art equipment provokes an enormous disparity between the people who design it (perhaps a highly specialized team of Phd´s), and the maintenance technicians who have relatively little knowledge about any given device. Although at the present time many extremely difficult artificial intelligence issues remain to be addressed, we ultimately envision a versatile computer assistant aiding the technician in the trouble shooting and repair tasks.

An important component of such a system will be a graphics display capable of showing all manner of blueprints, text, parts, animated sequences, and so forth. In particular, when conveying 3-D information about a mechanical assembly (how pieces fit together), strategic orders for assembly and disassembly will be an important subclass of the visual information the graphics system will be called upon to present.

We postulate that people vary dramatically in their visual abilities such as spatial thinking, pictorial memory, and geometrical transformations. For example, it seems quite plausible to assert that as a group, architects are "more visual" than accountants but one is hard pressed to quantify the factors involved. Additionally, people can simply be accustomed or are more comfortable with certain representations of the same thing, ie, an engineer might prefer to examine the blueprints of a builing, whereas a sculptor would like a 3-D model he can look at and perhaps even touch. Systems must be built which can accomodate or "personalize" their presentation to each individual's needs and preferences.

This thesis is concerned with a specialized aspect of visualization: the subproblem of representing 3-D solids on a computer graphics display. Several different techniques have been implemented for comparison, in a manner allowing for expansion as new ideas arise. We have used a style of computer graphics we believe is exemplary of ones which will be used in the future, namely raster scan digital video, described in the next section.

1.2 Raster Scan Computer Displays

The type of display used in this project is often refered to as "raster scan", or "computer video". It is important to understand the distinction between this new technology and two earlier, "parent" technologies, vector computer graphics and standard commercial video.

In the beginning (of computer graphics), around the early 60's, the phrase "computer graphics" was interchangeable with "vector graphics". Vector graphics meant systems which could draw vectors (straight lines) between any two points on the screen. A table of all the vectors comprising a picture, known as the display list, was kept, and one vector after another was drawn, until the display list was exhasted. Then the entire process would be repeated, at rates fast enough to avoid flicker. This scheme had several serious disadvantages. The CRT beam had to be randomly positionable, requiring costly and complex electronics to drive it. As a picture became more detailed, it would take longer to go thru the display list, causing the frame repetition rate to drop, making flicker more noticeable (storage tubes avoid this particular problem but have other limitations). Most displays were limited to one color, a monotonous oscilloscope green. Finally, the worst sin of all to some people, was the inability to display areas, textures, and photographic images.

In fairness however, vector displays had one saving grace; without it they would surely be extinct by now. Real time dynamic effects, ie changing the positions of lines without redrawing the entire display, is a trival operation with vector graphics. One simply modifies the appropiate part of the display list. The ease with which this can be done is independent of the complexity of the picture. In raster scan systems the exact opposite is the case.

Commercial video technology overcomes many of the previously mentioned disadvantages. The beam scans the screen in horizontal sweeps, going from top to bottom, its intensity modulated as it scans. This pattern is always the same from frame to frame, simplfying the deflection eletronics. The scan is repeated 30 times a second regardless of picture complexity, thus eliminating flicker (although stroboscopic effects can be seen on rapidly moving objects). Therefore realistic, full color, photographic images can be shown. However, commercial video operates entirely in the analog domain, that is, picture information is carried by signals of continously varying amplitude. This makes it very difficult to arbitrarily make or modify images.

Computer generated digital video, by constrast, offers complete, explict control over every resolvable picture element, or pixel, on the screen. There is an n-bit number in the computer memory (sometimes called a frame buffer) corresponding to each pixel specifying what color should be displayed. Thus, by mathematically operating on the frame buffer, the video image can be generated or modified in any manner.

The raster scan system used for this thesis is based around an Interdata Model 85, a fast, 16 bit minicomputer with dynamic control store which permits microcoding of critical subroutines for improved speed. The 85 was modified and augmented extensively in house by the Architecture Machine Group to give it its video capabilities, some of which are as yet unmatched in commercially available equipment. Its frame buffer is

206k bytes large, and can run at one, two, four, eight, or sixteen bits per pixel. It can also perform translation and zooming on an image in real time.

CHAPTER TWO

3-D Display Techniques

2.1 Introduction

One division of 3-D display techniques is that of section 1.2, namely, vector vs. raster scan. Vector machines are limited to vector 3-D algorithms, while raster scan machines can use both classes of algorithms. Of course, vector machines are more efficient than raster scan machines at displaying lines, but for our purposes this difference is irrelevant, so we use raster scan throughout.

Several different schemes are described in this chapter, and four of them are currently implemented. They are:

1) Simple wire frame
2) Hidden lines
3) Hidden surfaces, no shading, with transparency effects

4) Hidden surfaces, shading, and no transparency.

2.2 Vector 3-D

A wire frame display of a 3-D solid is created by projecting each edge of the 3-D solid onto the 2-D plane of the viewing screen. Wire frame displays are easy to implement and their computational burden is relatively small. Also, an object's internal decription is merely a list of its edges, unlike hidden surface algorithms which require "topological" knowledge about an object. Generating this additional information can prove to be a substantial burden.

The main catch with wire frames is that they are generally unacceptable for viewing anything but the simplest objects. There are basically two disadvantages. One, as mentioned earlier in reference to vector displays, is that areas (faces) are outlined instead of being explicitly displayed. More serious still is the display of every edge in the solid, from any viewing angle. That is, the lines that should be obscured, or "hidden", remain visible, an x-ray view in a sense. Thus it is difficult to interpret these pictures, and they can sometimes be ambiguous enough to permit two different, equally valid perspective orientations of the solid.

2.2.1 Z Intensity Depth Cueing

An easily implemented aid to depth perception, of somewhat limited value, is modulation of the intensity of drawn vectors with respect to depth. The further the vector in z, the darker the intensity assigned to it. The drawbacks to the method are twofold. First, the effectiveness drops with increasing picture complexity. Second, if the vectors all fall within a narrow range of z values, the change in brightness may be too small to be of use.

2.2.2 Hidden Line Removal

Dissatisfaction with the clutter and interpretation problems of simple wire frame displays led to the development of so called "hidden line" algorithms. These routines eliminate any edges or portions of edges which should be obscured by the perspective orientation. Hidden line removal causes an enormous increase in the "visual intelligibility" of an object.

However, hidden line packages per se are typically complex to implement and quite expensive computationally. On a vector scope, one has no choice. The analog to "hidden lines" on a raster scan machine, and substantially easier to come by, is known as a "hidden surface" algorithm. We shall see shortly that a hidden surface routine can be used to simulate a hidden line routine as a special case. This is how hidden lines were implemented for this thesis.

2.3 Raster Scan 3-D

Hidden surface algorithms exploit the raster scan's ability to display areas. In this case, the areas are the faces of the solid we wish to be rendered. Faces, or portions of faces that should not be seen from a given viewing angle are in fact not displayed. This produces a convincing, readily interpretable image, in full color. Advanced systems of this sort can produce extremely detailed and realistic images, and are finding uses in many fields such as scientific animation and real time flight simulators.

One immediate advantage of the hidden surface approach is the ability to render a top layer of solids transparent relative to an opaque bottom layer. Transparency is a helpful method for seeing relations between adjacent solids that are normally blocked. Transparency effects are easily created by using a weighted average between the color of the top layer and the bottom layer at each point. In equation form we have

displayed_color=alpha*bottom+(1-alpha)*top.

Alpha is the percentage of transparency running from 0 to 1, where a value of 1 makes the top layer invisible, and a value of 0 makes the top layer opaque.

2.3.1 Hidden Surfaces without Light Source Shading

If we display an object with all its faces the same color, we may have trouble distinguishing faces, which is crucial to achieving a 3-D effect. One way around this problem is to enhance the edges of the object. This looks like a hidden line display superimposed on the hidden surface display (by setting the faces to black, we get hidden lines from hidden surfaces, alluded to in section 2.2.2). This is an appropiate technique when a limited number of colors are available.

2.3.2 Hidden Surfaces with Light Source Shading

Instead of using the same color for each face of a solid, we may add a very effective depth cue by simulating a light source, and shading each face according to the angle its normal makes with the light rays striking it. If the light source is at infinity, then each face will have a uniform shade, since all light rays hitting are parallel to each other. Putting a light source at a finite distance from an object causes the shading of a face to be non-uniform, and thus more difficult to compute. The only disadvantage of using shading is that it requires many more distinct colors than not shading.

2.3.3 Hidden Surfaces with Shading and Shadows

A refinement of the shading technique of the previous section is to show shadows cast by parts of an object blocking a light source. This method is substantially more complex than simple shading. Its value in aiding 3-d visualization is difficult to assess at this time.

CHAPTER THREE

SOMA Puzzle Animation

3.1 The SOMA Puzzle

The SOMA puzzle was invented by Piet Hein, during a lecture on qúantum mechanics which failed to hold his attention. The intriging puzzle went on to achieve a following via Martin Gardner´s Mathematical Games column in Scientific American. For our purposes it has various properties, detailed in the Appendix, which make it an attractive medium for the study of 3-D visualization problems.

Briefly, the puzzle consists of seven different pieces, solid blocks, each constructed by joining together three or four unit cubes on their faces. Although the pieces themselves look quite irregular and unwieldy, they can be assembled into many simple and pleasing shapes. Perhaps most surprising is the hundred and forty different ways they can be put together to form a 3x3x3 cube (if you count symetries and rotations there are over a million ways). On the other hand, some figures may be built in only one way, or phrased differently, a given "target figure" may have only one "solution configuration".

3.2 Animation Techniques

So far, a collection of programs has been developed which will display arbitrary static configurations in several different styles. These are enhanced in order to produce animated sequences of the SOMA blocks in dynamic motion. This is accomplished by a simple application of a well known computer animation technique, key frame animation.

The idea in key frame is to specify a start frame, an end frame, and the number of intermediate frames in which the transition from start to end takes place. The computer interpolates between the two key frames to create the intermediate ones. In general, key framing is quite difficult, since shapes can change between key frames. In our case however, the shape of the SOMA pieces are fixed so interpolation need only be done on spatial position, orientation, and color.

A program was written which can find all solution configurations corresponding to the target figure. The output of the solution finding program is each block's spatial position and orientation in the target figure. This description can then be used as input to the display programs described in Chapter 2.

Therefore, showing a sequence of scattered pieces comming together only requires giving the parameters for the initial disjoint configuration and the completed target figure. Pieces can move simultaneously, in subgroups, or one at a time. One disadvantage (but fun to watch) of this simple interpolation scheme is that since blocks move in a straight line between start and end frame, they pass thru each other, ignoring real world physical constraints.

Currently real time sequences are made by using a video disk capable of recording a single frame of video at a time as they become available from the 85. The video disk is then played back at standard rates, and recorded onto video tape. For images of moderate complexity, the 85 does not posses the computatial power to generate frames in real time, ie 30 times a second.

Appendix

Section 3.3 of the ONR proposal,

Idiosyncratic Systems: Personalized Perspectives in the Explication of Mechanical Assemblies

The adoption of a "blocks" world is classic in AI paradigms for "computer vision".** A feasible paradigm for the study we are proposing is the advised adoption of a blocks-world paradigm, not for the sake of teaching people how to make block figures and to solve block puzzles, but because we wish to understand better how to aid people in their functions re of a 3-D mechanical world. The aim is not that of machine understanding of a blocks world, but insights into explication via machine.

Block puzzles are a rich paradigm for mechanical assembly, yet holding certain parameters constant so that study can focus on central problems of presentation and of aiding understanding.

**Cf. Winston, P.H. (Ed) The psychology of computer vision. New York: McGraw-Hill, 1975; Kuipers, B.J. "A frame for frames". In Bobrow, D.G. & Collins, A. (Eds.), Representation and understanding. New York: Academic Press, 1975

Consider the classic wooden blocks puzzle where the aim is to form a cube from about ten irregularly formed component pieces. As a paradigm, the puzzle has much to recommend it. The aim of the activity is well- defined: the assemblages of the puzzle. The process of assembly involves selection at every sub-stage of assembly. There is a component of serial dependency built into the puzzle: certain pieces will "fit" only into sub-assemblies formed earlier. There is structural dependency, in that the insertion of a piece can be impossible because part of the assemblage has an edge or jutting part that prevents it.

The placement of a part erroneously may or may not be signaled to the puzzle-solver by such direct cueing in that it is often possible to place a piece that doesn´t belong arise only later, out of global considerations from the "look" of the assembly thus far.

With classic wood puzzle there is usually only <u>one</u> solution: there is only one pattern of assembly that meets the end criterion of a <u>cube</u>. In constrast, consider the cube puzzle game, SOMA, invented by Piet Hein, and marketed in America by Parker Brothers Games, Inc. The cube to be assembled is a 3x3x3-unit cube, made up out of 7 irregular pieces which are combinations of the unit-cube. Again, there is one solution: a cube. But now there are many paths to that end. It has been calculated that there are exactly 1,105,920 ways to form the cube, if one considers all solutions that are reflections of each other or that can arise from each other by whole rotations of the cube or of its several parts.

By using both the simple classic wooden cube puzzle and the SOMA cube, one can systematically vary aspects of an assembly task. The assembly task can be one that admits of a unique solution, or of many. The interaction of variable means to task completion (cube) with mode and strategy of explication is at the heart of this proposal. A further bonus from the SOMA block pieces is that additionally, the end goal of the assembly can be systematically varied as well as there are many target/figures, ranging from the very simple to the very difficult, that can be constructed from the pieces.

In general, then, the wood blocks and the SOMA blocks comprise a problem set where there can be clearly defined ends, and either a unique route to any end, or many routes. Surely, the availability of <u>many ends</u> to a solution should modify and temper the methodology of explication. But all along, the block sets afford the experimenter with tangible criteria for evaluating subject performance: namely, a clearly defined end, i.e., an overall success/failure measure; a "time to completion" criterion measure, both for the total target figure, as well as for any defineable sub-assembly.

Given the basic suject task, cube assembly (or, possibly, "repair" of a "damaged" block configuration), variants in presentational strategies and modes of explication can be brought to bear: verbal inputs, ranging from hinting strategies, to explicit directions; diagramatics; animation sequences, perhaps bearing visual "Hints"; examples and counterexamples; the intersection of graphical and verbal strategies.

Apart from factors affecting explication of an on going assembly, the presentation of the problem initially, and re-presentation of the problem as it undergoes attempts at solution, bear study. The analogue here with "real-world" mechanical problems is that differing approaches of explication may be "indicated as assembly progresses". For example, the total configuration ("look" or gestalt) of the sub-assembly at stage a of assembly may be of a radically different order of visual complexity than at stage b. We are all familiar with the effect, toward the end of doing a jigsaw puzzle of "the puzzle beginning to solve itself", while recalling how the early stage was so easy (Do the edges first!), and how slow-going were the middle stages. The total "look" of the problem, how it is re-presented, can have a profound effect upon prospects for solution, (Cf. Schwartz, S. H. "Modes of representation and problem solving: well-evolved is half solved", Journal of Experimental Psychology, 1971, 91 (2), 347-350).

This proposal has alluded to "personal perspectives": in the visualization aspects of item assembly. This is no idle metaphor. In his recent doctoral dissertation in psychology at Brandeis University, John Banjafield showed dramtically how a seemingly small difference in problem presentation can dramatically affect prospects for solution. A simple "block puzzle" was used, two small identical pyramids of wood, which when put together in the proper way formed a larger pyramid of similar form.

The blocks were presented to different subjects in either of two manners: 1) where the blocks were presented on a table-top; 2) where they were supported above the tabletop by a small holder. Otherwise, the relative position of the blocks was the same in both conditions. The task was to examine the blocks without touching them, and then to pick them up and directly to put them together to form the larger pyramid. (i.e., they could not be "handled" prior to assembly, so that the only experimental difference was the "look" of the blocks.) The data showed a striking superiority for the above-the-tabletop, its "palpability" or look of firmness, seemed to prevent subjects from performing the necessary mental rotation and manipulation which involving part of one of the blocks extending down through the tabletop in order to form the proper figure for solution. Whatever the exact nature of the underlying phenomena, the functional difference was dramtic indeed, with much greater success for these subjects who had the presentational mode that was compatible with the sort of mental operation that was a sufficient condition for solution.

Problem variants for a SOMA blocks paradigm can include user-initiated and user-controlled operations such as variations in the transparency of the component pieces of the portrayed blocks. Block parts could be color-coded or not. The conjoint variation of color coding and varying level of transparency can be examined re its effects on performance.**

Perspective can be varied, literally providing personalization in this presentational dimension. Together with observer control of tranparency, color, rotation, the ability of the observer to control the aspect of view as well as subtle but possible critical differences in projection (axonometric; parallel; etc.), can be powerful aids to 3-D mechanical comprehension.

Lastly, with block puzzle paradigmatics, as well as with actual "real-world" assembleges, the issue of serial vs. parallel conceptualization is of more than passing interest. For example, Reed** has studied the problem of whether subjects can successfully judge whether or not the second of two sequentially presented patterns is a part of the first pattern.

**Reed, S.K. Structural descriptions and the limitations of visual images. Memory & Cognition, 1974 2(2), 329-336

His results suggested that subjects code patterns as structural descriptions, and subsequently find it hard to recognize a part of the pattern that does not match the description. Further, the results suggested that a structural description is comprised of a combination of visual and verbal codes, and that visual images lack detail when not supported by verbal detail. For our purposes these findings suggest two things: 1) that an operator's mental orientation during an assebmly task involving 3-D visualizable components may be an intimate mixture of both visual "pictures" and verbalized "procedures", and that explication protocols need to take both aspects into account; and, 2) that the direct visual aspect of any sub-assembly accomplished so far in the course of an overall task may not, of itself, model for the assembler what he has done thus far. The assembler's internal mental model of what he has thus far done, (i.e., "where he is") in problem space, may be a radically different compound of procedural descriptions and visual images. Exploration of the SOMA blocks world with our graphics facility may open that way for comprehension of what this amalgamation might be, how to model it, how to explicate the balance of the to-be-completed task in compatible modes and media.

(The author of this section is Dr. Richard Bolt, Architecture Machine Group.)

REFERENCES

Bouknight, W.J., Kelley, K.C. "An Algorithm for Producing Half-Tone Computer Graphics Presentations with Shadows and Movable Light Sources." SJCC 1970, AFIPS Press, Montvale, N.J.,1.

Entwisle, J. "An Image Processing Approach to Computer Graphics." IEEE Conference on Computer Graphics, July, 1974.

Negroponte, N.P. "Raster Scan Approaches to Computer Graphics." Computers and Graphics, 2, 1977, 179-193.

Negroponte, N.P., Bolt, R.A., and Tom, V. "Color Transparency Effects From Mosaics of Opaque Colors." Report to U.S. Army Research Office, Contract number DAAG 29-76-C-0037.

Newman, W.M., and Sproll, R.F. Principles of Interactive Computer Graphics. McGraw Hill, Inc. 1973.

Sutherland, I.E., Sproll, R.F, and Schumacker, R.A., "A Characterization of Ten Hidden Surface Algorithms." Report to the Office of Naval Research, contract number N 00014-72-C-0346.

Acknowledgements

Thanks to:

The Office of Naval Research for sponsoring this work,

Bill Donelson for helping me
use his hidden surface routines,

The Architecture Machine for being,

Nicholas Negroponte, my thesis advisor,
for making it all possible,

and Bonnie Biafore, who got me into to
all this in the first place,
and helped me get out as well...