6-1988

Using the Bat: a Six Dimensional Mouse for Object Placement

Colin Ware
University of New Hampshire, Durham, colin.ware@unh.edu

Danny R. Jessome
University of New Brunswick

Recommended Citation

USING THE BAT:
A SIX DIMENSIONAL MOUSE FOR OBJECT PLACEMENT

Colin Ware and Danny R. Jessome

School of Computer Science
University of New Brunswick
Box 4400, Fredericton, N.B.
CANADA E3B 5A4

(cware@UNB.bitnet)

ABSTRACT

Placement in space is inherently six dimensional. An object can be translated in x, y and z cartesian coordinates, and it can be rotated about three axes to change its orientation. A six dimensional sensor was configured as a one button six dimensional mouse (which we call a bat) and interfaced with an IRIS workstation thus creating an environment for investigating the problems of object placement. A software workbench was built which allows the manipulation of hierarchical scenes displayed on a monitor. The features of the workbench are described together with experiences using the bat in a variety of interaction modes.

KEYWORDS: Graphics, Input devices.

1. INTRODUCTION

The problem of interacting in three dimensional (3-D) space is still open. Placement in 3-D is actually a six dimensional (6-D) operation requiring three variables to specify location and three more to specify orientation. Therefore, to place an object in 3-D space with a single interaction requires a 6-D input device. In this paper we report on protocols for manipulating graphic objects using a 6-D device.

Previous studies of 6-D placement have adopted a variety of strategies and we briefly review a few of these to place our work in context. An obvious, but technically difficult solution, is to place the user inside the graphics environment. This has been done with complex helmets that transform the graphics environment with the user's movements so that the user feels he is locomoting in a space filled with synthetic virtual objects (Sutherland 1968; Fisher et. al, in press). A smaller scale implementation of the same idea allows users to place their hands in a graphics environment (Schmandt 1983). This is achieved with the use of Pieso electric goggles and a half silvered mirror. In the above systems hand held spatial sensors are used to allow the user to manipulate the graphic objects creating an interface which, as far as possible, mimics our normal interactions with solid objects.

Unfortunately, the illusion is far from complete and one of the more troubling artifacts is that the user's hand appears to pass through objects and is visible when it should be occluded. This is confusing and hinders manipulation of the environment.

A different approach to 6-D placement is to use simpler technology and by the use of clever software make an effective interface. Thus Evans et al.(1981) used a tablet to input rotations by mapping x and y hand translations to object rotations about y and x and mapping a stirring movement to rotation about z. This, of course, only provides 3 of the 6 dimensions needed for placement and therefore a change of state is necessary to provide positioning.

A crucial difference between the approaches listed above is whether the user's limbs are placed in the graphics environment. Concerning this point it is worth noting that the light pen which has a spatial correspondence with the displayed objects has found less favor than the mouse or the digitizing tablet which do not. Instead what seems to be more important is correspondence between the movement of the input device and the motion of the manipulated object.

We believe that for most applications there is little point in placing the user's limbs in the graphics environment. Nevertheless, this does not rule out the desirability of a 6-D sensor; if one can be readily obtained it is likely

to provide a better and more natural interface than a tablet or a mouse, no matter how sophisticated the inter-action protocol.

We have initiated a program of investigation into the problems of 6-D placement using a 6-D variant on the conventional mouse which we call the "bat" because it is like a mouse that flies. This encodes relative position, like the mouse, but delivers data in all six dimensions needed for object placement. Our goal is to evaluate the bat to determine how well it is suited to placement operations. However, a hardware device such as the bat cannot be evaluated independently of the task and the protocols implemented to allow the user to perform the task. Thus, we also are involved in investigating various interaction modes to find which allows the most natural and fluid dialogue with a 3-D world represented within the computer.

To focus our study we have isolated a single primitive operation, "6-D placement" which we feel embodies the most significant problems associated with spatial interaction. We use the term placement to cover the six dimensions of positioning and orienting. Also, since placement is inherently the placement of one object relative to other objects we chose to study placement in the context of a hierarchically constructed scene. In this environment, a child-object can be placed with respect to its parent object and the movement of an object causes the parallel motion of descendant sub-objects in the hierarchy. We feel that this generic task provides a rich paradigm for the study of scene manipulation.

There are two conceptually distinct parts to the problem of 6-D placement, namely visualization - how to make it possible to accurately perceive the spatial relationships of objects in the 3-D environment, and manipulation - how to make it possible to comfortably manipulate parts of the enviroment.

1.1 Visualization.

The traditional draftsman's tools for visualizing 3-D scenes are three orthogonal orthographic projections and a rendered oblique perspective view. This is a static arrangement which is not suitable for dynamic interaction since it requires a synthesis of the separate views to realize the scene. Accurate and rapid 6-D placement must depend on good rendering. However, it is unfortunately impossible with current technology to provide a fully accurate rendering of an arbitrary scene moving in real-time. For example cast shadows are important depth cues which are computationally expensive. However, the real time rendering of perspective, the elimination of hidden surfaces, stereopsis, and the kinetic depth relationships

have all been achieved in experimental systems and some commercial systems.

Of particular interest are kinetic depth and stereopsis. The kinetic depth effect is the name given to the phenomenon whereby a flat pictorial projection of a 3-D scene appears strongly 3-D when it is the projection of a rotating scene (Braunstein, 1976). Some systems have a feature whereby the depicted scene can be set in continuous smooth rotation for viewing purposes. Stereopsis is the name given to the brain's ability to extract 3-D relationships from the different views of the world provided by the two eyes. In a computer graphics system this is achieved by displaying the scene to one eye, rotating it about a vertical axis by a few degrees and redisplaying it. This necessitates some way of separating the images presented to the two eyes and in a workstation environment the viewer must wear special goggles, either containing shutters, polaroids or red and green filters (Foley and Van Dam, 1982).

1.2 Manipulation

The problem of manipulation is that of choosing a suitable 6-D interface device and an interaction protocol. The approach we have adopted is to create the bat; a logical extension of a mouse into six dimensions. With this device it is possible to achieve a natural 1:1 correspondence in both translations and rotations between the bat motion and the motion of the displayed object.

1.3 Research Strategy

Although the above issues of visualization and manipulation are conceptually distinct they are by no means independent in practice. The choice of visualization cues in many cases has a direct bearing on the choice of manipulation techniques. To take an extreme example; if the scene is to be visualized by three orthographic projections, then a conventional mouse may well be the best input device. On a more subtle level it is by no means clear that the best interface for a wire-frame scene will also be the best for a scene made up of filled polygons. From the perspective of studying the problem of placement, the lack of independence of variables means that there are far more variables involved in the placement problem than may be systematically studied in all possible combinations, even by a researcher most dedicated to "hardening" the discipline (Newell and Card, 1985).

To address this problem our strategy for studying placement has three levels. At the first level, we build the basic interface according to what seem to be uncontentiously sound principles without formal study of alternatives. Our protocol for dragging is an example which we simply copied from other successful interfaces. At the second level, we implement various modes of

interaction, assemble phenomelogical reports about their success, and assess their success or failure on that basis. At the third level, we plan formal empirical studies of different placement protocols in conjuction with various visualization aids. This paper is a report of first and second level activity.

We have constructed a hardware and software testbench specifically to address a number of key issues concerning 6-D placement. In designing this testbench we have been forced to make a large number of design decisions which contain implicit assumptions about the most effective protocols. What follows is a description of the environment we have constructed - which embodies the implicit assumptions - and also the variables we have chosen explicitly to investigate.

2. HARDWARE ENVIRONMENT

The workstation configuration is diagrammed in figure 1. The basic components are a Silicon Graphics IRIS 2400 workstation and a bat. The IRIS workstation provides the ability to do real-time spatial manipulations of objects which either consist of wire-frame figures, or filled polygonal objects.

We built our bat using a Polhemus 3SPACE[1] Isotrak which is a 6 degree of freedom spatial sensor. This device signals orientation and position relative to a fixed source. The source is mounted in a box to the right of the user's chair and the sensor is mounted in a rounded block of wood with a button on the top. The 3SPACE yields 9 bits of resolution in each of the 6 variables which translates to approximately 0.13 inches of position resolution and 0.7 deg of angular resolution. The position range is given in figure 1 as is the overall workstation configuration. The static accuracy of the three space is worse than its resolution and is distorted by metal objects (such as a monitor) in its vicinity (Schmandt, 1983). Fortunately, this distortion is not a problem for our application because only relative positioning information is used and over short distances this is negligible. A good features of the 3SPACE is that it can be configured to deliver quaternion output which simplifies the programming of rotations (Shoemake, 1985).

3. SOFTWARE ENVIRONMENT

3.1 Heirarchical Scenes and Kinaesthetic Correspondence.

A hierarchical scene can be constructed based on a special scene configuration file which contains the description of the objects and their relationships to each other. Code describing the objects is compiled into the program and these objects are placed in a display list at

---

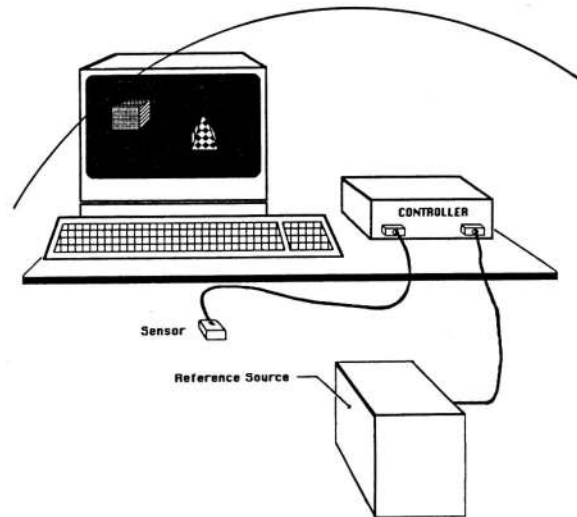1. 3SPACE Isotrak is a trademark of McDonnell Douglas

Figure 1. System layout. The six degree of freedom mouse is labelled "sensor" and its range of movement is indicated by the arc which cuts across the upper left corner of the screen.

execution time.

In a spatial manipulation system such as this it is essential that a natural correspondence be maintained between hand movement and the motion of the current object (or cluster of objects). This "kinaesthetic correspondence" in essence means that visually perceived motion and the motion of the limbs (in this case the arm and hand) as perceived through muscle and joint sensors, should be close to isomorphic. However, the exact ingredients which make kinaesthetic correspondence "natural" are not known. For example, as already mentioned, it does not seem to be necessary that the hand be in the same physical location as the object. However, it is well known that certain distortions of the eye-hand relationship - such as mirror reversal - are quite disorienting (Howard and Templeton, 1966). In our interface we provide kinaesthetic correspondence by ensuring that the direction of motion was always preserved for both rotations and translations.

Creating kinaesthetic correspondence is at variance with the most straightforward way of programming motion in a hierarchical scene, since each object in a scene hierachy has its translations and rotations defined in terms of the object immediately above it in the hierachy (Britton, et. al, 1977). We achieve the correct kinaesthetic correspondence relationships, at the cost of a restriction in the number of objects we can display in the hierarchy.

Fortunately, this number is still adequate for our goal of studying the human factors of object placement in a simple environment.

3.2 Display modes

A number of display modes are built into the system.

1) Objects can be constructed of filled polygons. These have their hidden surfaces removed by the method of not displaying backfacing polygons. A problem here is that the hierarchical scene description is traversed in a fixed order by a recursive subroutine and this may mean that an object can occlude another object when it should appear behind. This is a difficult problem to solve algorithmically, with our present hardware, while maintaining the essential real-time kinaesthetic correspondence. We have avoided rather than solved this problem by ensuring that in our experimental environment such inconsistent occlusions shall occur only rarely.

2) Wire frame outline figures can be displayed. In this case there is no hidden line elimination.

3) Both of the above modes are available with and without stereopsis. For the stereoscopic representation red-green anaglyphs are used and the colours are mapped into an 8 level brightness scale.

4. GENERIC INTERACTION MODE

We built the system with a "generic" interaction mode which contains the styles of interaction which we felt to be naturally right. Additional experimental modes are possible enhancements to the basic mode designed to enable us to evaluate their efficiency.

The interaction modes and the display modes are selected via a fixed menu. This is shown in figure 2 which shows the basic screen layout. The basic set of 6-D manipulations are selectable in arbitrary combinations using the following menu entries:

    all translations and rotations
    all translations
    translate x
    translate y
    translate z
    all rotations
    rotate x
    rotate y
    rotate z.

Generally, when using the bat, the "all translations and rotations" option is the most useful for inital object placement while some subset of the possible manipulations can be used for precise placement.
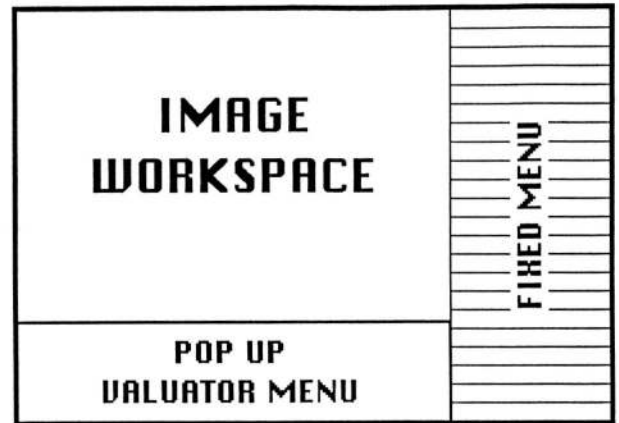

Figure 2. Menu layout (not to scale).

4.1 Basic Placement

The software interface to the bat allows it to be used in the following way for placement operations. With the button in the up position, a screen cursor displays the x, y position of the bat. The object to be moved is selected by positioning the cursor over it and pressing the bat button. Subsequently, pressing the button down and moving the bat - with button depressed - to a new location and orientation causes an identical movement of the current object (assuming that the "all translations and rotations" option is set). Moving the bat without the button depressed has no effect on the object. Thus, large movements can be made by "ratcheting", using the button as a clutch to connect the object.

If the root object of the scene is selected, the entire scene can be translated and rotated using the bat. If a subpart is selected then only that object and its descendents will move. The ability to rotate the scene allows the user "pick up" the scene and examine it from different angles. It also allows the user to select the most appropriate viewing angle for object placement.

4.2 Viewpoint Specification

There is an interesting point to be made here concerning the kind of protocol which is most natural for specifying a new view of the scene. We initially considered allowing the user to specify a new view by placing the bat at the physical location desired for the new viewpoint. This is the method advocated by Badler et. al. (1986) athough they point out that it has some drawbacks: "the lack of adequate spatial feedback made positioning the view a very consciously calculated activity instead of a simple and effortless process." However, there is an alternative model which arises naturally once the user recognises the scene to be defined heirarchically. In the context of a heirarchical scene it is natural to "pick up"

and position the entire scene (root object) in exactly the same way that a sub object in the heirarchy can be selected and manipulated. The viewer has the metaphor of a toy world which can be repositioned or reoriented by exact kinaesthetic correspondance with 6D mouse movements. Our initial experimentation has convinced us that this is far more intuitive than moving the viewpoint around. Two obvious advantages are 1) only one protocol is necessary for placing the objects and for repositioning and reorienting the viewpoint,. and 2) there is none of the disorienatation which can result from a jump to a new viewpoint.

5. SPECIAL MANIPULATION MODES

5.1 Autorotate

The kinetic depth effect causes a flat 2-D representation to appear as 3-D if the 2-D representation is the projection of a rotating scene. We created an option which causes the scene to rotate smoothly oscillating through 90 degrees about a vertical axis. During rotation, a sub-part of the scene can be moved relative to the entire scene. The question we are interested in is how easy is it to manipulate an object which is already moving? We find this mode of interaction is not difficult to master and does allow for an approximate object placement. However, it is necessary to stop the scene from rotating to achieve precise placement. Thus we do not find it to be a particularly useful or desirable enhancement to be used as an aid to placement.

Where autorotate is useful is when the user wishes to sit back and contemplate the scene. In this situation the rotation does much to enhance the 3-D percept and especially when used in conjunction with stereopsis results in a vivid spatial impression.

5.2 90 deg flip

An accurate placement in 3-space can be made using just two orthogonal views of the scene. Therefore, one of the menu options is a 90 deg flip switch which rotates the scene about a vertical axis. This is implemented as a toggle so that a second invocation returns the scene to its original orientation. This is extremely useful for object placement - the object can be positioned in the x,y plane of the screen, the scene flipped and a second x,y placement achieves the desired 3-D placement. Unfortunately the 90 deg flip is also disorienting, the abrupt switch to an orthogonal view of the scene leaves the observer struggling to find landmarks. A possible remedy for this disorientation might be a 90 deg slow rotation of the scene in which the observer would see a continuous transition from one view to another.

5.3 Dual Mode.

An excellent method for visualizing a scene is to "pick it up" using the bat and rotate it freely. This provides both kinetic depth effect cues and kinaesthetic correspondence cues because of the relationship of hand and object motion. We attempted to combine this visualization technique with a manipulation technique by implementing a special mode in which hand rotations rotate the scene while hand translations position an object within the scene. The idea is that the rotations allow easy visualization of the relationships between objects, while positioning of an object can be done simultaneously with translations of the hand. Our experience with this mode has not been such that it can be deemed a success. It is generally confusing and it is difficult to make accurate placements because rotations inevitably produce inadvertent translations.

5.4 Change of Eye-Hand Movement Ratios

Our interface allows the user to change the amount of hand motion required to perceive an object motion using a popup menu consisting of a set of valuators. The scaling for translates and rotates can be independently set or varied together. Our subjective experience with this variable suggests that for initial positioning a 1:1 ratio between hand motion and object motion produces a natural interface, one which works well for rough placement. However, due to the unsteadiness of the unsupported hand, it is impossible to obtain accurate placement using this ratio. For fine adjustments a ratio of up to 10:1 can be advantageous.

6. AN EVALUATION OF THE BAT

In the introduction we distinguished the problems of manipulation and visualization. We feel that the bat (with an appropriate interface) effectively solves the manipulation problem. Although as mentioned earlier, our bat has rather low resolution this is not a drawback since the unsupported hand is relatively unstable. When precise placement is required it is better to change the mapping from hand movements to object movement than to try to hold the hand very steady. Thus a "gear shift" or gain controller is essential.

A large number of visitors to our laboratory have tried out the spatial manipulation system described in this paper and we find that once they know how to select and attach an object - which takes about a minute if they are familiar with conventional mouse interfaces - they find the approximate placement of an object to be a trivial task. This undoubtedly due to the achievement of kinaesthetic correspondence between hand and object movement.

Previous investigators of the light pen have reported that arm fatigue is a problem due the the the necessity of holding the arm outstretched. We have not found that arm fatigue is a problem with the bat because it is like a mouse it in that it encodes relative motion and hence it can be held at waist level with the arm bent. This requires considerably less effort than holding the arm outstretched. Alternatively, the forearm can be rested on the arm of a chair and most of the object displacement can be achieved by wrist action.

Of the various manipulation aids which we describe, we find the simplest and most effective to be the 90 degree flip about a vertical axis. This is especially useful if z movement (into the screen) is disabled, allowing movement only in the x and y directions.

To conclude, the IRIS, used in configuration with the bat, provides a powerful and natural interface to 3-D scenes stored in a computer. Some modes of interaction are clearly more natural and effective than others and we have attempted to convey these findings in the present paper. Other issues are not clear cut on the basis of phenomenological evaluation. For example, it is not clear whether filled surfaces are superior to wire-frame renderings, especially when coupled with stereopsis. Our future plans involve explicit empirical testing of the accuracy of placement and the speed of manipulation using various visualization modes.

REFERENCES

Badler, N.I., Manoochehri, K.H., and Baraff, D., 1986, Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints. Proceedings of 1986 Workshop on Interactive 3D Graphics. Chapel Hill, NC, October.

Britton, E.G., Lipscomb, J.S., and Pique, M.E., 1977, Making Nested Rotations Convenient for the User, Computer Graphics, 12, 222-227.

Braunstein, M.L. 1976, Depth Perception Through Motion, Academic Press: New York.

Evans, K.B., Tanner, P.P., and Wein, M. 1981, Tablet-based valuators that provide one, two, or three degrees degrees of freedom. Computer Graphics, 15, 91-97.

Fisher, SS, McGreevy, M., Humphries, J., and Bobinet, W. Virtual environment display system. In Proceedings of the 1986 Workshop on Interactive 3D graphics, F. Crow and S.M. Pizar, Eds. (Chapel Hill, N.C. Oct.). ACM, 1986 To be published.

Foley, J.D., and van Dam, A., 1982, Fundamentals of Interactive Computer Graphics, Addison Wesley: Reading Mass.

Howard, I.D. and Templeton, W.B. 1966, Human Spatial Orientation. New York: Wiley.

Newell, A., and Card, S. 1985, The prospects for a psychological science in human-computer interaction. Human-Computer Interaction, 1, 209-242.

Schmandt, C. 1983, Spatial input/display correspondence in a stereoscopic computer graphic work station. Computer Graphics 17, 253-259.

Shoemake, K. 1985, Animating Rotation with Quaternion Curves. Computer Graphics, 19, 245-254.

Story, N.J., and Craine, 1985, Interactive Stereoscopic Computer Graphic Display Systems, Human-Computer Interaction - INTERACT'84/B. Shackel (ed.) Elsevier Science Publishers. (North Holland). 163-168.

Sutherland, I.E., 1968, A head mounted three dimensional display, FJCC, Thompson Books Washington DC., 757-764.