Controlling graphic objects naturally: use your head

Roger A. Browse[a,b], James C. Rodger[a], Ian Sewell[c], and Jeffrey Brooke[a]

[a]Department of Computing and Information Science,
[b]Department of Psychology, Queen's University, Kingston, Ontario, Canada
[c]Newbridge Networks Corporation Kanata, Ontario, Canada

ABSTRACT

During normal viewing of an object, a human observer will typically make small movements in the position of the head resulting in small *parallax-related image changes*. The significance of these changes is apparent when viewing a static stereographic display. Since the observer expects modifications in viewing direction that accompany side to side head movements, the lack of such changes in viewing stereographic displays creates the striking illusion that the static display is rotating in a compensatory direction. Using head tracking, we generate the appropriate pairs of images on a stereographic display device in order to maintain a stable virtual stereo object for the viewer. Unnatural, but learnable mappings from input devices such as a mouse or a joystick are typically used to bring about changes in the viewing direction and viewing distance in graphic displays. As an alternative to these techniques, we have extended the use of the monitored head position, resulting in a display system that permits control of graphic objects with subtle head movements. The device permits a zone of small head movements for which there is no rotation or scaling of the virtual object, but only parallax-related images changes as projected to each eye. A slightly exaggerated head movement initiates rotation and/or scaling of the scene that terminates when the head returns to a central viewing position. We are carrying out experiments to test the performance of human subjects in tasks that require head movements to control the rotation of graphic objects. A preliminary study that only examines rotation around a single axis suggests that it may be a very effective and natural technique.

Keywords: head-tracking, parallax, virtual object, interface techniques, human factors

1. INTRODUCTION

When the head of an observer moves, stationary objects at different distances from the observer appear to have different velocities of motion relative to the observer. The resultant sense of three-dimensional depth in the scene is referred to as a parallax depth effect. During normal viewing of an object, a human observer's small movements in the position of the head also results in small changes in the orientation and scale of the object as projected onto the observer's retinas.[1,2] These changes will be referred to as *parallax-related images changes* because they are a result of the same movements that often lead to parallax depth effects. The significance of these *parallax-related images changes* is apparent when viewing static stereographic displays. The observer expects modifications in viewing direction that accompany side to side head movements. These movements are actually rotations of the image as projected on the observer's retinas. The lack of such rotations creates the striking illusion that the static display is rotating in the compensatory direction, as it would have to in order to maintain a stable retinal image.
1

Using real-time six degree-of-freedom head tracking, it is a simple matter to calibrate the head position to the eye positions of the viewer, and to use the dynamic eye positions to generate the appropriate pairs of images on a stereo-equipped display device in order to maintain a stable virtual stereo object for the viewer. We have developed such a display system using an Ascension Technologies Bird sensor for head tracking, and a Silicon Graphics workstation

Further author information -
R.A.B. (correspondence) E-mail browse@qucis.queensu.ca; Voice: (613) 545-6069; Fax: (613) 545-6513
J.C.R. E-mail rodger@qucis.queensu.ca

equipped with CrystalEyes viewing glasses for processing and stereo display. Under lighting conditions that minimize the sense of the presence of the graphic display screen, such stereographic displays create a stronger illusion of viewing a real three-dimensional object than is available in stereographic images which are not adjusted for head movements.

There are three important aspects to the use of monitored head movements during the presentation of graphic displays. The first issue relates to the added sense of reality that accompanies a scene that appears to remain stable through viewer head movements. Naturally, one of the primary motivations in the development of virtual reality devices is to impart a sense of reality in the virtual objects. The strength of the added sense of reality is most easily demonstrated in adjustments that are made for movement of the viewer's head along the viewing axis. As the viewer's head moves toward the display screen, a virtual object positioned in front of the display screen would have to be increased in its display size in order to maintain a stable virtual object size, and conversely a virtual object positioned behind the screen would have to be decreased. A viewer of such a display will steadfastly claim that no image change is taking place, while to any observer whose head movements are not affecting the display, there is obvious shrinking and expanding of the object display. This is a prime example of the principal that under natural viewing conditions, humans perceive objects and ignore, if possible, the image formation conditions. It is more difficult to devise a situation for which the viewer may ignore the display when the full six degrees-of-freedom of the Bird sensor are exploited. In particular, it has not yet been determined what image refresh rates are necessary, what lag times are acceptable, and what resolution of position measurement is required in order for a true sense of object stability to be achieved.

The second issue relating to the *parallax-related images changes* concerns the perceptual performance that can be expected for head-free stereo displays that do not make display adjustments for head position. For tasks that require precise monitoring of motion, such as for teleoperated robots, it may be the case that the illusory movements that compensate for head movements may interfere in the judgement of the actual motion. Previously, we have used a robotic collision-detection task to evaluate subject's judgements of movement under different viewing conditions.[3] We have modified this experimental paradigm to test the interference effects of the illusory movements. The observer views a display in which one sphere moves towards another that is fixed in space. The trajectory results in either a collision, or near collision of the spheres, with a variety of actual smallest distances between the sphere centres. The subject's task is to report, as quickly as possible, whether or not a collision will take place. Stereo conditions with and without parallax related image changes are used as the two main conditions of the experiment.

The third issue is the possibility of controlling graphic objects with the monitoring of small head movements. This issue is the topic of the remainder of this paper.

2. USING HEAD MOVEMENTS TO CONTROL GRAPHIC OBJECTS

The purpose behind the use of computer graphic displays usually requires that depicted objects are rotated, scaled, and otherwise modified by the user. Unnatural, but learnable mappings from input devices such as a mouse or a joystick are typically used to bring about changes in the viewing direction and viewing distance. A less typical, but far more natural input is found in the use of sensor-equipped gloves that permit the user to manipulate virtual objects as if they were being held. In these cases, however, the use of such input precludes effective simultaneous use of the keyboard. As an alternative to these techniques, we have extended the use of the monitored head position, resulting in a display system that permits control of graphic objects with subtle head movements.

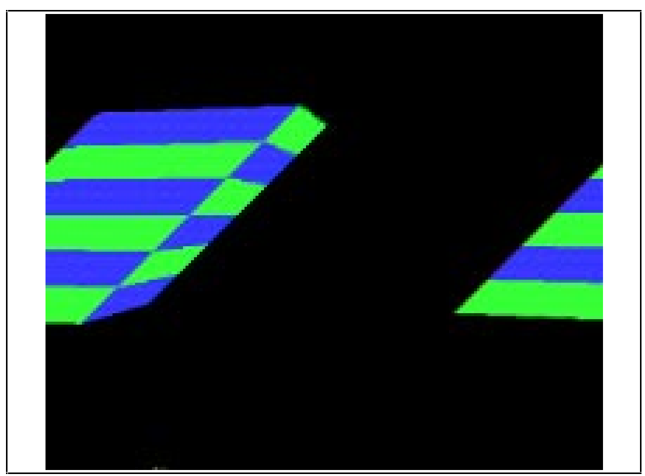
Almost continuously during normal waking life, people experience parallax related changes in the images that are formed on their retinas. It is a reasonable conjecture that during the course of this experience, a very precise, and natural mapping is learned between the direction of head movement and the direction of the apparent rotation of viewed scenes. The main purpose in experimenting with head controlled rotation of graphic objects is to determine if that natural mapping can be extended to apply in the case of viewing objects on display screens, and in particular, if the predictive knowledge of the direction of small rotational motion can extend to the control of objects rotating through large angles.

The most effective implementation of this device appears to be to permit a zone of small head movements for which there is no rotation or scaling of the virtual object, but only parallax-related images changes as projected to each eye, thereby creating an effect of a stable virtual object. For slightly exaggerated head movements, the component of the head movement within the same plane as the display is used to initiate a rotation of the object in exactly the same orientation as the parallax-related image changes that would result from the same but smaller magnitude head movement. Slightly exaggerated movements away from or towards the screen initiate a shrinking or expanding of the object. A three-dimensional ellipsoid is defined around the viewer's head within which only parallax-related images changes are made. When the head movement exceeds the surface of the ellipsoid, movement begins at a predesignated rate of rotation, scaling, or both. The movement subsides when the head reenters the ellipsoid. The parameters of the ellipsoid can be calibrated for an individual by spending a small amount of time making only movements that are not intended to initiate changes in the virtual object. Of course, explicit adjustment can be made later to suit the user's preferences.

There are a number of empirical questions that relate to the use of the head in controlling graphic objects. Can the user attain performance comparable to the use of a mouse, both in terms of the precision of obtaining desired orientations, and in terms of the speed with which such orientation changes can be made? Will users find the use of head movements to be tiring or awkward? In the long run, the answers to these questions will require extensive comparisons of the technique within demanding and complex graphical manipulation tasks. As a first step we have carried out an experiment that we believe justifies the development of some of the more extensive testing.

3. EXPERIMENTAL PROCEDURE

Within the experiment, observers sit with the back of their head positioned at about 120 cm from the display surface of a Silicon Graphics workstation as shown in Figure 1. After the subject initiates a trial with a keystroke, a cube appears on the screen. The cube is displayed with only two surfaces exposed such that the edge separating the surfaces forms a vertical line, as shown in figure 2. The subject is able to rotate the cube around its vertical axis only. The subject's task is to align the cube so that the vertical edge that separates the surfaces is in the middle, that is, such that the two visible surfaces are the same size. Each trial begins at an orientation that is from +- 35 to 10 from the target centre position. The subject is asked to perform the centring task as quickly as possible, and to indicate completion of the trial with a keystroke. A recording is made of both the amount of time taken and the accuracy with which the centring task is



accomplished.

There are four different conditions in the experiment, each involving a different mechanism by which the subject can rotate the cube on the display screen. There are two conditions that use the mouse device, and two use the Bird device. For either device, we refer to one of the conditions as a *locked* condition. For those conditions, the side to side movement of the device results in a direct and corresponding rotation of the cube on the screen. In each of these *locked* conditions, the range of natural side to side movement is sufficient to accomplish the rotation of the cube into its desired orientation. These conditions are included because they represent the simplest and most direct comparison of the use of head and hand movements in controlling the rotation of graphic objects. For the purposes of the experiment, we have informally established that normal viewing involves side to side head movements that are in the range of 6 cm to either side which is mapped to the 35 degrees of rotation necessary to accomplish the trial requiring the greatest rotation. For the mouse condition a translation of about 4.5 cm in either direction is mapped to an image rotation of 35 degrees.

We refer to the other condition for each device as *combined* condition. In this condition, a direct correspondence between movement and rotation is restricted to a central portion of the range of movement, and moving the device beyond that central portion initiates a rotation of the imaged cube at a rate of 36 degrees per second. The central portion of *locked* control is only sufficient to accomplish about 10 degrees of rotation of the cube, so for all but the smallest required rotations, the subject must exceed the central portion to bring the cube into rotation, and then return to the central portion to make a fine adjustment to centre the cube. It is interesting to note that this is really the only way that extensive rotations can be easily accomplished with the head movement control. It is unreasonable to assign the full range of rotation to a small range of movement because the rotation would become too difficult to control precisely. On the other hand, for mouse-based graphic control systems, there is a variety of possible techniques which can extend the range of rotation. Many systems utilize some form of reset condition, which allows the user to repeat mouse movements in order to effect larger rotations of the graphic objects. For the purposes of the experiment, the *combined* system was chosen in order to provide a direct comparison with the head control condition.

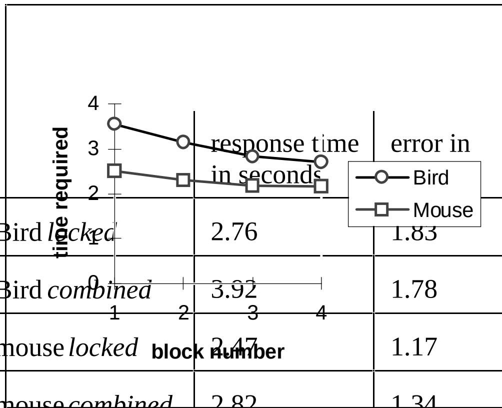
Five students at Queen’s University took part as subjects in the experiment, each reporting normal or corrected to normal vision. Each subject was provided with an explanation of the experimental task, and was given practice trials under the supervision of the experimenter. Each of the four conditions consisted of four blocks, with each block consisting of all 12 possible starting orientations of the cube shown in figure 2. The first of the 4 blocks was a practise block, designed to let the subject become familiar with the device and its operation. Alternating subjects did either all of the mouse-based trials first or all of the Bird-based trials first. The reason for wanting to keep all the trials for each device together is that the devices differ in the direction of rotation that accompanies side to side movement. A movement of the mouse to the right initiates a counterclockwise rotation as seen looking down the vertical axis of the cube. This is the normal mouse-based assignment of rotation, and it is as if the object were being spun around on its vertical axis by the hand movement. Movement of the head, and thus the Bird, initiates the opposite rotation, as if to provide a new view of the object from the new viewpoint that is attained through the movement. Thus there was a possibility of the subject being confused about the direction of movement that was required if the trials were in any way interspersed. In both situations, the *locked* condition preceded the *combined* condition.

Immediately after each trial, the subject was required to return to a centred position. In doing so, the subject viewed a small square on the screen which would be coloured red if the subject was not centred, and coloured green if the subject was. If centring was retained for one second, the centring box would vanish and the subject could initiate the next trial.

4. EXPERIMENTAL RESULTS AND DISCUSSION

This experiment evaluated the use of a mouse and head-mounted Bird senor in the control of one rotational dimension of a cube presented on a CRT display. The mean response times and mean errors for each of the two control methods are shown in Table 1. There was considerable variation in response time across subjects, and so in order to present the data in a more meaningful fashion, Figures 3 and 4 show the response times and error for which each subject’s means have been normalized so that performance on the *mouse locked* condition is set to 1.0 before an average is taken across the subjects. Thus Figures 3 and 4 show the average difference in performance among the four conditions.

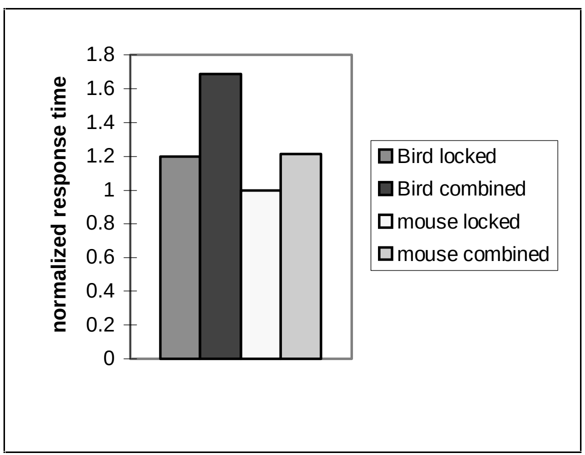
The basic conclusions that can be drawn from this preliminary study are that the use of the head-mounted Bird

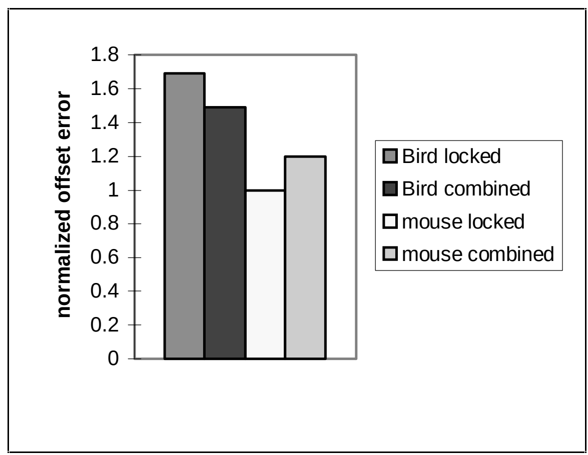


	response time in seconds	error in
Bird *locked*	2.76	1.83
Bird *combined*	3.92	1.78
mouse *locked*	2.47	1.17
mouse *combined*	2.82	1.34

provides the user with rotational control in the same range that is provided through the use of the mouse. While the averages show a small advantage for the mouse conditions, this was not true for all the subjects tested, neither for speed nor error. No subject reported that the head-mounted control seemed awkward or unnatural. No subject reported any fatigue or strain in carrying out the task.

It would be a simple matter to devise a test that demonstrated a clear advantage for either the mouse or the bird. If subjects were required to continuously type in material at the keyboard while controlling the orientation of graphic objects, the Bird would no doubt be superior. If subjects were required to control graphic objects in a way that required many quick movements, then no doubt the mouse would be superior due to the greater agility in the hand. The very slight advantage for the mouse within the preliminary task that we have devised, should not be generalized.





For both devices, we have included a *combined* condition in which movements outside of a central area result in free rotation of the graphic object being controlled. The reason why this condition was included is that it is the only obvious way that a head-mounted sensor could be used to control the full range of rotation of graphic objects. This method of rotation control did not appear to offer any serious difficulty in either the mouse or Bird conditions. In considering the response time differences for the *combined* and *locked* conditions, it is important to note that the free rotation rate was only 36 degrees per second, which was considerably slower than the speed at which the subjects rotated the cube in the *locked* mode. The average *combined* trial required about 12 degrees of free rotation which in turn required about 0.33 seconds to accomplish. With this consideration, the response times appear even more similar.

Figure 5 shows the response times for the use of the mouse and the Bird in each of the four successive blocks of the experiment, averaged across all subjects. From this data it is apparent that the experiment was for the most part testing subjects who had not yet attained peak performance in the task. The graph in Figure 5 suggests that if more blocks had been included, that the gap between response times for the Bird and mouse may have narrowed even further.

Error was measured as the number of degrees that the final attained rotation of the cube was from the required alignment. In all conditions this average was between 1 and 2 degrees. There is some issue as to whether or not small differences from the required central alignment may be perceived by subjects at all. The two faces of the cube must be of different hue or intensity in order for the subjects too perceive the boundary between them. We discovered that such differences could slightly alter the subjects' impressions of the centre, and thereby bias the results. Thus we arrived at the use of the cube surfaces as shown in Figure 2, which use the same intensities and hues in each of the cube faces. Overall, however, we suspect that the small differences in error between the mouse and the Bird conditions may be due to the precision of the devices, rather than the subjects' ability to control graphic objects with them.

The results of this preliminary experiment have encouraged us to proceed with some of the more complicated tests that will examine the use of head-controlled rotation in detail. These extensions include the use of two degrees of rotation, as well as scaling. In addition, we intend to carry out similar tests using stereo viewing. We are particularly interested in the effect of having the virtual object positioned at a distance from the viewer so that the *locked* portion of the rotation of the objects is exactly the natural parallax-related images changes that would accompany head movements. In the preliminary experiment reported here, the object rotated as if it were a lot closer to the viewer than the plane of the viewing screen. The correct distance from the virtual object may be necessary to take full advantage of the viewer's natural understanding of the relation between head movement and rotation direction.

The ultimate advantage of using one's head to control graphic objects is that the hands remain free for other tasks. Another important direction for this research is to establish the extent to which subjects may carry out manipulation tasks while using head-control of graphic objects. Again, observations of the use of head movements in everyday life suggest that such interference would be minimal.

ACKNOWLEDGEMENTS

The research reported in this paper was conducted with financial support from the Institute for Robotics and Intelligent Systems and the Natural Science and Engineering Research Council of Canada.

REFERENCES

1. B. Rogers and M. Graham, "Motion parallax as an independent cue for depth perception", *Perception*, **8**, pp. 125-134, 1979.
2. C. Caudek and D.R. Proffitt, "Depth perception in motion parallax and stereokinesis", Journal of Experimental Psychology, **19**, pp. 32-47, 1993.
3. R. A. Browse and S. Little, "Performance evaluation of viewpoint-controlled displays for teleoperated robots", *Proceedings of Stereoscopic Displays and Applications V*, San Jose, CA, pp.169-177, 1994.