Exaggerated Head Motions for Game Viewpoint Control

Robert J. Teather
rteather@cse.yorku.ca

Wolfgang Stuerzlinger
wolfgang@cse.yorku.ca

Department of Computer Science and Engineering,
York University, Toronto, Ontario

ABSTRACT

In this paper, we present an evaluation of exaggerated head-coupled camera motions in a game-like 3D object movement. Three exaggeration levels were compared to determine if the exaggeration was more beneficial than a realistic 1:1 mapping.

The results suggest that there is some user preference for this type of exaggeration; however, no significant differences by the experimental conditions were found, other than a learning effect.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *virtual reality.* H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *input devices, interaction styles, standardization, theory and methods.*

General Terms

Experimentation, Human Factors.

Keywords

Head-coupled perspective, head-tracking amplification, motion exaggeration.

1. INTRODUCTION

The greatest recent innovation in game design has undoubtedly been motion control. Popularized largely by the Wii, gamers can now use body gestures to perform game actions, rather than relying solely on pressing buttons. At present, this motion control is primarily isolated to the hand, but future trends appear to be moving more towards full-body motion control. Already, the Wii Balance Board™ allows players to control games by shifting their body weight, and technologies such as the Eye Toy™ and Microsoft's Xbox LIVE Vision Camera may go further still. We speculate that head tracking may be one such future innovation and is already becoming commercially available [5].

Coupling the virtual camera to the user's head position can give a realistic sense of perspective, and is commonly used in virtual reality. Sometimes, however, it may be more beneficial to exaggerate head motions when calculating virtual perspective. This is especially true of environments where the user is sitting, (e.g., desktop VR), that do not appear to benefit from traditional head tracking. This is also true of gaming, which normally occurs on a television or monitor. Moving one's head large distances is often infeasible due to space or comfort limitations.

We propose that software exaggeration of head movements could provide a reasonable technique to overcome these limitations. By exaggerating the movement of the user's head, subtle motions could be mapped to much larger viewpoint movements. Note that this is different from exaggerating head rotations. Instead, it is best thought of as "orbiting" the scene, similar to arcball rotation.

We present a system using exaggerated head-coupled motions to control the camera viewpoint, while performing 3D object movement tasks a mouse. We conjecture that this is similar to how one might use this in a game as the software used is similar to the "design mode" present in many games. The advantage of a system like this is to allow the user to play the game with the primary input device, while simultaneously performing useful viewpoint control operations with their head.

1.1 Related Work

Previous work has focused on amplifying head *rotations* with head-mounted displays [3]. Since HMDs tend to have a limited field of view, users must physically exaggerate head movements to achieve desired scene rotations. The authors instead used software to amplify head rotation, and consequently users did not have to exaggerate their motions. They found that participants made significant improvements in a visual search time when both their head and virtual hand rotations were amplified. Furthermore, subjective questionnaires found that users preferred the amplified state to the normal one, as it made turning one's head in the virtual environment more similar to doing so in reality.

Several 3D movement techniques also rely on exaggerating user motions. Two examples are Go-Go [7] and HOMER [2]. The Go-Go technique allows the user to interactively and non-linearly adjust the length of their virtual arm when manipulating an object in 3D. HOMER, on the other hand uses ray-casting selection, and automatically moves the user's virtual hand to the position of a selected object. Both of these effectively exaggerate the motion of the user's hand, thus allowing them to manipulate remote objects.

Several studies have attempted to quantify the benefit of head-coupled perspective (with or without exaggeration) in fish tank VR [1, 4, 8]. Like head-mounted displays, limited field of view can be a concern in fish tank VR systems. Since the display must always be in front of the viewer, one cannot physically turn around in these types of systems.

Mulder and van Liere [4] exaggerated head rotations in a fish tank VR system to account for this limited display space. The authors used the idea of rotation amplification to double the effective scene rotation. For example, if a user turned their head 10 degrees to the left, the scene would rotate 10 degrees to the right about their head. They compared this technique to a scene rotation technique using a 3D wand. However, their findings suggested

FuturePlay 2008, November 3-5, 2008, Toronto, Ontario, Canada.

that the head-coupled mode provided no significant improvements in speed over the wand version. There was also no subjective preference for one technique or the other. Furthermore, exaggerated rotations still do not overcome the problem that eventually, the screen will be out of the user's field of vision.

Other researchers conducted studies with 6DOF input devices for 3D manipulation tasks in fish tank VR systems, and compared stereoscopic to monoscopic display with and without head tracking [1, 8]. The first of these found that stereoscopic viewing allowed for significantly faster task completion times than monoscopic viewing. However, no significant effect was found for head tracking [1]. The authors reason that their tasks required only minimal head movement after the initial discovery of target locations. Similarly, the second found little benefit from using head-coupled perspective in a fish tank VR system [8]. However, it is possible this is due to the relatively low depth complexity of the 3D assembly tasks used in their experiment.

In general, the benefits of the extra depth cues provided by head-coupled perspective and stereoscopic graphics may be task dependent, as previous work has found that participants were able to more quickly trace a complex graph/tree structure when using the extra depth cues [9]. By comparison, the tasks used in other studies [1, 8] used much simpler scenes. It is possible that by exaggerating the head movements, a stronger effect may be detected in these types of tasks with lower depth complexity.

2. EXAGGERATED MOTIONS

Although turning one's head may be natural for an HMD or a CAVE, turning too far in a desktop VR or game system will result in only being able to view the screen through peripheral vision. This is why we instead propose to look at head *movement*, rather than *rotation*, as small lateral movements in front of the display still allow the user to look directly at the screen.

Since several of the above works indicated little to no benefit of head-tracking in fish tank VR systems, we use the idea of exaggeration demonstrated in previous work [2, 3, 4, 7] to determine if this will prove more useful. We believe that there are several innate benefits to controlling the viewpoint with one's head that extend to games. Controlling the viewing position with one's head should be very natural for users, since this is what they do every day in reality! By off-loading this control from an input device, such as a mouse, Wii remote or game pad, the input device can be used for performing the other tasks without requiring mode switching to control the viewing position.

Rather than amplifying head *rotations* [4], we propose to instead amplify head *movements*. The result is somewhat similar to creating a non-isomorphic, amplified head-controlled arcball rotation. For example, if the user moves their head up, the virtual scene rotates toward them, about its X-axis (relative to their viewing position). Moving their head left or right results in the world turning about its Y-axis. Zoom control could be mapped to head movements toward and away from the screen.

Using movement exaggeration simply treats the head as though it has moved further than it actually has. For example, if using a 1:2 exaggeration factor, the scene will rotate as though the user had moved their head 2 virtual units of distance for every 1 unit physically moved.

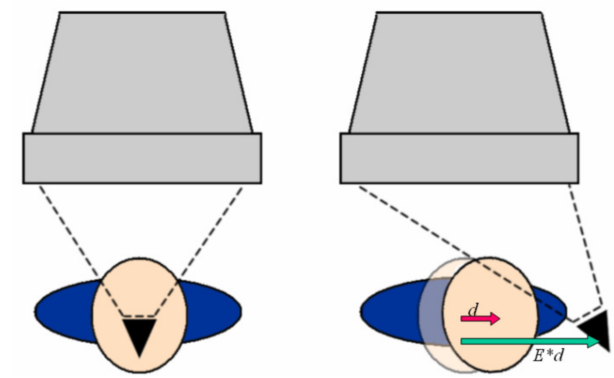

Figure 1. Virtual camera first positioned at head. When head moves distance d camera moves E times as far.

Because 3D tracking equipment can be extremely sensitive (e.g., Intersense IS900 is reported to have a positional accuracy of 0.75mm, www.isense.com) very slight head movements could cause unwanted rotations, especially when exaggerated. Since it is unlikely that users will be able to remain perfectly still, we opted to include a "home" region around their head, within which scene rotation is reduced to the minimal amount. We decided to make this region a 2.5cm radius sphere around the initial position of the 3D tracker. Within this region, the exaggeration was linearly interpolated from 0 at the centre of the sphere (i.e., no exaggeration, just standard head-coupled perspective) to 1 at the edge, and outside the sphere (i.e., the exaggeration was turned fully on). In other words, the closer to the centre of this sphere, the less extreme the effects of the exaggeration factor on any given trial. Upon moving out of the home region, the full exaggeration factor was used. Another possible solution to this problem would be to include a "clutch" button on the primary input device. Holding the clutch button would disable head motion control, to avoid unwanted viewpoint changes.

3. EXPERIMENT

We conducted an experiment to evaluate the effect of exaggerating the head movement. In particular, we were interested in determining if the exaggerated head motions technique was useful for moving objects in a 3D virtual environment. We were also interested in determining what degree of exaggeration was most beneficial (if any).

3.1 Hypothesis

We hypothesize that moderate levels of exaggeration will aid users in the object movement tasks. Specifically, we hypothesize that both the speed and accuracy of object movement will be improved. This is because participants will be able to get different viewpoints more quickly than with lower levels of exaggeration, hence improving their speed of task completion. Conversely, it is possible that high levels of exaggeration will be more difficult to control, and could negatively affect accuracy. Hence we believe that moderate levels will perform best overall.

3.2 Participants

Eight paid volunteers participated in the study, with ages ranging from 19 to 26, mean age 20.5 years. Seven participants were male, and one was female. Participants were asked about their previous experience with 3D games, since it is possible that gamers could outperform non-gamers in these types of tasks. They were also

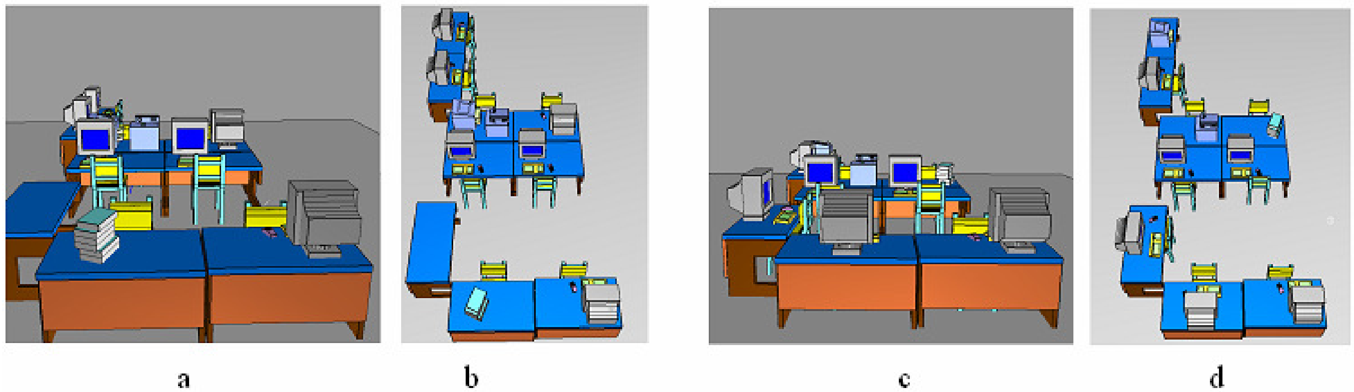

Figure 2. a) Starting scene, perspective view (this is what the participants saw), b) Starting scene, overhead view (provided for clarity of object positions), c) Target Scene, perspective view, d) Target Scene, overhead view.

asked about their previous experience with 3D modeling software (e.g., 3DS Max, Maya, etc.) for the same reason. Six indicated that they played games only several times per month or less often. The remaining two played games more frequently than this. Only one participant had any experience with 3D modeling software.

3.3 Equipment

Tasks were performed on an Intel Pentium 4 3.0 GHz with 512MB of RAM, and using Intel Extreme Graphics. A standard optical mouse was used as the input device. An Intersense IS900 was used for head tracking. The head tracking sensor was mounted on head-phones that the participants wore at all times during the experiment. We used custom 3D software written in C++ using OpenGL/GLUT for graphics.

The software includes a novel 2D to 3D mapping technique for object movement, based on the idea of front-face sliding. This algorithm ensures that all objects being moved remain in contact with other objects in the scene at all times. Objects are selected via mouse ray-casting, and the mouse is dragged to move them. Depth is handled automatically by the algorithm; objects simply slide across the closest surface to the viewer that their projection falls onto. Essentially, it reduces 3D positioning to a 2D problem, as objects can now be directly manipulated via their 2D projection. We opted to use this technique because previous research has suggested that novices have an easier time with it than other 2D/3D mapping techniques, such as 3D widgets. This is likely because the technique conforms better to the results of perception research and general best practices in 3D user interface design [8]. Full details of the technique and experiments evaluating it can be found in [6].

3.4 Procedure

After receiving an explanation of the experiment and signing informed consent forms, each participant was seated in front of the display and given the head-tracker to wear. At the beginning of each trial, the software would store the position of their head as the centre of the "home region", as described section 2, above. The participants were first allowed to experiment with the system to familiarize themselves with the 3D movement mapping technique used. The movement technique is similar to the "drag 'n' drop" metaphor and object movement techniques used in desktop computer interfaces and virtually every game which supports such operations. Hence, most participants took to it immediately. Nevertheless, they were allowed 5 minutes to practice, during which time the experimenter explained the task.

Following the practice session, participants performed the actual trials. Each block represented one level of exaggeration and consisted of 6 repetitions.

The experimental task used involved moving several pieces of furniture around a computer lab virtual environment. Participants were presented with a low-angle view of the lab, so that they would be required to use the head-tracking to see occluded objects in the distance. The task required that they move two computer stations (one from the back-most desk, and one from the second-row, right-most desk) to foreground desks, and a chair to go along with one of these computer stations. They also were required to move the printer from the second-row to the back-most desk, and a stack of books from the front-most desk to the second-row, right-most desk. Moving a computer station involved moving the monitor, keyboard and mouse, all separately. In total, each trial involved the movement of 9 virtual objects, of sizes ranging from very small (the mouse) to fairly large (monitor and printer). Figure 2 depicts the experimental task.

Finally, after completion of all 24 trials, participants were given a brief questionnaire to collect demographic information and their subjective preference of the exaggeration modes.

3.5 Design

The experiment was a single factor design with four levels. The factor was degree of exaggeration, and the levels were 1:1, 1:2, 1:3 and 1:5. Additionally, participants repeated each task 6 times in each exaggeration mode, for a total of 24 trials per participant. The orderings of the exaggeration conditions were counter-balanced according to a balanced Latin square to negate possible learning effects across conditions. Participants took approximately 1 hour total to complete the experiment.

4. RESULTS & DISCUSSION

Task completion time and accuracy were recorded. Accuracy was computed as the total error distance for the completed scene, compared to the target scene. This was calculated by computing the distance between the positions of each object in the completed scene and the position of where the object should have been placed in the target scene. These distances were then summed for each completed trial. Consequently, only objects that were moved in the scene could have any negative impact on accuracy. Note that this also includes accidental movements of objects that were not supposed to be moved. As a result, user errors in the task would negatively affect accuracy.

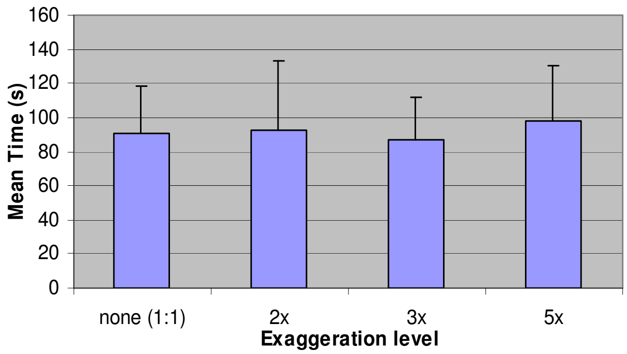

Figure 3. Task completion time by exaggeration level.

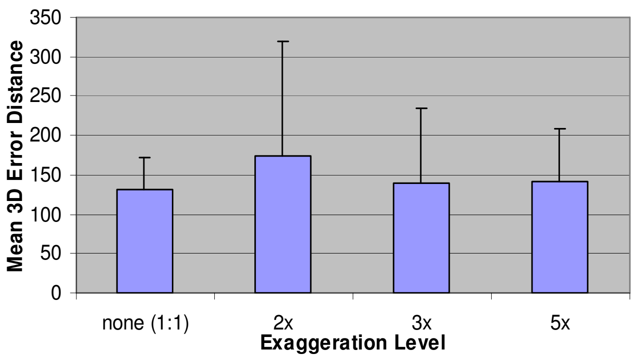

Figure 4. Mean error distance by exaggeration level.

Figures 3 and 4 depict mean task completion times and accuracy by exaggeration factor, with standard deviations. Task completion times and accuracy were analyzed with ANOVA. Overall, no significant differences were detected in task completion time ($F_{3,188} = 1.022$, $p = .38$) or accuracy ($F_{3,188} = 1.617$, $p = .19$) due to exaggeration factor.

A significant interaction was detected between trial and exaggeration ($F_{5,18} = 4.73$, $p < .01$) for task completion time. This is depicted as a power curve learning effect in Figure 5.

Data was also collected from subjective questionnaires given to the participants upon completion of the experiment. Seven out of eight of the participants indicated that they preferred the 3x level of exaggeration. The last indicated that they preferred the 5x level.

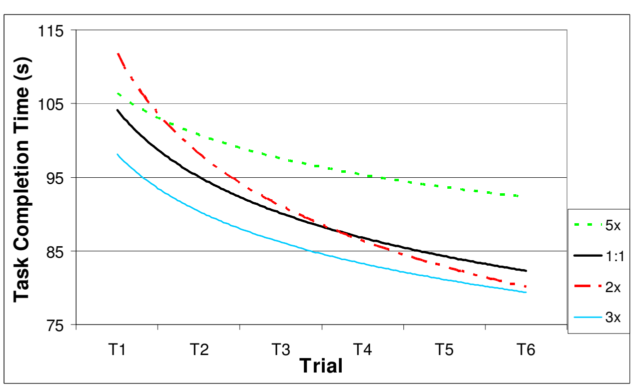

Figure 5. Fitted exponential learning curves.

To our surprise, the independent variable had no significant effect on either task completion time or accuracy of object placement. However, considering all participants preferred at least a moderate level of exaggeration, and the effects demonstrated in Figure 5, we suspect that further trials may have eventually demonstrated a significant difference in task completion time by exaggeration level. It appears that participants became proficient at the task more quickly with an exaggeration factor of 2. We believe this is the case, since in Figure 5, the line of the exaggeration factor 2 crosses both the line for exaggeration factor 5 and the 0 exaggeration level, indicating that with each subsequent trial, they became better faster under E2 than the other conditions.

5. CONCLUSIONS

In this paper, we suggested artificially exaggerating head movements to increase the rotation of the scene when using head-coupled perspective in a desktop VR setup. An experiment was conducted to determine if there was any benefit to doing this when performing 3D object movement tasks. Although no significant differences were found for speed or accuracy by level of exaggeration, subjective impressions from the participants suggested that they preferred at least a modest amount of exaggeration. In addition, under one level of exaggeration (an exaggeration factor of 2), users seemed to get better significantly faster than in the other conditions. We believe this indicates that a greater number of trials may have yielded significant differences, at least in task completion times.

6. ACKNOWLEDGMENTS

Thanks to John Bonnett of Brock U for the use of his lab and equipment. Thanks also to the participants for their assistance.

7. REFERENCES

[1] Boritz, J., and Booth, K. 1997. A Study of Interactive 3D Point Location in a Computer Simulated Virtual Environment. In *ACM VRST '97*, 181-187.

[2] Bowman, D, and Hodges, L. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *SI3D '97*, 35-38.

[3] Jay, C., and Hubbold, R. 2003. Amplifying Head Movements with Head-Mounted Displays. In *Presence: Teleoperators and Virtual Environments*, 12, 3, 268 – 276.

[4] Mulder, J. and Van Liere, R. 2000. Enhancing Fish Tank VR. In *IEEE VR 2000*, 91-98.

[5] NaturalPoint TrackIR. 2008. http://www.naturalpoint.com/trackir/. Accessed July 29, 2008.

[6] Oh, J.-Y. and Stuerzlinger, W. 2005. Moving Objects with 2D Input Devices in CAD Systems and Desktop Virtual Environments. In *Graphics Interface 2005*, 195-202, 2005.

[7] Poupyrev, I., Billinghurts, M., Weghorst, S., and Ichikawa, T. 1996. The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *UIST '96*, 79-80.

[8] Teather, R. and Stuerzlinger, W. Guidelines for 3D Positioning Techniques. In *ACM Futureplay 2007*, 61-68.

[9] Ware, C., Arthur, K., and Booth, K. 1993. Fish Tank Virtual Reality. In *Interchi '93*, 37-42.