EFFECTS OF TRANSITIONING BETWEEN PERSPECTIVE-RENDERED VIEWS

Jocelyn Keillor[1], Kevin Trinh[1], Justin G. Hollands[1], and Michael Perlin[2]
[1]Defence R&D Canada – Toronto
[2] CMC Electronics, Ottawa

Increasingly, multiple views on the same scene may be made available via a network of sensors in security and surveillance applications. Similarly, virtual reality and gaming applications may provide a user with more than one view on a scene in an attempt to improve user orientation. Previous work has demonstrated the utility of ‘visual momentum’ by showing an advantage in spatial judgments when users are provided with a transition between 2D and perspective views. However, no prior research has examined the utility of smooth transitions between different perspective-rendered views of the same scene. The present experiment compared two types of smooth transition between perspective-rendered views, and demonstrated that transitions generally provide an advantage on a spatial judgment task. Notably, transitions between perspective views did not need to include a bird’s eye view of the scene in order for users to make improved line-of-sight judgments.

INTRODUCTION

There is an increasing requirement to present the user with multiple perspective views on a three dimensional (3-D) scene in systems related to surveillance, security, spatial visualization, virtual reality, augmented reality, and gaming. The availability of multiple views may help the user build a veridical mental representation of the 3-D geometry of the scene. However, multiple views can also lead to disorientation and difficulties in making spatial judgments, particularly if the relative locations of the sensors and their respective fields of view are not understood by the user (Aretz, 1991; Keillor, Hodges, Perlin, Ivanovic, Hollands, 2002). It has been suggested that this type of disorientation may be minimized if an appropriate type of transition could be provided to show the user the relationship between views (Hollands, Pavlovic, Enomoto, & Jiang, 2004; St. John, Smallman, Cowen, 2002). This concept of “visual momentum” was first suggested to describe techniques used in film that help the audience maintain spatial understanding of a scene across discrete film cuts (Hochberg & Brooks, 1978).

Hollands et al. (2004) demonstrated the effectiveness of a continuous transformation between perspective-rendered and two- dimensional (orthogonal) views of a scene on spatial judgment tasks. It is often the case that more than one perspective view on a scene is available, yet to date no empirical research has examined the effectiveness of visual momentum between perspective-rendered views on a scene. Furthermore, a number of different types of transitions might be envisaged to demonstrate to the user the spatial relationship between perspective-rendered views. Indeed, if an orthogonal (bird’s eye) view is critical to the development of a veridical internal representation of the scene, one might expect superior performance on spatial judgment tasks when transitions include an orthogonal view relative to direct transitions. Thus, the utility of two types of smooth transitions (with and without an orthogonal view, see Figure 1) between perspective-rendered views was tested in a gaming-type environment, using spatial judgments as dependent measures.

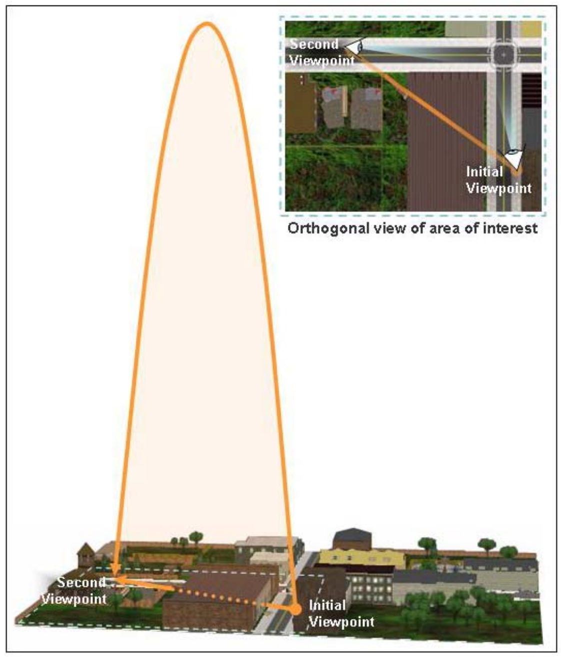


Figure 1. Linear and parabolic transition paths between an initial viewpoint and a final viewpoint. Parabola vertex at 180m.

The scenario we used to test the utility of transitions was one in which a dismounted soldier is attempting to determine whether it is safe to move to a visible point ahead. That is, would it be possible to move to that location without being seen by an enemy? The soldier in this simulated scenario has access to a feed from a camera placed on the other side of an object occluding the enemy, and Digital Terrain Elevation Data (DTED) is available with which to construct a wireframe view of the imaged area so that the transition to the camera view may be made explicit in a display. If it is the case that such transitions minimize disorientation, then it should also be the case that when a player in such a scenario is provided with a transition between an initial and final view (see Figures 2 and 3), performance might be improved on a related spatial judgment task.

Figure 2. Initial view.

Figure 3. Final view.

In order to determine the effects of providing transitions between views on spatial judgments we compared two types of transitions and a control condition as a function of the scene geometry (alignment and occlusion angles). We also manipulated the size of the area in which the graphics were displayed, with the thought that the utility of transitioning might interact with the availability of optic flow fields, or the potential requirement to make eye and head movements in order to examine disparate portions of the display. We also examined gender and mental rotation ability as between-subjects variables.

METHOD

Participants

20 (10 female, 10 male) volunteers, aged 19-51 , with normal or corrected-to-normal vision completed the experiment. Three additional participants who were unable to perform the task to a criterion (60% accuracy overall) were excluded. Participants were recruited from DRDC Toronto and the nearby community and were financially compensated for their participation.

Apparatus

The experiment was conducted in a room with normal temperature and dimmed lighting. The stimuli were presented on a Sony KDE-61XBR950 61" Plasma Wega HDTV (high definition television, native resolution of 1365 x 768 pixels), and responses collected via designated keyboard keys and mouse buttons on a Windows XP platform. The viewing distance was approximately 1 m.

Stimuli

A set of four simulated urban intersection models was created with OpenGL using Creator modeling tools. For each model, common city objects (e.g., houses, sidewalks, trees) were used to provide depth cues (see Figures 2 and 3 for examples).

Urban structures were modeled to scale for each corner of the intersection. An oil drum (marker) and enemy soldier were placed in the intersection. There were four possible sets of surrounding architecture, two types of occluders (a set of boxes or a bus of equal size), two directions of approach (from the north or south side) and two transition directions (left or right). These nuisance factors were balanced across visibility conditions.

The experiment consisted of a series of trials that were defined, in part, by the placement of several objects: a camera, an enemy, an occluder, and a marker (see Figure 4). For each of these objects there were seven possible positions where they could be placed, with 2041 (7^4) possible configurations. For each trial, two angles of interest were defined from the underlying geometry by employing a dichotic categorization according to the size of the angle between the camera, enemy and marker and the size of the angle between the enemy, occluder and marker. These angles were designated as "alignment angle" and "occlusion angle" respectively (see Figure 5). The alignment angle was constrained to range from 0 degrees (perfect alignment) to 40 degrees (a large mis-alignment). The occlusion angle ranged from 160 degrees to 200 degrees excluding the ranges between 175 degrees and 185 degrees in which the marker was partially occluded.

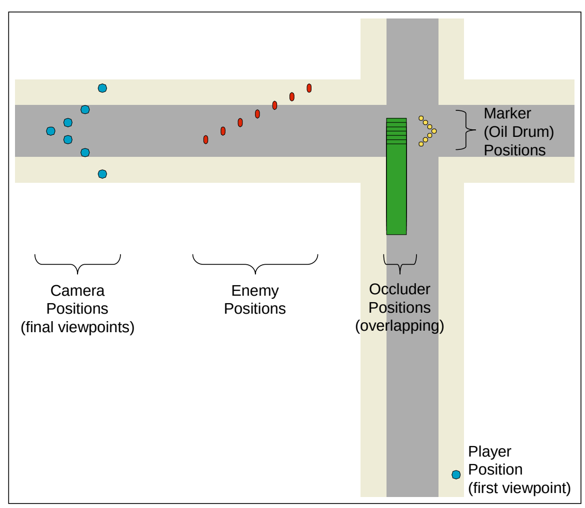


Figure 4. Object positions.

Sixty-four object configurations were pre-selected based on the criteria described below to create challenging visibility determinations for task 1 (described in Procedure section).

Criteria:

1) The camera had no line of sight to the marker.
2) The camera could not have a more favourable viewing angle of the marker than the enemy.
3) The marker could not be partially visible or partially occluded.

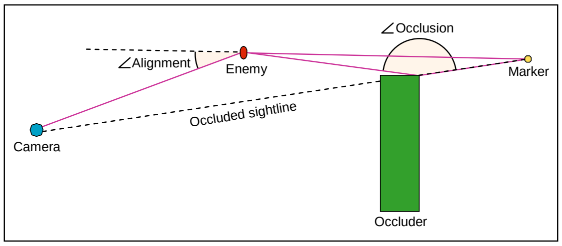


Figure 5. Alignment and occlusion angles. (Figure not to scale)

Of the object configurations, 32 had a line of sight from the enemy to the marker (i.e. the enemy could see the marker) and 32 did not.

Design

The experiment had a between-within design with display size (full screen @ 1360 x 768 vs. 3/8 full screen @ 510 x 288 pixels) as a between-subjects factor, and transition (direct vs. parabolic vs. discrete), alignment angle (small alignment angle vs. large alignment angle) and occlusion angle (small deviation from 180 degrees, large deviation from 180 degrees) as the within-subjects factor. Of the 10 participants per gender, 5 viewed the graphics in a large display format, and 5 viewed the graphics in a small display format.

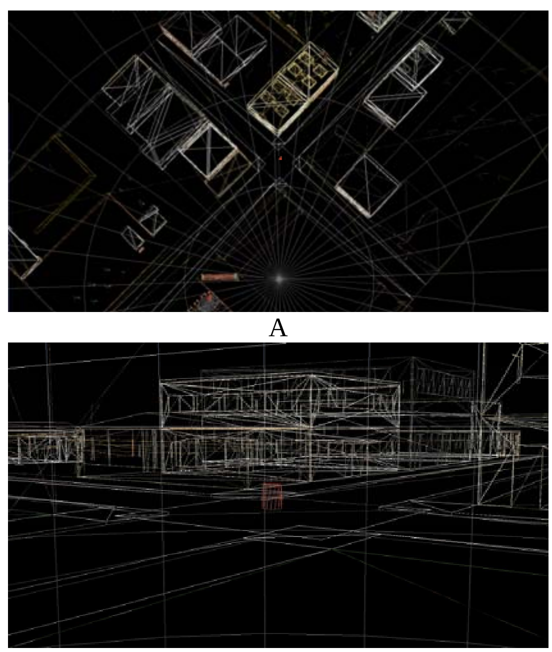
Two tasks were conducted for all conditions. Dependent measures were response time and accuracy (percent error for the visibility task and absolute distance for localization task). There were 12 (3 Transitions x 2 Alignment Angles x 2 Occlusion Angles) conditions. There were 384 (12 conditions x 32 iterations to accommodate nuisance factors) trials. Though nuisance factors, as described in the Stimuli section, were included in the experiment design to improve the generalizability of our results by providing subjects with a broader range of scenarios, we do not report on effects of these factors. These 384 trials were divided over three sessions of 128 trials each. Each session contained approximately the same number of trials of each type.

Procedure

Participants took part in three sessions (one session per day) of about 90 minutes each (including practice trials and breaks). The three sessions were completed over a maximum period of two weeks. In session one, participants completed a short demographics questionnaire on their video game experience. This was followed by a paper and pencil Mental Rotation Test (MRT) (Peters, 1995). At the outset of each session, participants completed a set of 6 representative practice trials. Following the practice trials, the participants completed a set of 128 trials with 1-2 minute breaks in between each of the 4 blocks of 32 trials. Each trial comprised two tasks, a line-of-sight judgement and a spatial localization task. No feedback was given during the experiment.

Each trial in the simulation began with an automated motion of the camera towards its initial resting position to provide the participant with some parallax cues to depth in the environment as they approached a 4-way street intersection. In addition to this viewpoint, participants were informed of a second (final) viewpoint (a view from a camera) of the intersection from the street perpendicular to their first view. Participants were instructed to judge whether it was safe to traverse to a point in the intersection, marked by an oil drum, as seen in Figure 2, given that the final viewpoint shows an armed enemy who might or might not have a sightline of the marker around the occluder. This is depicted in Figure 3 for the same trial as was shown in Figure 2. The marker was never visible from the final viewpoint.

To transition from the initial to final viewpoint participants were instructed to press the space bar. Two transition types were used: linear and parabolic. Transition paths for a typical trial are illustrated in Figure 1. During the transition, the environment was rendered in wireframe mode (see Figure 6), to simulate a computer model of the environment as opposed to a real viewpoint. Each transition took 4 seconds to complete. In the control conditions there was no transition and the second view was drawn to the screen immediately after the space bar was pressed.



A

B

Figure 6. Transition views - camera was always directed towards centre of intersection. a) Parabolic transition view from vertex of parabola. b) Linear transition view from midpoint of path.

Given both viewpoints and the transition, the first task was to determine whether the enemy could see the marker. Participants used the left or right mouse button to respond "yes" or "no" respectively. Response time (RT) commencing from sight from the second viewpoint and accuracy were recorded.

Upon completion of the line-of-sight task, participants were shown a bird's eye view of the terrain without the enemy and occluder as seen in Figure 7, and asked to complete a localization task. Participants were informed that the first viewpoint was at the bottom of the map and the second viewpoint was on the horizontal street to the left or right side. In this localization task they used the mouse to move the crosshairs and mark, with either mouse button, the location (in 2D) of the final viewpoint (camera location). This completed one trial. RT and accuracy were recorded for both the line-of-sight and localization tasks. Following completion of both tasks the next trial began; the imagery for the next trial was loaded within 3 seconds.

Figure 7. Localization task view.

RESULTS

An omnibus analysis of variance (ANOVA) was conducted for the mixed between-within model for each task (line-of-sight and localization). Gender and display size were between-subject variables and transition type, alignment angle, and occlusion angle were within-subject variables. One three-way interaction was found between transition, gender and display size whereby across transitions, and greatest for the discrete transition (control condition), response times in the line-of-sight task for females were slower while males improved with the small display compared to respective response times with the large display, $F(2,32) = 3.857$, $p < .032$. However, no main effects or two-way interactions (even with the within-subject variables) were found for the gender and display size factors for either response time or percent error for the two tasks. Therefore all subsequent analyses were conducted having collapsed across the between-subjects variables.

The ability to judge the line of sight was improved by having either the linear or parabolic transition rather than no transition, $F(2,38) = 17.347$, $p < .001$. This effect was greatest in the (more difficult) large alignment angle condition, $F(2,38) = 32.253$, $p < .001$ (see Figure 8). The pattern of results was the same for the response-time measure for the line of sight task, in that the advantage for transitions relative to control conditions was greatest for the (more difficult) large alignment angle conditions $F(2,38) = 5.769$, $p < .007$ (see Figure 9).

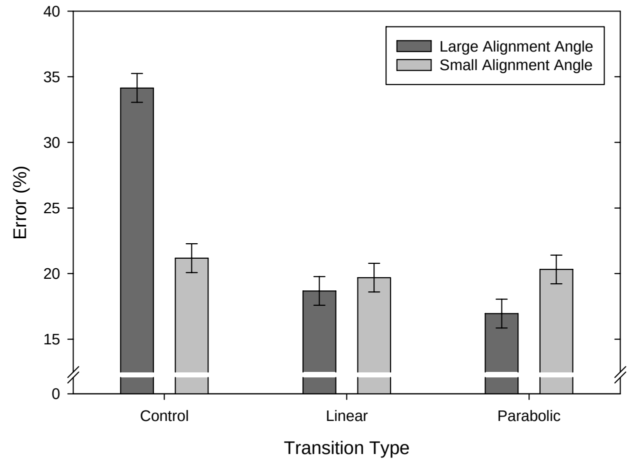


Figure 8. Effects of transition-type and alignment angle on percent error. Error bars indicate within-subjects standard error (Jarmasz & Hollands, in prep.).

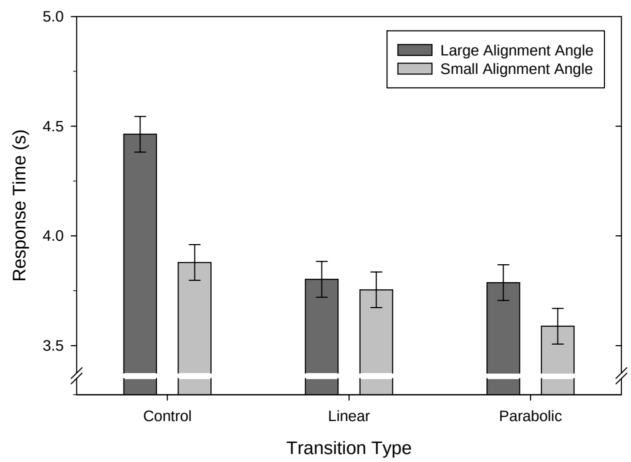


Figure 9. Effects of transition type and alignment angle on response time for correct trials. Error bars indicate within-subjects standard error (Jarmasz & Hollands, in prep.).

In another ANOVA, participants were divided into High- and Low- mental rotation ability groups based on a median split of their Mental Rotation Test scores. MRT score was a between-subjects variable and transition type, alignment angle, and occlusion angle were within-subject variables. The ability to benefit from transitions differed for the two groups as a function of the alignment angle $F(2,36) = 3.59$, $p < 0.015$ (see Figure 10). Specifically, post-hoc Fisher's LSD tests revealed that mental rotation ability affected participant's ability to correctly judge the line of sight only in the (more difficult) large alignment angle. That is, those with higher MRT scores benefited more from the provision of a linear transition than did those with lower MRT scores, ($p < .05$).

When the role of occlusion angle (of the marker beyond the occluder) was examined, it was evident that overall performance in a highly deviated occlusion angle was worse than in the case of a small deviation in occlusion angle, $F(1,19) = 96.461$, $p < 0.001$. This effect interacted with alignment such that the performance benefit derived from having a large alignment angle was greatest when the occlusion angle was also large, $F(1,19) = 17.473$, $p < 0.001$.

Thus, the ability to judge line of sight was best in the condition with a large deviation in occlusion angle and large alignment angle, and worst when the occlusion angle deviated minimally (from 180 degrees) and a large alignment angle was present. These effects did not differ significantly across transition types, $F(2,38) = 1.568$, n.s.

For the localization task, there was no effect of transition type on performance in judging the ground position of the second viewpoint, $F(2,38) = 2.658$, n.s.

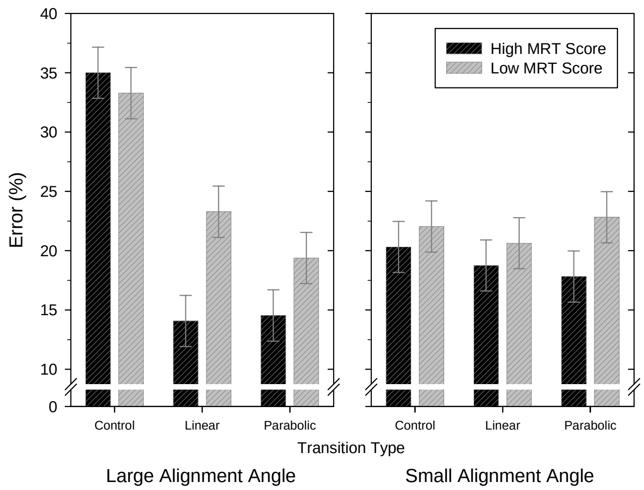


Figure 10. Effects of transition-type, alignment angle and Mental Rotation Test (MRT) score on percent error. Error bars indicate standard error for a mixed between/within- subjects interaction (Jarmasz & Hollands, in prep.).

DISCUSSION

The primary finding of this experiment is that the use of a smooth transition from one viewpoint to another can improve the ability of users to make a difficult line-of-sight judgment within a perspective-rendered scene. This advantage of transitions was largest for alignment angles that made the line-of-sight judgment more difficult. Interestingly, the type of transition employed was not important. That is, performance was similarly improved whether or not the transition included overhead views of the scene. This finding is of interest because, in general, such bird's eye views are thought to be important to the development of an accurate three-dimensional representation of a scene. Neither the direct nor parabolic transition contained the actual objects required to make the line-of-sight judgment, so it may be that orthogonal views must specify the locations or the objects in question in order to facilitate line-of-sight judgments. Orthogonal views might not contribute to the construction of a veridical spatial representation of the scene in any special way. Instead, the particular effectiveness of orthogonal views may be limited to the facilitation of relative position judgments only when the points to be judged are indicated in the display.

In both conditions involving a transition between views there was a four-second time interval before the final (camera) view was presented. Participants could not begin to prepare a response until they looked at the final view and saw the position of the enemy, so this extra four seconds could not be used for response preparation. It might be argued, however, that this time interval could give the participants time to somehow develop a better representation of the scene whether or not a transition was presented. However, this type of explanation has been excluded for transitions between two-dimensional and perspective views (Hollands et al., 2004) in that visual momentum can be demonstrated even when preparation time is available in the control case.

The finding that transitions between perspective views can benefit spatial judgments suggests that smooth transitions between views may be beneficial in a number of application domains including gaming and medical imaging. The rapidly expanding number of sensors being employed in military and surveillance settings suggests that the technique may be particularly important to the design of these systems. In military or search and rescue applications sensors could be mounted in unoccupied aerial vehicles (UAVs) as well as in multiple air and ground locations. Increasingly, information from air-support may be shared with those on the ground, and the correct interpretation of this information may help prevent friendly-fire incidents. In addition, sensors may provide degraded information; thus, when no transition is provided it might be difficult to determine if an unclear object in one view is the same as an object in a different view, particularly if the sensors provide very different perspectives on the scene. The use of visual momentum may prove vital to many safety-critical applications in which an accurate representation of a complex scene is required.

ACKNOWLEDGMENTS

The authors thank Julian Spillane for technical assistance, and Justin Anisman for data collection. This work was funded in part by a National Search and Rescue Secretariat grant to the first author.

REFERENCES

Aretz, A. J. (1991). The design of electronic map displays. *Human Factors, 33,* 85-101.

Hochberg, J., & Brooks, V. (1978). Film cutting and visual momentum. In J. W. Senders, D. F. Fisher, & R. A. Monty (eds.), *Eye movements and the higher psychological functions,* pp. 293-313. Hillsdale, NJ: Erlbaum.

Hollands, J. G., Pavlovic, N. J., Enomoto, Y., & Jiang, H. (2004). Advantage for visual momentum not based on preview. *Proceedings of the Human Factors and Ergonomics Society – 48th Annual Meeting,* pp. 1803-1807. Santa Monica, CA: Human Factors and Ergonomics Society.

Jarmasz, J., Hollands, J.G., Confidence intervals in repeated-measures designs: a "numbers of observations" principle. In preparation.

Keillor, J., Hodges, K.J., Perlin, M., Ivanovic, N., & Hollands, J.G. (2002). Imaging systems in search and rescue: Implications for geographic orientation. *Proceedings of the Human Factors and Ergonomics Society – 46th Annual Meeting,* pp. 170-174.

Peters, M. (1995). Revised Vandenberg & Kuse Mental Rotations Tests: forms MRT-A to MRT-D. Guelph ON, Canada: Technical Report, Department of Psychology, University of Guelph.

St. John, M., Smallman, H.S., & Cowen, M.B. (2002) Evaluation of schemes for combining 2-D and 3-D views of terrain for a tactical routing task. Pacific Science & Engineering Group Technical Report.