Vestibular cues and virtual environments

Laurence Harris[1,3] Michael Jenkin[2,3] Daniel C. Zikovitz[3]
Departments of Psychology[1] , Computer Science[2], and Biology[3]
York University
4700 Keele St., North York, Ontario, Canada

Abstract

The vast majority of virtual environments concentrate on constructing a realistic visual simulation while ignoring non-visual environmental cues. Although these missing cues can to some extent be ignored by an operator, the lack of appropriate cues may contribute to "cybersickness" and may affect operator performance. Here we examine the role of vestibular cues to self-motion on an operator's sense of self-motion within a virtual environment. We show that the presence of vestibular cues has a very significant effect on an operator's estimate of self-motion. The addition of vestibular cues, however, is not always beneficial.

1 Introduction

Virtual environments have been proposed for many, varied applications including the treatment of phobias[16], providing control for a mobile robot[1] and even providing a safe place for children to practice crossing the street[7]. In many virtual simulations effort has concentrated on constructing an accurate, realistic, visual simulation while other environmental and motion cues tend to be ignored.

There are many cues that normally contribute to the perception of motion. Consider the common virtual environment in which an operator equipped with a head-mounted display is physically limited to the tracking region of the head tracker and "flies" about an environment using a joystick or some other pointing device to signal desired motion. Such a subject is then presented with a wealth of visual cues to self motion, but the non-visual cues that are found in the real-life version of the experience are sorely missing.

There are several visual cues that contribute to telling operators of VR simulation that they have moved in their environment. These include judging their change in position against the visually-provided position and visual direction of objects in the environment. Another visual motion cue is optic flow; when moving through a three-dimensional environment the components of the retinal image stream across the retina and the resulting optic flow contains information about the direction and velocity of the operators' movement[3, 18]. Optic flow alone can generate a very strong impression of self-motion which is called vection [6].

A virtual environment with only visual cues ignores the physical cues to linear self motion which are largely signaled by the acceleration-sensitive otolith division of the vestibular system[10]. Periods of constant velocity cannot be registered by this acceleration-sensitive system; but for motions with changing velocities, position can be obtained by double integration of the acceleration signal. Humans are able to use vestibular information to assess a position change [2, 8, 9, 11] and direction of travel [17, 12, 19].

So an operator "moving" within a vision-only virtual environment is presented with at least two contradictory cues to self motion; the visual stimulus is carefully arranged to signal motion of the operator, while the vestibular system provides a signal which is consistent with no motion or motion at constant velocity.

There are many possible outcomes of this unnatural combination of visual and vestibular cues. It may be the case that one or the other of the two cues dominates, or the operator may average the two cues in some way. Disparate cues, generating an inter-sensory conflict have been postulated as a major reason for motion sickness [13, 15]. The associated nausea can be expected to reduce the efficiency of an operator working in a virtual environment.

This paper describes an experiment in which we pitted optic flow against non-visual cues to linear self motion to reveal which was the dominant sense in determining the perception of how far one has moved.

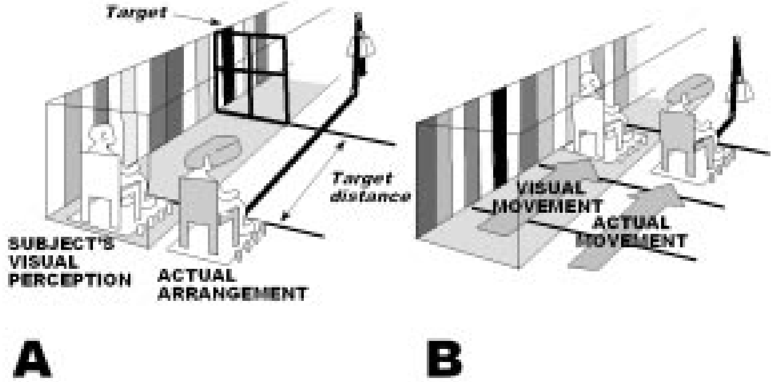

Figure 1: Experimental setup. Panel A shows the starting conditions and panel B shows typical movements experienced during the experiment. The left side of each panel shows the subject's visual perception and the right side shows the actual motion for the visual plus vestibular condition. For vision-only trials, the chair did not move but the subject's visual perception was as in panel B.

2 Methods

Participants

Ten subjects took part in this experiment. They were drawn from the graduate student population at York University and all had normal visual and vestibular functioning. The experiments were approved by the York University ethics committee. Subjects were paid at a standard rate for experimental subjects.

Apparatus

In order to examine the relationship between optical flow and non-visual cues to self motion we constructed an apparatus capable of generating different visual-vestibular conditions and controlling for the possibility that the subject could track outstanding visual features in the environment. To achieve this we presented visual information by in a virtual reality (VR) system and, at the same time vestibular cues by physical movement. Figure 1 shows the experimental arrangement. Subjects sat in a chair on a cart and viewed a computer-generated display in a VR helmet. This was a single-screen device so the two eyes viewed exactly the same optic flow. The armrest on the chair was equipped with a set of buttons to record operator responses.

Virtual Reality Equipment

Imagery was generated using an R10000 SGI Indigo^2 computer. Output video was converted to NTSC format and presented to the subject via a Liquid Image MRG3 head-mounted display. The display provided binocular viewing of a single colour 768x556 pixel liquid crystal display with a 84^o (horizontal) by 65^o (vertical) field of view. The head-mounted display was equipped with a six-degree-of-freedom Flock of Birds head tracker. The Flock of Birds provided simultaneous measurement of head position and orientation over a range of $\pm$3ft. Motion was tracked to accuracies of 0.5^o and 0.07in. at rates up to 144Hz. In order to extend the range of the head tracker, the transmitter was mounted on the cart and thus only measured subject motion relative to the cart. The cart was attached by a rope to a weight hung from pulleys. When this weight was released, it pulled the cart at a constant acceleration (0.1 - 0.4 ms^{-2} depending on the subject's weight and the pulley arrangement) along a 4 meter track. This is shown in the part of Figure 1A labeled "actual arrangement". Motion of the cart was measured by sliding an optical encoder (resolution over 3000 encoder counts per meter) along an earth-fixed cable. Head position in space could then be deduced by adding the head-on-cart signal from the Flock of Birds to the cart-on-earth signal provided by the optical encoder.

Visual Stimulus

The part of Figure 1A labeled "subject's visual perception" illustrates the virtual environment in which subjects perceived themselves during this experiment. Subjects appeared to be located in a virtual corridor 50m long, 2m wide and 2.5m high, roughly modeled on the standard cross-section of the corridors at York University. The walls were textured with vertical stripes 0.5m wide which changed colour on a random schedule approximately once a second. The duration of a particular colour on any stripe was chosen randomly from a uniform distribution. The coloured stripes are indicated by grey shades in the black-and-white diagram. This flickering between colours was added in order to reduce the possibility that the subject could track a feature on the wall. We wanted to force subjects to use only optical-flow information and not to track objects. The video imagery was updated at approximately 6Hz.

Experimental Conditions

Condition 1: Vision alone

For the vision-only condition, the subject did not move physically at all but moved only visually down the corridor. The pattern of visual motion was derived from pre-stored information obtained in other trials when the subject actually did move. Thus we could ensure that visual stimulation was identical in the vision-only and vision-plus-vestibular trials.

Condition 2: Vestibular alone

The video signal was turned off and the subject physically moved in complete darkness.

Condition 3: Vision plus Vestibular

The vision-plus-vestibular condition was the "natural" condition. The subject's visual position in the virtual corridor was derived from the simultaneous physical motion of the subject.

Procedure

Subjects were positioned within the virtual corridor so that they were facing down its length. At the beginning of the experiment (panel A in Figure 1) subjects were presented with a visual target at a distance of between 1 and 8m in front of them. Subjects were encouraged to move their heads about and get a good idea of how far away the target was using parallax and perspective cues. When they were ready, subjects pressed a button which made the target disappear and started them moving down the corridor (panel B in Figure 1) under one of the three possible conditions described above (vision only, vestibular only or vision plus vestibular). The subject did not know which condition was about to be experienced while they were judging the target distance. The subject was instructed to press a button when they felt that their nose would have touched the target had it been visible.

Data Analysis

The computer recorded the subjects' position in the corridor when they pressed the button. These distances were then averaged and subjected to a linear regression analysis to determine if the relationship between the actual target distance and the point at which the button was pushed varied between the three conditions.

3 Results

Figure 2 shows the distance that a subject had to travel in order to feel that the target had been reached. The means and standard deviations of 10 subjects are plotted against the actual target distances. Regression lines are forced through the origin. Perfect performance is shown as the shaded bar with a slope of 1.0. Subjects were very accurate at judging movement under vision-only conditions (slope=0.95) but consistently and dramatically over-estimated the distance traveled in the dark (slope=0.30).

When visual and non-visual cues were presented in their natural combination, that is, when optic flow and vestibular cues were present, subjects still consistently over-estimated their motion (slope = 0.45, Figure 2) for all target distances tested (up to 8m). This is despite the fact that the very same visual cue when presented alone was associated with accurate performance. In the vision-plus-vestibular condition subjects' judgments were significantly earlier than in the optic-flow-only condition (**F**=148.2; dof=106; $p<0.01$) and were not significantly different from the judgments made in the vestibular-only condition (**F**=0.27; dof=151; ns).

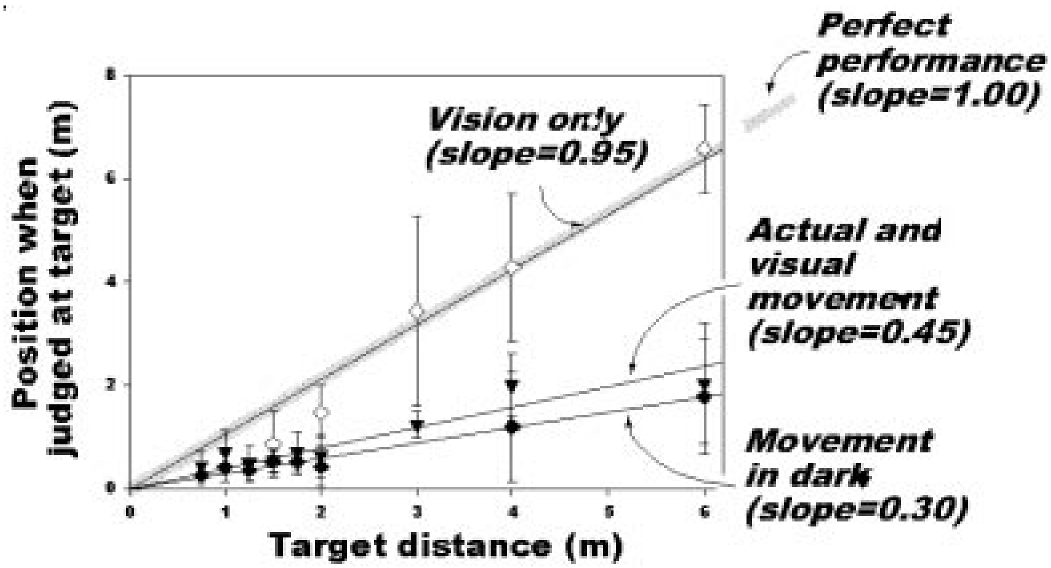

Figure 2: Judged target distance. Graph showing the relationship between actual target distance (horizontal axis) and perceived target distance (vertical axis) for three conditions. When judgements were made to "vision only" stimuli (open diamonds), performance was very close to veridical (slope = 1.0); when movement was in the dark and only vestibular information was available (filled diamonds), performance was poor (slope = 0.3) and physically-traversed distances were judged to be much longer than they actually were (1m was judged to be 3m); when both visual and vestibular cues were present (filled triangles), performance was much closer to the vestibular alone than the vision alone condition and distances continued to be overestimated.

4 Discussion

The results of this study showed that while subjects are capable of judging the distance of their self-motion accurately using our visual display alone, their performance was seriously degraded when they were moved at the same time. When given a target at between one and eight meters down the corridor, our subjects could accurately estimate when they had traveled the specified distance under the visual alone condition. In the movement-in-the-dark (vestibular only) condition, however, they pressed the button when they had traveled only about one third of the required distance. That is subjects *overestimated* their distance traveled when constantly accelerating in the dark. If the distance to match was 3m, for example, subjects felt they had traveled through this distance when they had gone only 1m. That is 1m traveled at constant acceleration in the dark feelt like 3m.

When subjects were really, physically moved, providing both visual and vestibular cues in the natural arrangement, they *again* pressed the button much too soon; close to the distance they had indicated under the vestibular-alone condition. That is under constant acceleration in the light, with only optic flow and vestibular cues, subjects consistently over-estimated their motion. We conclude from this study that optical flow alone is not the dominant factor in assessing distance traveled, although it can be used accurately when there are no competing cues. Instead vestibular cues, when present, *capture* the perception and dictate the perceived distance traveled.

Why is the vestibular-only condition so powerful?

Other studies (eg [4, 8]) have also reported subjects overestimate their perceived distance traveled when accelerating in the dark. This might be due to an overshoot in otolith activity to a step in acceleration [11, 21]. It might also be viewed as an open-loop response of a system whose loop is usually closed by active control of the movement of the head.

Why vestibular capture?

Perceptual thresholds for object motion are raised and object velocity is underestimated during head movement [14] (see [5] for a review). The reduction in the use of visual movement clues under our vision-plus-vestibular condition might represent another example of this. Anecdotally, our subjects often reported that the visual stimulus did not appear to be working during combined trials.

How generalizable are our findings?

At first glance, our findings are rather counter-intuitive: why, when vision and vestibular cues are both present, are subjects not accurate at judging the distance they travel? It is important to realize that our experiments were designed to pit only *two* of the many cues normally available to help in the estimation of self motion. Our visual cues were deliberately not as rich as the natural visual cues available during self motion. We did not have position cues or stereoscopic cues for example. Also, the visual system may not be well able to deduce the meaning of an optic flow generated by a constant acceleration of the person [20]. In contrast to this impoverished visual signal, our vestibular cue was particularly powerful. Continuous, passive acceleration is an unusual cue under natural circumstances.

Can we generalize from our experiments using optic flow only, passive movement, and constant acceleration to situations containing other combinations of sensory stimuli? Our experiments suggest that we should expect the unexpected and that the only way to explore the role of multi-sensory contributions to the perception of self motion, so important if we are to simulate self-motion convincingly, is to design and carry out further experiments.

Applications of our findings.

Understanding the conditions under which the vestibular system dominates in the perception of self motion might provide an underpinning for developing powerful techniques in virtual reality technologies. The virtual reality experience is at present often unconvincing and even nauseogenic. This has been attributed to a lack of appropriate vestibular cues. But how much extra-visual motion should be provided? This study has shown that just presenting vestibular cues during passive VR simulation can lead to unwanted effects. Our results show that when vestibular cues are present they can capture the visual system with the *undesired* and *unexpected* result of rendering the combined system less accurate than the vision-alone system. Designers of immersive environments which attempt to provide vestibular information in order to overcome some of the nauseogenic properties of VR environments must take special care that the introduction of even “correct” vestibular information may degrade operator performance in the resulting environment.

Acknowledgments

The financial support of NSERC Canada through Research Grants and Collaborative Grant program and the Institute of Space and Terrestrial and Science (ISTS) of Ontario. We would like to acknowledge the contributions of a number of York undergraduates who have contributed to this study notably Emre Onat who helped us start the whole thing off. We would also like to thank Aastra Aerospace for their contributions to early parts of this project, Doug Crawford and Carolee Orme who made useful comments on an earlier draft of this paper and Agnieszka Kopinska who helped with the statistics. The authors would like to thank Nicole Aucoin for assistance with the programming.

References

[1] N. Aucoin, O. Sandbekkhaug, and M. Jenkin. An immersive 3D user interface for mobile robot control. In *IASTED Int. Conf. on Applications of Control and Robotics*, pages 1–4, Orlando, FL, 1996.

[2] A. Berthoz, I. Israel, P. Georges-Francois, R. Grasso, and T. Tsuzuku. Spatial memory of body linear displacement: what is being stored? *Science*, 269:95–98, 1995.

[3] J. J. Gibson. *The Senses Considered as Perceptual Systems.* Houghton Mifflin, Boston. Mass., 1956.

[4] S. Glasauer and I. Israel. Otolithic thresholds influence the perception of passive linear displacement. *Acta Oto-Laryngologica*, S520:41–44, 1995.

[5] L. R. Harris. Visual motion caused by movements of the eye, head and body. In A. T. Smith and R. J. Snowden, editors, *Visual Detection of Motion*, pages 367–435. Academic Press, London, UK, 1994.

[6] I. P. Howard. *Human Visual Orientation.* John Wiley, New York., NY, 1982.

[7] D. P. Inman, K. Loge, and J. Leavens. VR education and rehabilitation. *JACM*, pages 53–58, 1997.

[8] I. Israel, N. Chapuis, S. Glasauer, O. Charade, and A. Berthoz. Estimation of passive horizontal linear-whole-body dispacement in humans. *J. Neurophysiol.*, 70:1270–1273, 1993.

[9] J. M. Loomis, R. L. Klatzky, R. G. Golledge, J. G. Cicinelli Pellegrino, J. W. Pellegrino, and P. A. Fry. Non-visual navigation by blind and sighted: assessment of path integration ability. *J. Expt. Psych. (General)*, 122:73–91, 1993.

[10] O. E. Lowenstein. Comparitive morphology and physiology. In H. H. Kornhuber, editor, *Handbook of Sensory Physiology, Vestibular System*, pages 75–124. Springer Verlag, New York, NY, 1974.

[11] A. Mayne. A systems concept of the vestibular organs. In H. H. Kornhuber, editor, *Handbook of Sensory Physiology, Vestibular System*, pages 493–580. Springer Verlag, New York, NY, 1974.

[12] M. Ohmi. Egocentric perception through interaction among many sensory systems. *Cog. Brain Res.*, 5:87–96, 1996.

[13] C. Oman. Motion sickness - a synthesis and evaluation of the sensory conflict theory. *Canadian J. of Physio. and Pharmacol.*, 68:294–303, 1990.

[14] B. Pavard and A. Berthoz. Linear acceleration modifies the perception of a moving visual scene. *Perception*, 6:529–540, 1977.

[15] J. T. Reason. Motion sickness adaptation: a neural mismatch model. *J. Roy. Soc. Med.*, 71:819–829, 1979.

[16] D. Strickland, L. Hodges, M. North, and S. Weghorst. Overcoming phobias by virtual exposure. *JACM*, pages 34–39, 1997.

[17] L. Telford, I. P. Howard, and M. Ohmi. Heading judgements during active and passive self-motion. *Expt. Brain Res.*, 104:502–510, 1995.

[18] R. Warren. The perception of egomotion. *J. Exp. Psych: Human Percept. and Perf.*, 2:448–456, 1976.

[19] W. H. Warren, M.W. Morris, and M. Kalish. Perception of translation heading from optical flow. *J. Exp. Psych.: Human Percept. and Perf.*, 14:640–660, 1988.

[20] P. Werkhoven, H. P. Snippe, and A. Toet. Visual processing of optic acceleration. *Vis. Res.*, 32:2313–2329, 1992.

[21] L. R. Young and J. L. Meiry. A revised dynamic otolith model. *Aerospace Medicine*, 39:606–608, 1968.