First steps with a rideable computer

Robert S. Allison[2], Laurence R. Harris[1,3], Michael Jenkin[2],
Greg Pintilie[2], Fara Redlick[3], Daniel C. Zikovitz[1,3]
The Centre for Vision Research,
and
Departments of Psychology[1], Computer Science[2] and Biology[3]
York University
4700 Keele St., Toronto, Ontario, Canada, M3J 1P3

Abstract

Although technologies such as head mounted displays and CAVEs can be used to provide large immersive visual displays within small physical spaces, it is difficult to provide virtual environments which are as large physically as they are visually. A fundamental problem is that tracking technologies which work well in a small enclosed environment do not function well over longer distances. Here we describe Trike *– a 'rideable' computer system which can be used to generate and explore large virtual spaces both visually and physically. This paper describes the hardware and software components of the system and a set of experiments which have been performed to investigate how the different perceptual cues that can be provided with the Trike interact within an immersive environment.*

1. Introduction

Advances in computer graphics and hardware display technology over the past five to ten years has resulted in the ability to generate visual displays of very high fidelity. At the same time tracking technology has also seen considerable advancement although fundamental problems remain. With a few exceptions, existing tracking systems are designed around technologies which limit the user to a small physical footprint. For example, mechanical tracking systems such as Fake Space's BOOM, and Puppetworks trackers, physically connect the wearer to a fixed position and thus the user is mechanically tethered to a point in space. Wireless electromagnetic systems such as Ascension Technology's Flock of Birds tracking system and Polhemus Fastrak remove the mechanical link, but only offer a limited operational range. For example the Flock of Birds has an operational range of ±4 feet with the standard transmitter and ±10 feet with the extended range unit. Vision-based trackers such as Northern Digital's POLARIS have the potential of a longer operating range but their accuracy degrades with distance due to the use of triangulation to measure position. Acoustic tracking systems are limited by the decay in the audio signal as a function of distance and the sound absorption properties of air. Existing virtual reality systems may provide a large visual space but existing tracking technology limits the user to a relatively small physical world.

As individual tracking units limit the user to a small physical space, one solution is to use a battery of trackers to track over larger spaces. For example, a network of Flock of Bird transmitters could be used to provide coverage over a large space but at significant cost and still requires a connection from the receiver to the base station.

Although the Global Positioning System (GPS) – especially in its differential mode – could be used to provide tracking information over larger ranges, it is not without its problems. For example, it cannot be used in environments in which no clear line of sight exists to the satellites such as indoors, near mountain ranges, in forests, and builtup urban areas. It also has a relatively slow update rate. (See [4] for a description of GPS and its limitations.) That being said, GPS can be used to correct drift in other sensors.

Various alternatives have been proposed to extend the range over which an immersive visual display can operate. Perhaps the most extreme version of these has been the development of wearable computers (e.g. [15, 17, 20]). The goal of these systems is to provide a computer and interface which is 'worn' by the user and which augments their normal senses as they move. Although most existing wearable systems are designed to provide two-dimensional visual overlays on the wearer's normal visual field, some wearable computer systems provide more immersive visual displays. For example, [12] describes a wearable computer system that in part mimics a pair of binoculars but with both eye's views the same. The wearer looks through the 'binoculars' and is presented with a computer-enhanced version of the environment. [5] describes a backpack-based system in which the user views the world through a head-mounted display in which synthetic or computer generated elements can be superimposed over the normal view. The lack of a static base station and some mechanism for measuring displacement relative to the base make estimating absolute position very difficult. Thus only relative orientation information is available for the generation of the visual field in this type of system.

An alternative to providing long-range physical motion is to allow the user to simulate long-range motion through various mechanisms which involve only limited physical movement of the operator. Simple systems such as joysticks, mice, keypads, etc., permit users in first-person games like Doom to move throughout large visual spaces while remaining essentially stationary. More sophisticated mechanisms have also been used, including stationary bicycles[3] and hang-gliders to simulate their mobile versions, large motion bases coupled with aircraft cockpit mockups to simulate aircraft, and active treadmills[18, 11]. One issue with these types of approaches, especially stationary ones, is that the operator experiences sensory conflict between a visual display which indicates motion and a stationary physical experience. This can be nauseogenic and is certain to restrict the comfort and performance of the user after a while.

Consider what a person experiences as they walk down a street. Self-motion cues will be picked up by the various sensory systems. These cues include visual flow, vestibular cues and proprioceptive cues – the latter being information about the relative positions of the various parts of the body. In stationary virtual reality systems only visual cues are provided, other cues to self-motion are not simulated appropriately. This inter-sensory conflict may result in discomfort and poor performance.

In order to provide both visual and non-visual cues over a large physical space, we have developed a 'rideable computer system' (*Trike*) based on a commercially available tricycle, which can be used by a user to work within a large scale virtual environment. Standard virtual reality technology is used to provide an immersive visual display relative to *Trike* itself, and the physical motion of *Trike* is used to generate non-visual cues. *Trike* is instrumented so that its motion relative to some initial base frame can be computed.

In addition to providing a natural mechanism for navigation in a large-scale virtual world, *Trike* can also be used to investigate the relative importance of different visual and non-visual cues to the perception of self-motion in virtual environments. This can be accomplished by manipulating the various sensory cues that *Trike* can present to the rider.

This paper is broken down into two main sections. Section 2 describes the mechanical and software design of *Trike*, while some initial experiments into the relative importance of various motion cues to the operator's perception of self-motion are examined in Section 3.

2. Designing a rideable computer

As the goal is to construct a device which can be physically moved to generate non-visual cues to motion, it is important that *Trike* is tethered no more than necessary to a base station. Computing and tracking is therefore performed on board the vehicle. The device is tethered for power only. Batteries could be attached to the vehicle at some point in order to remove this tether as well. As *Trike* has a relatively high payload and is tethered for power, weight and power consumption are not significant constraints on *Trike's* design. This permits standard off-the-shelf components to be used on the vehicle. *Trike* (see Figure 1) is based around a standard size adult tricycle. The drivetrain of the stock vehicle has been modified so that there is only a single gear and so that pedaling backwards causes the vehicle to go backwards. The slack in the drive chain which is normally present in a bicycle or tricycle has been reduced in order to improve the vehicle's response to pedaling.

In order to estimate the vehicle's motion relative to some initial frame, the vehicle has been instrumented with a potentiometer connected to the steering axis that senses the steering direction, and the drive wheel has been instrumented so that its rotation can be measured. A small onboard microprocessor monitors these sensors as well as two buttons mounted on the handlebars. The microprocessor communicates with the main onboard computer via a standard serial link.

The main onboard computer is an SGI O2 with dual display (the "two-headed option"). The two video outputs of the machine are fed to a binocular head-mounted display which is equipped with a Polhemus head tracker. The tracker's transmitter is mounted on the bicycle. The Polhemus tracker reports the position of the helmet relative to *Trike* and *Trike's* instrumentation yields motion and direction information. Thus the rider's position can be calculated from the combination of these two tracking systems and an appropriate visual display is generated.

2.1. Tricycle Kinematics

In order to update the rider's position with respect to the world it is necessary to construct a kinematic model of the vehicle. Fortunately bicycles and tricycles have a straightforward kinematic model. Full details of the kinematic models of these and other wheeled vehicles can be found in [4] but the basic concept is sketched here.

Front View

Side View

Figure 1. The TRIKE rideable computer

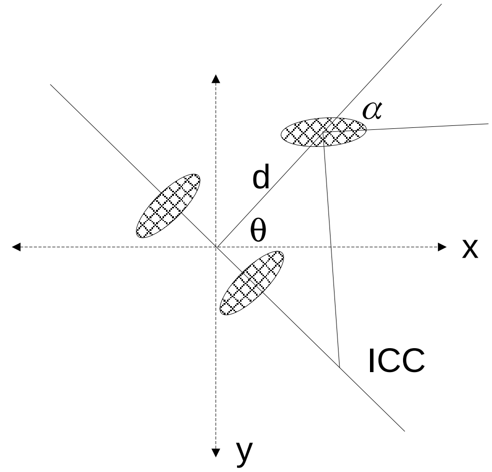

Figure 2. Bicycle/tricycle kinematics. The shaded ovals represent the three wheels of *Trike*. The instantaneous centre of curvature (ICC) must lie at the intersection of lines orthogonal to, and passing through the rotation axis of each wheel.

Tricycles are steered vehicles in which changes in wheel orientation or steering direction are used to change the trajectory of the vehicle. For a wheeled vehicle to move without slippage each wheel must follow a circular course around the vehicle's instantaneous centre of curvature (ICC) and must roll on the ground with a velocity which is consistent with the geometry of the wheel placement. As power is provided to the wheels, the entire vehicle will then rotate about the ICC. The task of determining the kinematics of any wheeled vehicle, and the tricycle in particular, reduces to the task of determining the ICC of the vehicle.

The ICC must lie at the intersection of lines drawn through and perpendicular to the rotational axis of each wheel (see Figure 2). Thus for a tricycle, the ICC must lie on a line passing through the rotational axis of the rear wheels which is perpendicular to the body of the tricycle. The front wheel can be steered and thus the ICC lies on that part of the line passing through the rear wheels which intersects the line drawn along the axis of the front wheel. Given a maximum steering angle of the front wheel, a tricycle has a minimum turning radius and rotates about a point on the line passing through the rear axle.

If the steered front wheel is set at an angle α from the straight-ahead direction and moves with ground contact velocity v, the tricycle will rotate with angular velocity ω about a point lying a distance R along the line perpendicular to and passing through the rear wheels, where R and ω are given by

$$R = d\tan(\pi/2 - \alpha), \quad \omega = v/(d^2 + R^2)^{\frac{1}{2}}$$

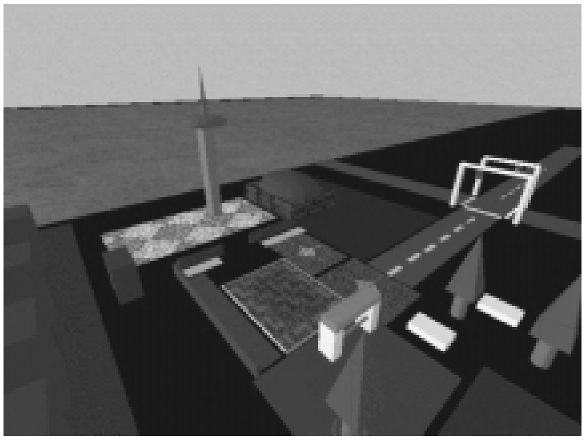

Figure 3. A sample rideable environment. The environment is a full 3D textured environment. The user's view of the environment is generated based on the current state of the head tracker and the bicycle kinematic model (see text).

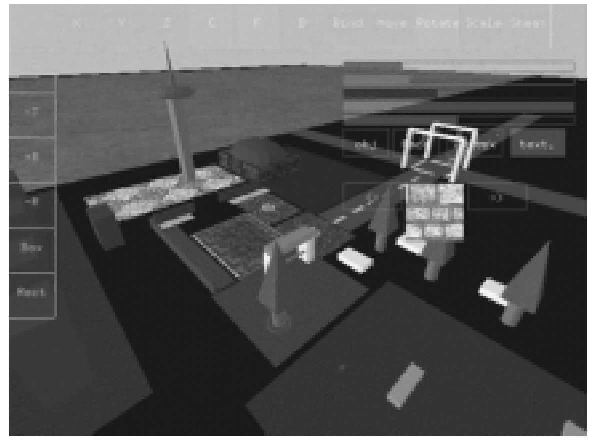

Figure 4. The user interface: transparent buttons are overlaid over the rendered 3D environment. These buttons can be used to manipulate existing objects in the environment and to initiate the creation of new objects. The 'brick' pattern in the middle of the image is the currently selected texture for texturing existing surfaces.

and d is the distance from the front to the rear axle as shown in Figure 2.

Suppose that the tricycle is at some position (x, y) and "facing" along a line making an angle θ with the x-axis at time t, i.e. it has pose $[x\ y\ \theta]^T$. Then the ICC is given by

$$\text{ICC} = (x - R\sin(\theta), y + R\cos(\theta))$$

and after a short interval δt the pose of the trike is given by

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} \cos(\omega\,\delta t) & -\sin(\omega\,\delta t) & 0 \\ \sin(\omega\,\delta t) & \cos(\omega\,\delta t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x - \text{ICC}_x \\ y - \text{ICC}_y \\ \theta \end{bmatrix} + \begin{bmatrix} \text{ICC}_x \\ \text{ICC}_y \\ \omega\,\delta t \end{bmatrix}$$

This describes the motion of a bicycle rotating at distance R about its ICC with an angular velocity given by ω.

2.2. Constructing a visual world

Given the kinematic model of *Trike* and the state of the head tracker, the next step is to generate a visual world in which to ride. Figure 3 shows a sample virtual environment visible from *Trike*. The environment was constructed using a point and click interface (described below) and consists of polygonal structures which can be texture mapped. The entire environment is properly shaded and the operator's view is updated based on the state of *Trike* and the operator's head position and orientation.

The current 3D environment does not support rider-environment interactions, although these could be added in a trivial fashion. For example, rider-obstacle collisions and rider-environment interaction via the buttons mounted on the bicycle handlebars are possible.

Environments are constructed through the 3D graphical interface shown in Figure 4. The user interface was designed to allow the user to interactively create a virtual environment encompassing all 3 spatial dimensions. It allows the creation of simple geometrical objects such as boxes and rectangles through a point and click mechanism. Objects can also be imported following a VRML-like file format. The interface allows a user to specify operations such as move, rotate, scale, and shear to act upon existing objects. For many operations the construction interface manages the parameters necessary to allow such implicitly 3D operations to be performed as sequences of 2D tasks. For example, an object may be moved along the XY, YZ, or XZ planes according to the settings on the interface. The interface also allows setting object properties such as color, transparency, and specularity. Operations through the interface can also take the form of assigning and manipulating object textures. Textures can be set on polygonal faces and then scaled, moved or rotated.

3. Initial experiments with the Trike

Trike is capable of being driven through complex 3D environments and can be used to investigate fundamental questions of human perception. In an ongoing series of ex-

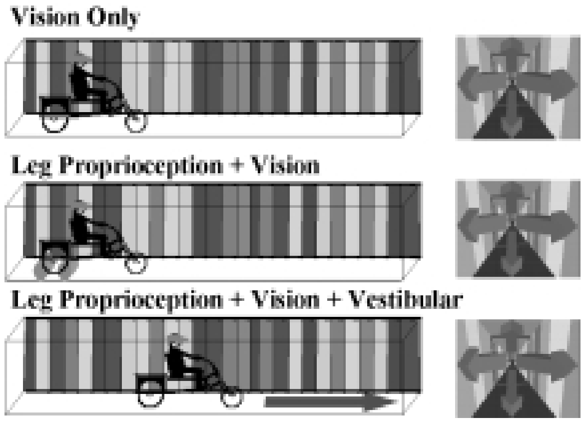

Figure 5. Three experimental conditions. Subjects were either presented with either vision only (top), leg proprioception and vision (middle), or leg proprioception and vision and vestibular cues to their self-motion (bottom). The right hand picture shows the rider's view.

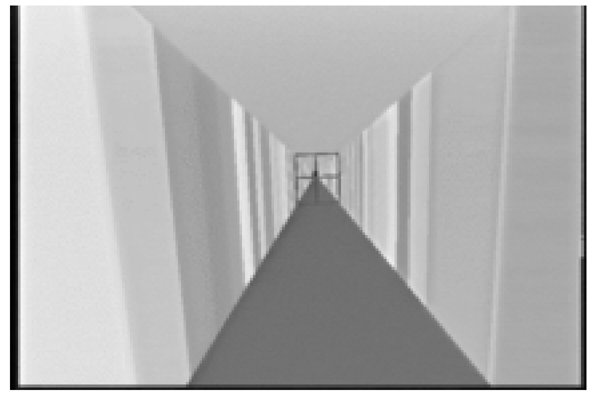

Figure 6. Experiment Display. Subjects were positioned in a corridor with a red target (cross) positioned some distance in front of them. Parallax and scale cues were available to help judge the target distance.

periments (see [7, 8, 19, 9]) we have been investigating the sensory cues that contribute to the perception of linear self-motion in real and virtual environments. In these experiments the steering column was locked and only straight-ahead motion was permitted. In the experiments described in [7, 8, 9] subjects were passively transported in either visual, vestibular or visual-vestibular space. Splitting apart the visual and vestibular cues, and manipulating the relative strengths of the two cues permitted an examination of their relative contribution to the perception of self-motion.

One question which was unanswered by these earlier studies is the relative contribution of active versus passive cues to movement. In [19] we examined the role of active vs. passive locomotion on the perception of self-motion. By mounting *Trike* on rollers it was possible to compare active vs. passive motion without vestibular cues. When *Trike* is actually moved on the ground, vestibular cues were added as well. Figure 5 illustrates the three conditions reported here. The subject can be exposed to only visual cues to movement (Figure 5 top) while sitting on the stationary *Trike*. Proprioceptive and visual cues can be combined in the absence of vestibular cues by riding *Trike* while it is mounted on rollers (Figure 5 centre). Finally all cues can be combined by allowing *Trike* to move under the rider's control (Figure 5 bottom).

The effectiveness of the cues presented in each of these three conditions on the perception of self-motion was examined by having subjects indicate when they felt that they had moved through a visually presented target distance. For these experiments, *Trike* was constrained to move in a straight line and subjects were presented with a visual display that was patterned on hallways at York University. The texture on the walls of the hallway was augmented with a striped texture which was changed on a random schedule in order to ensure that subjects could not track environmental features.

In each trial the subject was presented with a visual target (Figure 6). After a subject-controlled period, the target was extinguished and the subject moved down the hallway. The subject was to indicate when they felt they had reached the previously presented target by pushing a button located on the handlebar of the Trike. In the vision-only condition the subject was translated down the hallway at a constant acceleration in software. In the leg proprioception conditions the subject pedaled at a previously-trained constant acceleration and this was used to drive the visual display. In the leg proprioception + vision + vestibular condition the trike was unmounted from rollers and the subject pedaled down a real hallway while viewing a virtual one as shown in Figure 1.

The experiment was repeated at a number of different target distances and the mean responses for target distances up to 12m are shown in Figure 7. Figure 7 provides data for 10 volunteer subjects. Data from the three conditions are plotted as well as the veridical response. Subjects were very accurate in the vision-only condition. In the stationary-pedal condition subjects pushed the button too late, that is they moved farther than the target before they pushed the button. In the moving condition, subjects pushed the button too early, that is they had not moved as far as the target when they pushed the button.

When a subject presses the button too early it indicates that they perceive that they have moved farther than they

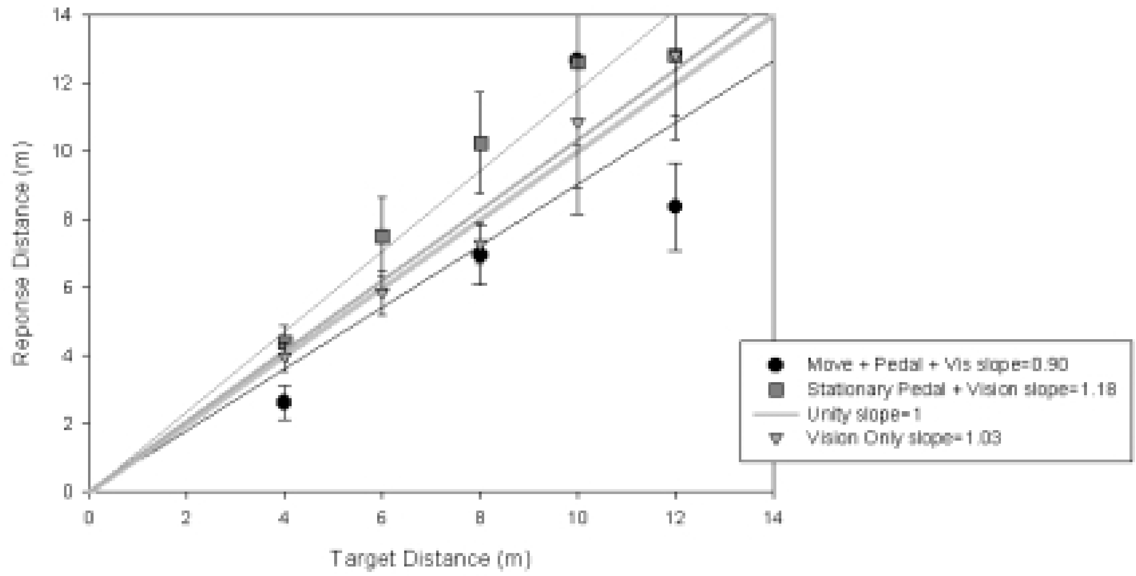

Figure 7. The point at which subjects pressed the button indicating that they perceived that they had travelled through a given distance. The perceived distance depends on the condition.

have really moved. When a subject presses the button too late, it indicates that they perceive that they have not moved as far as they have really moved. Thus in the vision + active + stationary condition, subjects did not perceive that they had moved as far as they had really travelled. In the passive vision-only condition subjects were quite accurate in estimating the distance that they had moved. Addition or removal of the cues which are normally present during self-motion have significant effects on subject's perception of self-motion.

4. Discussion

Human perception is multi-modal. During self-motion we receive a number of different cues from a variety of sensory systems. These cues include: optic flow; auditory motion; vestibularly sensed angular and linear acceleration; proprioceptive and somatosensory information about the position and movements of the limbs; and knowledge that instructions have been sent to the various muscles (efference copy). Each sensory system has different response properties and information from these sensory systems interact in complex and subtle ways to generate our perception of self-motion. It might be thought that virtual reality primarily seeks to mislead the visual sense. In fact the process of providing an image stabilized relative to an imaginary space that a person wearing a head- slaved VR helmet is able to explore, seeks to fool the totality of the various systems that contribute to spatial awareness.

Successful navigation requires keeping track of one's current position in order to compare it with the expected position and to provide a reference from which further movement can be planned. It is possible to calculate position completely from scratch at regular intervals throughout a motion. Such a navigation strategy is called *piloting*. However, the idea that humans regularly use piloting runs counter to our intuition of what happens as we move around the everyday world. Although occasionally we do need to pause and take our bearings, there is normally a sense of continuity during a movement. When we arrive at a location, we have usually already anticipated the position of some of the landmarks, especially key, task-related ones, relative to ourselves. Such anticipation implies being able to use the history of the movement, not relying on a new survey.

Calculating one's path by updating earlier positions using sensory information about the self movement as it progresses can provide an efficient continuity that fits well with introspection and, by extrapolation into the future, allows anticipation. Cumulatively updating one's position in this way is called *path integration*[16] and see Loomis et al.[14] for a recent review. Path integration is part of a so-called dead reckoning navigation strategy in which no access to external landmarks is required between planning a movement and reaching the goal. *Trike* has been designed explicitly to investigate the effectiveness of path integration and answer the question: "How successfully then can a virtual reality system replace the natural experience of exploring an environment?". Humans are able to use path integration[13, 6, 10, 1] and it is known that navigation performance is improved and cybersickness minimized when real walking is allowed[2]. Using *Trike* we can precisely monitor and manipulate the sensory information available during navigation tasks.

Virtual Reality systems, which augment or modify one of these sensory systems, may confound our overall sensation of motion in various ways. A common result of this confusion is nausea (cyber-sickness) and a resulting degradation of performance. Designers of immersive visual systems must take great care that their augmentation of the visual field or other sensory inputs does not interfere unpredictably with the normal perceptual processes. *Trike* utilizes standard virtual reality technologies to generate an immersive visual display while utilizing real motion to generate compatible non-visual cues. This permits operators a wider operational range than is found in more traditional virtual reality systems. Limited only by the length of the power cord and the available free space, subjects can explore large virtual environments and obtain appropriate visual and non-visual cues to their motion.

Initial experiments suggest that subjects do perform differently when presented with different combinations of visual, non-visual and active cues to their motion. In order to construct perceptually-equivalent virtual environments, it is essential that a better understanding is obtained of the relative contributions of the perceptual systems which are active in virtual environments. The experiment reported here investigates perception in a visually impoverished environment and for only a limited set of motion profiles. Ongoing work is extending this to two-dimensional environments and more general motion profiles.

References

[1] M. A. Amorim, S. Glasauer, K. Corpinot, and A. Berthoz. Updating an object's orientation and location during non-visual navigation: a comparison between two processing modes. *Percept. Psychophys.*, 59:404–418, 1997.

[2] S. S. Chance, F. Gaunet, A. C. Beall, and J. M. Loomis. Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration. *Presence*, 7(2):168–178, 1998.

[3] H. Distler and H. H. Bülthoff. Psychophysical experiments and virtual environments. In *Virtual Reality World'96*, Stuttgart, Germany, 1996.

[4] G. Dudek and M. Jenkin. *Computational Principles of Mobile Robotics*. Cambridge University Press, New York, NY, 1999.

[5] S. Feiner, B. MacIntyre, T. Höllerer, and A. Webster. A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban evironment. In *Proc. IEEE 1st Int. Symp. on Wearable Computers*, Cambridge, MA, 1997.

[6] S. Glasauer, M. A. Amorim, and A. Berthoz. Linear path integration during locomotion in normal and labyrinthine-defective subjects. *Europ. J. Neurosci.*, page 210, 1992.

[7] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Vestibular cues and virtual environments. In *IEEE VRAIS'98*, pages 98–105, Atlanta, GA, 1998.

[8] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Vestibular cues and virtual environments: choosing the magnitude of the vestibular cue. In *IEEE VR'99*, pages 229–236, 1999.

[9] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Vestibular capture of the perceived distance of passive linear self motion. *Archives Italiennes de Biologie*, 138:63–72, 2000.

[10] I. Israel, R. Grasso, P. Georges-Francois, T. Tsuzuku, and A. Berthoz. Spatial memory and path integration studied by self-driven passive linear displacement. I. Basic properties. *J. Neurophysiol.*, 77:3180–3192, 1997.

[11] H. Iwata. Walking about virtual environments on an infinite floor. In *IEEE VR'99*, pages 286–293, 1999.

[12] S. A. Lewis, G. D. Havey, and B. Hanzal. Handheld and bodyworn graphical displays. In *Proc. IEEE 2nd Int. Symp. on Wearable Computers*, Pittsburgh, PA, 1998.

[13] J. M. Loomis, R. G. Golledge, and R. L. Klatzky. Navigation system for the blind: Auditory display modes and guidance. *Presence: Teleoperators and Virtual Environments*, 7:193–203, 1998.

[14] J. M. Loomis, R. L. Klatzky, R. G. Golledge, and J. W. Philbeck. Human navigation by path integration. In R. G. Golledge, editor, *Wayfinding, mapping and spatial behavior*, pages 125–152. John Hopkins Press, Baltimore, MA, 1999.

[15] E. Matias, I. S. MacKenzie, and W. Buxton. A wearable computer for use in microgravity space and other non-desktop environments. In *Companion of the CHI'96 Conference on Human Factors in Computing Systems*, pages 69–70, New York, 1996. ACM.

[16] H. Mittelstaedt. Homing by path integration in a mammal. *Naturwissenschaften*, 67:566–567, 1980.

[17] J. J. Ockerman and A. R. Pritchett. Preliminary investigation of wearable computers for task guidance in aircraft inspection. In *Proc. IEEE 2nd Int. Symp. on Wearable Computers*, Pittsburgh, PA, 1998.

[18] A. Pelah and H. B. Barlow. An illusion of accelerated self-motion following treadmill jogging. *Invest. Ophthal. and Vis. Sci.*, 37:2400, 1996.

[19] F. Redlick, L. R. Harris, and M. Jenkin. Active motion reduced the perceived self displacement created by optic flow. *Invest. Ophthal. and Vis. Sci.*, 40:4199, 1999.

[20] A. Smailagic and D. Siewiorek. User-centered interdisciplinary concurrent system design. *IBM Systems Journal*, 1999.

Acknowledgments Financial support from the Centre for Research in Earth and Space Technology (CRESTech) and NSERC Canada are greatfully acknowledged.