The intent to move; generating spatial memory in Virtual Environments.

Uwe Rosebrock and Peter Vamplew

School of Computing, University of Tasmania, GPO Box 252-65, Tasmania, Australia
Email: U.Rosebrock@utas.edu.au, Peter.Vamplew@utas.edu.au

Abstract

The devices used with virtual environments for data entry aim to provide the user with a means to interact with the virtual environment in a natural, multi-modal manner. One particular issue faced by these devices is allowing to user to navigate through the virtual space in a manner similar to their real-world experience. Within the last few years a number of attempts to meet the challenge of providing this functionality have been made. These attempts have used a variety of different approaches to the problem, with significant differences in terms of the size and price of the input device, safety issues and the sense of immersion experienced by the user.

The extent to which spatial memory is produced in the user has been commonly studied in real world experiences, but has rarely been investigated using different input devices within virtual environments. In this pilot study a number of device types are investigated based on different approaches to capturing the user's intent; pointing, waist tether and foot tracking. Runtime data as well as user response to questions examining spatial memory are used to make a statement about the comparative suitability of the tested devices. These result are used to motivate a proposal for a future approach for full-body-motion-capture input devices.

1. Introduction

Virtual environment (VE) interface technology, that is, the devices and requirements that are imposed on the user in order to interact with a VE, have grown in importance with the increasingly popularity of VE systems. In this report the term VE is used synonymously with *virtual reality* and *synthetic environment*. There is no widely accepted definition of the term, and the approach chosen here is to describe a VE system as a computer-generated world with which the user can interact with the purpose of altering the state of the user or of the computer (Durlach and Mavor, 1995). The intention is to provide the user with a meaningful environment with which he or she can interact in a natural, multi-modal manner.

For example a virtual prototyping application might surround a designer with the visual representation of a new Space Station design which they could then move through to determine the ease of access to critical maintenance hatches. In this case, the major form of interaction would arise through the user's body movements, not only in moving to different parts of the space craft, but in seeing whether they could reach a given bolt with enough maneuvering space to exert the necessary torque to release it (Tanner, 1993).

The above example is representative of immersive VE systems, where the user is essentially surrounded by the virtual world to the exclusion of the real world. VE systems may also be non-immersive. In this case, the user views the virtual world indirectly through a computer monitor or some other display and, typically, interacts with the VE using more traditional keyboard, mouse, and trackball interfaces. A third alternative is augmented reality systems where the virtual world is superimposed over the real world. Here the intention is to supplement the real world with useful information, for example, guidance in performing a real world task. Draper et al. (1998) provides a more detailed description and definition of immersive and non-immersive VE as well as other descriptors used such as telepresence. They not only points out that there seems to be some confusion about a clear definition of any of the above mentioned, but also agree with the problem stated by Sheridan in 1988 (Sheridan, 1988) that the experience of telepresence is not even a well formulated research problem.

Regarding entry devices, two major groups can be distinguished. The first finds its origin in the non-immersive VE system but may also be applied in immersive VE systems as well. Generally most pointing devices belong to this group, such as mouse, joystick or related devices. The second group belongs more to the immersive VE systems and reflects the tracking of limbs or the entire body. As such, body motion is translated into VE, e.g a 'data glove' represents a device limited to tracking hand and finger movements. Tracking, also called Position and Orientation Tracking or Position Tracking and Mapping, is used in VEs where the orientation and the position of a real physical object is required. Specifying a point in 3-D requires the transition position, that is, the Cartesian coordinates x, y, and z. However, many VE applications manipulate entire objects and this requires the orientation to be specified by three angles known as pitch (elevation), roll, and yaw (azimuth). Thus, six degrees of freedom (DOF) are the minimum required to fully describe the position of an object in 3-D.

While full-body motion is commonly viewed as the most challenging VE interface technology to be developed, it is important to note that some types of full-body motion are feasible with current technology. Consider those cases where a user is passively moved through a VE in a vehicle (eg. CyberMotion Interactive Motion Seat).
However the more general case of interfaces where the actual body motion of the user is tracked and translated into a VE system are still pose implementation challenges.

Self-motion interfaces or tracking are defined as those cases where the user moves him or herself through a VE, as opposed to being passively moved in some type of vehicle. Currently in many VE systems the illusion of self-motion through the environment is supported by generating visual displays that represent some concept of “flying” when the user points a finger or some other type of pointer in the direction he or she wishes to travel. Undoubtedly, there are many types of application for which such interaction is ideal, but flying through an environment may well give a different perspective and less detailed knowledge of the environment than that which can be acquired by preparing the body to 'walk' through it.

A variety of attempts have been made to build full body motion capture devices to allow this type of navigation. The US Defense Department supported a variety of these projects with the aim to introduce a training device for military purposes. These devices are very costly and some require extensive training to prevent injury. Some others are still on the drawing board and incredibly bulky (Treadport - Christensen, Hollerbach, 1998; Tristano, 1999, IWATA).
There are also some commercially available devices, which are unfortunately either restricted to certain kinds of movement such as the 'Virtual motion' (Global Entertainment Systems, Ltd. England) to a surfing motion or are hi-tech body tracking devices which do not provide any safety precautions to limit the motion of the user and avoid collision with real-world object (Ascension USA; InterSense USA).

Whilst it has been demonstrated that these devices can be used to navigate through VEs, relatively little study has been done on the effects of each style of device on the user's perception and spatial memory of the environment. (Hollerbach, 1999).

This pilot study describes the first in a series of investigations in which the suitability and usability of a selection of available devices to navigate in an immersive VE is to be compared. After reviewing the results a proposal shall be made for a body motion capture device which allows the user to move freely and safely. Also an attempt shall be made to establish what degree of freedom in movements is essential to be described as 'to move freely'.

Another objective for this stage is to identify the importance of the variety of aspects of the human motion with the aim to propose a low cost device covering the most important features of human motion.

It is also proposed, that the intention of moving with the entire body provides sufficient grounds to produce a better spatial memory than simple pointing devices. Where intention means that the actual movement does not have to be performed to its full extent, e.g. having the intention of walking would prepare the body to lean forward and initiate the first step. A body's point of gravity would have the 'intention' to move. This approach transferred into body motion capture could mean that provision made to enable the user to actually walk would not be necessary. Hence the system would still give a sufficient impression of movement in physical space compared to movement in VE without the difficulty of preventing injury.

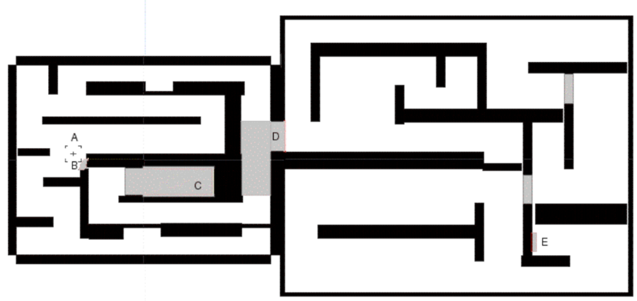
2. Procedure

2.1. The Virtual Environment

In this experiment a number of volunteers were asked to perform simple tasks involving navigation within several VEs using several different styles of entry device.to test spatial memory. Three devices were used representing some of the basic entry devices in VE. Statistics about their performance of these tasks, and a followup interview were used to assess the impact of the entry device on their spatial memory of the VE.

Each VE used in this experiment consisted of a maze of a similar degree of difficulty, with a different VE used for each entry device (to prevent learned experiences about the maze from carrying over from one trial to the next). Each maze had the same number of landmarks and the same task had to be performed in the maze. From an obvious starting point, two ramps or one bridge and a ramp had to be found to enter the second part of the maze. Here a button was to be found and once the button had been pressed the subject had to return as directly as possible to the starting point. Before starting, the subject was asked to stay free of walls and introduced to the fact that they would be questioned about the dimensions and orientation of the maze and its landmarks. The order in which each subject used the alternative input devices was randomly determined.

The mazes were created using the shareware level editor QuArK Armyknife ^{TM} for the ID^{TM} 3dEngine used in Quake ^{TM} II and others games. The freely available source code for the Quake ^{TM} II dynamic linked libraries was altered to measure the desired output of the subjects current position and collision detection. The game specific attributes were removed.



Picture 1:
Example of one maze used, square [A]; starting point; landmarks [B to E], gray; either ramp [C], pathways or bridge [D] or target button [E]. The two rooms were on different height levels.

The subjects wore a 3D helmet in all tests (VFX1, Forte Technologies). As pointed out by Strauss (1995) and Kolasinski (1995) a person exposed to VE can experience severe side effects which could be described as 'Cybersickness' (Strauss, 1995). Every session was therefore kept as short as possible and extensive information regarding possible side effects was provided to subjects.

2.2. Spatial Memory

Spatial memory that is the knowledge about the spatial organization one might gain from experiencing a 3-dimensional environment.
After each walk through a VE maze the subject was asked to sketch the maze on a blank piece of paper and locate landmarks. Each subject was given approximately the same amount of time to finish the task. A black and white plan in birds-eye view of the maze was then provided and the subject was asked to reconstruct their way through the maze and to point out landmarks.
The score was calculated from a total of 28 points. From the total, 14 points were assigned to sketching the maze and 14 to reconstructing the path through the maze. 2 points were given for every landmark for accuracy and orientation in either task (8 total). Another 0 to 3 points were given for the accuracy with which the path was reconstructed when the map of the maze was provided. The last 0 to 3 points were given for the accuracy with which the map was sketched or the path reconstructed.

2.3. Entry Devices

Amongst the great number of entry devices there are only a few types of devices which can be used for navigating in VE. There are the simple pointing devices, which include mouse, joystick, trackball, keyboard, eyeball tracking etc. and more sophisticated devices tracking movement in physical space by relating to movements of certain body parts or the entire body. In this pilot study a standard mouse was used.

Over the last three years a foot-tracking device (the Circumnavigator) was built by a number of postgraduate and Honours students at the University of Tasmania (Denny 1997, Davidson 1998). The user's feet are strapped into two diametrically opposite foot holders. The foot holders are connected via circular bearings to a top disk, which can in turn pivot relative to its base. The user's motion has some similarity to walking on skies. The current angle of the top disk is measured by photoreceptors which read this information from a grey-encoded circular pattern fixed to the underside of the disk. A software interface tracks the changes in these values over time and allows the Circumnavigator to imitate a mouse interface for connection to existing VE software. As the subject‘s feet are strapped to the Circumnavigator and they wore a headset while performing their tasks, a harness is used to protect the subject from any accidents resulting from loss of balance.

To prove the earlier made proposal, that the intent of moving with the entire body provides sufficient grounds to produce a reasonable spatial memory, a design similar to the locomotion interface created by Sarcos Inc (Christensen, Hollerbach, 1998; Tristano, 1999) was built. Sarcos's Treadport comprises a treadmill and a mechanical tether with some haptic feedback. In this study a simple mechanical tether was attached to a subject's waist. Movements to either side forward or backwards are translated into a turning movement in VE or forward/backwards movement respectively. The treadmill and haptic feedback elements are not implemented. The device was linked to the software via joystick interface, used to provide a common connection point to the existing software.

3. Results

The data obtained from each of the test runs through a maze provided the cartesian coordinates in every frame displayed by the software, the frame time, and the current state of the user's location in the VE. The latter was used for collision detection.
It was found that the cartesian coordinates did not provide any more information than using the actual time between certain events and the state the 'user' was in at any time in VE. An example of the raw output data is given in Table 1.

Table 1: Example of the raw output data

	x		z		Y		ms		index
origin x	4572	z	1010	y	346	frame time	23	touches	4
origin x	4575	z	980	y	346	frame time	23	touches	2
origin x	4577	z	954	y	346	frame time	23	touches	2
origin x	4579	z	935	y	346	frame time	22	touches	4
origin x	4579	z	935	y	346	frame time	23	touches	2

Table 1: x,z,y - scales in size in x, y and z direction; frame time reflects the display time of one frame; touches -> index provides a quality indicator in which state the user is in this particular frame (2- ground only, 3 -slope, 4 -wall, 5 -corner, etc.).

Given this data and knowledge of the location of the landmarks within each maze, the exact time taken to reach certain landmarks could be calculated. It was also possible to detect if the subject collided with any walls. Consecutive collisions were treated as a single incident.

The absolute time taken to carry out the task was not considered as valuable data for comparison as the individuals had different preferences of speed in the VE. A more useful measure is the ratio between the time the subject needed to find the final landmark (a button) and the time to return to the starting point. Even this has its limitations in analysing the spatial memory developed by the user on the outward leg of the journey, as there were significant variations in the care with which the return journey was made, as indicated by a relatively high number of collisions. These collisions had the effect of artificially inflating the ratio of time between the two legs of the task.

Therefore the best indicator of the users' spatial memory were the results of the post-task interview process described in Section 2.2. After each VE trial, the user's spatial memory of that environment was rated at a value between 0 and 28, with higher scores reflecting a greater accuracy of spatial memory.

Table 2 shows the results of the mostly raw data received from this setup. It provides the results for each user for each input device. Total time taken, the ratio of time taken on the two legs of the task, the total number of touches (collisions with obstacles), the frequency of the those collisions (to account for user variations in speed), and the score on the spatial memory test. Note that test subjects 4 to 6 did not finish the maze using the Circumnavigator (CN) due to considerable frustration in using the device.

	Input device	Total time [ms]	Ratio in/out	Touches total	Touches every[ms]	Score
Test Person 1	Pointer	232549	3.35	269	864	13
	Waist	799110	3.01	851	939	19
	CN	1041705	2.06	1525	683	21
Test Person 2	Pointer	237360	3.23	241	985	9
	Waist	819172	2.50	477	1717	22
	CN	1173060	3.47	1949	602	22
Test Person 3	Pointer	110940	1.24	75	1479	7
	Waist	558829	2.15	879	635	16
	CN	987721	3.06	3020	327	19
Test Person 4	Pointer	144844	1.89	139	1042	4
	Waist	482352	2.06	605	797	19
Test Person 5	Pointer	147465	1.56	142	1038	10
	Waist	496713	2.03	471	1055	19
Test Person 6	Pointer	177547	2.78	232	765	3
	Waist	698197	2.86	435	1605	22
Average	Pointer (6)		2.34		1029	7.67
	Waist (6)		2.44		1125	19.50
	CN (3)		2.86		537	20.67
Standard Deviation	Pointer (6)		0.90		246	3.78
	Waist (6)		0.42		440	2.26
	CN (3)		0.73		187	1.53
T-test /1 unequal Variances /2 equal Variances	Pointer vs Waist		0.82[/1]		0.32[/1]	0.00003[/2]
	Waist vs CN		0.41[/1]		0.033[/2]	0.22[/2]
	Pointer vs CN		0.39[/1]		0.010[/1]	0.00008[/1]

Table 2. Results of the experiment (CN = Circumnavigator; (n) = n number of test persons; t-test = unpaired, two tailed; touches every [ms]= total time /touches total; ratio in/out = time used to find target / time to return to start; score see Section 2.2 'Spatial Memory'). F-tests were applied to determine if equal or unequal variances had to be used.

From the data in Table 2 it can be seen that the total time taken varies greatly between different users and different devices. As mentioned earlier part of the data has to be seen in context, e.g. the ratio of entry time to return time clearly provides a better indication than total time as it eliminates personal preferences of speed. Additionally, a subject 'rushing carelessly' through the maze will have an increased number of impacts with obstacles, which explains the variants in the number of 'touches' between individuals.

Regardless of these individual variations, there were some clear patterns related to the entry devices. All of the users recorded their fastest time using the pointer device. Their second fastest with the waist-tether and their slowest time with Circumnavigator (users 4 to 6 did not complete the Circumnavigator trial, but at the point of quitting they had already exceeded their times for the other input devices).

In terms of accuracy of control, the waist driven tether and the pointing device navigated with equal accuracy of movement (touches every [ms]) ($p > 0.05$)) through the maze. The CN compared to the pointer and waist driven tether revealed significant differences in accuracy of movement ($p < 0.05$), which was probably the main factor in the frustration experienced in using this device, and would also account for at least some of the decrease in speed. Possibly this effect was exaggerated by unfamiliarity with the device, although as all subjects using the CN persevered at least until they reached the return landmark, this did not seem to be an issue.

Despite significant differences in the accuracy of navigation, there seemed to be little difference in the entry - to return time ratio ($p>0.5$) between all three devices. However the scores for the spatial memory test indicate clear differences between the devices. When compared to the pointing device, the devices involving extensive body movement (the waist-tether and the CN) clearly generated a better spatial memory of the virtual maze ($p<0.001$).

In overall summary, the waist-tether proved to be the most suitable entry device for these tasks, providing a better spatial memory and similar accuracy to the pointer device, whilst being both easier to use and more accurate than the Circumnavigator.

4. Discussion

There have been a reasonable number of studies comparing real world tasks to tasks in VE. Most of these used a desktop setup to impose the VE onto a subject. Few have used headsets to immerse the subject in the VE. Amongst others Wilson (Wilson et al, 1997, 1996) and Bliss (Bliss et. al., 1997) have found evidence that learning about spatial details is possible in VE. On the other hand there seems to be little research about how different entry devices might influence this behavior. In this experiment a number of simple entry devices were tested.

The results of this experiment indicate that the choice of entry device does influence the spatial memory of the environment learnt by the user. Both of the devices which involved movement of the entire body (the waist-tether and the Circumnavigator) greatly increased the user's spatial memory of the VE when compared to the performance of the same users with the pointer entry device which did not require full-body motion.

The observation that the waist-tether produced similar results to the Circumnavigator in terms of spatial memory supports the hypothesis that the intention of moving the body in a direction sufficiently stimulates the spatial senses to enhance spatial memory, without the requirement to actually allow a simulated walking action. In fact, the additional complexities required for a device to mimic such an action may in practice be detrimental to user experience – half of the users in this small trial found the Circumnavigator too frustrating to use and were unwilling to complete the trial run using it.

The question therefore arises what are the means, which have to be considered to assure a safe usage of a full body motion tracking device and what is the degree of freedom necessary to provide a productive experience in VE?

Undoubtedly, if the user is immersed in VE via a head-mounted display there is a need to restrict their movement to a particular location, otherwise they risk colliding with objects in physical space. Some form of safety harness or other barrier to real-world locomotion would therefore be necessary.

Determining the degree of freedom necessary to provide a suitable experience in VE demonstrates a greater problem to answer as it is also exposed to some variations due to subjective preferences. Providing a 'feel' for walking and running currently demonstrates the greatest problem technically, and so it is reasonable to ask; does a low cost device need to provide this functionality if a similar experience can be provided without the technical problem of providing actual walking capabilities? It seems more beneficial to provide greater freedom of movement for the entire body rather than to satisfy one sense of motion with great effort.

It is therefore proposed to develop a device which allows the user to move freely with arms, legs and upper body (including turning movements) but restricting the movement to a finite volume which would also cater for safety issues. A flexible tether setup allowing for the above mentioned features could provide this freedom.

References

Ascension,
http://www.ascension-tech.com

Bliss, J.P., Tidwell, P., & Guest, M. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. Presence, 6(1), 73-86.

Christensen, R., Hollerbach, J.M., Xu, Y., and Meek, S., "Inertial force feedback for a locomotion interface," Symp. on Haptic Interfaces, Proc. ASME Dynamic Systems and Control Division, DSC-Vol. 64, Anaheim, CA, Nov. 15-20, 1998, pp. 119-126

Davidson, C., Walking in Virtual Reality with: The Circumnavigator,
Honours Thesis, October 1998, University of Tasmania, Department of Engineering

Denny, M., Walking in Virtual Reality: The Circumnavigator, 1997, pp. 52-84

Draper, J.V., Kaber,D.B., Usher, J.M. (1998). Telepresence. (Special Section: Virtual Environments: Models, Methodology, and Empirical Studies) Human Factors, v40 i3 p354(2)

Durlach, N. I., & Mayor, A. S. (Eds.). (1995). Virtual reality: Scientific and technological challenges. Washington. DC: National Research Council, National Academy Press.

Hollerbach, J.M., personal communication re the Treadport, July 1999

InterSense,
http://www.isense.com

IWATA Lab, University of Tsukuba, Japan
http://intron.kz.tsukuba.ac.jp/TRDML/torus.html

Kolasinski, G. (1995). Simulator sickness in virtual environments. (Technical Report 1027). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Sarcos™ Treadport
http://www.cs.utah.edu/~jmh/Locomotion.html

Sheridan,T.B. (1988), Teleoperation scorecard: Déjà vu and really new. Proceedings of the International Symposium on Teleoperation and Control (pp.5-9). Bristol, England:IFS.

Wilson, P.N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. Human Factors, 39(4), 526-531.

Wilson, P.N., Foreman, N., & Tlauka, M. (1996). Transfer of spatial information from a virtual to a real environment in physically disabled children. Disability and Rehabilitation, 18(12), 633-637.

Tanner, S. 1993. `"The Use of Virtual Reality at Boeing's Huntsville Laboratories." In Proc. 1993 IEEE Virtual Reality Annual International Symposium, September 18_22, pp. 14_19.

Tristano, D., Hollerbach, J.M., and Christensen, R., "Slope display on a locomotion interface," Intl. Symp. Experimental Robotics, Sydney, March 26-28, 1999.