ACQUISITION OF DIRECTIONAL KNOWLEDGE IN VIRTUAL ENVIRONMENTS CREATED BY PANORAMIC VIDEOS

by

Arefe Dalvandi
Bachelor of Engineering in Information Technology, Sharif University, 2007

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
School of Interactive Arts and Technology

SIMON FRASER UNIVERSITY

Spring 2011

APPROVAL

Name: **Arefe Dalvandi**

Degree: **Master of Science**

Title of Thesis: **Acquisition of Directional Knowledge in Virtual Environments Created by Panoramic Videos**

Examining Committee:

Chair:

Dr. Marek Hatala
Associate Professor,
School of Interactive Arts & Technology

Dr. Tom Calvert
Senior Supervisor
Professor Emeritus,
School of Interactive Arts and Technology

Dr. Bernhard Riecke
Co-Supervisor
Assistant Professor,
School of Interactive Arts and Technology

Dr. Brian Fisher
External Examiner
Associate Professor,
School of Interactive Arts and Technology

Date Defended/Approved:

ABSTRACT

This thesis documents the creation and analysis of virtual environments generated using panoramic video. The proposed virtual environments offer greater visual realism, but are expensive and time consuming to produce. Consequently, experiments were needed to assess how efficiently they support directional tasks or sense of presence. In this study, participants' ability to locate specific places in the environment and their subjective sense of presence were compared across three conditions: panoramic video, regular video and slide show. Participants reported a stronger sense of presence in the panoramic video condition, although none of the techniques demonstrated a greater efficiency in providing directional knowledge. Thus, it does not appear that the costs of creating panoramic video are warranted, especially for those applications involving only the sequential learning of specific landmark locations. However, the current experimental design was found not revealing differences between the three different locomotion techniques, as the tasks were too difficult for participants.

Keywords: Directional knowledge; Navigation; Panoramic Video; Photo-realistic Techniques; Pointing Test; Virtual Environment.

ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to my supervisor, Dr. Tom Calvert, who has supported me in every possible way throughout my thesis with his patience and knowledge, whilst allowing me the room to work in my own way. He has been my inspiration, as I hurdled all the obstacles in the completion of this research work. One simply could not wish for a better or friendlier supervisor.

This thesis would have not been possible without the encouragement and support of my family, friends, and loved ones, especially Parjad Sharifi who performed the hard work of capturing all the video materials for this research.

I am grateful to Bernhard Riecke for introducing me to the field of navigation and wayfinding. I also appreciate the guidance of Dr. Tom Loughin, and Dr. Brian Fisher in the statistical analysis of this study. Finally, I would like to thank Gordon Pritchard and other staff of SIAT graduate laboratory who constantly aided me in preparing and running the computing and video capturing equipment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1: INTRODUCTION

1.1 Problem

Every day there is increasing use of virtual reality techniques in applications such as game, virtual worlds, training, simulation and more. Virtual reality (VR) is a collection of hardware and software coordinated to allow a user to communicate with a computer-simulated environment (Steuer, 1992). This computer-simulated environment is a simulation of the real world (e.g., street views) or an imaginary world (e.g., video game worlds) and is called a virtual environment (VE). The most primary way of presenting a virtual reality is to display the virtual environment's visuals on a computer screen or head-mounted displays. However with more technology involved, such as speakers, gloves, and motion-captures, a greater range of sensation and interactivity can be provided in the virtual reality experience (Brooks, 1999).

Conventionally, virtual environments are created by literally building a 3D model of every object in the scene. Using current technologies, 3D modelling techniques can produce reasonably realistic VEs, but creating a realistic VE comprising the objects found in a real scene, for example a street scene, can involve a lot of detailed manual modelling work and requires high speed computers and expensive display systems to support real-time animation of multiple moving objects (Park & Calvert, 2008).

In most cases, users only want to interact with a small number of foreground objects. However, the 3D rendering engine has to deal with all the geometrical objects in the scene. This significantly lowers the efficiency and creates a serious need for high-speed computers to keep the response time reasonable.

Instead of using a VE with synthetic objects, the environment could also be created with images captured from a real environment. The most important benefit of the image-based approach is that the required amount of work and the complexity of the technique are independent of the amount of detail and complexity that exists in the scene. This approach also provides a photo-realistic view of the environment without requiring special hardware systems. Therefore, it is cost effective and does not need intensive computation compared to 3D modelling techniques (Liu, Sun, Georganas, & Dubois, 2003).

Following the advent of the image-based approaches for creating virtual reality, several techniques have emerged for using images and video to facilitate different applications of virtual environments. Applications such as games, virtual tours, training and experimental VEs, have started to benefit from the photo realistic VEs to improve their usefulness and efficiency.

Image based techniques basically use still images or video with limited or panoramic fields of view. Usually, sequences of images or videos from a real site are utilized in different ways to form the realistic background scene of a VE or simulate a type of interaction with the VE such as looking around, walking, or flying in it.

The area that seems to be receiving the most benefits from image-based methods involves virtual reality created for the purpose of navigation, such as virtual maps, virtual tours, etc. The most popular example of these types of VEs may be Google Street View (Behind The Scene, n.d), which provides sequences of panoramic images from street views and gives users the ability to virtually and interactively navigate in a remote area. An other popular navigation-oriented application of virtual reality is virtual tours, which are designed to provide a sense of orientation in a remote site and create a memory of the scene and marketing targets in it (such as hotels, restaurants, historic sites) for future recall. Such tools can facilitate way-finding tasks for users if they are appropriately designed and implemented. Depending on the way they represent the real world to the user and provide navigation facilities, virtual navigation tools offer a certain level of usefulness. Therefore, these VEs need to be evaluated for their target application.

In this research we focus on using photo-realistic VEs for the purpose of navigation. We propose a method for creating a virtual tour of an urban environment using panoramic images and panoramic video. Then, we evaluate our virtual tour with respect to users' performance in acquisition of some specific types of navigational knowledge.

Navigational knowledge is the knowledge required to develop permanent representations of the location and appearance of significant objects in an environment as well as being able to stay oriented with respect to these objects. This knowledge is acquired in the forms of object (landmark) identity, route knowledge, and survey knowledge which are discussed in more detail in Section 3.2.

The accuracy of the navigational knowledge obtained from an environment as well as the time required to access this knowledge affects travellers' performance in wayfinding and orientation in that environment. This performance, usually called "navigation performance", has been used as the major factor for evaluating the usefulness of navigation-oriented virtual environments.

We evaluate the effectiveness of using realistic images and video (panoramic and non-panoramic) in creating virtual tours, by analysing users' performance in acquiring directional knowledge. We examine and compare this performance in three different situations: a virtual tour implemented with sequences of discrete panoramic images captured at certain distances , a virtual tour implemented by transition videos with limited available field of view, and a virtual tour implemented with panoramic video. After being exposed to these virtual tours and travelling during several learning sessions, users are tested on their abilities to make directional choices.

This study aims to discover comparison information about the effectiveness of discrete versus continuous VE travelling techniques in creating a sense of orientation by comparing the image sequence and video simulations of navigation. In addition, the study examines the role of a panoramic view versus a limited view for transition videos in acquiring directional knowledge.

1.2 Applications

The results of this study can be applied to different fields of virtual environment navigation, such as navigation in virtual tours, virtual maps (e.g., Google street view), virtual game environments, training using virtual reality (e.g., wayfinding training), and simulators (e.g., driving simulators). These results will help in the design of more efficient VEs in terms of giving a sense of orientation to the users and facilitating the way-finding tasks for them.

In cases with VEs that are designed for completing navigation-based tasks in realistic environments, such as some games, the higher the navigation performance, the higher the performance of the main task in the VE (e.g., chasing a target in the game). In addition, for applications such as experimental VEs used to study human behaviour in navigation-involved activities such as studying humans' fear of crime in urban environments (Park & Calvert, 2008), the more realistic the navigation in the VE, the higher the chances that users' natural reactions will be provoked and that experiments will provide valid experimental results.

However, the main application for the results of our study, is building cost effective virtual tours and virtual navigation tools that are specifically designed for providing navigational knowledge about a remote site in a way that improves users' performance in way-finding and orientation when they are exposed to the real site.

Although panoramic images and video seem to be very powerful and popular tools for building virtual environments, they can be very expensive to produce with high quality. Thus, it is necessary to evaluate the role of panoramic imagery in improving the usefulness of the VE before allocating resources to create it.

1.3 Thesis Outline

In this thesis we address the problem of assessing panoramic video-based VEs in regards to how effectively they support acquisition of directional knowledge and sense of presence. Chapter 2 provides an overview of the technologies and methods for capturing panoramic images and panoramic video as well as a review on the previous works and applications of the panoramic imagery. Finally, research motivations, and the proposed technique and materials for creating a panoramic video based virtual tour are explained. Chapter 3 briefly reviews concepts related to navigation and presence. It describes definitions for these concepts and the metrics used for assessing navigational knowledge and presence. This chapter ends with the research motivations and the proposed experiment for assessing the panoramic video-based virtual tours described in Chapter 2. In Chapter 4, research questions, variables of study and the experimental method (participants, materials, and design) are described. Chapter 5 and 6 provide the results of the statistical tests performed for behavioural data, and introspective data respectively. These results are discussed and related to the literature in Chapter 7; a possible path for the future work is also described. Finally, the conclusions are drawn in Chapter 8.

2: BACKGROUND ON USING PANORAMIC IMAGERY IN VIRTUAL ENVIRONMENTS

2.1 Overview

The use of photographic imagery as part of environments created with computer graphics is an innovative and popular technique. Still imagery can be used in a variety of ways, including the manipulation and compositing of photographs inside video based simulations such as movie maps, and the texture mapping of still photographs onto 3D graphical models to achieve photorealism.

Recently, with the advent of numerous VE applications, and advanced photography and image manipulation methods, several techniques for using images in creating VEs have been proposed and investigated. While some image-based methods implement virtual tours and video maps of the real environments, other techniques augment images with 3D modelled VEs for different purposes. The most obvious advantages of using images include providing a photorealistic view and reducing the labour involved in modelling

The Aspen Movie Map (Lippman, 1980) was a very early example of using images to create a VE. To create this virtual map, sequences of television frames were captured at constant distance along a street. Also, at every intersection similar sequences of frames were captured along all the possible turns. Using two optical video disks, travel in this virtual map is simulated so that one of the disks played the sequence corresponding to travel along a street and the other one was positioned at the start of the sequences for upcoming deviations from that street. Appropriate images are selected from one optical disc at a time depending on the user's interaction with the map. In this example, locomotion takes place in discontinuous hops along streets. In addition, view alternation is not continuous. Only two views other than straight down the road are provided (at 90 degrees to the right or left from the straight view) and view alternation can occur as a switch between those two.

Panoramic images and panoramic video are recent techniques used in creating VEs to address the problem of discontinuous view alternation and discrete movement. The use of orientation-independent images allows a greater degree of freedom in interactive viewing and navigation.

2.1.1 Panoramic Imagery

Miller (1996) describes the word panorama as "a neologism taken from the Greek and meaning all seeing" (p. 1). The Oxford dictionary defines panorama as "an unbroken view of the whole region surrounding an observer" and "a picture of a landscape or other scene, arranged on the inside of a cylindrical surface, to be viewed from a central position" (Oxford English, n.d) .

Usually an image with a field of view of the human eye (about 160 by 75 degrees) is termed panoramic (Panoramic photography, n.d.). Some panoramic images cover a 360 degrees field of view. Today panoramic photography is being used in both art and technology.

If an environment is simulated by full panoramic images, when users are virtually placed in that environment they smoothly change their view up to 360 degrees horizontally and 180 degrees vertically. This resembles how people observe the real world surrounding them and increases the realism in the users' experience with the VE. This advantage has motivated much research and development into techniques for making panoramic images as well as using panoramic images for making VEs.

2.1.2 Techniques for Capturing Panoramic Images

Panoramic images are usually captured using special cameras and equipment. Stitching together many images captured by regular cameras can also create panoramas. Currently the following techniques and equipment are being used for panoramic photography:

Lenses with a very wide angle of view such as Fish eye lenses: fish-eye lenses are lenses with very large field of view (e.g., 120 degrees) which makes them powerful tools for creating panoramas (Xiong & Turkowski, 1997). However, the resulting images are distorted and need computational reconstruction to become normal. The most important disadvantage of using a fish-eye lens is that capturing a large scene with many details in one image gives fewer pixels for each detail of the scene. In addition, because of the distortion, captured pixel density is not constant for different parts of the scene. Therefore, the reconstruction method has to approximate many of the missing pixels for the image – especially for the parts with low pixel rate - for example, in order to generate a rectangular image out of a fish-eye image. This decreases the quality of the final panorama.

Parabolic mirrors: curved mirrors can reflect a concentrated image of the environment around them. Instead of directly capturing images of the environment with a camera, the camera can capture the reflection from the surface of a parabolic mirror, which is a distorted reflection of a large part of the environment (Benosman & Kang, 2001). As a result, a large amount of information is captured in a single image but this provides fewer pixels for every visual detail of the scene. This method, similar to fish-eye lenses, has problems with limited image resolution and position dependency, because these sensors capture an omnidirectional scene with a single camera. Images captured using curved mirrors need computational reconstructive processes.

Rotating a single camera or using multiple regular cameras: a panorama can be created by stitching many images from a regular camera. This method has the advantage of offering a high-resolution panorama because it adds up the pixels of all the images, but needs stitching software programs (Neumann, Pintaric, & Rizzo, 2000). Rotating cameras and multi-camera systems are two ways of taking the images required for a panorama. These methods can obtain a high resolution image with uniform resolution although exact camera calibration is necessary to accurately generate a panoramic image from multiple images. Stitching methods and software are described in the next section.

Panoramic cameras: recently, panoramic cameras have been introduced to the world of advanced photography. These cameras can have 5 to 11 lenses (usually around 6 or 8) and can directly capture a cylindrical panoramic image or panoramic video. While they provide the easiest way to generate panoramic images, they are currently very expensive (Point Grey, n.d.).

2.1.3 Panorama Stitching Methods

Techniques for stitching many overlapping images can significantly increase the field of view of the constructed panorama and remove the need for expensive fisheye lenses or panoramic cameras. These techniques apply mathematical transformations to the images to prepare them for being appropriately attached to each other (Irani, Anandan, & Hsu,1995; Szeliski, 1996; Szeliski & Shum,1997).

Stitching techniques for carefully controlled camera motion put constraints on how the images are taken and only produce cylindrical images whereas techniques for uncontrolled 3D camera rotation do not.

2.2 Previous Works with Panoramic Still Images

A very recent and useful example of image-based VEs for navigation is Google Street View (Behind the Scenes, n.d.). Started in 2007 in five cities in the US, it provided panoramic views of the streets so that people could explore the world remotely. For collecting images Google utilized nine cameras covering the 360 degrees horizontal field of view and mounted them on a car. The car was driven down the city streets and cameras captured images at certain distances.

These images were then stitched together to form panoramas and were matched to specific location using GPS data. Using Google street view tools, travelling along streets occurs by stepping from each panoramic position to the next one. At each position, users can smoothly change their view for up to 360 degrees horizontally and 290 degrees vertically. While this view alternation is a great opportunity, navigation is only possible along the specific routes that are captured by the camera. For example, users can choose the routes at each intersection but they cannot move in the direction perpendicular to the route (e.g., they cannot cross the street).

Another similar approach was used to create a Virtual Museum (Miller et al., 1992). In this example at selected points in the museum, a 360-degree panning movie was rendered to let the user look around. Walking from one of the points to another was simulated with a bi-directional transition image sequence, which contained an image for each step in both directions along the path connecting the two points.

Chen (1995) attempted to overcome the moving direction limitation by capturing panoramic images at all the intersection points on a grid map of the environment. He composed a virtual environment by connecting a number of these images to form a walkthrough sequence. Walking in this VE is accomplished by "hopping" to different panoramic points on the grid. Unlike Google street view and the virtual museum proposed by Miller et al., in which stepping occurs in only two directions along the designated paths, using the grid approach hopping can happen in eight directions at each panoramic point on the grid. Consequently users are able to move along a desired path in the environment by connecting short hops in different directions.

However in all these examples using still images and moving in discontinuous consecutive hops results in a lack of realism in the interaction experienced by the user. In addition, most of these VE's can work very well if there are no moving objects in the scene. If there are moving objects, or to have a smooth navigation experience, the sequence of images can be replaced by video.

2.3 Previous Works with Panoramic Video

Moving the point of observation while capturing panoramic images creates panoramic video at different frame rates. Panoramic videos have been used with or without augmented elements (e.g., 3D modelled objects) to serve for different applications such as virtual maps, immersive virtual worlds, video surveillance etc. Research on the panoramic video-based tools can be classified by:

- The technology (e.g., multiple camera capturing system, fish-eye lenses, etc.) and method (e.g., stitching method) used for creating panoramic video;
- The application of the panoramic video such as for video surveillance, remote office, immersive worlds, or navigation tools;
- The proposed interactions and interface devices for communicating with panoramic video such as the display (e.g., head mounted,

desktop, or large screen) or tracking devices (e.g., head tracker, treadmill); and

- The method for improving the perception of panoramic views or navigation in them such as imposing panoramas on 3D modelled surfaces or adding 2D or 3D graphical objects to provide sense of depth.

Utilizing multiple regular cameras is a popular technique for capturing panoramic video. It is inexpensive and not too difficult to accomplish. Neumann et al. (2000) used an array of five video cameras recording images at a 30 Hz frame rate. The images from neighbouring cameras overlap slightly to facilitate the merging process and the result is high-resolution (3K x 480) panoramic images. The video streams feed into a digital recording and playback system which maintains precise frame synchronization. Panoramic recordings took place in an outdoor mall with the camera in a static position and in different lighting situations a well as on a truck moving at speeds between 0-40 mph. Similarly Sato, Kanbara, Yokoya, and Ikeda (2004), acquired movies of outdoor scenes by a multi camera system mounted on a car moving at a constant speed to obtains six 768x1024 images at 15 fps. They introduced a tele-presence system which enables users to move by actual walking and change their view point in a photo realistic virtualized environment using a high resolution omnidirectional movie. In other research, Tang, Wong, and Heng (2002) used live video streams from ordinary CCD cameras installed in working sites and proposed a software system called the immersive cockpit which stitched multiple video streams and recreates a panoramic immersive view at the remote site. At the working site the whole camera set can be moved as long as the relative position and orientation of the cameras remain consistent. Ono et al. (2005) captured panoramic videos for their driving view simulation system by a using a vehicle whose roof was equipped with nine video cameras and ran along a targeted road to capture video of the real world.

However, omnidirectional cameras are available today and although they are relatively expensive, some research has benefitted from their availability and being easy to use. Peri and Nayar (1997) used an omnidirectional camera to capture video at 15 fps frame rate and produced a single video stream with a hemispherical field of view. They have also proposed a real time software system called omniVideo that can generate multiple perspective and panoramic video streams from such an omnidirectional video stream. Kimber, Foote, and Lertsithichai (2001) proposed a virtual reality system called FlyAbout which used spatially indexed panoramic video for navigation simulation. Panoramic videos were captured from continuous paths by moving an ominidirectional camera along those paths.

Depending on the type of interactions provided and the way the navigation is simulated in a panoramic video-based system, different interface devices are being used. The most popular interface devices for interacting with panoramas are head mounted displays and head tracking devices which allow users to turn their head freely and observed the desired portion of the panoramic view (e.g., Neumann et al., 2000). Navigation in the omniVideo system, proposed by Peri and Nayar (1997), can be performed by modifying camera parameters such as pan, tilt, zoom, and roll. Users can modify these parameters using interactive devices such as a mouse or a joystick or a head tracker. In a novel approach, Sato, Kanbara, Yokoya, and Ikeda (2004) used a treadmill to detect user's locomotion speed. Video was played back at the frame rate corresponding to this speed. They projected the omnidirectional movie on an immersive multi screen that covered the front, right and left views of the user. In summary interaction with panoramic video-based environments can be performed by selecting the desired view, choosing the moving path --although options are restricted to the paths that have been captured-- (e.g., the FlyAbout system by Kimber et al., 2001) or zooming in and out on specific objects of interest in the environment.

The already discussed systems and virtual reality techniques, present opportunities to capture live views of a remote site or simulate environments that are difficult or labour intensive to produce using traditional computer graphics modelling methods. Although a number of applications can benefit from realistic scene capture and presentation, interactions with the objects are not possible, and there are still some limitations on how users can vary their viewpoints or achieve sense of depth. Several approaches mix panoramas with 2D or 3D graphic elements to improve perception of depth and distance or to increase the interactivity. Verbree & Anrooij (2004) added 2D graphic shapes such a vanishing lines or circles which varying sizes (based on their distance from the camera) to the images to improve pedestrians' perception of perspective, distance and depth. Ono et al. (2005) synthesized the geometric model and real video images to produce user's view for a driving simulation application. Geometric models were two simple walls along the roadside, which were split into slits. After some image processing, the omnidirectional image is mapped to these slits. Based on where user's viewpoint is located — if it is located out of the capturing path -- the slits rotate towards the user's visual line direction. For example, left view from a point can be composed of forward left, left, and backward left from an omnidirectional image captured at t1, t2, and t3 respectively. Other works include a virtual city simulation system using panoramic images and superimposed geometrical models (Kolbe, 2003), and applications of panoramic images as the background for descriptive and geometric data in location-based services (Hoeben & Stappers, 2006; Teodosio & Mills, 1993).

2.4 Evaluation of Photo-realistic Virtual Environments

Image and video representations of environment have been assessed in a limited way in different application domains such as navigation, electronic commerce, cognition etc.

Gale, Golledge, Pellegrino, and Doherty (1990) compared children's performance on learning routes in a field environment and in a laboratory watching videotape. Participants' ability to recognize scenes was similar in the two conditions. However, their navigation performance (ability to follow the same route, and draw sketch maps of the route) was better in the field environment. Meijer, Geudeke, and van den Broek (2009) demonstrated that visual realism improves participants' knowledge of spatial layout and routes in a virtual environment. They compared participants' map identification, viewpoint recognition, route drawing, and route reversal abilities in a photo-realistic supermarket environment made of panoramic images and in a similar but nonrealistic environment. For all the tasks, performance was higher in the photo realistic environment. Tan, Timmermans, and de Vries (2006) evaluated a stereoscopic panoramic navigation system as a method of data collection about pedestrian route knowledge by comparing it with the traditional paper map-and-pencil technique. Navigation in their VE was simulated by displaying panoramic views of successive intersections along the route of travel. They observed that data collected through maps better describe the route choice behavior of pedestrian than the data collected through stereoscopic panoramic interactive navigation sessions. Through a somewhat similar study, Tan, de Vries, & Timmermans (2006) assessed the potential value of stereo panoramic VR systems in triggering participants to reenact their travel behavior (e.g. stops, turns, shopping, navigation method and duration). They hypothesized that this visual trigger can lead to more valid data compared to a paper-and-pencil diary, which in fact it did based on their experiment.

Rizzo et al. (Rizzo et al., 2004) studied the effect of panoramic images on memory. In their experiment participants listened to a news story while watching the video related to it in three different conditions: single frame video, panoramic video on the flat screen and panoramic video displayed by HMD. They hypothesized that the added engagement provided by panoramic video would improve participants' ability to recall. Accordingly, for the immediate recall of the story items, panoramic situations worked significantly better than the single frame video. However, the HMD was did not work any better than the flat screen.

Howes, Miles, Payne, Mitchell, and Davies (2001), compared QTVR panoramas with hypertext and pictorial images for shopping using an electronic commerce system. According to their results, the QTVR and picture-based environments led to enhanced recall of products over the hypertext environment. Their participants found the QTVR environment more navigable than either the picture-based or the hypertext environment. Participants using QTVR also took significantly shorter routes in shopping for items than participants in the other conditions.

Finally, Macedonio, Parsons, DiGiuseppe, Weiderhold, and Rizzo (2007) found relationships between the immersiveness and physiological correlates of anger arousal (i.e., heart rate, blood pressure, galvanic skin response, respiration, and skin temperature) in panoramic video based environments. They indicated that over time, panoramic video-based virtual scenarios can be physiologically arousing.

Besides the above studies, however, not many research experiments have been conducted to evaluate photo-realistic VEs. The literature still lacks studies that can comprehensively answer questions such as: How effectively photo-realistic VEs supports acquisition of navigational knowledge compared to the non-photorealistic VEs? Or, how effectively different photo-realistic techniques support performance in different types of virtual environment tasks?

2.5 Research Motivations and the Proposed Technique

2.5.1 Motivation

Compared to 3D modelling techniques, photo-realistic techniques require lower costs in terms of modelling labour and high performance technologies to create and represent virtual environments. They also provide higher visual realism. Photo-realistic techniques utilize sequences of images or videos captured from a real environment to computationally simulate motion in that environment. However, different types of photorealistic techniques vary in terms of the cost and amount of effort that is required for their implementation.

Depending on the level of interactivity to be provided, the difficulty involved in creating a photo-realistic virtual environment can range from very simple to very difficult and/or costly. The simplest situation involves recording a sequence of regular images (with limited FOV of less than 90 degrees) along specific paths in an environment. In this case, no interactivity in terms of view alternation or path selection is offered, plus the locomotion is discontinuous and hardly resembles the natural mode of transportation. A higher level of realistic experience can be provided by substituting the sequence of images with linear video or single frame images with panoramic images.

Finally, the most difficult VE model to implement is one which gives the user the ability to navigate in any direction and at any desired speed with the ability to change their view at any time during the navigation. Nevertheless, it is practically impossible to capture all the possible perspectives along all the possible paths in an environment.

The idea of panoramic video is that it allows for all possible view alternations from a specific viewpoint located on a specific path that is already captured. Therefore, if it is captured with sufficient granularity of directional choices, and played at interactively selected frame rates, it can lead to creation of a highly naturalistic VE.

Here, we describe a procedure for making a virtual environment using panoramic video, which can be utilized for implementing highly interactive video-based virtual environments. This system is composed of three main components: (1) the panoramic video capturing system, (2) software for creation of panoramas and the virtual environment, and (3) an interactive chair-based interface which is provided to make the interaction more natural. The components of our proposed system and the procedure for creating a panoramic video-based virtual environment are explained in the following sections.

2.5.2 Proposed System

2.5.2.1 Panoramic Video Capturing System

A system comprising eight regular video cameras connected to a pc is designed and implemented so that the combined field of view of the cameras covers the whole 360-degree horizontal field of view. Cameras are "Sony CXD3172AR", each with a 90 degrees horizontal field of view. Cameras are placed at a uniform height on the outer surface of a cylinder. The cylindrical box contains necessary electronic elements for powering the cameras and connecting them to the PC. Consecutive cameras on the cylinder have 45-degrees difference in their viewing direction and 22.5 degrees overlap in the view angle. This ensures that the resultant images have enough overlap to be stitched to each other.

Using BNC cables, all the 8 cameras are connected to a video card that can handle multiple video inputs and is placed in a regular Windows PC. The PC is also placed on the cart. A software program works with this video card to control the camera capture settings such as video frame rate, resolution, compression etc., and manages the start and stop of the video capture. Although cameras are individually capable of capturing video at 15 fps and 640 X 480 pixel resolution, when they work together the optimum resolution and frame rate decrease to 422X316 and 10 fps respectively. This is probably because of the limitations in the input data bandwidth of the video card.

The cameras and the software program managing them were acquired as a package, but some modifications have been applied to the software program as part of this thesis work. These modifications which are implemented using the C# programming language are basically for attaining the highest possible resolution from the cameras.

Since the cameras and the PC had to be powered during the video capture, two 12-volt batteries and an inverter (12v dc to 120 v ac) are also placed on the cart. Power is supplied for cameras directly from one of the batteries, and for PC through a voltage inverter.

Figure 1: Panoramic video capturing system. It includes 8 cameras mounted on a trolley which carries a PC and batteries for powering the cameras and the PC.

2.5.2.2 Software for Developing Virtual Tours from Panoramas

After the video is captured and stored in the computer a Java program performs the following operations on the video files in order to prepare them for the stitching process:

- Video streams are split into still frames,
- Frames are organized into directories so that the synchronized frames of all the cameras were collected in a single directory.

Given a template panorama which needs to be produced with some manual work involved, the PTGui software program (Photo stitching, n.d.) stitches images in each directory to each other and forms a single panoramic image for every directory. Further, panoramic still frames are transformed to the 3X2 video format of the Pano2Vr software and sequenced and encoded into a Flash movie. The video format of the Pano2Vr program (Flash Panorama, n.d.) is used to prepare panoramic video frames that could be mapped onto the inner surface of a 3D virtual model similar to a cube.

The final step is to use ActionScript code to map the flash movie onto a 3D surface (i.e., similar to the inside surface of a cube, but not with clear edges). This mapping removes the intrinsic distortions found in panoramic images such as inclined horizontal edges. The ActionScript code also controls the projection parameters such as pan, tilt and zoom. Any forms of interactivity in the manipulation of the video such as controlling the frame rate, playing back the selected video, changing the view in the panorama ant etc., can be simply implemented at this stage by ActionScript programming.

2.5.2.3 Interactive Chair Interface

In order to provide an intuitive interaction with the virtual tour system, a rotating office chair is modified so that users can change their view in the panorama by rotating the chair while they are sitting on it (See Figure 2). This is implemented by attaching an optical mouse to the central rotating pivot of the chair to detect the chair's relative direction of rotation and to match the view of the panorama with the mouse cursor position. A user can sit on this chair having a laptop placed on his/her lap on top of a laptop holder. The laptop holder is a simple box designed to keep the laptop fixed in place. It also has straps which go around the user and fasten him/her to the chair.

Figure 2: Interactive chair interface for supporting body-based rotations in the panoramic video. The chair is a regular office chair which is modified by attaching an optical mouse to its rotating pivot. The mouse detects user's rotations in 360 degrees and the display system displays the corresponding part of the panoramic view.

2.5.3 Problem

With the current technology, creating long, high quality panoramic videos requires expensive, special cameras and takes a considerable amount of time, computer memory, and manual work. Also, a highly interactive VE requires that many sequences of panoramic video be captured. To reduce the processing time and memory requirements, panoramic video can be replaced with sequences of images. To remove the cost of expensive cameras and cut down the amount of manual work, panoramic video can be replaced with regular video (single frame video). Each of these simpler techniques carries one of the exclusive benefits of the panoramic video. The former provides a panoramic view and the ability to change the view while the latter provides smooth locomotion.

Considering all the effort that could be saved, it is important to question whether or not either one of these simplified versions can efficiently substitute for panoramic video in a specific application area. Although there are quite a number of research projects focused on techniques for creating panoramic video, the literature lacks comprehensive studies on how effectively the virtual environments created by panoramic video works for different purposes such as navigation.

To make a contribution to this research question, we suggested evaluating a basic implementation of VEs created by panoramic videos before paying the costs for creating highly interactive ones (which require high granularity of video sequences). For this purpose, we prototyped some simple virtual tours with restricted interactivity (e.g., constant speed and predetermined paths) from part of the Surrey Central area using our proposed system. Three versions of these virtual tours were implemented using panoramic video, regular video and panoramic image sequences (we call it “slide show” in this thesis). These virtual tours are similar in terms of being photo-realistic and the quality of technique implementing them. However, they vary in terms of the type of locomotion technique they offer, and the implementation costs they require. In this thesis we conducted an experiment to evaluate these three different locomotion techniques. Inspired by the extensive attempts in using photo-realistic techniques for navigation purposes (Gale et al., 1990; Hoeben & Stappers, 2006; Kimber, Foote, & Lertsithichai, 2001; Kolbe, 2003; Lippman, 1980; Naimark, 1997; Tan et al., 2006; Teodosio & Mills, 1993; Verbree & Anrooij, 2004; Behind the Scenes, n.d), we evaluated these techniques, mainly regarding directional knowledge acquisition. In order to broaden our evaluation criteria and obtain information about the participants' quality of experience using these locomotion techniques we compared their sense of presence in each of the virtual tours.

In the next chapter we explain the basic concepts in the navigation studies and presence studies; then we review the factors that affect navigation performance and sense of presence in a VE. Using the existing theories and experimental results we then, build up our experimental methodology and hypotheses.

3: BACKGROUND ON NAVIGATION AND PRESENCE IN VIRTUAL ENVIRONEMTNS

3.1 Overview

Pedestrian navigation has received a lot of attention from the creators of the photo-realistic virtual environments. Panoramic image and video based environment are extensively used in navigation tools (e.g., Google street view), commercial virtual tours, as well as for training and test purposes such as observation of humans' navigational behaviour (Tan et al., 2006; Waller, 2005; Waller, Beall, & Loomis, 2004). VEs are usually evaluated to see if they efficiently serve their users towards completing a successful task based on the application domain of the VE. The two most popular evaluation criteria for VEs are navigation performance and immersivness of the VE. This is because most of the applications of the VEs (e.g., games, driving simulators, street maps, virtual tours) one way or another need the users to perform successful navigations as well as feel present in the environment to some level.

In this chapter we briefly review the literature on navigation and sense of presence in virtual environments. Finally, we propose an experiment for evaluating panoramic video efficiency regarding these two criteria.

3.2 A Review of Navigation Research

3.2.1 Definitions and Theories

Navigation is an inherently cognitive process that determines paths in an environment and manages travelling in that environment. The actual application of navigation knowledge is called '*wayfinding',* which includes exploration and search tasks (Darken & Sibert, 1993). Navigational awareness is defined as having complete spatial knowledge of an environment in order to orient oneself in that environment. A person is oriented when s/he knows her/his own location relative to other important objects in the environment, and can locate those objects relative to each other. In other words, to stay oriented in an environment one needs to develop a permanent internal representation of the location and identities of significant objects in the environment (McNamara, Sluzenski, & B. Rump, 2008). Spatial knowledge, as described in Siegel and Whites' theoretical framework (Siegel & White, 1975), consists of three types of knowledge: object identity or landmark knowledge, route knowledge, and survey knowledge.

Object Identity or Landmark Knowledge is knowledge about the identity and appearance of the highly salient objects in the environment that are important to navigational memory; such objects include the goals of navigation and indicators of where there is a change in direction. These objects are sometimes called landmarks and are used as anchors for orienting oneself in a new environment. Landmark knowledge is a building block for other spatial knowledge (McNamara et al., 2008; Ishikawa & Montello, 2006; Siegel & White, 1975).

Route Knowledge is procedural knowledge about how to travel along a known route, and can take the form of a sequence of actions (e.g., turn right or left) taken at a sequence of landmarks. With sufficient route knowledge, a navigator can successfully travel from one landmark to another on a known route but does not indentify alternative routes. S/he might know about the approximate distances between landmarks. However, the knowledge about the relationships of places on the route is formed by sequential travel and is unidirectional. Consequently, a person will be better at recalling these relationships when it is in the direction they learned the route (McNamara et al., 2008; Ishikawa & Montello, 2006; Siegel & White, 1975).

Survey Knowledge is knowledge of the overall configuration of an environment from an exocentric viewpoint. It is characterized by the ability to estimate Euclidian distances and infer alternative routes that have not been travelled before. Survey knowledge is sometimes referred to as a cognitive map (McNamara et al., 2008; Ishikawa & Montello, 2006; Edwards, Thompson, & MacGregor, 1998; Siegel & White, 1975)

Based on the theoretical framework proposed by Siegel and White (1975), the process of developing spatial knowledge in a new environment is a stage-wise transition between the landmark knowledge, route knowledge, and survey knowledge acquisition. People progress over time from the basic stage of having landmark knowledge to ultimate stage of having complete survey knowledge. Montello (1998 as cited in Ishikawa & Montello, 2006) however, explained the spatial knowledge acquisition process as a continuous transition between the stages while suggesting the consideration of huge individual difference in this process. For survey maps to develop both the stage-wise transition and continuous transition frameworks require a person to perform multiple route navigations in an environment and metrically scale and integrate routes into a global allocentric reference system. However, Ishikawa and Montello (2006) tested participants over a course of 10 sessions and found contradictory results. Participants were driven along a route in an unfamiliar environment and their progress in attaining knowledge about the environment was recorded after each session. Surprisingly, some participants obtained metric knowledge about the distances and directions (related to the survey knowledge) from the early sessions. Most of the ones that did not achieve this knowledge in the early sessions, never obtained it. Only a few of them progressed continuously over the time.

3.2.2 Navigation Performance

Richardson, Montello, and Hegarty (1999) suggest that performance of learning the spatial layout of the virtual environment is predictive of performance of learning in the real environment as the same cognitive mechanisms are involved. Darken and Sibert (1993) also demonstrated that people use real environment wayfinding strategies in the virtual environments. Therefore, they support the application of environmental design principles in the virtual world in order to improve wayfinding performance.

Navigation performance in virtual environments can be influenced by several factors related to the virtual environment and the users. These factors are discussed as follows:

Interface attributes: attributes of the communication (input and output) devices between users and virtual environment determine how naturally people interact with the virtual world. How strongly these communication tools replicate interactions with the real world can affect participants' ability to successfully navigate in the virtual environment, because, people are more able to apply real world strategies to the navigation. Field of view, photo-realism, desktop vs. immersive and vestibular feedback are among those interface attributes that have been shown to influence spatial cognition (Nash, Edwards, Thompson, & Barfield, 2000). For example, visual realism is demonstrated to improve participants' knowledge of spatial layout and routes in a virtual environment (Meijer et al., 2009). The experiment leading to this result, compared participants' map identification, viewpoint recognition, route drawing, and route reversal abilities in a photo-realistic supermarket environment and in the similar but nonrealistic environment. For all the tasks, performance was higher in the photo realistic environment. Moreover, sensory information about linear and angular acceleration (i.e., inertial information) or body-based cues (i.e., vestibular, proprioceptive, and efferent information) that result from active movement have been shown to be useful for maintaining orientation and facilitating the acquisition of spatial knowledge (Waller, Loomis, & Haun, 2004). Alfano and Michel (1990) have shown that the reduction of peripheral vision impairs self-orientation during the locomotion and decreases performance in forming a cognitive map of a room. This experiment was done in the real world; however, it should be applicable to the virtual environments as people can be assumed to apply the same cognitive mechanisms. Others (Arthur, 1996) suggest that while a very small field of view (less than 40 degrees) weakens performance in tasks such as visual search, navigation, perception of size and space, and spatial awareness, a very wide field of view can cause simulator sickness.

User attributes: individuals' spatial ability, sense of direction, age, gender, and experience affect their spatial cognition and learning performance in a virtual environment. In a wayfinding experiment (Prestopnik & Roskos-Ewoldsen, 2000) participants mentally followed directions for a route from one location on a campus site to another and at the final location they pointed to the origin. Results demonstrated that the accuracy of pointing was predicted by sex and familiarity and that the response latency was predicted by participants' sense of direction. There are reliable differences between males and females' strategies and cognitive abilities in wayfinding tasks as described in (Cushman, Duffy, Stefenella, & Vaughn, 2005; Lawton, 1994). Also, individuals vary extensively in spatial learning abilities (Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006).

Environmental design: complexity, integration, size, density, landmarks, paths, and districts can affect people's process of spatial knowledge acquisition (Nash et al., 2000).

Navigational aids: maps, signs, verbal direction, and photographs can help navigation in the virtual environments. Navigation aids used in the real world have often been adopted for use in virtual environments (Darken & Sibert, 1993; Nash et al., 2000).

Navigation method or metaphor: navigation metaphor or method such as active versus passive (Gaunet, Vidal, Kemeny, & Berthoz, 2001; Witmer & Kline, 1998), or continuous versus teleport (Witmer & Kline, 1998) is important as an improper metaphor can lead to disorientation. Before quantifying the effects of navigation methods or comparing the efficiency of different methods or environments, participants need to take on a particular strategy (Nash et al., 2000).

3.2.3 Measuring Navigation Performance

Obtaining each level of the directional knowledge is associated with specific cognitive abilities; therefore, specific behaviours can be indicators of the level of one's acquired spatial knowledge. Considering these associations, several observable metrics have been developed in order to assess people's spatial knowledge.

3.2.3.1 Landmark Knowledge

The most common approach for measuring landmark knowledge is to use recognition and recollection tasks in which participants recall specific landmarks in the environment they have explored. This is usually done by having participants select pictures of the places or objects they remember from the environment among a sequence of random images. Landmark knowledge is then quantified as the number of correctly identified objects (Gaunet et al., 2001).

Alternative methods have included asking participants to write down the names of all the objects without time constraints and counting the correct number of objects included (McCreary & Williges, 1998), or having participants place landmarks in their proper positions on a map (Edwards et al., 1998).

3.2.3.2 Route Knowledge

Considering a route as sequence of landmarks, route knowledge is related to the participants' ability to make the right decision about how to get from one landmark to the next landmark along a route. Therefore, it is conventionally measured by directional pointing tasks in which participants stop while navigating along a specific route and point to the previously explored as well as the next to be explored landmarks on the route (Mallot & Gillner, 1999; Edwards et al, 1998). Landmark sequencing is also a method for measuring route knowledge (Ishikawa & Montello, 2006) where participants in the task would be required to locate landmarks on a route so that they identify which landmark would be encountered first.

A route imitation task is another method for assessing route knowledge where participants navigate a particular route in the virtual environment several times and then their knowledge is examined by measuring the accuracy of traversing the equivalent route in the real world (Witmer, Bailey, Knerr, & Parsons, 1996; Gale et al., 1990). In another study (Edwards et al., 1998) researchers measured participants response time for locating a particular object in the virtual environment after they learn about the identification of objects during a guided tour on the environment.

3.2.3.3 Survey Knowledge:

Because survey knowledge produces a cognitive map, researchers assess participants' internal depiction of the environment as representative of their survey knowledge. Sketch maps have been proved to be an external measure of participants' orientation and are used for assessing cognitive maps. In particular, they are more efficient for assessing topological knowledge rather metric knowledge (Billinghurst & Weghorst, 1995).

Edwards et al. (1998) used the number of objects participants correctly recalled and placed on a paper-based top view map to measure their survey knowledge following exposure to a virtual environment. In another experiment (Darken & Sibert, 1993) researchers generally determined the extent of participants' spatial knowledge of a virtual environment they were exposed to by having them draw a map of that environment. However, sketch maps carry the risk of being over analyzed especially because they are not independent of the participants' memory, drawing abilities, or the ability to project a 3D environment onto a 2D map. In addition, sketch maps are hard to evaluate quantitatively (Billinghurst & Weghorst, 1995).

Distance and angle estimations are common methods for measuring metric knowledge. Commonly researchers (Ishikawa & Montello, 2006; Koh, 1997) attempt to evaluate survey knowledge by measuring participants' accuracy and response times as they pointed to the locations of particular objects in a virtual environment and/or by measuring participants' estimations of the euclidean distances from themselves to other objects and between pairs of locations.

Another indication of having efficient survey knowledge is the ability to infer shortcuts or take routes not previously taken. Utilizing this metric, participants of an experiment completed the task of locating a particular room in a virtual environment where the two obvious (and more likely to be previously traversed) routes to that room were blocked. In another experiment (Darken & Sibert, 1993), researchers had participants completed a series of searches with no priori information regarding their location. Then, they assessed participants' survey knowledge by measuring the distance they traveled and the ratio of space searched to the total virtual environment space.

However, survey knowledge or cognitive maps are hard to assess because it is difficult to find an external representation of a participant's internal map. Also, cognitive maps are known to be highly subject-specific (Billinghurst & Weghorst, 1995).

3.3 A Review of Sense of Presence in Virtual Environments

3.3.1 Definitions of Presence

The term "Presence" despite its common usage, has been controversial among researchers. Several researchers have attempted to provide a scientific and practical definition for it using different perspectives and theories. A numbers of these definitions and theories that have been mostly used in the practical measurements of presence are selected from several surveys (den Dekker & Delleman, 2007; Ijsselsteijn, a, de Ridder, Freeman, & Avons, 2000; Schuemie, van der Straaten, Krijn, & van der Mast, 2001) and described here.

Ijsselsteijn et al. (2000) distinguished between the physical presence, the sense of being physically located in a virtual space and social presence, the feeling of being together. Lombard and Ditton (2006) identified six different explanations of presence: realism, transportation, immersion, social richness, social actor within medium, and medium as social actor. However, most often in the literature of immersive VE presence is conceptualized as transportation: people are usually considered present in an immersive VR when they report a sensation of being in the virtual world rather than operating it from outside (Schuemie et al., 2001).

A well-known perspective on the nature of presence, which is the basis for several techniques of measuring presence, distinguishes between the subjective presence as a person's judgment of being physically present in a remote environment, and objective presence, as the possibility of effectively completing a task in a virtual environment (Schloerb, 1995 as cited in Schuemie et al., 2001). Another commonly used definition is the degree to which a person feels suspension of disbelief in what he or she is experiencing (Schuemie et al., 2001).

Slater (2004) questioned the value of subjective presence as a factual or authentic entity outside the mind of researchers. Slater and Wilbur (1997) applied the term presence to only the subjective phenomenon and distinguish between "presence" and "immersion". In this perspective, presence is defined as the "subjective sensation of being in a VE", whereas, "immersion" is "an objective description of aspects of the system such as field of view and display resolution".

Similarly, Witmer and Singer (1998) defined presence as "the subjective experience of being in one place or environment even when one is physically situated in another”. However, the word “immersion” has sometimes been also used to define a stage in which a person feels that they are being cut off from the reality, which closely resembles the subjective definition of presence.

3.3.2 Factors Affecting Sense of Presence in Virtual Environments

Several factors related to the virtual environment, the user, the task, or the external world can influence people’s level of presence in a virtual environment. Comprehensive surveys provided by Ijsselsteijn et al. (2000), Schuemie et al. (2001), and Nash et al. (2000) classified these factors related to the virtual environment into:

Resolution: resolution is defined as the connectedness and realism of the information presented. Higher resolutions contribute to higher sense of presence in VEs. (Schuemie et al., 2001; M. Slater, Usoh, & Steed, 1994; Witmer & Singer, 1998)

Consistency: the level of consistency of the information presented to different senses so that they lead the user to the overall perception of the mediated environment. Consistency also refers to the fact that the user can anticipate the effects of actions (Slater, Usoh & Steed, 1994 as cited in Schuemie et al., 2001). Predictability is theorized to allow for better adaptation and lead to more presence (Witmer & Singer, 1998; Barfield & Weghorst, 1993 as cited in Nash et al., 2000).

Breadth: the number of different sensory modalities used synchronously to mediate the virtual experience. More sensory modalities increase the amount of information being provided to the user and isolate the individual from the external distractions; therefore, it leads to a better sense of presence (Sheridan, 1994; Bob G. Witmer & Michael J. Singer, 1998).

Depth: depth is defined as the resolution of each sensory modality provided. Providing more depth to an essential sensory modality allows for better presence (Sheridan, 1994; Witmer & Singer, 1998).

Content: the meaningfulness of the objects, actors and events represented by the medium (Lombard & Ditton, 1997) or the social realism of the environment help the user to achieve a greater sense of presence. (Witmer & Singer, 1998)

Speed: speed is the update rate of the control and visual display. Slow update rates remind participants of the artificial nature of the environment and decrease their sense of presence (Barfield & Hendrix, 1995).

Range of interactivity: this is the extent to which the attributes of the form or content of the mediated environment can be manipulated (Sheridan, 1994). Greater ability to change and modify the environment is predicted to increase presence (Witmer & Singer, 1998).

Motion: mobility of the user and dynamic objects in the environment improves sense of presence (Witmer and Singer, 1998).

The communication medium: this refers to the ways (e.g., the hardware) in which users interact with the virtual environment. For example, stereoscopy, spatialized sound and head tracking devices have been shown to significantly increase presence (Nash et al., 2000; Schuemie et al., 2001)

Besides the virtual environment's features, the adaptability, motivation and experience of the participants play a great role in indicating their level of presence (Lombard & Ditton, 1997). Witmer and Singer (1998) suggest that the willingness of subjects to interact and believe the realness of the environment increases their sense of presence. Generally, users with more experience in virtual environments may achieve a greater sense of presence (Barfield & Weghorst, 1993).

The type of the task and its degree of difficulty and automation also affects presence. If a task needs a great amount of attention it increases the sense of presence (Witmer & Singer, 1998; Barfield & Weghorst, 1993 as cited in Nash et al., 2000). Finally, there is a higher chance of feeling present in a virtual environment if there are fewer distractions in the external world surrounding the user (Witmer & Singer, 1998).

3.3.3 Measuring Presence

Presence is measured using subjective and objective methods. Subjective measurements rely on participants' judgements of their sense of presence. Objective measurements measure participants' sense of presence using metrics that can predict participants' level of immersion based on participants' behaviours, or performance in specific tasks.

3.3.3.1 Subjective Measurements

Post-test rating scales: subjective questionnaires are the most commonly used method for measuring presence. These subjective ratings are argued to be the primary method of measuring presence because presence is essentially a subjective sensation. Witmer and Singer's PQ (1998), the Igroup presence questionnaire (Schubert & Friedmann, 1999) and the ITC sense of presence inventory (Lessiter, Freeman, Keogh, & Davidoff, 2000) are among the most well known presence questionnaires formed on the basis of different theoretical views on the concept of presence. The theoretical basis of a presence questionnaire determines its scope and relevance to certain application domains. The drawback of the questionnaires is that they are highly dependent on participants' judgements of their sense of presence as well as their memory of the experience with the VE. Slater (2004) argues that researchers cannot rely heavily on questionnaire results because the idea of presence may be something that researchers bring into the mind of participants.

Continuous presence assessment: instead of rating the sense of presence after the virtual environment exposure, this method measures presence during the experience of virtual reality. It is usually implemented by asking participants to move a slider in order to identify their level of feeling present. The benefit of this method is that it is sensitive to the temporal variations in the sense of presence and is not dependent on the participants' ability to recall their sensation of the experience. However it has been argued that participants need to divide their attention between the virtual reality experience and the slider-controlling task; therefore they cannot reach the belief of being present in the virtual reality (IJsselsteijn et al., 2000).

Psycho-Physiological measurements: cross-modality matching is a psycho-physiological method for measuring presence relying on the fact that presence cannot be easily stated verbally. Using this method, participants are required to express their judgement of a subjective sensation in one modality by responding through adjusting a parameter in another modality. For example they are required to make a sound as intense as they feel present in a virtual environment (IJsselsteijn et al., 2000).

Breaks in presence: using this technique, after participants start feeling present in the virtual reality they report on a break in their presence whenever they become aware of the reality (Slater & Steed, 2000). This method is highly dependent on the participants' judgments about whether they are present in the virtual world or the real world. Especially in the cases that a person has partially concentrated on both of the mediated and the real world, such judgements can be very unreliable or difficult to make.

3.3.3.2 Objective measurements

Behavioural measures: the fundamental assumption in this method is that if people are highly present in a virtual environment they react to the virtual stimuli as if they are real. Reflexive responses such as avoiding a rapidly approaching virtual object (Sheridan, 1994), socially conditioned responses such as smiling, and postural adjustments such as leaning in the counter-direction of a mediated movement (Ijsselsteijn et al., 2000) are representatives of presence in this method. A benefit of this method is that these responses are triggered without the participant having any control on them. However, misinterpretation of the responses by the observer is possible in this method (Ijsselsteijn et al. 2000).

Dual task measures: reaction time to a secondary task is claimed to be a measure of presence as it is assumed that when a greater part of the attention is dedicated to the primary task, a smaller part is focused on the secondary task (Witmer & Singer, 1998; Barfield & Weghorst 1993). Therefore, performance in the secondary task is related to the amount of immersion in the primary task. In an example of this method (Rudolph P. Darken, Bernatovich, Lawson, & Peterson, 1999) participants were exposed to a VE and a movie (as the secondary task) and were asked to recall from both. Their presence in the VR was measured by the amount of the narrative they could recall from the movie. In another experiment, while participants played a video game, an alert signal broke their presence. Participants' response time to the signal was considered as the measure of their presence.

Adjustable distraction method: inspired by Lombard and Ditton's definition of presence as ''the degree to which inputs from the physical environment are shut out'' (Lombard & Ditton, 1997), the underlying supposition of the adjustable distraction method is that presence is as strong as the minimum amount of an external stimuli required to break it. In this method a distracter primitive stimulus (visual or audible signal), which varies from undetectable to unavoidable, is emitted and participants are required to react to this signal whenever they notice it. The degree at which the signal starts to become noticeable to the participants demonstrates their level of presence (Nordahl & Korsgaard, 2009).

3.4 Research Motivations and the Proposed Experiment

The purpose of our virtual tours described in Section 2.5 is to introduce a remote neighbourhood and its landmarks to a user so that the users can successfully maintain a sense of orientation while visiting that area in the real world. In order for an individual to stay oriented in an environment, s/he needs to develop internal representations of the location and identities of significant objects in the environment. This is so that s/he is able to identify her/his own location relative to these objects as well as their locations relative to each other (McNamara et al., 2008). Therefore, an experiment, which assesses these participant abilities, can measure the efficiency of virtual tour implementation techniques.

To assess participants' abilities to identify significant objects in the environment, a landmark recognition task is a clear option and easy to implement as discussed in Section 3.2.3.1.

To assess higher levels of spatial knowledge, there are several well-known methods proposed by the literature (as described in Section 3.2.3), such as distance and angle estimations, map drawing, searching, and wayfinding. Distance estimation techniques are not appropriate for our case, as knowledge about the distances between objects does not necessarily identify their relative places. Other methods, such as assessing wayfinding or search strategies, are also clearly not appropriate for evaluating our virtual tours, as they need participants to actively explore the environment. Among the techniques used for assessing participants' knowledge of the relative locations of objects, sketch maps and angle estimation methods seem to fit our question best as they mainly assess participants' knowledge of "what is where" in the environment. Due to the difficulties of assessing sketch maps and all the risks they carry (as discussed in Section 3.2.3), we relied on the direction (angle) estimation method for our experiment.

The direction estimation method has been extensively used in the literature of navigation to assess participants' accuracy of allocentric and egocentric images of an environment in different experimental conditions (Mou, Timothy P. McNamara, Björn Rump, & Xiao, 2006; Riecke, Veen, & Bülthoff, 2002; Wraga, Creem-Regehr, & Proffitt, 2004). Using this technique, participants point to specific objects in an environment and researchers record the amount of error they make by subtracting the angle of the pointed direction from the actual angle of the target object's direction. Researchers then utilize this data to calculate three types of error, each indicating a different aspect of a participant's spatial knowledge: ego-orientation error, absolute (or signed) pointing error, and configuration error. Ego-orientation error represents participants' misunderstanding of their self-orientation, absolute pointing error indicates participants' misunderstanding of their location relative to the other objects in the environment, and configuration error indicates their misunderstanding of the location of other objects in the environment relative to each other (configuration of the objects). Therefore, these errors overall, indicate how well a person is oriented in an environment.

Due to the extensive amount of information we can extract from the participants' pointing errors about their sense of orientation, we were strongly motivated to use the direction estimation test in our study. Using this test, we can measure participants' performance (accuracy and response time) in pointing from one significant location in the environment to the other significant locations.

To perform a direction estimation test, several methods have been used in the literature such as drawing the direction on the paper (Koh, 1997; Satalich, 1995), pointing in the real environment, and pointing in a panoramic image using HMD, desktop display, or printed images. Waller, Beall, and Loomis (2004) conducted research on the accuracy of different tests for measuring pointing errors. They asked participants to point from one landmark in a familiar environment to the other known landmarks in that environment using four different situations: a real environment, panoramic images on a desktop PC, panoramic images in a HMD, and with pen and paper. Based on their results, pointing tests performed using panoramic images displayed on a HMD and desktop computers are not significantly less accurate than pointing tests performed in the real environment. This makes panoramic images a powerful tool for measuring pointing errors or directional perception. Therefore, we used panoramic images on a laptop screen for our pointing tests.

In addition, Waller et al. (2004) show that participants' pointing abilities using panoramic images of a real environment in a virtual setting reflect their abilities in the corresponding real environment. Taking advantage of this, we ran our pointing tasks with participants who were exposed to panoramic images of the environment in a laboratory after they have travelled on the three different virtual tours. Consequently, we claim that our results can be extended to real environments. This means that significantly better pointing performance in any of the virtual tours would demonstrate the effectiveness of the corresponding simulation technique in providing required directional perception for participants' future navigation in the real environment.

Besides the directional performance, we assessed participant-acquired sense of presence in the virtual tours created by the panoramic video, regular video and slide show techniques. Sense of presence is important in cases where a virtual tour is used not only to provide directional knowledge, but also to give a sense of being in the remote environment (e.g., some games). Besides, knowing about the participants' quality of experience can provide additional information about why a certain locomotion technique is possibly more efficient than the others. For example, the added engagement generated by a strong sense of being in a virtual environment can also affect participants' task performance in some cases. We used subjective questionnaires for measuring sense of presence because objective measures can interfere with the assessment of task performance and they are also more difficult to analyse.

Because of the increased range of interactivity (i.e. view alternation), the additional sensory information (i.e. body-based cues provided by chair-based interface), and the more naturalistic locomotion (i.e. continuous locomotion) offered by panoramic video, we expected that a person's awareness of the environment and their self-orientation would be enhanced using this technique. Following a similar discussion, we also predicted a higher level of sense of presence in the panoramic video. These rationales are more broadly investigated in the next chapter where we explain the details of our research question and experimental design.

4: RESEARCH METHODS

4.1 Overview

In this chapter we explain the details of our research method for assessing participants’ directional knowledge and subjective sense of presence in virtual tours created by panoramic video, regular video and a panoramic image slide show. Research questions and hypotheses are described followed by the experimental method we used for studying our participants’ behaviours. The experimental method is described in terms of participants, experimental settings, experimental design, and experimental procedure.

4.2 Research Question

The fundamental question for this research was:

“How does the locomotion technique offered by a photo-realistic virtual tour of an environment affect human acquisition of directional knowledge and sense of presence in that environment?“

In the other words, the question was how fast and accurately can people obtain directional knowledge about an environment when they learn about it by navigation in virtual tours of that environment using different travelling techniques. More specifically, this question aimed to compare human directional knowledge acquisition performance in three different virtual tours implemented using the following techniques:

Panoramic slide show: locomotion in this virtual tour involved abrupt transitions between spatially separated locations and was simulated by displaying a slide show of panoramic images captured at these locations in the environment. The traveller could navigate in the virtual tour by hopping from one position to another and at each position could look around by smoothly alternating their view in the associated panoramic image.

Regular video: in this virtual tour locomotion was simulated by displaying a 15 fps front facing video recorded while moving through the environment. The video was a regular video with a limited field of view (FOV) - approximately 90 degrees horizontally. The resulting navigation was smooth and continuous but the view was front facing and it was not possible to change it during the locomotion.

Panoramic video: this technique was similar to the last one except that the recorded videos were panoramic with a 360-degree field of view. In this version of the virtual tour, not only was the movement continuous and smooth, but also, smooth view alternation was possible at any point during the navigation.

Table 1: Three virtual tour locomotion techniques compared in this study and their characteristics

Condition	View alternation is possible	Transition is smooth
Panoramic slide show	YES	NO
Normal video	NO	YES
Panoramic video	YES	YES

Following these different approaches to simulation of travelling, our research question could be divided into the following concise sub-questions:

- How does continuous movement vs. discrete movement in a virtual tour affect human directional knowledge and sense of presence acquisition?
- How does the ability to change the view direction during navigation in a virtual tour affect human directional knowledge and sense of presence acquisition?
- How does learning by repeating navigation tasks in the virtual tour affect human directional knowledge acquisition in the three different locomotion techniques?

To answer the above question, we assessed participants' subjective sense of presence as well as their directional knowledge. We believed that participants' performance in orienting themselves in the environment indicates their directional knowledge about that environment. As discussed in Sections 3.2 and 3.4, to orient oneself, one needs to be able to recall significant places in the environment and recognize one's location relative to these significant places as well as their locations relative to each other. Performance is then defined as the accuracy and response time in completing a task. Thus, we measured participants' accuracy and response time in completing recollection and orientation tasks to understand their directional knowledge of the environment following their navigation in the virtual tour. We assessed participants' recollection performance using the following two measures:

Recollection error: this variable measured the accuracy of participants' recollection of places. It was obtained by calculating the percentage of the number of places a participant cannot correctly recall from an environment after navigation in the virtual tour of that environment.

Recollection time: this variable measured participants' response time is recollecting places. It indicated how quickly participants can access their memory to recall a place they have previously visited in a virtual tour.

Consequently, participants' performance in the recollection task was considered to be improved if the average value of at least one of the recollection error or recollection time is decreased.

To assess participants' performance in orienting themselves we used the following measures inspired by the work of (Riecke, Cunningham, & Bülthoff, 2007):

Absolute ego-orientation error: this error measured the constant error in pointing judgments. Ego-orientation error generally represents the difference between the participants' actual heading and the heading assumed by the participant while pointing. Ego-orientation error at a certain location in our experiment was obtained by calculating the circular mean of signed pointing errors made at that location.

Configuration error: the configuration error is generally defined as the standard deviation across target objects of the signed pointing errors and indicates the accuracy of the localization of each target in relation to the others. In other words, configuration error is a measure of the internal consistency of pointing judgments. At a certain location in our experiment, it is obtained by calculating the mean angular deviation of signed pointing errors made at that location. We used circular statistic for calculating ego-orientation and configuration errors as pointing errors are intrinsically circular data.

Absolute pointing error: this variable demonstrates how accurate the participants' knowledge of their location is relative to a target object or locations in the environment. It is the unsigned difference between the pointed direction and the actual direction of the target.

Pointing time, this variable measures how quickly participants can access their directional knowledge about the environment and its objects.

In our experiment, absolute ego-orientation error, configuration error, and absolute pointing error evaluated the accuracy of participants' estimations of the self-heading, object-to-object and self-to-object localization respectively. Pointing time, evaluated the response time in making those estimation. An improved performance in the orientation task was consequently defined as a decrease in at least one of the absolute ego orientation, configuration or absolute pointing error while the others are not impaired, or a decrease in the response time while the error values are not increased. For example, if the average value of ego-orientation error decreased in one condition while the average value of configuration error and absolute pointing error did not increase, we would be observing an improvement in the orientation performance.

Finally, the main research questions were: (1) is a participant's performance in orienting oneself in an environment affected by the locomotion technique they used when learning the environment? (2) is a participant's sense of presence in a virtual tour affected by the locomotion technique that simulates that virtual tour?

4.3 Discussion and Rationales

In order to build up our hypothesis we looked at possible reasons why a participant would make different types of errors and how the three different locomotion techniques could affect these errors. In the following, we discuss the relationships between specific types of errors and locomotion techniques. Then we discuss how these facts led us to the hypotheses.

Regarding the recollection task, as panoramic views provide a larger amount of visual information, it was reasonable to expect people to remember significant locations in the panoramic conditions more conveniently. If participants are required to recall significant locations such as intersections (not just a single significant object) in an environment, panoramic views give them more possible visual cues for remembering the place. As in real physical environments, landmark objects are not always available in the front view, so having access to peripheral view helps participants to find objects that stand out in the environment and remember them as landmarks.

Regarding the pointing task, participants make pointing errors if they misjudge their self-heading, or objects' locations in the environment. A participant makes an ego-orientation error possibly because s/he cannot correctly identify their self-heading relative to the objects available in the scene at the pointing location. As there are more objects in any single panoramic view than non-panoramic views, participants are likely to be able to identify and remember their self-orientation more accurately. Therefore when they are returned to the pointing location they can examine their heading relative to more objects available in a panoramic view (or real world views).

Configuration error would possibly come from a participant's misjudgement of the distances between where the turns happen or misunderstanding of the turning directions along the route. Judging distances is apparently easier in video conditions compared to the slide show condition because in the slide show condition participants are teleported from one panoramic point along the route to the next panoramic point without traversing the in-between path. The study of James and Craig (1995) on the perception of distance in a VE demonstrated that participants had better performance in estimating distances between themselves and a target when they were permitted continuous viewing of the environment and the target during the movement, compared to when they only watched the target prior to the movement. Witmer and Kline (1998) also showed that in a VE when participants actively traversed the route segment using a joystick or treadmill interface, their distance estimation was improved compared to when they viewed the environment from the beginning point of a route segment and were then teleported to end point of that route segment. The slide show condition of our experiment was similar to the teleport condition and the video conditions were somewhat similar to the active traversal condition except that the speed of forward movement is not actively controlled by the participant. Therefore, we expected the video to help participants' perception of distances between route segments.

Judging turn angles was expected to be easier in the panoramic video condition compared to the other two conditions because, as explained in the following arguments, both the panoramic view and the video can help better understanding of turn directions:

> A turn is indicated by a rotation of view of the displayed images, not the actual rotation of the participant in all three of our virtual tours. However, in panoramic conditions participants have the chance to actively rotate themselves during the turn. For example, when a 90-degree turn is complete, participants can rotate their body and their view in the panorama to their right (or left) and examine the turn direction by looking at where they just came from. A study (Wraga et al., 2004) on the effect of active and passive rotations (display vs. participant) indicated that the unsigned error of the perceived turn direction is significantly less in the

active condition. Participants were also faster in this condition. Therefore, we expected a lower average configuration error and pointing error in the panoramic conditions.

On the other hand, studies show that when there are no body movement cues, optic flow plays an important role in understanding directions (Kearns, Warren, Duchon, & Tarr, 2002). As there is no optic flow indicating turns in the slide show condition we consider that there is a higher chance of misjudging turn angles in this condition, which leads to a higher average configuration and absolute pointing error. Correspondingly, Gaunet et al. (2001) tested participants' ability to reproduce the shape of a route after navigation along that route in a passive smooth transition mode (similar to our video conditions but with computer-generated graphics) and a passive snapshot exploration mode (similar to our slide show but with computer-generated graphics). They observed significantly greater errors in the snapshot mode than the smooth transition mode.

In addition, in the slide show condition 'passing a turn' results in several discrete and abrupt alterations in the orientation. At each step of alteration the participant does not know which direction s/he is going to face at the next step. This condition is similar to the 'disorientation' condition of the experiments run by Mou, McNamara, Rump, and Xiao (2006) in which a participant does not know how much s/he has rotated and then is asked to estimate their rotation angle. In this condition the participant does not have any visual flow or body based cues to estimate the angle of rotation. On the other hand, video conditions, at the turning locations, are similar to the ‘update’ condition, in which a participant is ordered to rotate with opened eyes and then estimate the turning direction. In the ‘update’ condition participants can use visual and body-based information to understand the angle of rotation. Mou et al. (2006) observed that ego-orientation error is significantly more in the ‘disorientation’ condition than in the ‘update’ situation. Although this result is found in the real world condition, it helps us build up our hypothesis for the virtual conditions. Therefore, for the slide show tour we anticipate misunderstanding of the ego-orientation at each orientation alteration step during a single turn. This can lead to a large total misunderstanding of the angle turned. Consequently when all the turns along a route are combined, participants’ understanding of the configuration of the significant locations in the environment will be distorted.

To predict which locomotion condition would provide a stronger sense of presence, we looked at the literature for factors affecting sense of presence in virtual environments (discussed in Section 3.3.2). Consequently, we expected our participants to experience a greater sense of presence in the panoramic conditions (panoramic video and slide show) compared to the non-panoramic condition (regular video) for the following reasons:

(1) A VE that provides more sensory modalities is expected to increase the sense of presence (Steuer, 1992; Witmer & Singer, 1998). Apparently more sensory modalities are engaged in interaction with the

panoramic views, as participants would have both visual and body based senses involved when changing their view. (2) In the panoramic conditions participants would have a greater range of interactivity, which should help their sense of presence (Steuer, 1992; Bob G. Witmer & Michael J. Singer, 1998). (3) Participants were required to devote more attentional resources to control a larger visual area and physical rotations in the panoramic conditions. Devoting more attentional resources to the VE has been shown to increase participants' sense of presence (Barfield & Hendrix, 1995; Riley, Kaber, & Draper, 2004; Bob G. Witmer & Michael J. Singer, 1998). (4) There was user motion involved in the interaction with panoramic views, which has been demonstrated to increase sense of presence (Witmer & Singer, 1998),

On the other hand, we relied on the following facts to predict a higher sense of presence in the video conditions (i.e. panoramic videos) compared to the non-video condition (i.e. slide show):

(1) Slow update rates would remind participants of the artificial nature of the virtual environment. Videos had a natural update rate compared to the slide show (Steuer, 1992; Witmer & Singer, 1998). (2) There was more motion of the components of the environment in the video. Motion of the components of the environment has been shown to increase sense of presence (Witmer & Singer, 1998).

To sum up, panoramic video had the benefits of both panoramic view and video as discussed above. Plus, it provided a more natural mapping between the controls and the actions in the VE (Steuer, 1992; Witmer & Singer, 1998), because the walking action was mapped to the video of walking and the looking around action was mapped to a physical rotation. Consequently, we expected it to grant a better sense of presence compared to the other two conditions.

4.4 Hypotheses

Hypothesis 1: Participant's performance in recollecting places will be improved after navigation in the panoramic video tour where user-controlled changes in viewing directions are possible compared to the regular video tour where user-controlled changes in viewing directions are not possible.

Hypothesis 2: Participant's performance in orientation (pointing) will be improved after navigation in the panoramic video tour where user-controlled changes in viewing directions and continuous transitions are possible compared to the normal video tour where user-controlled changes in viewing directions are not possible or the slide show tour where continuous transitions are not possible.

Hypothesis 3: participants' performance in both recollecting and orientation will be better after navigation tasks are repeated in all three locomotion conditions (normal video, panoramic video, and slideshow).

Hypothesis 4: Participants' average subjective sense of presence is greater during navigation in the panoramic video tour than in the regular video and slide show tours.

Hypothesis 5: Performing pointing tasks is subjectively easier after navigation in the panoramic video tour than in the regular video or slide show tours.

4.5 Variables

For testing the first hypothesis the independent variable was the locomotion technique having three levels of panoramic video, regular video, and slide show. The dependent variable was the recollection error and recollection time.

For testing the second hypothesis the independent variable was the locomotion technique. The dependent variables were the absolute ego-orientation error, configuration error, absolute pointing error and pointing time.

For the third hypothesis, the independent variables were the type of locomotion technique and the number of repeating trials, and the dependent variables were the recollection error, recollection time, absolute ego-orientation error, configuration error, absolute pointing error and pointing time.

In this thesis, for the sake of simplicity, sometimes we refer to the collection of absolute ego-orientation error, configuration error, and absolute pointing error generally as the pointing errors because their values came from the pointing data. We also call the collection of recollection error, recollection time, pointing errors and pointing time, behavioural dependent variables as they come from the behavioural tests.

For our fourth and fifth hypotheses, the independent variable was the type of locomotion technique and the dependent variables respectively were the difficulty of the pointing task and the sense of presence perceived by the participants. These variables were measured by introspective questionnaires; therefore, we refer to them as introspective dependent variables in this thesis.

4.6 Experimental Method

4.6.1 Participants

We used a quota-sampling method by recruiting whoever was available on SFU Surrey campus until we had enough of both male and female participants to meet our target of a total of 18 subjects. Gender equality was the only discrimination we made for selecting participants; this was applied due to the fact that other research has shown that males and females show significant differences in their navigational knowledge acquisition strategies (Cushman et al., 2005).

Our subjects were 18 adults, nine females and nine males, in the age range of 23 to 40 years old. They were recruited by being asked in person or by collective emails. Therefore, our participants were mostly undergraduate and graduate students in SIAT and the Mechatronics program at SFU, who were interested in the study or accepted to participate in the experiment in exchange for a standard payment. Table 3 describes some demographic information about the participants (See Appendix C).

4.6.2 Experimental Materials and Settings

In order to prepare an appropriate and valid experimental design to answer our research questions we had to consider several issues regarding the type of environment from which we captured our videos, the way we collected video and images, and the pointing method. Details of the materials and settings for our experiments are as follows:

4.6.2.1 Environments

We selected three regular residential environments from the area close to the SFU Surrey campus. The three environments were similar in terms of their general environmental look, the shape of the traversed routes, the number of turns in the routes, and the angles of each turn. Maps of selected routes for each condition are presented in Figure 3, Figure 4, and Figure 5. There are four turns during each route. All turns are at 90 degrees and there is a total of 180 degrees change in direction. Each route passes by at least one four-way intersection and three three-way intersections in the environment. We tried to do the pointing tests in the real environment prior to selecting it, in order to obtain a sense of the complexity of the environments.

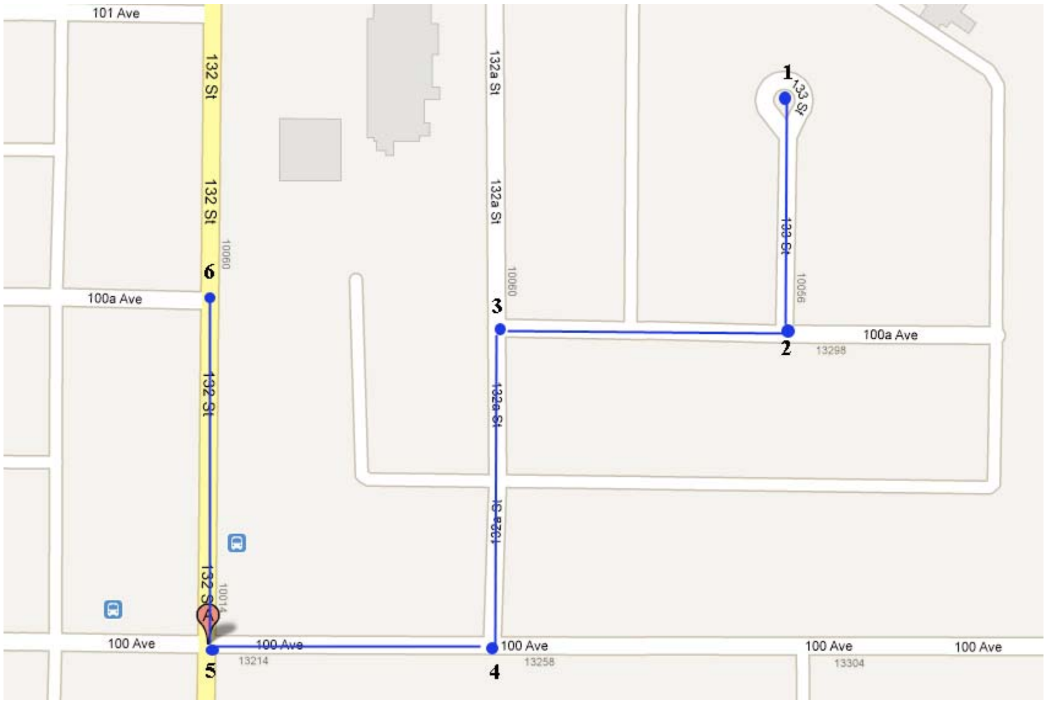

Figure 3: **Map of the first route captured. Numbers indicate intersection at which turns happened. Intersections 1, 3 and 6 were set as the first, second and third pointing locations respectively.**

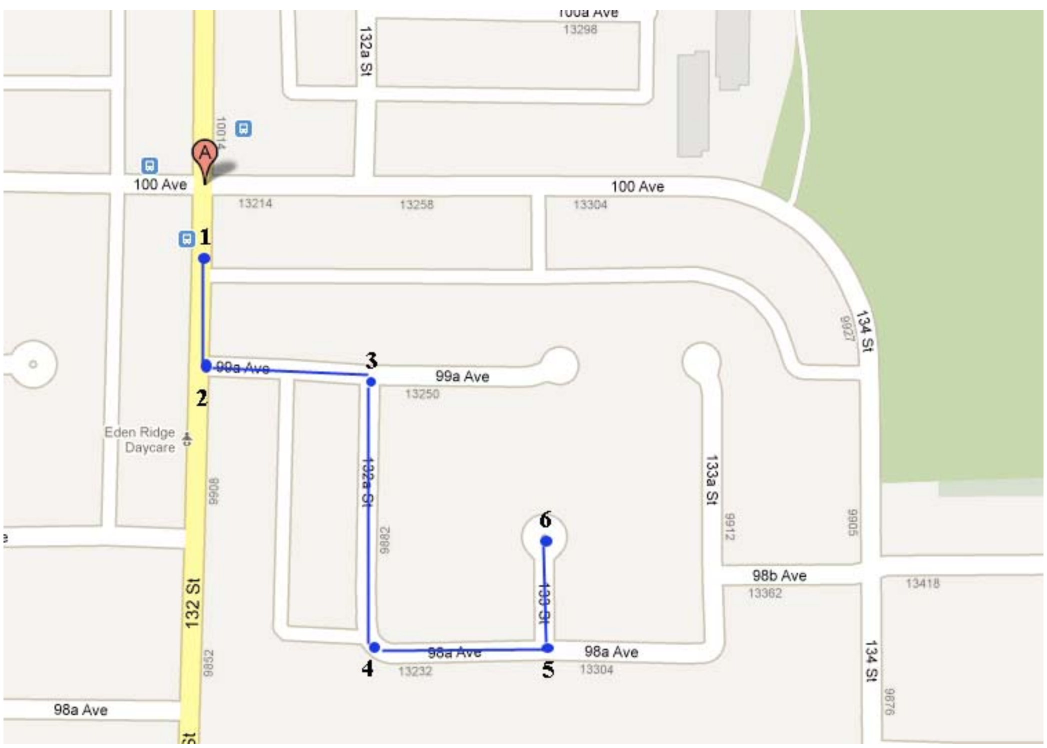

Figure 4: **Map of the second route captured. Numbers indicate intersection at which turns happened. Intersections 1, 3 and 6 were set as the first, second and third pointing locations respectively.**

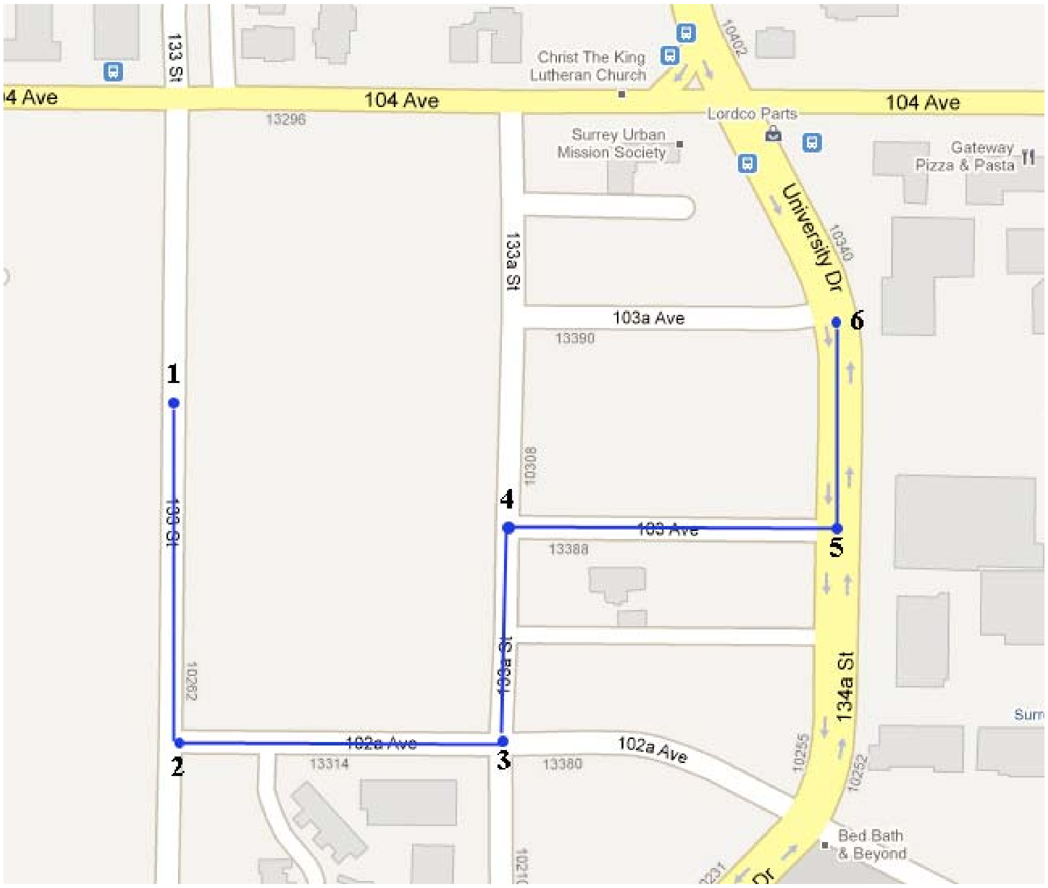

Figure 5: **Map of the third route captured. Numbers indicate intersections at which turns happened. Intersections 1, 3 and 6 were set as the first, second and third pointing locations respectively.**

4.6.2.2 Video and Image Materials

Using the camera system described in Section 2.5, we captured videos of about five minutes length from each of the environments by pushing the cart containing cameras down the selected routes at a speed of about 2-3 m/s for the straight paths and 1-1.5 m/s at the turns. For the panoramic slide show, video frames were sampled at 3-second intervals producing 1 panoramic image at about every 6-9 meters distance during the straight paths and about every 3-4.5 metres during the turns. This was to allow participants, in the panoramic slideshow condition, to know there is a turn as there is no optic flow indicating it. Assuming that a whole turn happened with a maximum of 90 degrees change in the direction and six meters change in the traversed distance, there were about three images available for a whole turn. Consequently, these images had about 30 degrees difference in their viewing angle and this provided for an almost smooth turn.

4.6.2.3 Experimental Settings

Using black drapes on four sides of a 2.3 square metre area in the laboratory (to block external light or distractions) an immersive dark cubical space was built as the platform for the experiments (See Figure 6). In this small cubical space participants sat on the interactive chair and we put a laptop and a laptop holder (to increase the height of the laptop and keep it fixed at the place) on their laps. They watched the videos on the laptop screen and changed their view angle in the panoramic videos or the slide shows by rotating their chair. They also received audio guides through a headphone during their navigations. The audio guide informed them about the upcoming intersections along their way and provided references to the places about which they would be asked later.

Figure 6: The immersive platform for running the experiments in the SIAT Virtual Reality lab. It contains the interactive chair in a dark space disconnected from the outside distractions by black drapes.

4.6.2 Experimental Design

Because of proven between-participant variability in spatial learning abilities (Hegarty et al., 2005), we decided on a within-participant experimental design so that all the participants were exposed to all three experimental conditions. For handling the possibility of one condition affecting or carrying over to another, we used a completely counterbalanced design approach. Therefore, each of the six possible orders of the three conditions was tested with three participants.

For all conditions navigation was partially passive and users could not control the speed of navigation, choose their path, or make stops. Also, it is only possible to go forwards, not backwards. For panoramic video and panoramic slide-show conditions, they could change their view and look around in the panoramic view. We explained this situation to our participants by using the wheelchair passenger metaphor: they imagined sitting in a wheelchair being pushed down the streets at a fixed speed. During this wheelchair ride they could look around in panoramic conditions. In this way we kept the amount of time participants spent in each condition equal so as not to let the time confound the effects of our independent variables on the results.

To assess the effect of learning on the participants' performance in the test, each participant did the navigation-and-test two times for every condition. Consequently, each participant performed 6 trials (three conditions x two repetitions for each condition).

Street intersections in the environments were chosen as the test locations. After every experimental trial, participants were tested for pointing from three of the N test locations (beginning and ending point of the route plus an intersection in the middle) to all the N-1 other test position in a randomized order.

4.6.2.1 Tasks

After each tour participants carried out a recollection task and a pointing task. For the recollection task, participants looked at a series of panoramic images of the places in the tour mixed with some other images. These images are images they had previously seen during the navigation. All the images were presented in a panoramic format even in the regular video condition. These images were selected from the environment so that they had at least one object in them that was contrasting the background. But these object were not necessarily placed in the front view of the image. For each image, if participants could recall the place in the image as belonging to the tour environment they just visited, they pressed the “Yes” key and if not they pressed the “NO” key on the keyboard. In cases where they had no idea if they had seen the place or not they pressed any other key on the keyboard. The space key and the enter key were marked as “YES” and “NO” respectively.

For the pointing task, participants were shown panoramic images of the first, third and last intersections they had passed by in the tour. For each intersection the experimental program asked them if they were ready for pointing and they pressed the “Ready” key after they explored the intersection area for a sufficient amount of time. Then the program asked them to point to all the other intersections in a random order. For each pointing, participants saw the panoramic images of the pointing locations in the standard view (same as the view in regular video) and a thumbnail image of the target location in the middle of the screen. To point, they rotated their chair until a red “+” sign at the centre of the panoramic image was at the direction they wished to point to, and they pressed the space key. Figure 7 shows the graphical interface of the experimental system.

Figure 7: **Pointing test interface. The background image is a panorama of the pointing location and the thumbnail in the middle of the screen show the target location to point to.**

4.6.3 Experimental Procedure

1. Introduction and consent form: participants read a description of the experiment and signed a consent form.

2. Demo and Learning: participants received verbal and written instructions on how to use the experimental system and do the tasks, before and while they navigated in a panoramic video training environment. They also practiced pointing tasks by seeing a demo of

how to point and pointing to a couple of targets in the learning environment. Then participants were left on their own to play with the system and the chair until they felt comfortable with them.

3. Test: participants watched the videos or slide shows and after each virtual tour trial the computer brought up the tasks and participants performed the recollection and the pointing tasks consecutively, as described in Section 4.6.2.

4. Presence questionnaire: after participant finished the second trial of each of the virtual tours, they answered a short questionnaire about their subjective sense of presence. Questions of this questionnaire are selected from the Igroup presence questionnaire (Schubert & Friedmann, 1999).

5. Post-experimental questionnaires: at the end of the experimental session participants answered a questionnaire about their general spatial ability, immersion ability, and how they solved tasks. They also provided comments about how difficult the tasks were or other details of the experiment. Questions about participants' spatial ability were selected from the Santa Barbara sense of direction questionnaire (Hegarty et al., 2002), and questions about participants' immersion ability were selected from igroup presence questionnaire (Schubert & Friedmann, 1999). Questionnaires can be found in the Appendices C and D.

4.6.4 Statistical Analysis Design

Our experimental design is a ‘split-split plot design with whole plot in a multiply blocked crossover design’. Our independent variables have categorical values (locomotion, trial, pointing location) and our dependent measures have continuous values (time and angle). Besides our designed independent variables, there were other factors such as the order of exposure, participants’ differences, and etc. that could affect the values of our dependent variables although we were not directly interested in studying their effects. Therefore, we used a mixed-model analysis of variance and enter those factors as random effects. The complete list of our fixed and random effects is represented in Table 2. The design of our mixed-model analysis is done based on our consultation with SFU statistical consultant, Dr. Tom Loughin (Following his advice we used the REML (REstricted or REsidual Maximum Likelihood) method using JMP statistics software program).

Table 2: List of fixed and random effects in our mixed-model analysis of variance

Source	Type
Route	Random
Order	Random
Participant	Random
Locomotion technique	Fixed
Order * Participant	Random
Trial	Fixed
Locomotion Technique * Trial	Fixed
Locomotion Technique * Trial * Order * Participant	Random
Pointing Location	Fixed
Pointing Location * Locomotion Technique	Fixed
Pointing Location * Trial	Fixed

5: RESULTS: ANALYSIS OF THE BEHAVIOURAL DATA

5.1 Overview

Behavioural data is the data we collected from participants performing pointing and recollecting tasks. For every single pointing, a participant's raw pointing error is calculated by subtracting the real direction of target from the pointed direction of target relative to the pointing location. At every pointing location, a participant's absolute ego-orientation error is obtained by calculating the circular mean of pointing errors at that location and configuration error is obtained by calculating the circular standard deviation of these pointing errors. At every pointing location, a participant's absolute pointing error is obtained by calculating the regular mean of absolute values of pointing errors after they are modified to reside between -180 to 180 degrees. Similarly a participant's pointing time at a certain pointing location refers to the average of response times in the pointing task at that location. So, for every participant there is a value stored for the absolute ego-orientation error, configuration error, absolute pointing error, and pointing time, per pointing location, per trial, and per locomotion technique. However, for every participant, recollection errors and recollection times are stored per trial per locomotion technique.

Our first hypothesis stated that the performance in the recollection task is better after navigation in panoramic conditions (i.e. panoramic video and panoramic slide show). This means that the recollection error is smaller on average in these conditions compared to the regular video condition or that the average recollection time is smaller while the recollection error is not bigger.

Our second hypothesis predicted that performance in the pointing task is better after navigation in the panoramic video tour than the slide show or regular video tours. This means that we expected that the average value of at least one of the ego-orientation error, configuration error or absolute pointing error to be smaller in the panoramic video condition compared to the other conditions, or the average pointing time to be smaller in this condition while none of the errors are larger.

The third hypothesis was that in all of the locomotion conditions both the recollection and pointing performance will be better in the second trial of navigation than the first one.

The averages of the pointing errors and pointing time are illustrated and compared in Figure 8 for each trial and locomotion technique. The averages of recollection error and recollection time for every locomotion technique are illustrated in Figure 9 in separate charts for each trial. Figure 10 and Figure 11 illustrate averages of pointing errors and pointing times at each pointing location. While Figure 10 is more efficient for comparing pointing locations in each locomotion condition, Figure 11 helps in comparing locomotion techniques at each pointing location. Based on these diagrams, the averages of behavioural dependent variables do not depict any clear pattern of difference in different locomotion conditions; however, they consistently decrease in the second trial.

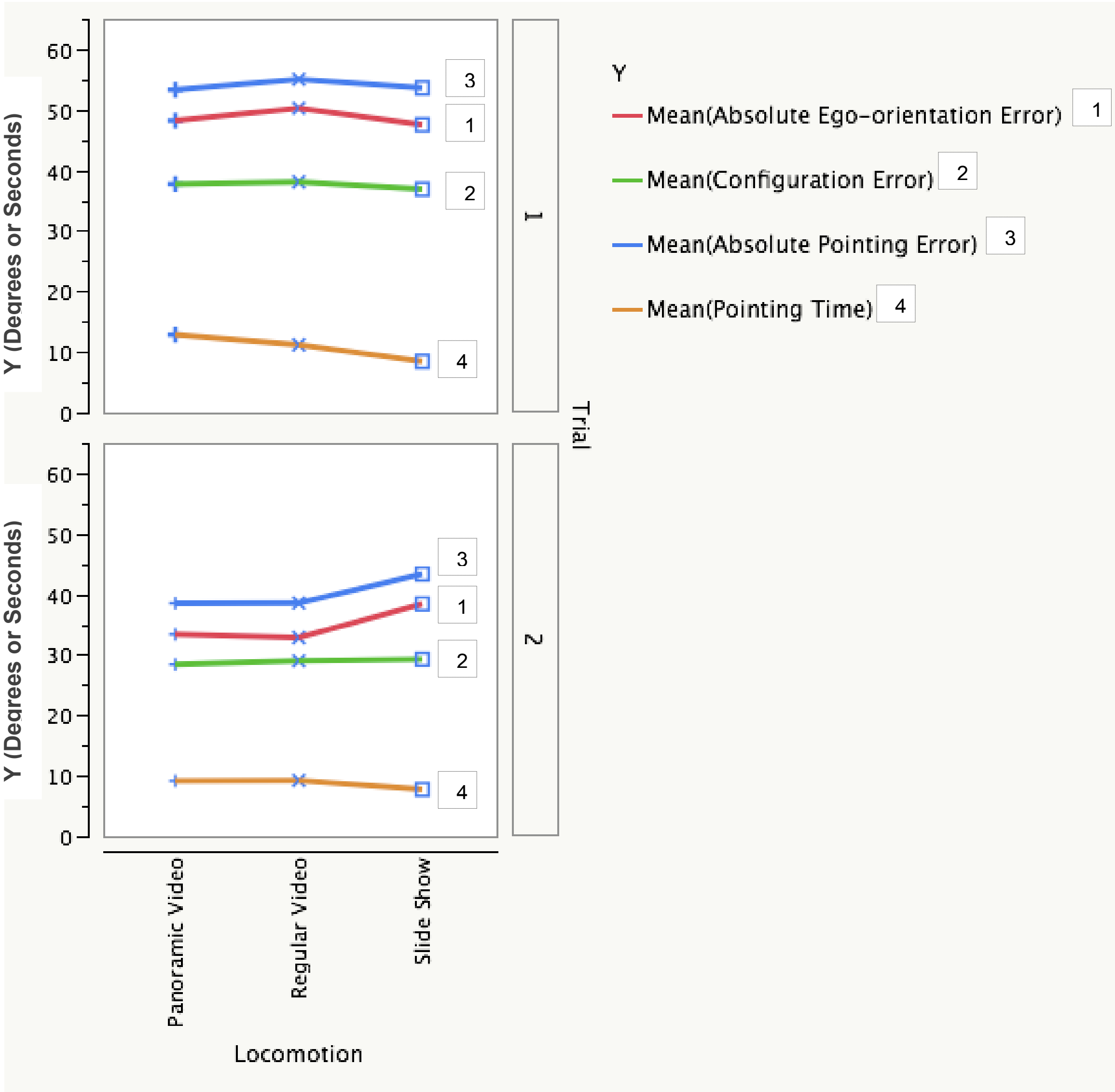

Figure 8: Comparisons of locomotion techniques regarding the mean pointing errors and pointing times in every trial.

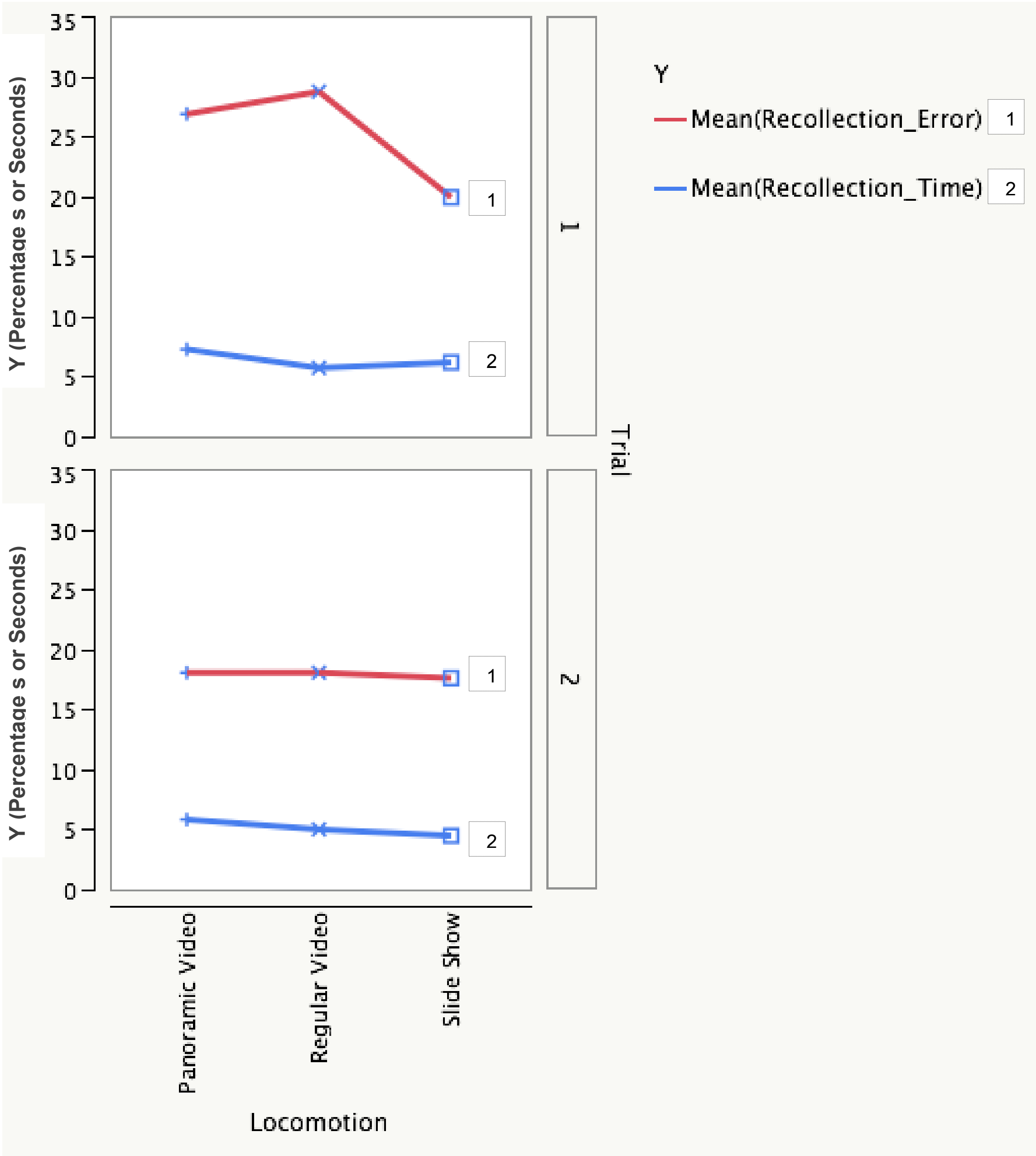

Figure 9: Comparisons of locomotion techniques regarding the mean recollection error and recollection time in every trial.

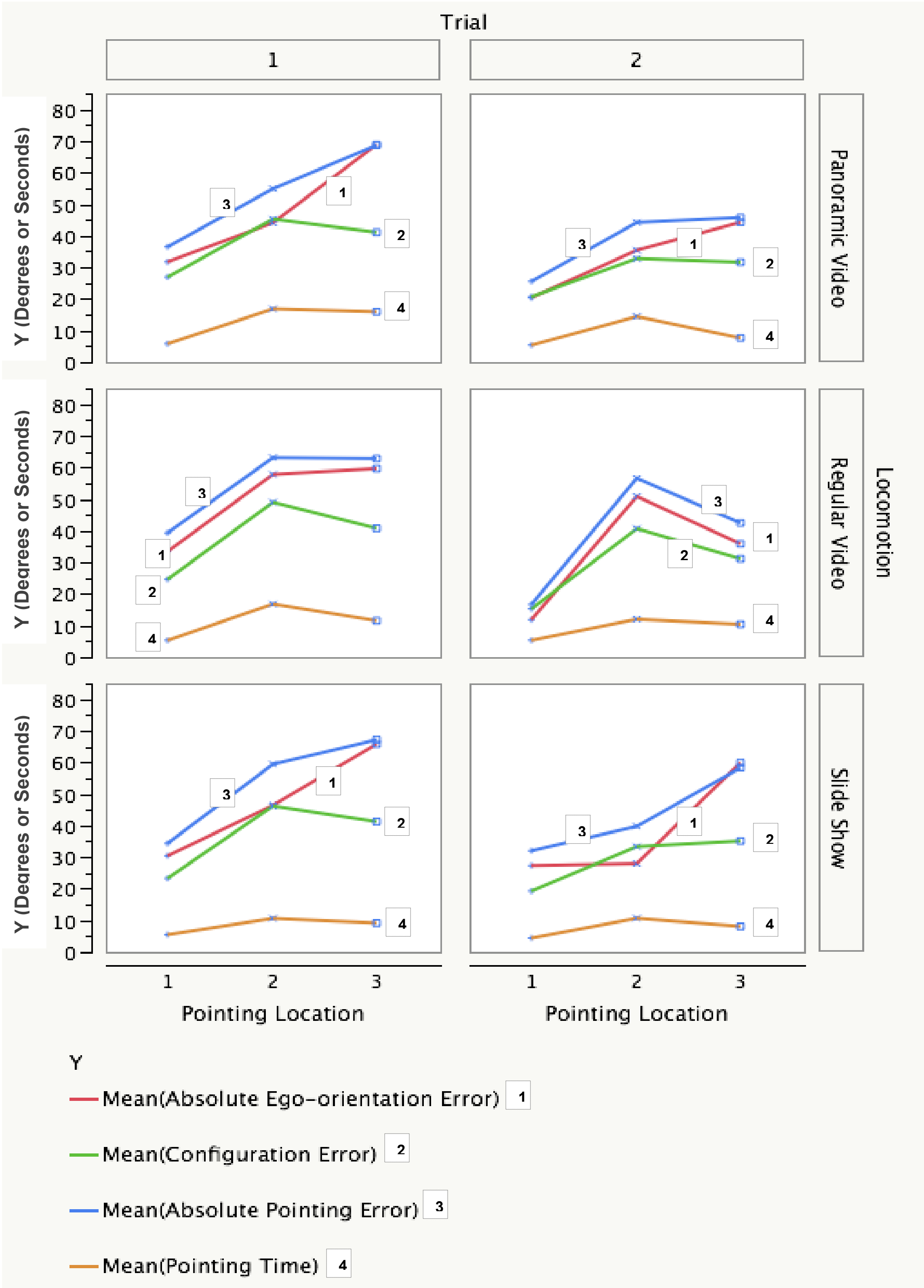

Figure 10: Comparisons of pointing locations regarding the mean pointing errors and pointing time in every trial of every locomotion technique.

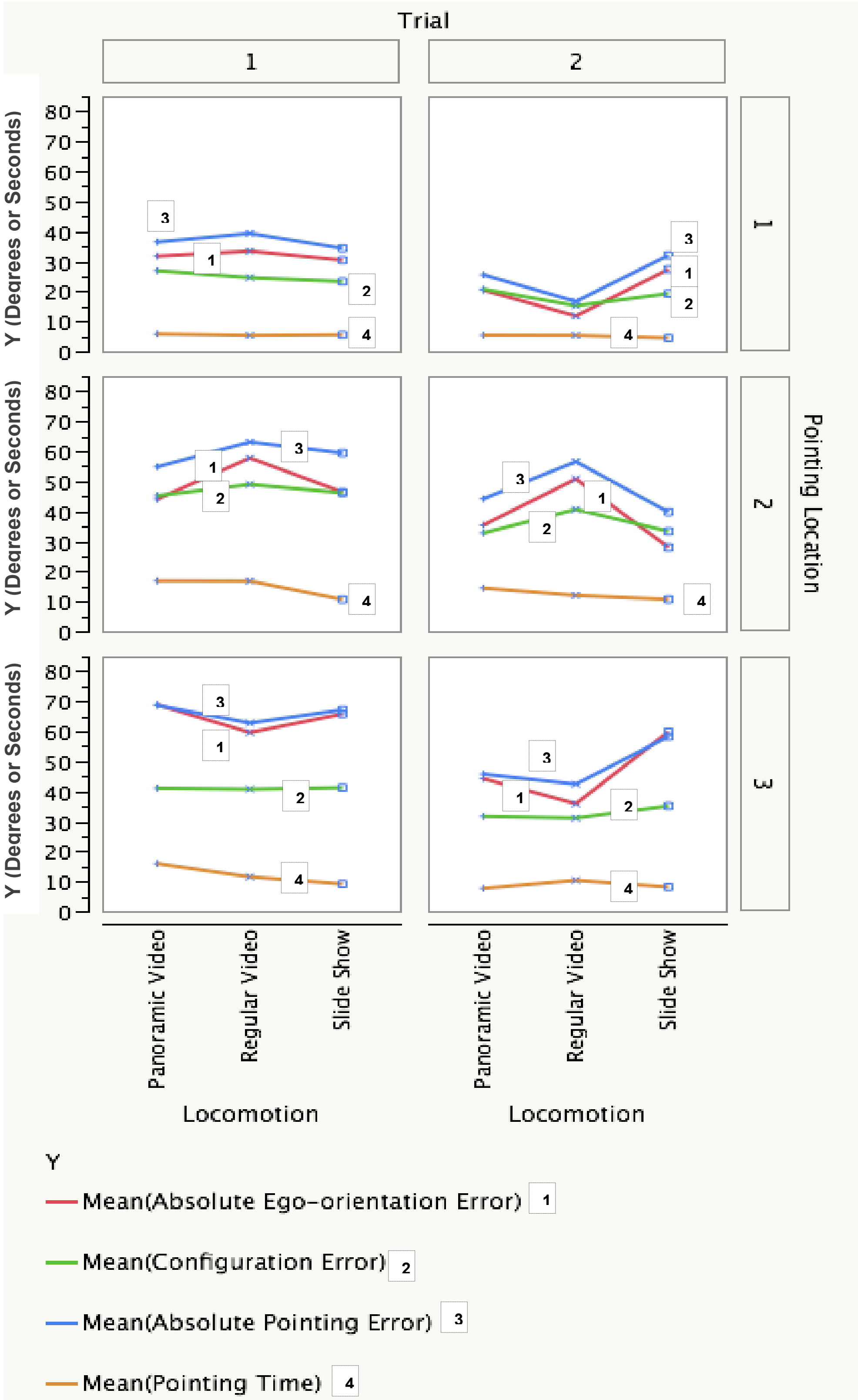

Figure 11: Comparisons of locomotion techniques regarding the mean pointing errors and pointing time at every pointing location in every trial.

5.2 Inferential Analysis

The analysis is done using a mixed-model analysis of variance having three levels of locomotion (panoramic video, regular video, slide show), two levels of trial (one, two), three levels of pointing location (first, second, third) as the fixed effects. The order of exposure to the different locomotion techniques, participants, and the routes were entered into the model as random effects. We used an alpha level of .05 for all statistical tests. The outcomes of this analysis are set out in the following sections for every independent variable, and summarized in the tables provided in Appendix D.

5.2.1 Absolute Ego-orientation Error

As described in Table 4, in the first trial, the mean absolute ego-orientation errors are 48.2, 50.3, and 47.5 for panoramic video, regular video and slide show respectively. These means decrease to 33.3 for panoramic video, 32.8 for regular video, and 38.4 for slide show in the second trial.

Data medians presented in Table 4 demonstrate that in the first trial of all the locomotion techniques, half of the ego-orientation errors have an absolute value smaller than 33.5 degrees. This number decreases to 20 degrees in the second trial. In both of the trials, means and medians there is a large difference between the means and medians. This indicates the existence of huge outliers in the data.

Results of the mixed-model ANOVA on the absolute ego-orientation errors demonstrated the following results:

The main effect of the trial yielded an *F* ratio of $F(1,51)=11.77$, $p=.001$, indicating that the mean absolute ego-orientation error was significantly smaller in the second trial ($M=40.75$, $SD=18.96$) than the first trial ($M=47.97$, $SD=19.63$), $p<.05$.

There was a significant main effect of pointing location, $F(2,204)=14.86$, $p<.001$. Post-hoc analysis using Tukey's HSD criterion (described in Table 11) indicated that the mean absolute ego-orientation error was significantly smaller at the first pointing location ($M=25.83$, $SD=32.78$) than the second pointing location ($M=43.72$, $SD=39.52$), $p<.05$, and it was smaller at the first pointing location than the third pointing location ($M=55.68$, $SD=54.90$), $p<.05$. There was no difference between the mean absolute ego-orientation errors at the second pointing location and the third pointing location, $p>.05$.

However, the main effect of locomotion technique was not significant, $F(2,31.52)=.088$, $p>.05$. A summary of the corresponding mixed-model analysis is provided in Table 10.

5.2.2 Configuration Error

As described in Table 5 in the first trial, the mean configuration errors are 37.7, 38.1, and 36.9 for panoramic video, regular video and slide show respectively. These means decrease to 28.4 for panoramic video, 29 for regular video, and 29.2 for slide show in the second trial.

Results of the mixed-model ANOVA on the configuration errors demonstrated the followings:

The main effect of trial yielded an *F* ratio of $F(1,51)=42.05$, $p<.001$, indicating that the mean absolute configuration error was significantly smaller in the second trial ($M=28.85$, $SD=18.96$) than the first trial ($M=37.56$, $SD=19.63$), $p<.05$.

There was a significant main effect of pointing location, $F(2,204)=55.37$, $p<.001$. Post-hoc analysis using Tukey's HSD criterion (described in Table 13) indicated that the mean configuration error was significantly smaller at the first pointing location ($M=21.64$, $SD=13.5$) than the second pointing location ($M=41.17$, $SD=20.24$), $p<.05$, and it was significantly smaller at the first pointing location than the third pointing location ($M= 36.79$, $SD= 19.34$), $p<.05$. There was no difference between the mean configuration errors at the second and third pointing locations, $p>.05$.

However, the main effect of locomotion technique was not significant, $F(2,31.27)=.01$, $p>.05$. A summary of the corresponding mixed-model analysis results is provided in Table 12.

5.2.3 Absolute Pointing Error

As described in Table 6 in the first trial, the mean absolute pointing errors are 53.3, 55, and 53.6 for panoramic video, regular video and slide show conditions respectively. These means decrease to 38.5 for panoramic video and regular video, and 43.3 for slide show in the second trial.

Data medians presented in Table 6, demonstrate that in the first trial half of the absolute pointing errors have a value smaller than 51 degrees in regular video and slide show conditions while half of the absolute pointing errors have a value smaller than about 34 degrees in the panoramic video condition. Considering the large difference (i.e., 9 to 17 degrees) between means and median in the second trial, it appears that there are huge outliers, which have affected the means of absolute pointing errors in the second trial. In this trial panoramic video has the smallest median (median=21.5) that indicates half of the absolute pointing errors are smaller than about 22 degrees in panoramic video while they are smaller than about 28 and 35 degrees in the regular video and slide show respectively. Results of the mixed-model ANOVA on the absolute pointing errors demonstrated the followings:

The main effect of the trial yielded an *F* ratio of $F(1,51)=24.30$, $p<.001$, indicating that the mean absolute pointing error was significantly smaller in the second trial ($M=53.98$, $SD=37.05$) than the first trial ($M=40.09$, $SD=33.92$), $p<.05$.

There was a significant main effect of pointing location, $F(2,204)=27.40$, $p<.001$.Post-hoc analysis using Tukey's HSD criterion (described in Table 15) indicated that the mean absolute pointing error was significantly smaller at the first pointing location ($M=30.7$, $SD=30.48$) than the second pointing location($M=52.95$, $SD=31.21$), $p>.05$, and it was smaller at the first pointing location than the third pointing location ($M=57.47$, $SD=40.32$), $p<.05$. There was no difference between the man absolute pointing error at the second and third pointing locations, $p>.05$.

The main effect of locomotion technique was not significant, $F(2,28.39)=.16$, $p>.05$. Summary of the corresponding mixed-model analysis results is provided in Table 14.

5.2.4 Pointing Time

As described in Table 7 in the first trial, the mean pointing times are 12.8, 11.1, and 8.4 seconds for panoramic video, regular video and slide show conditions respectively. These means decrease to 9.1 seconds for panoramic video, 9.2 seconds for regular video and 7.7 seconds for slide show in the second trial. Results of the mixed-model ANOVA on the absolute pointing times demonstrated the followings:

The main effect of the trial yielded an *F* ratio of $F(1,51)=5.92$, $p=.018$. indicating that the mean pointing time was significantly smaller in the second trial ($M=8.67$, $SD=7.35$) than the first trial ($M=10.8$, $SD=10.8$), $p<.05$.

There was a significant main effect of pointing location, $F(2,204)=32.25$, $p<.001$. Post-hoc analysis using Tukey's HSD criterion (described in Table 17) indicated that the mean pointing time was significantly smaller at the first pointing location ($M=5.26$, $SD=3.52$) than the second pointing location ($M=13.5$, $SD=11.13$), $p<.05$, and it was significantly smaller at the first pointing location than the third pointing location ($M=10.43$, $SD=9.4$), $p<.05$. There was no difference between the mean pointing time at the second and third pointing locations, $p>.05$.

The main effect of locomotion technique was non-significant, $F(2,31.5)=2.33$, $p>.05$. A summary of the corresponding mixed-model analysis results is provided in Table 16.

5.2.5 Recollection Error

As described in Table 8 in the first trial, the mean recollection errors are 26.8 %, 28.7%, and 19.9% for panoramic video, regular video and slide show conditions respectively. These means decrease to 18% for panoramic video, 18.1% for regular video and 17.6% for slide show in the second trial.

Results of the mixed-model ANOVA on the recollection errors showed that he main effect of trial was significant, $F(1,67.56)=8.61$, $p=.005$, indicating that the mean recollection error was significantly smaller in the second trial ($M=17.9$, $SD=13.3$) than the first trial ($M=25.15$, $SD=16.8$), $p<.05$. However, the main effect of the locomotion was not significant $F(2,57.38)=1.19$, $p>.05$. A summary of the corresponding mixed-model analysis results is provided in Table 18.

5.2.6 Recollection Time

As described in Table 9 in the first trial, the mean recollection times are 7.3, 5.7, and 6.2 seconds for panoramic video, regular video and slide show conditions respectively. These means decrease to 5.8 seconds for panoramic video, 5 seconds for regular video and 4.5 seconds for slide show in the second trial.

Results of the mixed-model ANOVA showed that the main effect of trial was significant, $F(1,64.15)=7.73$, $p=.007$ indicating that the mean recollection time was significantly smaller in the second trial (M=5.1, SD=3.26) than the first trial (M=6.37, SD=2.87), $p<.05$. However, the main effect of locomotion was non-significant $F(1, 69.37)=1.12$, $p>.05$. A summary of the corresponding mixed-model analysis results is provided in Table 19.

5.3 Gender Differences

Data for the pointing errors of males and females are compared in Figure 12. As illustrated, in the first trial males have a considerably lower average error (absolute ego-orientation error, configuration error, and absolute pointing error) in the regular video and slide show conditions. However, in the panoramic video condition males and females have similar average errors as males perform slightly worse and females perform slightly better than in the other two conditions.

In the second trial differences of average errors between males and females decrease in regular video and slide show conditions as females improve and males perform almost the same. But in panoramic video, males improve and females perform no better, therefore, their differences increase noticeably.

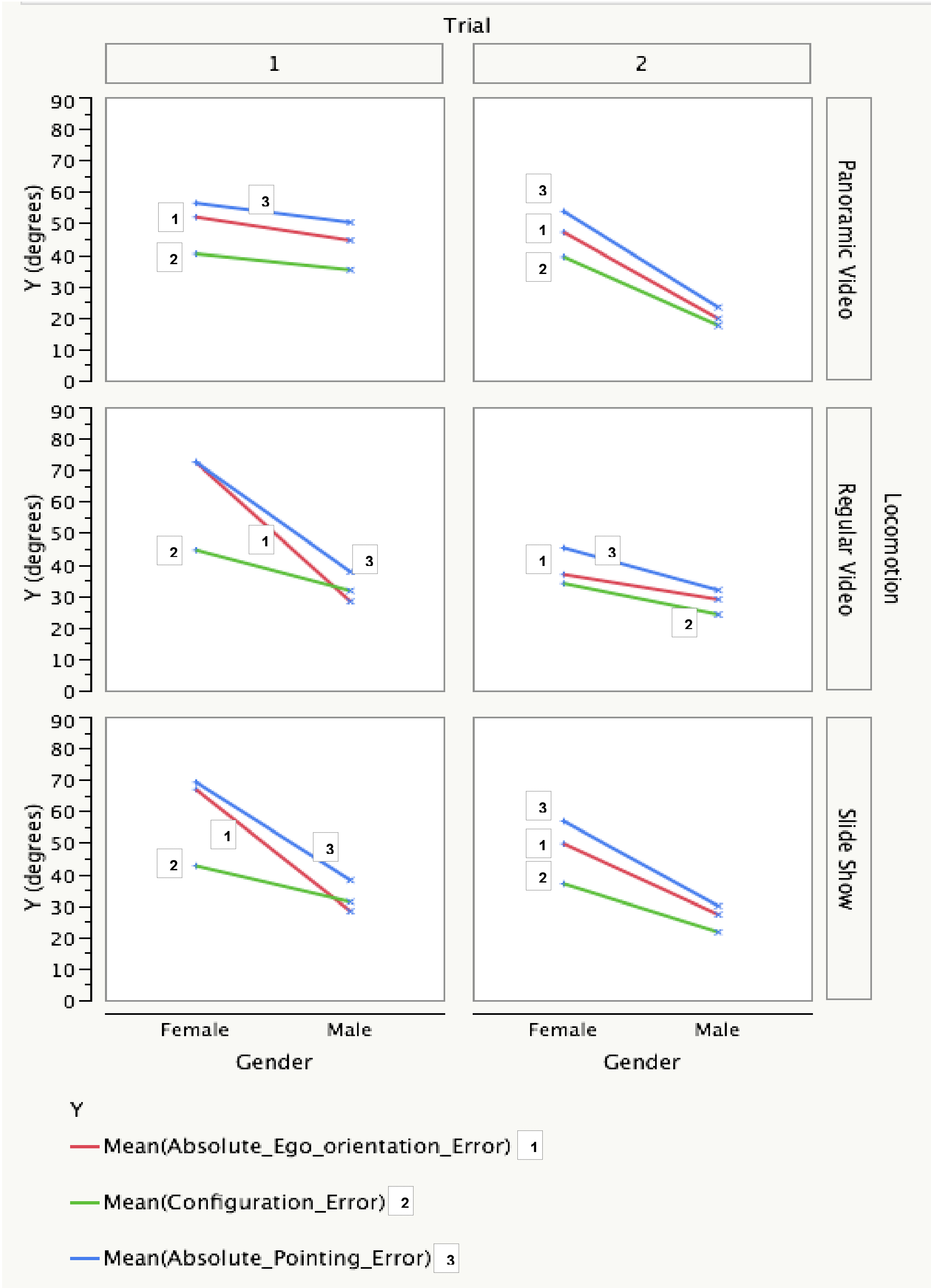

Figure 12: Comparisons of males and females performance regarding pointing errors for every locomotion technique and trial.

5.4 Summary

As there was no significant difference between the averages of any of the behavioural dependent variables in three different conditions, locomotion techniques did not have any significant effect on the participants' performance in both recollection and pointing tasks. Consequently panoramic video did not significantly improve participants' directional knowledge acquisition compared to the slide show or regular video. Therefore, we cannot accept our first and second hypotheses. However, the average value of all the behavioural dependent variables (i.e. errors and times) significantly decreased in the second trial. Consequently, we accept our third hypothesis.

Besides locomotion technique and trial, we tested the effect of pointing location on participants' pointing performance by including it in the mixed-model analysis. Participants performed pointing tasks more accurately and faster at the first pointing location as indicated by the significantly smaller mean absolute ego-orientation error, configuration error, absolute pointing error and pointing time at the first pointing location than the second and third point locations.

6: RESULTS: ANALYSIS OF THE INTEROSPECTIVE DATA

6.1 Overview

Participants' introspective assessments of the locomotion techniques and the experiment in general are collected through questionnaires (provided in Appendix A and B). These data are collected in three categories: participants' quantitative assessments of their subjective sense of presence in every locomotion condition, participants' quantitative assessments of the difficulty of pointing tasks performed in every locomotion condition, and participants' qualitative assessments of the experimental settings, locomotion techniques, and tasks. In this chapter we analyse both the quantitative and qualitative introspective assessments.

6.2 Subjective Sense of Presence

Our fourth hypotheses predicted that participants' reach the highest sense of presence in the panoramic video condition compared to the other conditions. Reported presence scores were subjected to a mixed-model analysis of variance having three levels of locomotion technique (panoramic video, regular video, slide show). The order of exposure to the different locomotion conditions, participants, and routes were entered into the model as the random effects.

The main effect of locomotion technique was significant, $F(2,28.61)=11.98$, $p<.001$. Post-hoc analysis using Tukey's HSD criterion indicated that the mean sense of presence was higher in the panoramic video condition (*M*=5.4, *SD*=0.5) than regular video condition (*M*=3.7, *SD*=0.4), *p*<.05, and it was higher in the panoramic video condition than the slide show condition (*M*=3.2, *SD*=0.3), *p*<.05. There was no significant difference between the mean sense of presence in the regular video condition and the slide show condition, *p*>.05 (see Appendix G). The average senses of presence in different locomotion conditions are compared graphically in Figure 13.

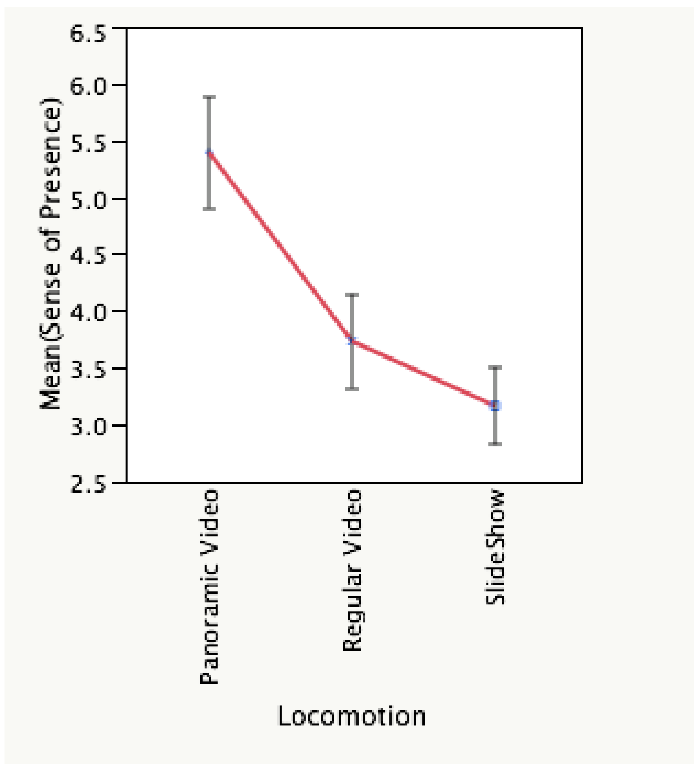

Figure 13: Comparisons of locomotion techniques regarding mean participants' subjective sense of presence

6.3 Difficulty of the Pointing Task

Our fifth hypotheses predicted that performing directional tasks is easier for the participants in the panoramic video condition than the other conditions.

In the post-questionnaire participants reported on how difficult the tasks were in different locomotion conditions. They rated the difficulty of tasks in each locomotion condition by a number from 0 to 10. We conducted a mixed model analysis of variance on these reported difficulty scores, having three levels of locomotion (panoramic video, regular video, slide show) as the fixed effects. The participants, paths, and the order of exposure to the different locomotion conditions were entered into the model as the random effects. The results demonstrated that there is a significant main effect of locomotion $F(2,49.63)=7.32$, $p=.002$. Post-hoc analysis using Tukey's HSD criterion indicated that the mean difficulty of the pointing task is smaller in panoramic video condition ($M=5.6$, $SD=.6$) than the slide show condition, $p<.05$, and it is smaller in the regular video condition ($M=5.7$, $SD=.6$) than the slide show condition ($M=7.2$, $SD=.5$), $p<.05$. There was no difference between the mean difficulty of pointing task in the panoramic video condition and regular video condition, $p<.05$ (see Appendix G). Comparison of average difficulty of pointing task for each locomotion technique is illustrated in Figure 14.

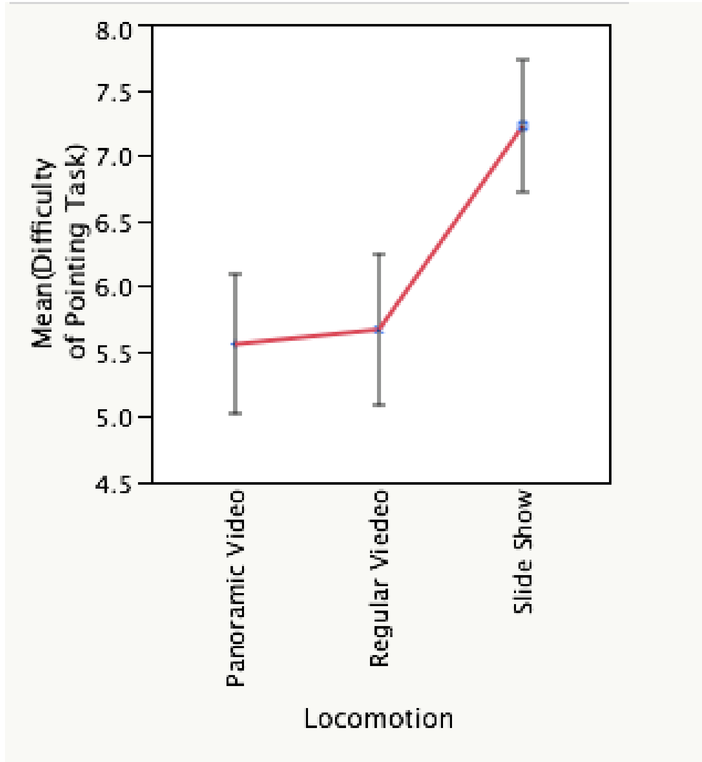

Figure 14: Comparisons of locomotion techniques regarding the average difficulty of tasks reported by participants

6.4 Participants' Qualitative Assessments

Participants' answers to the post-experiment questions are classified content-wise and presented as follows.

6.4.1 How Participants Solved the Task

Based on their reports, it appears that the participants' basic strategy for learning the routes and solving the pointing tasks was to memorize the turning direction (left or right) at each turning intersection. Then, mostly in the second round of navigation, they tried to draw a mental map of the route based on this information. Many of the participants drew the map on the side of the computer surface using their fingers. They reported having their thumb as a reference to the beginning of the route and drawing turns using their other fingers. A few of the participants, though, only used their mind to imagine the route maps.

There was one participant who found it easier to use her physical position against her virtual position as a reference of direction especially in panoramic video condition. So, she rotated her chair to the turning direction at every intersection and finally understood the map of the route. Another one reported using his right and left hands for keeping track of turns.

One strategy participants used in the panoramic conditions in order not to lose track of their moving direction was to let the video or slide show run through and not to change their view a lot.

6.4.2 What Participants thought of the Chair-based Interface

Many participants thought that the chair-based interface was tiring because it did not let them rotate very fast and they had to step over foot by foot. Some participants also found it uncomfortable, unnatural and distractive, so that it kept them away from feeling present in the virtual environment. However, some of participants used it as a physical reference to keep track of their straight ahead view and the relative turning directions. A few people also found it interesting and playful.

6.4.3 What Was Difficult About the Task and Why

Participants related the difficulty of the tasks to the following problems:

6.4.3.1 Unrealistic Images

Low quality videos and pictures plus misalignment of adjacent frames of panoramas produced a difficult condition for subjects and many of them found the settings unrealistic and non-immersive. Poor graphic quality also reduced the contrast of the scenes and landmarks so that it was hard to recognize places and learn the routes.

6.4.3.2 Difficulty of Keeping Track of the Routes

Because the navigation was passive with no possibility of stopping and going back and forth, participants had difficulty keeping track of the route. Especially in the panoramic conditions when they tried to look around they lost sense of direction. Participants generally commented that it was hard to keep track of the route, memorize the turns, and realize the distances between the consecutive turning points. Consequently it was difficult to imagine the map of the route and the relative positions of places.

6.4.3.3 Difficulty of the Pointing Task

Participants reported that the pointing task was very difficult when pointing to intersections 1 or 2 from the 6^{th} intersection (end point of the route), but it was pretty easy to point forward (e.g. from first intersection to second or third).

6.4.3.4 Slide Show Problems

In the slide show condition it was hard to understand the direction of turns as well as to keep track of straight view and approximate the length of each route. In addition the slide show was annoying because it was not smooth and loading images was sometimes slow. Participants complained about missing the straight direction.

6.4.3.5 Panoramic Video Problems

The only problem with panoramic video was that when participants tried to look around, due to the lack of control over the forward movement they lost track, of the direction they were moving in. In a few cases participants became dizzy from rotating to see what is happening around them.

6.4.4 What to Change

Participants suggested the following changes, listed in order of popularity:

- Enhancing the quality of the videos and images.
- Slowing down the speed of the panoramic slideshow.
- Making the slides closer to each other (increasing the number of images).
- Enhance the quality of stitching
- Providing the ability for users to pause the videos and go forward and backward.
- Providing the ability for users to turn and change view as fast as they turn their head.
- Putting some limitations on the horizontal field of view so that as participants can see their surrounding they cannot turn all the way around and get lost.

6.5 Summary

Participants reached the highest sense of presence on average during the panoramic video tour confirming our fourth hypothesis. However, participants complained about how distractive the panoramic view and interactive chair are in the panoramic video condition.

Pointing tasks were significantly easier after navigation in the panoramic video and regular video tours than the slide show tour. This partially confirms our fifth hypothesis; however, pointing tasks were not any easier in the panoramic video than the regular video. Participants reconfirmed having the most difficulty in keeping track of the straight-ahead view, turning directions, and the traversed distances during the slide show tour.

Participants noted several problems about the experiment in general, the most highlighted of which are the low quality of images, inability to control the movement, difficulty of memorizing all the turns, and the difficulty of the pointing task itself.

7: DISCUSSION

7.1 Overview

In this chapter we discuss the outcomes of this study in the context of its limitations and the previous works reported in the literature. The outcomes of this study are put into two categories: behavioural and introspective. While behavioural results demonstrate participants' task performance, introspective results exhibit their quality of experience in the three different locomotion conditions.

The behavioural results indicated an average satisfactory performance in the recollection task although no significant difference was observed between locomotion techniques rejecting our first hypothesis. Accordingly, all the locomotion conditions were equally capable of supporting the recall of landmarks. On the other hand, the behavioural results also indicated an average poor performance in the pointing task. No significant difference was observed among the locomotion techniques regarding any of the absolute ego-orientation error, configuration error, absolute pointing error or pointing time. Therefore, we could not accept our second hypothesis. However, in both recollection and pointing tasks, participants' performance significantly improved after repeating the tour navigation in all the locomotion conditions, confirming our third hypothesis. Men generally performed better than women in all conditions except for the first exposure to panoramic video, where men and women performed relatively equally.

The introspective results indicated that participants reached a higher level of sense of presence in the panoramic video tour than the other two conditions confirming our fourth hypothesis. Participants also rated the pointing task less difficult in the panoramic video and regular video conditions than in the slide show. This partially supported our fifth hypothesis. Participants related the difficulty of the task to the difficulty of adjusting to the panoramic view, the low image quality, and the interactive chair interface.

In the next sections we first outline the limitations of the study and discuss how these limitations possibly affected the results of study. Secondly we explain the techniques for handling the reliability of the data. Then we discuss in detail the outcomes of the study compared to the literature followed by a discussion of possible future work.

7.2 Limitations of the Study

The major outcome of our study was that there was an average poor performance in acquiring directional knowledge as well as no significant difference between different locomotion techniques in providing directional knowledge. It appears that the subjects found the pointing task to be very difficult and this was the major reason for their poor performance. As a result these tasks do not allow us to discriminate between locomotion modalities. The obvious solution is to devise easier tasks, but they must not be too easy since these also will make it difficult to discriminate between locomotion modalities.

We are aware that our conclusions regarding task difficulty are influenced by the study limitations and we discuss these limitations here. Those limitations are related to participants, poor image quality, passive movement, characteristics of the environments, and dependency of task on memory, all of which could have contributed to participants' poor directional performance and made it difficult to see differences between the locomotion modalities. Others, which are related to carry-over effect and laboratory settings, describe the limitation of the applied analysis and the extent to which we can generalize our results. The following paragraphs explain these limitations in detail:

Participants: The participants in this study were mostly graduate students at Simon Fraser University studying applied science and interactive arts. As described in Appendix C, the majority of the participants did not have much previous exposure to different examples of virtual reality, game environments and panoramic images. Therefore, the equally high average errors in all the conditions might relate to the participants' compatibility issues. In addition, the number of participants was limited to 18 people. Although participants' subjective reports revealed a wide range of spatial abilities (ranging from 2.6 to 8 out of 10), the sample size was still small enough to leave the chance of having a type 2 error which can result in observing no significant difference while there is, in fact, a difference.

Poor video quality: Because of budget and technology restrictions, videos generated for the experiments suffered from a low pixel rate and poor stitching quality. The original single frame images had 322 pixels for a 90-degree horizontal field of view, which resulted in 3.6 pixels per degree. These were stitched together to form a 360-degree panorama. Poor graphics and visual realism may have affected participants' recognition of places and their consequent performance in the tasks as well as their engagement in the navigation in all the locomotion conditions. However, the most important advantage of panoramic conditions over the regular video condition is in providing greater visual information. If participants fail in using this available visual information because of its poor quality, the difference between the panoramic and non-panoramic conditions declines.

Environment characteristics: The environments used for the virtual tours were residential neighbourhoods with flat visual appearances, thus lacking significant landmarks or other contrasting urban elements. Routes were long (around 500 metres) and had moderately complex structures (5 turns along the routes). These characteristics could have made the tasks unnecessarily difficult, contributing to participants' confusion and misidentification of locations and consequently increased the chance of having large amount of random data that does not reveal differences between locomotion techniques.

Dependency of the tasks on memory: The performance of participants in both the pointing and recollection tasks was not dependent only on their survey knowledge but also on their memory abilities. Although most survey knowledge experiments depend somewhat on memory, the dependency in our study was exaggerated due to the, long routes, low video quality and passive navigation. These characteristics may have impaired the ability to acquire survey knowledge and sense of presence in the virtual environment, causing participants to rely more heavily on memory.

Passive movement: Because moving in the virtual environments was passive, participants could not go backwards, control the speed of their movements, or stop. If participants had had active control over their navigation, the different locomotion techniques might have resulted in participants developing dissimilar navigation strategies for different locomotion modalities. For example, if a navigation method (such as wayfinding) requires visual search, it seems likely that the panoramic conditions could be more efficient as participants are able to search a larger visual field more easily.

Carry-over effect: Although the order of exposure to different locomotion techniques was counter-balanced, participants experienced each successive environment immediately after they had completed the task for the previous environment. Lack of washout time in between consecutive exposures brings the risk of a carry-over effect in the results, meaning that performance in one condition can be partly dependent on the performance in the other conditions (Breakwell, 2006).

Laboratory settings: In order to control all the possible confounds, the experiments were done in a restricted laboratory environment. Although there are benefits to creating uniform experimental conditions, some fixed settings (such as the height of the chair regardless of the height of the participants) could have confounded participants' performances. In addition, the restricted laboratory setting affects the application of the results to real world situations.

7.3 Reliability

During the design and conduct of the experiments we avoided causes of unreliability by applying the following techniques:

- **Using adequate measures that have been previously used and validated by others:** Pointing and recollection tests are common tests in the literature of navigation for assessing survey and landmark knowledge. Recollection error, absolute ego-orientation error, configuration error, absolute pointing error and response time are measures that have been used in the literature (discussed in Chapter 3). Items on the presence questionnaire were selected from the igroup presence questionnaire (Schubert & Friedmann, 1999) that has been widely used by virtual reality researchers.
- **Consistent measurement using a software program:** The experimental procedure is implemented as a software program so that the pointing and recollection data is recorded consistently across all the participants and conditions.
- **Keeping the testing situation as free from the contaminated influences as possible:** Experiments were done in silent

laboratory environment with participants sitting in an area isolated from the external distractions by black drapes.

- **Using a precise measurement tool:** The pointing test program was capable of measuring pointing errors of about 1 degree. The questionnaires had rating scales that ranged from 0 to 10 to increase sensitivity to a range of subject responses.

7.4 Discussion

7.4.1 Participants Recollection Performance

Participants on average performed recollection tasks equally well in all the locomotion conditions with a low recollection error of about 20%. This means that all the techniques are equally capable of providing enough landmark knowledge. This result is compatible with the work of Gaunet et al. (2001) as it demonstrated that participants' scene recognition ability is not affected by the passive smooth transition (similar to our video condition but with computer-generated graphics) or passive snapshot exploration (similar to our slide show condition but with computer-generated graphics) of the environment.

The similar recollection performance in all the three locomotion conditions can be because participants are more likely to pay attention to the front view than side views so there is no practical difference between the panoramic and non-panoramic conditions. Or, as Mallot and Gillner (Mallot & Gillner, 1999) explained, participants do not recognize places as panoramic views or a configuration of objects for route reconstruction, but they recognize an individual object.

7.4.2 Participants General Performance in the Pointing Task

Participants on average performed pointing tasks very poorly in all the locomotion conditions. The average values of absolute ego-orientation error, configuration error, and absolute pointing error vary but are approximately 45 degrees in all the conditions. If a participant performed all the pointings on a completely random basis, the expected value for his/her absolute pointing errors would be 90 degrees respectively. Therefore, an average error of about 45 degrees does not demonstrate total disorientation. However, it does not show a good sense of orientation either. It should be noted that one or two individuals performed much better than the average.

As noted above, the participants' poor average performance and their post-experiment comments, led us to believe that the task was generally difficult relative to the quality and type of directional knowledge that our virtual tours provided for the participants. The pointing task included pointing to all the other intersections from three locations on a relatively long route involving a sequence of six intersections. Therefore, in addition to landmark and route knowledge, participants were required to obtain partial survey knowledge in order to complete the pointing task well. For example, when participants were pointing from the end point of the route to the beginning points of the route, they had to refer to their mental map of the whole route. This is why they attempted to imagine the whole route or draw it using their fingers on the laptop surface as mentioned in their comments. However, due to several problems, participants had a difficult time obtaining and picturing this map.

One important problem causing poor directional performance could be related to the low graphic quality. Meijer et al. (2009) demonstrated that visual realism can significantly affect participants' knowledge of the spatial layout and routes in an environment. Although our proposed virtual tours utilized photo-realistic techniques, many participants reported a failed attempt to perceive the environments realistically because of the low image resolution. The other problem can be that the routes were long and lacking in distinctive landmarks and so the task was highly dependent on the memory. Anyhow, this poor average directional performance demonstrates that most of the participants did not acquire survey knowledge.

7.4.3 Comparison of Participants' Pointing Performance from Different Pointing Locations

We looked at the pointing data at each pointing location separately because the choice of pointing location had a significant effect on all of the pointing errors. Apparently, participants performed significantly better at the first pointing location, which is at the beginning of the route, than at the other two pointing locations in the middle and at the end of the route. A low-medium ego-orientation error (about 25 degrees) indicates that participants had a more accurate understating of their ego-orientation at the first pointing location. A low-medium configuration error (about 22 degrees), on the other hand, indicates a lower level of variability in pointing errors and a more accurate image of the locations of other intersections relative to each other. Significantly lower mean pointing times at the first pointing location on the other hand, suggest that, on average, participants were more confident about their pointings at this location. Therefore, we can assume that it was easier for the participants to reflect their understanding of the route structure and the relative locations of the intersections from the beginning of the routes. Two of the participants actually noted this fact by commenting that pointing forward was easier than pointing backwards (refer to Section 6.4.3.3). This is not surprising as the following possible reasons suggest.

One possible reason, as Satalich (1995) suggested, is that the knowledge about the relationships of places on the route (also known as route knowledge) is formed by sequential travel and is unidirectional. Consequently, a person will recall these relationships better when it is in the direction in which they learned the route.

Another possible explanation can be that when participants were at the beginning point of the route, they recalled the order of intersections and actions in a forward manner (e.g. turned right at the next intersection and turned left at the intersection after that) in order to picture the structure of the route. However, at the end point of a route, they tried to recall this sequence in backward order. Empirical support has showed that recalling a memory list in backward order is more cognitively demanding than recalling it in forward order; thus, performance in a concurrent task decreases when the memory list is recalled in backward order (Vrij et al., 2008).

As a third explanation, some research illustrated that the spatial mind is orientation-dependent; that is, people recognize spatial relations between objects more efficiently from some perspectives than from others (McNamara et al., 2008).

Finally, specifically in the case of our study, there is a higher chance for participants to confuse their self-orientation at the second pointing location because it is one of the four intersections along the route. However, they were most likely better at recognizing the location and orientation at both the beginning and end point of the route, as there was only one beginning and end point for each route. The slightly smaller – though not significant – pointing errors at the third pointing location (i.e., end point of the route) compared to the second pointing location (i.e., a middle point of the route) can be taken as support for this assumption. This phenomenon may also be explained by the primacy-recency effect (Sousa, 2006) which asserts that, during a learning episode, people best recall items which come first (primacy), second best the items which come last (recency), and third best the items which come in the middle.

7.4.4 Comparison of Participants' Pointing Performance with Different Locomotion Techniques

In contradiction to what we predicted in our second hypothesis, participants on average did not perform pointing tasks any differently in the three different locomotion conditions. However, considering the difficulty of the task (discussed earlier) and the high average absolute ego-orientation error and configuration error at the second and third pointing locations, we suspect that the pointing data is mixed with a considerable amount of random data. If so, this part of the data only represents participants' inability to complete the pointing task in any of the locomotion conditions, but does not provide any reliable comparisons between the different techniques. Thus, a replication of this study that carefully removes this problem could identify differences between locomotion techniques. This possible future study is explained in more details in Section 7.5.

Despite the similar behavioural results in all the locomotion conditions, participants reported having a different quality of experience in each locomotion condition considering the sense of presence, difficulty of performing the pointing task and their personal evaluations.

Introspective results demonstrated that participants reached a significantly higher sense of presence during navigation in the panoramic video tour as predicted by our fourth hypothesis. Participants' sense of presence was assessed from their answers to the presence questionnaire which contained questions about the reality of their experience, its consistency with the real world experience, feeling of being surrounded by the virtual environment, and sense of being in that environment. On the other hand, in their comments, participants complained about the unrealistic looking images and the unnatural chair-based interface. However, panoramic video with the surrounding view images and smooth video movement seems to have had reduced the effects of these barriers better than the other two locomotion techniques.

Although we observed a higher sense of presence in panoramic video, participants' navigational knowledge acquisition performance in this condition was not any better. This can be explained by the fundamental assumption of the dual task measurements of presence (see Section 3.3.3 for more details): since participants needed to devote a greater part of their attentional resources to vision control, physical rotation, and distraction handling in panoramic video, they reached a higher sense of presence, but had fewer attentional resources left for learning the specific structure and composition of route elements. However, there is no strong argument regarding the type and strength of the relationship between the subjective sense of presence and navigation performance in the literature (Nash et al., 2000).

On average, participants rated the pointing task to be easier in the panoramic video and regular video conditions than in the slide show condition partially confirming our fifth hypothesis. However, in a somewhat contradictory manner, they complained about having serious problems with the panoramic views. They related these problems to the difficulties keeping track of the straight-ahead direction, which resulted in difficulty in understanding the turn directions and learning the routes. Therefore, it appears that the high update rate of the panoramic views (10 fps) provided by the video display compensates for the difficulties of tracking the straight-ahead view. Consequently, the slide show technique is identified as the most difficult locomotion technique for obtaining survey knowledge in passive navigations. This is not surprising as the problem of losing track of the straight-ahead view can produce misunderstanding of the turning directions, especially when there is no optic flow indicating the turn.

Some researchers such as Witmer and Singer (1998) suggest that if a task is more difficult and needs greater attention, it increases the sense of presence. However we observed a contradictory pattern in our experiment. Participants had the most difficulty learning the environment in the slide show tour, but they had the most sense of presence in the panoramic video tour. This can be because when the difficulty of the task exceeds a specific level, participants lose their motivation and engagement. This can also be because participants' answers to the 'difficulty' question are not as reliable as their answers to the presence questionnaire as there was only one question about the difficulty of task in the post questionnaire, while there were several questions (i.e. six questions selected from Igroup presence questionnaire) in the presence questionnaire.

7.4.5 Gender Difference in Pointing Performance

Finally, we observed a gender difference in the pointing performance so that males performed generally better than females. This is consistent with the reliable differences between males and females' strategies and cognitive abilities in wayfinding tasks reported in the literature (Kim et al., 2007; Cushman et al., 2005; Lawton, 1994). However, we cannot rely on the details of gender effects observed in this study, since the number of participants was too limited for this purpose.

7.5 Future Work

This thesis has reported on an initial exploration of the use of panoramic video to create virtual environments. The work has focused on how experiments can be designed to evaluate the effectiveness of such virtual environments. Although the particular experimental setup that we designed and implemented did not prove to be directly useful for evaluating the virtual environments we attempted to create, we learned a lot about the problem and we are now in a position to set out a more effective approach. We will first discuss the problems with our current experimental approach and then propose a better approach that could be implemented in the future.

7.5.1 Problems with the Current Experimental Design

As discussed above, our experiment attempted to get measures of the spatial knowledge that subjects gained from three different types of virtual tour of the physical environment. The biggest problem with the experiment in general was that the pointing task used to evaluate this knowledge was just too hard. The task was perceived to be too difficult by the participants for several possible reasons: (1) Participants had to depend highly on their memory in order to learn the long and complex routes, (2) Although the pointing task required survey knowledge, people with average spatial abilities are not likely to obtain survey knowledge after two identical route traversals in a large-scale environment, (3) Putting participants in a random place on the route is not natural, since, in the real world, people actually move through the environment in order to reach any location.

Another problem which could have contributed to not finding performance differences between panoramic video and the other modalities, results from the failure of our assumption that increased engagement, visual information, and realness in panoramic video leads to an increase in participants' performance in acquiring directional knowledge. In other words, we assumed that the panoramic view would provide an increased amount of visual information; in combination with the interactive chair and the smoothness of movement in video, we expected it to increase engagement and the realness of the experience. However, in reality, several major problems worked against this assumption. First of all, the low quality of the images (especially panoramas because of the stitching traces) highly reduced the value of increased visual information. Secondly, the interactive chair was perceived to be more distractive and unnatural than engaging. Additionally, the difference between the participants' physical FOV (45 degrees) and virtual FOV (90 degrees) may have contributed to an unnatural perception of the panoramic video experience and confusion in using body-based cues for sensing directions. It appears that participants failed to take advantage of the panoramic video features in general, as they reported using it in the same way that they used regular video.

The third problem involves the strategies that participants developed for completing the task. Apparently the only strategy that most participants utilized was to remember the intersections and the associated actions as they reported in the post experiment questionnaire. This is most likely because the virtual tour navigation was passive and did not involve any searching or path selection. Also the pointing task was predictable and did not necessarily require any searching or interaction with the panoramas. We observed and participants reported that except for when they were passing by the turns they did not interact with the panoramic view that much. Considering the issues involved with the interactions (discussed earlier) it appears that the costs of using the interactions provided in panoramic video were too high to let people get engaged in the interactions.

The final problem relates to the insufficiency of the information that our experimental results provide to illuminate: (1) the relationship between the participants' behaviour in the tasks and the final conclusions, (2) the relationship between the different levels of directional knowledge and the efficiency of the three different locomotion techniques.

The relationship between the participants' behaviour in the tasks and the final conclusions is not clear to us. For example, in our experimental design the recollection task was separated from the pointing task. Therefore, when participants made a high pointing error, we could not judge if it is because they did not even recognize where they were or because they had obtained a distorted image of the environment. Similarly, we have no means to identify if a huge pointing error is the result of a random pointing or a misunderstanding of the environment.

The relationship between the different levels of directional knowledge and the efficiency of the three different locomotion techniques is not clear as our experiment was designed to discretely assess the lowest and highest levels of navigational knowledge, landmark recognition and survey knowledge. Therefore, our experiment did not clearly assess the other possible levels of navigational knowledge in-between landmark and survey knowledge such as route knowledge. Based on the results, we know that all the three techniques were equally good for landmark recognition and equally poor for survey knowledge. However, it is not transparent at what level during this wide range of navigational knowledge, a specific technique possibly started or stopped being more efficient than the others. After we observed significantly lower errors participants made when pointing from the first pointing location (as discussed in Section 7.4.3) we performed separated inferential analysis on the pointing data collected at each specific pointing location. Results of these analyses (described in Appendix F) showed that at the first pointing location --where the pointing task is easier--participants performed significantly better in the regular video condition than the other two conditions. At the second and third pointing locations –where the pointing task is harder to do -- no significant difference between the locomotion techniques was found. We do not emphasize the significantly better efficiency of regular video as a reliable outcome of this study because this analysis carries a high chance of type 1 error. However, we are confident that an improved experimental method that increases both the directional information and task difficulty in a transparent stage-wise manner would find significant differences between the three different locomotion techniques.

7.5.2 A Possible Follow-on Study Design

In order to design a study that would allow us to effectively evaluate panoramic video in comparison to panoramic slides and regular video as the basis for virtual environments it is necessary to devise tasks substantially easier than that in the present study but not so easy that all subjects achieve high performance in all conditions. There are a number of aspects that will affect the difficulty of the tasks.

To address the problem with the difficulty of memorizing routes and developing survey knowledge, routes need to be shorter and less complex. Also, we can use multiple (e.g. two) routes that overlap, interconnect, and/or make a loop. In this way, participants have the chance to pass by some landmarks from multiple directions and update their mental image of the environment. It would also be helpful if there were clear landmarks that are easy to identify and remember on all routes. To address the issues with the difficulty of pointing task a solution is to move participants through the environment to reach the pointing location.

To address the final problem discussed in the previous section, we need to design tasks for each level of directional knowledge as well as to combine recollection and pointing task. Also, the virtual tours need to give participants room to take advantage of the available features in each locomotion technique and develop knowledge acquisition strategies. Considering all these concerns a possible experimental design solution would:

- Have both passive and active navigations. This can be a between-subject variable;

- Use two short routes that have common clearly identifiable landmarks (or intersections) such as the ones in Figure 10; and
- have a supplementary data collection method such as map drawing.

Based on this, a possible experimental procedure for the participants in the passive group would be:

1. Participants traverse the route '*abcde'* (see Figure 10). During the traversal, at each intersection the video (or slide show) stops for few seconds to let participants learn the place.
2. Participants perform the recollection and pointing tasks so that they start from the beginning of the route. They are then shown four landmark pictures on top of the screen. Two of the pictures are from the landmarks placed at the next interaction, one in the front view and the other in the side views. The other two landmarks are from other places. Participants are supposed to select the next landmark as they remember it. They then need to specify the direction to take in order to get to the next landmark and also estimate the direction to all the other landmarks. After completing a similar task at each intersection, participants are moved to the next intersection on the route until they get to the end of the route. Instead of pointing randomly, participants can skip pointing if they cannot make any estimate of directions. Finally they draw the map of the route.

3. Now, participants traverse route *'abcde'* and '*efcga'* and perform the similar recollection and pointing tasks for this route. Then they draw the map of the two routes integrated.

The procedure for the participants in the active group would be:

1. Participants start from the beginning point '*a*'. They travel along the route *'abcde'* so that they can control the speed and direction of navigation on the route (they are not yet given options to move to the route *'agcfe'*).

2. Then participants perform the recollection and pointing task as explained for the passive group.

3. Now participants can travel along both routes and select any of the possible directions (i.e. four directions at the point '*c'* and two directions at the points '*a', 'b'* , *'d', 'f'* and *'g'*). After they finish navigation they perform the recollection and pointing tasks again. Finally they draw a map of the environment.

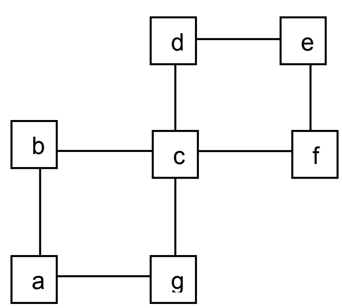

Figure 10: **route design for a possible future study.**

In order to keep both groups spending almost the same amount of time learning the environment, participants in the active group need to be carefully instructed about how many landmarks they need to visit in total, how much time they have and how much time on average they can spend on making decisions at each intersection.

Since we are interested in the differences between the locomotion techniques, the study can be broken into two studies, one for passive navigation, and the other for the active navigation. Another alternative to this design would be to have three groups of participants for passive, semi-active and active navigation. In the semi active navigation they cannot control the speed of video but can stop at the intersections and choose their next step. In the active navigation, they can control the speed at any moment during the navigation. Other detailed logistics of the experimental design have to be addressed by the future researcher.

Results of such an experiment should be analysed at several levels. Locomotion techniques can be assessed by looking at landmark knowledge, route knowledge and survey knowledge data separately. For example, by looking at how participants select the next landmarks on the route (using front view landmark or side view landmark), we realize how useful a panoramic view is for the obtaining landmark knowledge. Locomotion techniques can also be compared by looking at how many participants progressed from landmark knowledge to route knowledge and to survey knowledge in each condition.

Finally, for any similar future study, we suggest removing the interactive chair and using a HMD with equal physical and virtual FOV presented in each view in order to remove the barriers in experiencing a natural interaction with the panoramas. If the budget permits, we strongly advise increasing the quality of images and the frame rate of video. These will automatically improve stitching quality with the current image stitching programs.

7.5.3 More Possible Future Studies

Besides the experiment explained in the last section, other paths to take for future research include:

- To remove the slide show conditions as it raised the most complaints of participants and so we know it as the worst technique based on the results of this study. Then, panoramic video and regular video can be compared in more details. Or
- To compare the three conditions of panoramic video, regular video, and regular video with panoramic views only at the intersections (or landmarks).

8: CONCLUSIONS

All three locomotion techniques are equally capable of providing sufficient landmark knowledge; however, on average, they afford inadequate directional knowledge relative to what is required for participants to perform well in the orientation tasks.

Regardless of the type of locomotion technique, participants formed an orientation-dependent spatial knowledge of the environment, as their pointing performance was better at the first pointing location than the second and third pointing locations. Apparently, most of the participants did not develop successful survey knowledge possibly because of the long routes and the lack of active exploration or good visual realism. This incomplete survey knowledge is highly evident in participants' poor and random pointings performed at the middle and end points of the routes. Therefore, no strong conclusion about the distinctions between different locomotion techniques in providing survey knowledge can be drawn based on this study.

In the scope of our study, panoramic video surprisingly offered no advantage over the regular video or slide show. However, it is much more labour intensive and expensive to create. This implies that, for tasks involving only the passive learning of a specific route, there is no need to spend time, money and special technologies for creating panoramic video.

However, when it comes to more complex tasks (obtaining survey knowledge), panoramic video did not work significantly worse than slide show or regular video despite participants' lack of previous experience with panoramic video. Although most of the participants were familiar with regular video (e.g., from watching regular movies) and slide shows (e.g., from Google street view), neither of them worked better than panoramic video. This demonstrates the high compatibility of the panoramic video and its potential for future uses. This potential plus the power of panoramic video to offer sense of presence, provides a motivation for the assessment of panoramic video in active exploration tasks or navigational tasks that require searching and engagement.

APPENDICES

Appendix A: General Post-questionnaire

Q1. Please answer the following in the space provided

- Name (optional):
- Age:
- Gender:
- Occupation (or field of study):
- Experiment code:

Q2. How often do you use the computer? (in hours/day)

Q3. Do you play 3D video games? (Especially 3D games)

- Yes
- No

Q4. If yes, how many hours/day do you play on average?

Q5. Have you used Google street view?

- Yes
- No

Q6. If yes, how often? If no, have you used anything similar?

Q7. Have you been exposed to panoramic images? If yes, how often?

Q8. How did you solve the task? Did you use any strategies?

Q9. What did you think of the chair-based rotation interface?

Q10. How difficult was the pointing task in the panoramic video condition for you?
(0=very easy, 10=very difficult)

Q11. How difficult was the pointing task in the normal video condition for you?
(0=very easy, 10=very difficult)

Q12. How difficult was the pointing task in panoramic slideshow condition for you? (0=very easy, 10=very difficult)

Q13. What was difficult about the task? Why? Please elaborate.

Q14. How would you rate your every day spatial orientation/sense of direction?
(0=rather poor, 10= quite good)

Q15. How would you rate your ability to visualize an environment based on a 2D image? (0= rather poor, 10=quite good)

Q16. How well do you usually remember a new route after you have traveled it only once? (0=not so well, 10=very well)

Q17. Are you easily disturbed when working on a task? (0=very easily, 10=not easily)

Q18. How good are you at blocking out external distractions when you are involved in something? (0=poor, 10=quite good)

Q19. Do you easily become deeply involved in movies or TV dramas?

Q20. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

Q21. Are you familiar with any of the routes followed?

- Yes
- No

Q22. Did anything bother you during the experiment?

Q23. Which of the locomotion techniques did you like the most for virtual tour navigation?

- Panoramic Video
- Panoramic Slide Show
- Regular Video

Q24. Any suggestions about what to change?

Q25. Any other comments?

Appendix B: Sense of Presence Questionnaire

Q1. How well do you agree with this statement: “In the virtual tour I had a sense of being there”. (0=disagree, 10=totally agree)

Q2. How well do you agree with this statement: “Somehow I felt that the virtual world surrounded me”. (0= disagree, 10= totally agree)

Q3. How much did your experience in the virtual environment seem consistent with your real world experience? (0= inconsistent, 10=Quite consistent)

Q4. I had a sense of acting in the virtual space, rather than operating something from outside? (0=disagree, 10=totally agree)

Q5. How real did the virtual world seem to you? (0= unreal, 10= quite real)

Q6. How aware were you of the real world surrounding while navigating in the virtual world? (i.e., sounds, room temperature, other people, etc.)?

Appendix C: Participants' Demographic Description

Table 3: Demographics of the study participants

Participant Number	Age	Gender	Occupation	How often do you use computers?	Do you play 3D video games? How often?	Do you use Google street view? How often?	Have you been exposed or worked with panoramic images?	Spatial Ability (out of 10)	Immersion Ability (out of 10)
1	26	Female	SIAT Grad Student	5	No	Just a couple of times	No	2.7	3.8
2	25	Male	Education Grad Student	4	Yes1 hr per day	Once a month	No	3	4
3	23	Female	Computing Science Grad Student	5	No	Most of the times	No	4.3	6.5
4	23	Male	Science Undergrad Student	5	Yes1 hr per day	More than 50 times	More than 100 times	5	4.3
5	24	Female	Mechatronics Grad Student	14	No	Four or five times a month	Rarely	5	5.5
6	25	Female	Computing Science Grad Student	7	No	Rarely	Rarely	5.3	3.5
7	25	Male	Mechatronics Grad Student	7	No	Rarely	No	6.3	7.3

8	25	Female	SIAT Undergrad Student	10	No	Only couple times	Rarely	6.3	5.3
9	39	Male	Computers/ Theatre/Misc	8	No	Once a month	Yes, created some myself	6.3	5.3
10	40	Male	SIAT Grad Student	8	No	Once per two-three months	Rarely	6.3	5.5
11	25	Female	Education Grad Student	5	No	At least once a week	Some times	7	6.8
12	26	Female	Human Geography Grad Student	5	No	Just tried it once or twice at all	Rarely	7	4
13	25	Male	Computer Science Grad Student	10	No	Once or twice a week	Some times	7.3	7.8
14	30	Female	SIAT, PhD Student	6	No	Twice a month	Some times	7.7	6.3
15	26	Male	SIAT Grad Student	5	Yes1 hr per day	Once or twice a week	Some times	7.7	7
16	24	Male	Computer Science Grad Student	15	No	Only couple of times	Rarely	8	5.5
17	26	Male	Mechatronics Grad Student	6	No	Very often	Yes, Very often	7.5	6
18	29	Female	SIAT Grad Student	10	No	Rarely	Rarely	4.5	5.5

Appendix D: Tables for Descriptive Statistics of Behavioural Data

Table 4: Descriptive statistics of the absolute ego-orientation errors

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantile
1	Panoramic Video	54	48.2	41.6	33.5	15
1	Regular Video	54	50.3	54.4	25.5	10
1	Slide Show	54	47.5	48.0	28.5	8
2	Panoramic Video	54	33.3	42.4	15	5
2	Regular Video	54	32.8	34.2	20	8
2	Slide Show	54	38.4	45.4	17	7

Table 5: Descriptive statistics of the configuration errors

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantile
1	Panoramic Video	54	37.7	19.5	33	21.5
1	Regular Video	54	38.1	21.2	34.5	19.8
1	Slide Show	54	36.9	18.4	34	20
2	Panoramic Video	54	28.4	19.8	21	13.8
2	Regular Video	54	29.0	19.3	25	12
2	Slide Show	54	29.2	18.1	27	13

Table 6: Descriptive statistics of the absolute pointing errors

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantile
1	Panoramic Video	54	53.3	33.7	42.5	25.5
1	Regular Video	54	55.0	39.7	51	19.8
1	Slide Show	54	53.6	38.2	48	21
2	Panoramic Video	54	38.5	34.7	21.5	13
2	Regular Video	54	38.5	31.3	27.5	12.8
2	Slide Show	54	43.3	36.0	34.5	13

Table 7: Descriptive statistics of the pointing times

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantile
1	Panoramic Video	54	12.8	13.4	7.5	3.9
1	Regular Video	54	11.1	10.8	6.4	4.3
1	Slide Show	54	8.4	6.7	6.2	4.5
2	Panoramic Video	54	9.1	9.2	6.3	3.8
2	Regular Video	54	9.2	6.7	7.2	4.3
2	Slide Show	54	7.7	6.0	6.1	3.6

Table 8: Descriptive statistics of the recollection errors

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantiles
1	Panoramic Video	18	26.8	17.1	25	14.6
1	Regular Video	18	28.7	19.4	25	16.7
1	Slide Show	18	19.9	12.8	16.7	8.3
2	Panoramic Video	18	18.0	17.0	12.5	6.2
2	Regular Video	18	18.1	11.5	20.9	8.3
2	Slide Show	18	17.6	11.4	16.7	8.3

Table 9: Descriptive statistics of the recollection times

Trial	Locomotion Technique	N	Mean	Std Dev	Median	%25 Quantiles
1	Panoramic Video	18	7.3	4.6	5.8	3.9
1	Regular Video	18	5.7	2.4	5.2	4.2
1	Slide Show	18	6.2	2.3	6	4.4
2	Panoramic Video	18	5.8	3.6	5.1	3.3
2	Regular Video	18	5.0	2.5	4.1	3.3
2	Slide Show	18	4.5	2.4	3.5	2.9

Appendix E: Tables for Inferential Analysis of Behavioural Data

The following tables summarize the results of the mixed-model analysis and post-hoc tests on the pointing errors. All the statistics are performed at the .05 alpha level. DF is the degrees of freedom of the numerator and DF Den is the degrees of freedom of the denominator of the F ratio.

Table 10: Tests of fixed effects on the absolute ego-orientation errors using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	51	11.766	.001
Locomotion	2	31.52	.088	.916
Pointing Location	2	204	14.858	.000
Trial * Locomotion	2	51	.371	.692
Trial * Pointing Location	2	204	.228	.796
Locomotion * Pointing Location	4	204	1.633	.167
Trial * Locomotion * Pointing Location	4	204	.498	.737

Table 11: Pairwise comparisons of pointing locations regarding the mean absolute ego-orientation error, using Tukey's HSD criterion

Pointing Location (I)	Least Sq Mean	Std Error	Pointing Location (J)	Significance (p value)
1	25.833	5.907	2	< .05
2	43.722	5.907	3	>.05
3	55.685	5.907	1	< .05

Table 12: Tests of fixed effects on the configuration errors using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	51	42.048	.000
Locomotion	2	31.27	.013	.987

Pointing Location	2	204	55.373	.000
Trial * Locomotion	2	51	.155	.856
Trial * Pointing Location	2	204	.726	.485
Locomotion * Pointing Location	4	204	1.317	.265
Trial * Locomotion * Pointing Location	4	204	.323	.863

Table 13: Pairwise comparisons of pointing locations regarding the mean configuration error, using Tukey's HSD criterion

Pointing Location (I)	Least Sq Mean	Std Error	Pointing Location (J)	Significance (p value)
1	21.639	2.873	2	< .05
2	41.176	2.873	3	>.05
3	36.796	2.873	1	< .05

Table 14: Tests of fixed effects on the absolute pointing errors using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	51	24.299	.000
Locomotion	2	28.39	.163	.850
Pointing Location	2	204	27.402	.000
Trial * Locomotion	2	51	.435	.649
Trial * Pointing Location	2	204	.308	.735
Locomotion * Pointing Location	4	204	1.387	.239
Trial * Locomotion * Pointing Location	4	204	.967	.426

Table 15: Pairwise comparisons of pointing locations regarding the mean absolute pointing error, using Tukey's HSD criterion

Pointing Location (I)	Least Sq Mean	Std Error	Pointing Location (J)	Significance (p value)

1	30.69	5.30	2	< .05
2	52.95	5.30	3	>.05
3	57.47	5.30	1	< .05

Table 16: Tests of fixed effects on the pointing times using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	51	5.92	.018
Locomotion	2	31.5	2.33	.114
Pointing Location	2	204	32.25	.000
Trial*Locomotion	2	51	.99	.377
Trial* Pointing Location	2	204	1.06	0.349
Locomotion * Pointing Location	4	204	.82	.510
Trial * Locomotion * Pointing Location	4	204	1.26	.285

Table 17: Pairwise comparisons of pointing locations regarding the mean pointing time using Tukey's HSD criterion

Pointing Location (I)	Least Sq Mean	Std Error	Pointing Location (J)	Significance (p value)
1	5.26	1.26	2	< .05
2	13.52	1.26	3	< .05
3	10.42	1.26	1	< .05

Table 18: Tests of fixed effects on the recollection errors using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	67.56	8.607	.005
Locomotion	2	57.38	1.191	.311

Trial * Locomotion	2	67.56	1.047	.357

Table 19: Tests of fixed factors on the recollection times using mixed-model analysis

Source	DF	DF Den	F	Sig.
Trial	1	64.15	7.734	.007
Locomotion	2	69.37	1.122	.332
Trial * Locomotion	2	64.15	.408	.666

Appendix F: Tables for Inferential Analysis of Behavioural Data for the First Pointing Location

The following tables summarize the results of the mixed-model analysis and post-hoc tests for pointing errors collected at the first pointing location. All the statistics are performed at the .05 alpha level. DF is the degrees of freedom of the numerator and DF Den is the degrees of freedom of the denominator of the F ratio.

Table 20: Tests of fixed effects on the pointing errors and pointing times, for the first pointing location using mixed-model analysis

Error	Source	DF	DF Den	F ratio	Sig.
Absolute Ego-orientation Error	Trial	1	85	6.913	.010
	Locomotion	2	85	69.760	.000
	Trial*Locomotion	2	85	1.383	.256
Configuration Error	Trial	1	51	10.895	.002
	Locomotion	2	31.34	.937	.402
	Trial*Locomotion	2	51	.554	.554
Absolute Pointing Error	Trial	1	85	8.744	.004
	Locomotion	2	85	44.931	.000
	Trial*Locomotion	2	85	2.115	.127
Pointing Time	Trial	1	51	1.421	.239
	Locomotion	2	30.92	.319	.729
	Trial*Locomotion	2	51	.566	.571

Table 21 : Pairwise comparisons of locomotion techniques regarding the mean absolute ego-orientation error at the first pointing location, using Tukey's HSD criterion

Locomotion Technique (I)	Least Sq Mean	Std Error	Locomotion Technique (J)	Sig.
Panoramic Video	31.14	4.17	Regular Video	< .05

Regular Video	20.15	4.84	Slide Show	> .05
Slide Show	26.21	5.54	Panoramic Video	> .05

Table 22: Pairwise comparisons of locomotion techniques regarding the mean absolute pointing error at the first pointing location, using Tukey's HSD criterion

Locomotion Technique (I)	Least Sq Mean	Std Error	Locomotion Technique (J)	Sig.
Panoramic Video	33.30	4.35	Regular Video	< .05
Regular Video	26.44	4.83	Slide Show	> .05
Slide Show	32.34	5.34	Panoramic Video	> .05

Appendix G: Tables for Inferential Analysis of Introspective Data

Table 23: Pairwise comparisons of locomotion techniques regarding the mean participants' subjective sense of presence, using Tukey's HSD criterion

Locomotion Technique (I)	Least Sq Mean	Std Error	Locomotion Technique (J)	Sig.
Panoramic Video	5.486	.418	Regular Video	< .05
Regular Video	3.745	.415	Slide Show	> .05
Slide Show	3.249	.418	Panoramic Video	< .05

Table 24: Pairwise comparisons of locomotion techniques regarding the mean difficulty of tasks, using Tukey's HSD criterion

Locomotion Technique (I)	Least Sq Mean	Std Error	Locomotion Technique (J)	Sig.
Panoramic Video	5.556	.551	Regular Video	> .05
Regular Video	5.667	.598	Slide Show	< .05
Slide Show	7.222	.517	Panoramic Video	< .05

REFERENCE LIST

Alfano, P. L., & Michel, G. F. (1990). Restricting the field of view: perceptual and performance effects. *Perceptual and Motor Skills*, *70*(1), 35-45.

Arthur, K. (1996). Effects of field of view on task performance with head-mounted displays. In *Conference companion on Human factors in computing systems common ground - CHI '96* (pp. 29-30). Presented at the Conference companion, Vancouver, British Columbia, Canada. doi:10.1145/257089.257116

Barfield, W., & Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, *1*(1), 3-15. Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environment: A conceptual framework. *Proceedings of the International Conference on Human-computer Interaction, 5*, 699-704. doi:10.1007/BF02009709

Behind the Scenes | Google Maps with Street View. (n.d.). *Google Maps*. Retrieved November 16, 2010, from http://maps.google.com/intl/en_us/help/maps/streetview/behind-the-scenes.html

Benosman, R., & Kang, S. B. (2001). *Panoramic vision: sensors, theory, and applications*. Springer.

Billinghurst, M., & Weghorst, S. (1995). The use of sketch maps to measure cognitive maps of virtual environments. In *Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)* (p. 40). IEEE Computer Society. Retrieved from http://portal.acm.org/citation.cfm?id=836005

Breakwell, G. M. (2006). *Research methods in psychology* (3rd ed.). London: Sage Publications.

Brooks Jr., F. P. (1999). "What's Real About Virtual Reality?", IEEE Computer Graphics And Applications, 19(6), 16

Chen, S. E. (1995). QuickTime VR: an image-based approach to virtual environment navigation. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (pp. 29-38). ACM. Retrieved from http://portal.acm.org /citation.cfm?id=218380.218395 Computer-Mediated Communication (on-line serial) 3(2). Available: http://jcmc.indiana.edu/vol3/

Cox, A. L., Cairns, P., Berthouze, N., & Jennett, C. (n.d.). The Use of Eyetracking for Measuring Immersion. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.137.9617

Cushman, L. A., Duffy, C. J., Stefenella, T., & Vaughn, W. K. (2005). Gender differences in navigation strategy, not performance. *Alzheimer's and Dementia*, *1*(1, Supplement 1), S22-S23. doi:10.1016/j.jalz.2005.06.108

Darken, R. P., & Sibert, J. L. (1993). A toolset for navigation in virtual environments. In *Proceedings, ACM User Interface Software & Technology* (pp. 157–165).

Darken, R. P., Bernatovich, D., Lawson, J. P., & Peterson, B. (1999). Quantitative Measures of Presence in Virtual Environments: The Roles of Attention and Spatial Comprehension. *CyberPsychology & Behavior*, *2*(4), 337-347. doi:10.1089/cpb.1999.2.337

den Dekker, E., & Delleman, N. (2007). *Presence*. Retrieved from http://stinet.dtic.mil/oai/oai?&verb=getRecord&metadataPrefix=html&identifier=ADA474136

Edwards, G., Thompson, J.,&MacGregor, C. (1998). Investigating the role of guided tours in virtual reality environments. *Proceedings of the Human Factors Association of Canada*, *Missausauga, Ontario, 1*,1–4.

Flash Panorama Player and QTVR Converter - Pano2VR - Garden Gnome Software. (n.d.). *Garden Gnome Software*. Retrieved November 16, 2010, from http://gardengnomesoftware.com/pano2vr.php

Gale, N., Golledge, R. G., Pellegrino, J. W., & Doherty, S. (1990). The acquisition and integration of route knowledge in an unfamiliar neighborhood. *Journal of Environmental Psychology*, *10*(1), 3-25. doi:10.1016/S0272-4944(05)80021-0

Gaunet, F., Vidal, M., Kemeny, A., & Berthoz, A. (2001). Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory. *Cognitive Brain Research*, *11*(3), 409-420. doi:10.1016/S0926-6410(01)00013-1

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151-176. doi:10.1016/j.intell.2005.09.005

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, 30, 425-447.

Hoeben, A., & Stappers, P. (2006). Taking clues from the world outside: navigating interactive panoramas. *Personal and Ubiquitous Computing*, *10*(2), 122-127. doi:10.1007/s00779-005-0017-8

Howes, A., Miles, G. E., Payne, S. J., Mitchell, C. D., & Davies, A. J. (2001). Incidental Memory and Navigation in Panoramic Virtual Reality for Electronic Commerce. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*(2), 239 -254. doi:10.1518/001872001775900869

Ijsselsteijn, W. A.,   a, A. I., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants and measurement. Retrieved from http://citeseerx.ist.psu.edu /viewdoc/summary?doi=10.1.1.13.4919

Irani, M., Anandan, P., & Hsu, S. (1995). Mosaic based representations of video sequences and their applications. In *Proceedings of the Fifth International Conference on Computer Vision* (p. 605). IEEE Computer Society. Retrieved from http://portal.acm.org/citation.cfm?id=840102issue2/lombard.html.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93-129. doi:10.1016/j.cogpsych.2005.08.003

James, K. R., & Caird, J. K. (1995). The effects of optic flow, proprioception, and texture on novice locomotion in virtual environments. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting,* 1405–1409.

Jennett, C., Cox, A., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, *66*(9), 641-661. doi:10.1016/j.ijhcs.2008.04.004

Kearns, M. J., Warren, W. H., Duchon, A. P., & Tarr, M. J. (2002). Path integration from optic flow and body senses in a homing task. *PERCEPTION-LONDON-*, *31*(3), 348–374.

Kimber, D., Foote, J., & Lertsithichai, S. (2001). Flyabout: spatially indexed panoramic video. In *ACM Multimedia* (pp. 339–347).

Koh, G. (1997). *Training spatial knowledge acquisition using virtual environments* [Online]. Available: http://18.62.2.144/VELbIndx.html

Kolbe, T. H. (2003). Augmented Videos and Panoramas for Pedestrian Navigation. *GEOWISSENSCHAFTLICHE MITTEILUNGEN*, *66*, 45--52.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, *30*(11-12), 765-779. doi:10.1007/BF01544230

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282-297. doi:10.1162/105474601300343612

Lippman, A. (1980). Movie-maps: An application of the optical videodisc to computer graphics. *SIGGRAPH Comput. Graph.*, *14*(3), 32-42.

Liu, P., Sun, X., Georganas, N.D., Dubois, E. (2003). Augmented reality: a novel approach for navigating in panorama-based virtual environments. *Proceedings of The 2nd IEEE Internatioal Workshop on Haptic, Audio and Visual Environments and Their Applications*, 2003, 13-18, Ottawa, Ontario, Canada.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2), 0-0. doi:10.1111/j.1083-6101.1997.tb00072.x

Macedonio, M. F., Parsons, T. D., DiGiuseppe, R. A., Weiderhold, B. K., & Rizzo, A. A. (2007). Immersiveness and physiological arousal within panoramic video-based virtual reality. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *10*(4), 508-515. doi:10.1089/cpb.2007.9997

Mallot, H. A., & Gillner, S. (1999). Route navigating without place recognition: What is recognised in recognition-triggered responses? *PERCEPTION-LONDON-*, *29*(1), 43–56.

McCreary, F. A.,&Williges, R. C. (1998). Effects of age and field-of-view on spatial learning in an immersive virtual environment. *Proceedings of theHumanFactors and Ergonomics Society*, *42nd Annual Meeting, 2*, 1491–1495.

McNamara, T., Sluzenski, J., & Rump, B. (2008). Human Spatial Memory and Navigation. In *Learning and Memory: A Comprehensive Reference* (pp. 157-178). Oxford: Academic Press. Retrieved from http://www.sciencedirect.com/science/article/B9633-4S1TM42-55/2/76c860d7811f8a0383eb4297d5440dce

Meijer, F., Geudeke, B. L., & van den Broek, E. L. (2009). Navigating through Virtual Environments: Visual Realism Improves Spatial Cognition. *CyberPsychology & Behavior*, *12*(5), 517-521. doi:10.1089/cpb.2009.0053

Miller, A. L. (1996). The Panorama, the Cinema and the Emergence of the Spectacular. *Wide Angle*, *18*(2), 34–69.

Miller, G., Hoffert, E., Chen, S. E., Patterson, E., Blackketter, D., Rubin, S., Applin, S. A., et al. (1992). The virtual museum: Interactive 3D navigation of a multimedia database. *The Journal of Visualization and Computer Animation*, *3*(3), 183-197. doi:10.1002/vis.4340030305

Mou, W., McNamara, T. P., Rump, B., & Xiao, C. (2006). Roles of Egocentric and Allocentric Spatial Representations in Locomotion and Reorientation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1274-1290. doi:10.1037/0278-7393.32.6.1274

Naimark, M. (1997). A 3D Moviemap and a 3D Panorama. In *SPIE Proceedings* (Vol. 3012).

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A Review of Presence and Performance in Virtual Environments. *International Journal of Human-Computer Interaction*, *12*(1), 1. doi:10.1207/S15327590IJHC1201_1

Nordahl, R., & Korsgaard, D. (2009). Distraction as a measure of presence: using visual and tactile adjustable distraction as a measure to determine immersive presence of content in mediated environments. *Virtual Reality*, *14*(1), 27-42. doi:10.1007/s10055-009-0140-3

Neumann, U., Pintaric, T., & Rizzo, A. (2000). Immersive panoramic video. In *Proceedings of the eighth ACM international conference on Multimedia* (pp. 493-494). Marina del Rey, California, United States: ACM. Retrieved from http://portal.acm.org /citation.cfm?id=376408

Ono, S., Ogawara, K., Kagesawa, M., Kawasaki, H., Onuki, M., Honda, K., & Ikeuchi, K. (2005). Driving View Simulation Synthesizing Virtual Geometry and Real Images in an Experimental Mixed-Reality Traffic Space. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality* (pp. 214-215). IEEE Computer Society. Retrieved from http://portal.acm.org /citation.cfm?id=1104996.1105218

Oxford English Dictionary Sign In. (n.d.). *Oxford English Dictionary: The definitive record of the English language*. Retrieved November 18, 2010, from http://dictionary.oed.com/cgi/entry/50170445?single=1&query_type=word &queryword=panorama&first=1&max_to_show=10

Panoramic photography - Wikipedia, the free encyclopedia. (n.d.). *Wikipedia, the free encyclopedia*. Retrieved November 18, 2010, from http://en.wikipedia.org/wiki/Panoramic_photography

Park, A., Calvert, T. (2008). A Social Agent Pedestrian Model. *Computer Animation and Virtual Worlds, G. J. Kim, H. Qin, N. Magnenat-Thalmann (eds)*. John Wiley & Sons, Ltd, West Sussex, UK. Vol. 19, Nos. 3-4. pp. 331-340.

Peponis, J., Zimring, C., & Choi, Y. K. (1990). Finding the building in wayfinding. *Environment and Behavior*, *22,* 550–590.

Peri, V. N., & Nayar, S. K. (1997). Generation of Perspective and Panoramic Video from Omnidirectional Video. Retrieved from http://citeseerx.ist.psu.edu.proxy.lib.sfu.ca/viewdoc/summary?doi=10.1.1.41.5746

Photo stitching software 360 degree Panorama image software - PTGui. (n.d.). *Photo stitching software 360 degree Panorama image software - PTGui*. Retrieved November 16, 2010, from http://www.ptgui.com/

Point Grey - Spherical - Ladybug2 - CCD FireWire Camera. (n.d.). *Point Grey CCD and CMOS Digital Cameras for Industrial, Machine, and Computer Vision*. Retrieved November 18, 2010, from http://www.ptgrey.com/products/ladybug2/index.asp

Prestopnik, J. L., & Roskos-Ewoldsen, B. (2000). The Relations Among Wayfinding Strategy Use, Sense Of Direction, Sex, Familiarity, And Wayfinding Ability. *Journal of Environmental Psychology*, *20*(2), 177-191. doi:10.1006/jevp.1999.0160

Richardson, Montello, Hegarty, S. R. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, *27*(4), 741–750.

Riecke, B., Cunningham, D. W., & Bülthoff, H. H. (2007). Spatial updating in virtual reality: the sufficiency of visual information. *Psychological Research*, *71*(3), 298–313. doi:http://dx.doi.org/10.1007/s00426-006-0085-z

Riecke, B. E., Veen, H. A. H. C. V., & Bülthoff, H. H. (2002). Visual Homing Is Possible Without Landmarks: A Path Integration Study in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, *11*(5), 443-473.

Riley, J. M., Kaber, D. B., & Draper, J. V. (2004). Situation awareness and attention allocation measures for quantifying telepresence experiences in teleoperation. *Human Factors and Ergonomics in Manufacturing*, *14*(1), 51-67. doi:10.1002/hfm.10050

Rizzo, A., Pryor, L., Matheis, R., Schultheis, M., Ghahremani, K., & Sey, A. (2004). Memory assessment using graphics-based and panoramic video virtual environments. In *Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech*.

Schloerb, D.W. (1995). A quantitative measure of telepresence, *Presence: Teleoperators and Virtual Environments*, 4(1), 64-80.

Satalich, G. A. (1995). *Navigation & wayfinding in virtual reality: Finding proper tools and cues to enhance navigation awareness*. Unpublished thesis dissertation, University of Washington, Seattle, WA.

Sato, T., Kanbara, M., Yokoya, N., & Ikeda, S. (2004). An Immersive Telepresence System with a Locomotion Interface Using High-Resolution Omnidirectional Movies. In *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4 - Volume 04* (pp. 396-399). IEEE Computer Society. Retrieved from http://portal.acm.org/citation.cfm?id=1021244

Schubert, T. W., Friedmann, F., & Regenbrecht, H. T. (1999). Decomposing the sense of presence: Factor analytic insights. In *2nd International Workshop on Presence* (pp. 3–23).

Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. (2001). Research on presence in virtual reality: a survey. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *4*(2), 183-201.

Sebe, I. O., Hu, J., You, S., & Neumann, U. (2003). 3D video surveillance with Augmented Virtual Environments. In *First ACM SIGMM international workshop on Video surveillance* (pp. 107-112). Berkeley, California: ACM. Retrieved from http://portal.acm.org /citation.cfm?id=982466#citedby

Sheridan, T. (1994). Further musings on the psychophysics of presence. In *Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology'., 1994 IEEE International Conference on* (Vol. 2, pp. 1073-1077 vol.2). Presented at the Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology'., 1994 IEEE International Conference on. doi:10.1109/ICSMC.1994.399986

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, *10*, 9-55.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence-Teleoperators and Virtual Environments*, *3*(2), 130–144. doi:10.1162/1054746041944849

Slater, M. (2004). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *13*(4), 484-493. doi:10.1162/1054746041944849

Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators and Virtual Environments, 9*(5), 413-434.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, *42*(4), 73-93. doi:10.1111/j.1460-2466.1992.tb00812.x

Szeliski, R. (1996). Video Mosaics for Virtual Environments. *IEEE Comput. Graph. Appl.*, *16*(2), 22-30.

Szeliski, R., & Shum, H. (1997). Creating full view panoramic image mosaics and environment maps. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (pp. 251-258). ACM Press/Addison-Wesley Publishing Co. Retrieved from http://portal.acm.org/citation.cfm?id=258861

Tan, A. A., de Vries, B., & Timmermans, H. J. (2006). Using a stereo panoramic interactive navigation system to measure pedestrian activity scheduling behaviour: a test of validity. *ENVIRONMENT AND PLANNING B PLANNING AND DESIGN*, *33*(4), 541.

Tan, A., Timmermans, H., & de Vries, B. (2006). Route Knowledge in Complex Environments: An Analysis of Pedestrian Recall Using Stereoscopic Panoramic Interactive Navigation. *Spatial Cognition & Computation: An Interdisciplinary Journal*, *6*(3), 279. doi:10.1207/s15427633scc0603_4

Tang, W., Wong, T., & Heng, P. (2002). The immersive cockpit. In *Proceedings of the tenth ACM international conference on Multimedia* (pp. 658-659). Juan-les-Pins, France: ACM. Retrieved from http://portal.acm.org/citation.cfm?id=641141

Teodosio, L. A., & Mills, M. (1993). Panoramic overviews for navigating real-world scenes. In *Proceedings of the first ACM international conference on Multimedia* (pp. 359-364). Anaheim, California, United States: ACM. Retrieved from http://portal.acm.org /citation.cfm?id=168422

Verbree, E., & Anrooij, A. V. (2004). Interactive navigation through distance added valued panoramic images. In *ISPRS, Panoramic Photogrammetry Workshop*.

Waller, D. (2005). The WALKABOUT: Using virtual environments to assess large-scale spatial abilities. *Computers in Human Behavior*, *21*(2), 243-253. doi:10.1016/j.chb.2004.02.022

Waller, D., Beall, A. C., & Loomis, J. M. (2004). Using virtual environments to assess directional knowledge. *Journal of Environmental Psychology*, *24*(1), 105-116.

Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review*, *11*(1), 157-163. .

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45,* 413–428.

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2007). Increasing Cognitive Load to Facilitate Lie Detection: The Benefit of Recalling an Event in Reverse Order. *Law and Human Behavior*, *32*(3), 253-265.

Witmer, B. G., & Kline, P. B. (1998). Judging Perceived and Traversed Distance in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *7*(2), 144-167. doi:10.1162/105474698565640

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225-240. doi:10.1162/105474698565686

Wraga, M., Creem-Regehr, S. H., & Proffitt, D. R. (2004). Spatial updating of virtual displays during self- and display rotation. *Memory & Cognition*, *32*(3), 399-415.

Xiong, Y. & Turkowski, K. (1997). Creating image-based VR using a self-calibrating fisheye lens. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on* (pp. 237-243). Presented at the Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on.