Investigation of Eye Movements during Walking in Real and Simulated Environments

Ming Zhao

A thesis submitted to the University of York, Electronics Department for the degree of Doctor of Philosophy

October 2012

Abstract

This thesis reports experimental research on locomotion simulators, improving design methods and investigating human performance in simulator environments (SE) in comparison with real-world environments (RE). Locomotion simulators are used in gaming, training and medical applications, but in cases where motion capture of human movement is required they can be difficult to apply effectively without a great deal of trial-and-error. A new design tool is introduced, tested and successfully applied to this problem. In addition, eye movement performance is investigated and compared for a natural locomotion task requiring counting and reporting on ball targets of different sizes and colours while walking along a 38m corridor in both RE and its simulated counterpart. It is found that the number of eye saccades and other measures in the real environment are different and consistently better (lower) in performance compared with a carefully replicated simulated environment. Experiments that isolate the possible factors for these differences find that neither walking itself, neither walking speed nor visual optic flow speed alone can explain the differences. Rather, it is found that the underlying factor is a mismatched combination of walking and visual speed during locomotion that can occur in SE, revealing also that eye movements are involved in the complex perceptual interactions taking place during locomotion and are not due to mechanical effects of physical movement. An efficiency metric of eye movement performance during locomotion is devised as a model metric, and used to compare and tune the different combinations of parameters in locomotion simulators. Using this metric, results show that SE conditions can be tuned to achieve levels of performance comparable to the real world by matching optic flow and walking speeds at middle speed levels. Overall, the research contributes to a better understanding of human interactions in locomotion, and introduces new methods and approaches for investigating, tuning and optimizing the design and development of virtual reality locomotion environments.

List of Contents

List of Figures

List of Tables

List of Equations

List of Publications

P. Jepson, M. Gilbert, M. Zhao, W. Lunniss and A. Pelah, 'Visualization and Sonification of Human Locomotion Data for Rehabilitative Biofeedback', accepted for publication in Proceedings of the SPIE, SPIE/IS&T Human Vision & Electronic Vision, San Jose, California, January 2010.

M. Zhao, A. Pelah, 'Eye Movement Patterns during Locomotion in Real-World and Simulated Environments', In *Poster presented at AVA Christmas meeting, Journal of i-Perception, Volume 3 (4).* York, 19th December, 2011.

Acknowledgements

Firstly, I would like to express my sincere appreciation to my supervisor Dr Adar Pelah. Four years ago when I just graduated, I was on the crossroads in my life; it was Adar who showed and instructed me to the most correct way. I still remember his patient suggestion at that time that a PhD degree would help with my future career and life. In the past seven years (including my three years undergraduate study), Adar did his best to help me as much as he could with not only my academic, but also some personal problems I met, internship applications, funding applications etc. For me, he is not only a respected supervisor, but a great friend, I am so lucky to study with him in my student life.

I would also like to thank my thesis supervisor Dr Andy Hunt for his recommendations on research directions, and his exceedingly helpful suggestions on time management and organisation.

I am very grateful to my colleague Matthew Gilbert for his valuable support; improving my programming skills would not have been possible without his continuous support, enthusiasm and many useful discussions.

I wish to express my sincere gratitude to Professor Aude Billard and Basilio Noris for kindly lending us the use of the eye tracking apparatus which they developed at the Learning Algorithms and Systems Laboratory in the Department of Microengineering at EPFL, Lausanne, Switzerland.

I also would like to thank my parents for their love and financial support over the many years I have been studying in UK, for which I am deeply grateful. Without them I could not have completed my PhD studies.

I very much appreciate my girlfriend Chengjun Cao for her encouragement and love. She has been waiting for me in China for four years, and her love has never changed during these years. She gave me the motivation to continue my PhD study and also gave me every support she could both financially and emotionally.

I would like to take this opportunity to thank Mr Clive Smith and my sister Lulu Smith for their great help and kindnesses and for looking after me as they did.

I am also grateful to all my friends in these four years, thanks indeed for their consistent participation in my experiments, and also my life in here would not have been nearly as colourful without them. Some of them have already left York, such as YinlinXu, Qingyuan Meng, Lei Xu, Jilin Zhao, Ran Wang, Bin Pan, Xiaoyou Liu, Mengmeng Zhang, Cheng Chen, Chen Geng, Hua Ling, Tengjiao Li, ChenhuiBao, Ruoan Ma and XinZheng. Some of them are still in York with me, like Jiazhuo Cheng, Yuan Zhang, Shuo Zhao, Wen Cai and Cong Lu. My time in York has been made the more pleasant for being with them. Thank you all!

Declaration

I undertake that all the material presented for examination is my own work and has not been written for me, in whole or in part, by any other person. I certify that all material in this thesis which is not my own work or is the outcome of work done in collaboration has been identified and that no material has been or is being concurrently submitted for any other qualification other that the degree of Doctor of Philosophy at The University of York.

Ming Zhao

York, UK

18 October, 2012

Chapter 1. Introduction

1.1 General Background

A locomotion simulator is a device or system that allows humans to experience walking as naturally as possible in a controlled environment by presenting the illusion of locomotion activities while remaining stationary in a room or laboratory. Locomotion simulators are used and applied in many areas such as entertainment, military and clinical applications. Compared to the real world, a locomotion simulator can provide a user with a safe and convenient environment for training, rehabilitation or other purposes. Technologies involved in a locomotion simulator are treadmills (or other interfaces) and virtual reality displays. A virtual reality is a simulated environment, often similar to the real world (RE), intended to create a lifelike experience for users, whose movements can be captured using cameras or other forms of motion capture technology. One problem currently is that in locomotion simulators the full locomotion space can sometimes be difficult to capture without a great deal of trial-and-error and experimentation in the positioning of the motion capture cameras.

In a number of previous studies, a locomotion simulator was used to help understand people's locomotion by comparing subjects' walking performance during tasks in real and virtual environments (see special issue on Walking in Real and Virtual Environments, Pelah & Koenderink 2007). The locomotion tasks include distance estimation, obstacle avoidance and walking and visual optical flow matching etc. The navigation equipment used in the locomotion simulator are normally a joy-stick, mouse, keyboard (or other interfaces) and, for actual walking interfaces, a treadmill (e.g. Pelah et al., 1998; Durgin, Gigone & Scott, 2005; Rieser et al., 1995); the display technologies in the simulator used either head mounted display (HMDs) or a large-screen immersive display (LSID) (e.g. Fink, Foo & Warren, 2007).

In the obstacle avoidance studies, Fink & Warren (2002) found that there are differences between walking in the matched real environment and a locomotion simulator. Experiment results showed that the large deviation of walking paths could slow down subjects' walking speed when avoiding obstacles, the reason being that people felt uncertain about the egocentric distance when walking in the virtual simulator. These findings echo Fajen & Warren (2005), who also found the behaviour differences between real world and HMD simulated locomotion. A study on travelling distance estimations by Frenz & Lappe (2005) found that when subjects walked in an optic flow sequence in a locomotion simulator, they consistently underestimate the distance by 25% compared with the walking in the real environment; it suggested that in the distance estimation tasks in virtual environments, subjects should undertake an adaptation session (e.g. Rieser et al., 1995; Pelah & Barlow, 1996) to re-calibrate their visuomotor system while physically walking in the simulated locomotion. Another interesting topic in simulating locomotion is to match the visual perceived speed to the real walking speed on the treadmill (e.g. Pelah et al. 1998). An experiment by Thurrell et al (1998;

2005) suggested that walking on the treadmill would slow down perceived optic flow speed proportionally to walking speed. This finding was supported by other studies (e.g. Banton et al., 2005; Durgin, Gigone & Scott, 2005;Thurrell &Pelah, 2005), which found that in the treadmill based simulator, subjects' perceived visual speed was slower when they were walking at very high speed.

In some aspects, a simulator has its disadvantages compared with a real environment. For example, a training pilot in the simulator may experience lack of expected sensations because the true environment is not mimicked effectively, which may eventually make the pilot not properly prepared for flying in the real environment. Also, participants' reactions in simulated environment may be unreliable as when they are affected by a participants' age and the amount of experience in the corresponding real environment (e.g. U.S. Army Research Institute for the Behavioural and Social Sciences, 2005).

In this PhD study, we use the Stroke Mobility Rehabilitation (StroMoHab) system as a locomotion simulator platform for the research experiments. StroMoHab was initially designed to simulate real world locomotion for rehabilitation of patients with gait problems, such as those who have had a stroke. Unlike the majority of simulator studied and technologies, StroMoHab therefore focuses more on the lower body, and detects people's leg position and movements via motion capture cameras while providing an avatar for real-time patient feedback on a virtual reality screen.

In general, the use and development of simulators is increasing, and the overall goal of most applications is to make virtual reality that can bring the user the same richness and interactivity of experience as they would experience in the real environment. The present study represents a contribution towards that larger goal.

1.2 Aims and Objectives

Locomotion is a fundamental human activity, including not only leg, arm and trunk movements but also eye movement; eye movements are also natural movements that humans apply during locomotion. In this thesis we study eye movements during normal locomotion, and investigate by comparing eye movement patterns during locomotion in a simulated environment versus a real environment, in order to 1) develop and test the methods for improving the design and configuration of motion capture in a locomotion simulator; 2) Improve understanding of human interaction within a locomotion simulator environment, and the differences between a locomotion simulator and real environment; 3) to increase understanding of the relationship between eye movements and locomotion; 4) to find the best conditions under which locomotion simulator environments perform as similarly as possible to real world environments.

1.3 StroMoHab Platform

This section provides the theoretical background and some key knowledge of the StroMoHab platform which should be understood before starting to develop this simulation system.

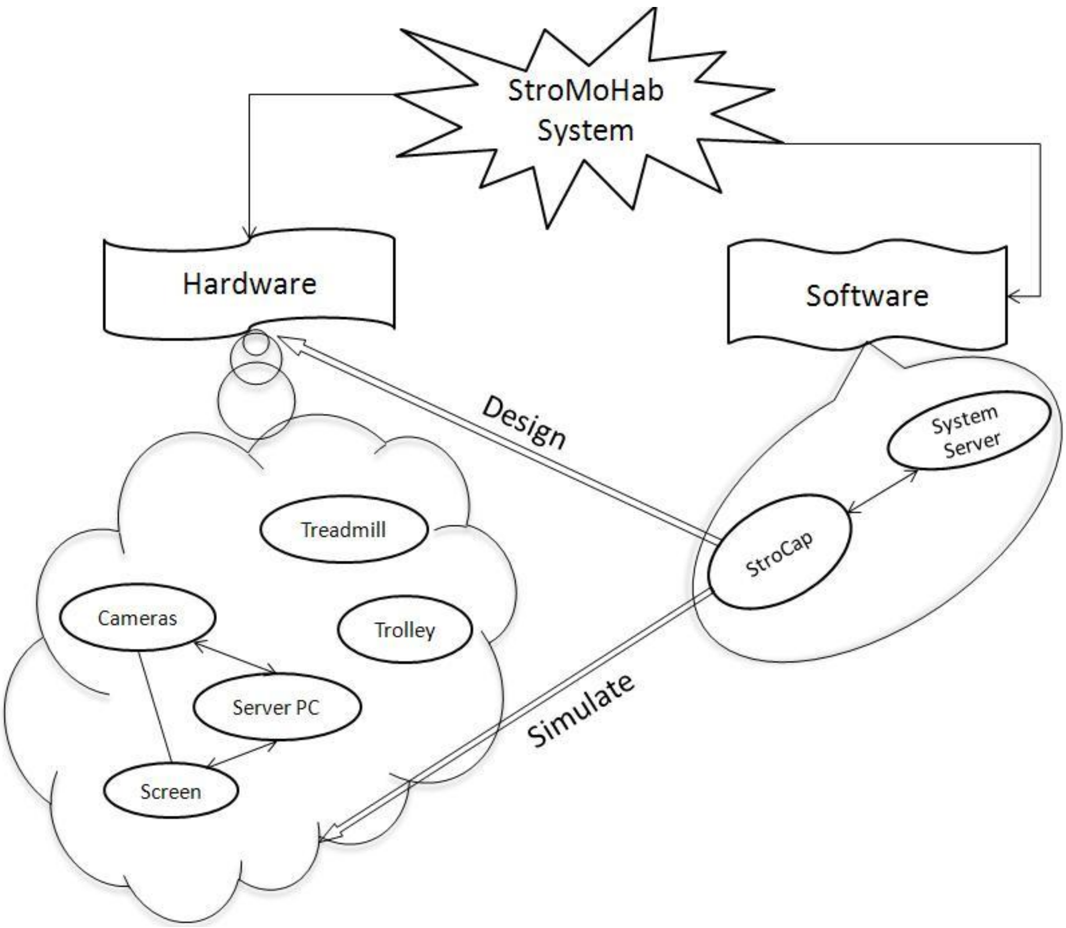

Figure 1.1 A schematic diagram of StroMoHab system, the components included in the system and the relationships between them. The full set of components shown is applicable to the StroCap research. Note that the locomotion experiments that were run on StroMoHab also included an eye tracker (not shown) and did not require use of the motion capture system (labelled 'cameras' above).

The StroMoHab system contains hardware and software sections (see Figure 1.1). In the hardware part, it includes a treadmill, between 6 and 12 motion capture cameras, a server PC, a trolley and a display screen. The cameras and screen are all connected with the server PC via USB cables, and also the cameras are mounted on the sides of the screen. In the software part, it contains a system server, which installed in the server PC for camera calibration; and also contains the StroCap software, which was designed, coded and tested by the author to assist in the StroMoHab system (and potentially other motion capture systems) configuration. As shown in Figure 1.1, the StroCap has two main functions: the first is used to design the StroMoHab system, and the second is used to simulate the whole system when it is connected to the system server. The full StroCap design and testing work will be described in detail in Chapter 2 and Chapter 3. The experiments reported in Chapter 6 did not require the use the StroCap system and motion capture, but did use an eye tracker (see Chapter 5).

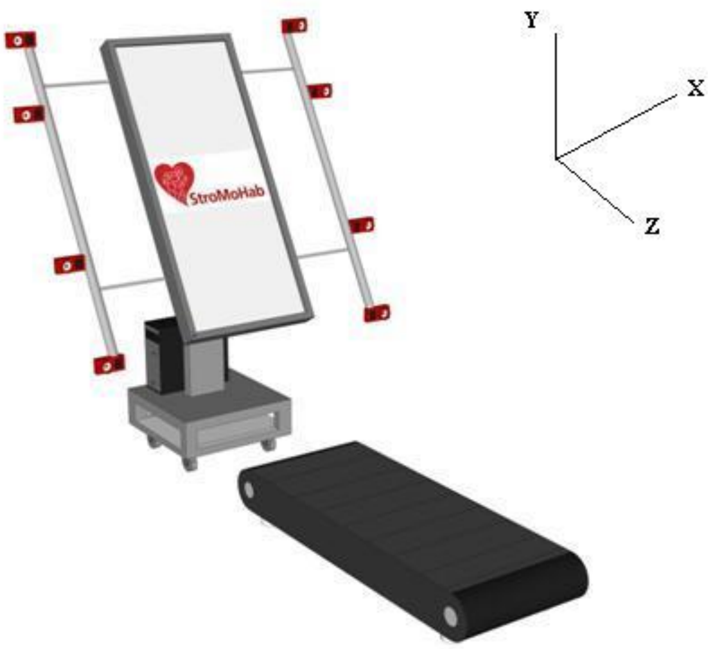

Figure 1.2 StroMoHab system design diagram. It contains eight motion capture cameras, a linear treadmill, a virtual reality screen with adjustable height, a PC and a trolley.

StroMoHab (see Figure 1.2) stands for the ‘Stroke Mobility Rehabilitation’ system, which was designed for people who have difficulty with their movement because of a brain hemorrhage such as stroke, but could also be applied to other movement and walking problems, for example in sports and Parkinson's disease. In the UK, stroke is a serious problem; it is a very common cause of severe disability and death. There are about 150,000 people in UK have a stroke every year. People who affected by stroke are from a young to an old age, defined as over 65 (e.g. Stroke Association, 2005). Stroke problems are well known and have existed for a long time, and patients benefit from specialized stroke care. Compared with the general hospital, the organized hospital stroke unit provides an 18% reduction in death by the benefits of stroke rehabilitation (NHS Greater Glasgow and Clyde, 2002). But there are still not many patients receiving the right level of, and enough treatment. An investigation shows that although 75% of trusts have a stroke unit, actually only about 25% of patients do rehabilitation in them (Rudd et al. 1999).

StroMoHab is a treadmill based system that combines motion capture technology with virtual reality to aid in the rehabilitation of stroke patients, but could also be applied as a locomotion simulator. The StroMoHab system consists of a computing speed treadmill, a server PC, six to twelve 'OptiTrack FLEX: V100' motion capture cameras (see Figure 2.19) and is provided with 'Arena Express' software for camera calibration. It also has a four wheel stand with adjustable height and a 50 inch 'Sony' plasma virtual reality display. By using this machine, people can receive visual and auditory feedback from their movements in real-time and can practice real world tasks. By using this machine, patients can practice on it with markers on their body, and their body's position and movements will be detected by the motion capture system via the cameras on the frames, and then the system gives them remedial suggestions on a virtual reality screen to help them move correctly.

Systems similar to StroMoHab have been used to study rehabilitation for stroke. For example Fung et al. (2006) created a simulator with self-paced treadmill for the locomotion training. Their experiments compared the walking temporal and distance measurements of gait when subjects are walking on the treadmill with different VE scenarios. Results showed that it is technically feasible to use a locomotion simulator for locomotor training, and stroke patients' locomotor capacity can be improved progressively by locomotor training with increasingly complex VEs. These findings echo Jaffe et al. (2004) who used the task of stepping over obstacles to improve stroke patients gait parameters. Twenty subjects participated in this experiment; experimenters measured and compared their gait velocity, step length, obstacle clearance capacity and walking endurance before and after two weeks of training. Subjects were instructed to step over obstacles either in a virtual environment when walking on the treadmill or on a 10m walkway. Results showed that through two weeks training of stepping over obstacles in VE, it appeared to have a significant improvement in gait velocity, stride length, walking endurance and ability to step over objects. These results provided evidence that obstacle training has clinical effectiveness for improving gait velocity post-stroke, and also demonstrated that the virtual obstacle training can enhance clinical performance.

Currently research facts shows there is no substitute for well trained physiotherapists to facilitate effective rehabilitation, but the rehabilitation process of a stroke patient can be improved.

1.3.1 Motion Capture

Motion capture is also named ‘motion tracking’, or ‘mocap’, and is used to record an object’s movements in three dimensions and translate that movement to a digital model. Motion capture has been used to animate characters on computers since the 1970s (e.g. Sturman 1994), initially used as a photogrammetric analysis tool in biomechanics research, and then expanded to other areas, such as gait analysis (e.g. Saboune & Charpillet, 2005; Simon, 2004), military, entertainment and computer animation.

Technically, motion capture can be classified into three broad categories: mechanical motion capture, the electromagnetic motion capture and the optical motion capture (e.g. Carranza et al. 2003). In mechanical motion capture technology, performers have to wear skeleton structured body suits to track human joints and angles, which are known as ‘exo-skeleton motion capture system’s due to the way the sensors are attached to the body (e.g. WorldPress.com, n.d.). An electromagnetic motion capture system produces low-frequency electromagnetic fields, and the actor / actress wear several magnetic receivers to pick up and calculate the movement (e.g. Metamotion, n.d.). Optical systems utilize at least one camera to provide overlapping projections to calibrate a subject, and then triangulate the subject’s 3D position data. Optical motion capture includes active optical systems, passive optical systems, semi-passive optical systems and markerless optical systems (Guerra-Filho, 2005). Active optical system triangulates the position of markers by using LEDs to generate their own light (e.g. CodaMotion, 2012; Qualisys, 2011). The active markers are identified automatically by the motion capture cameras and make the real time capture and display possible. Different to active LED markers, semi-passive markers use photosensitive marker tags to analyze the optical signal. The semi-passive markers rely on tags with photo sensors to be detected (e.g. Moeslund & Granum, 2001) which in order to track their locations and orientations. The markerless motion capture system does not require subjects to wear any equipments or devices for tracking, but some special computer algorithms are used in markerless technology to analyze optical input streams and identify the human subject's position and orientation, breaking them down into constituent parts for tracking (e.g. Computer Stroies, 2011)

The StroMoHab system uses a passive optical system. The performer has to wear with retro reflective material markers to reflect the camera light back in order to be sampled (see Figure 1.3), and his motion will be detected and identified by the position or angles between the markers as he is moving. Acoustic, inertial, LED, magnetic or reflective markers, or combinations of any of these, are tracked, optimally at least two times the frequency rate of the desired motion, to submillimeter positions (e.g. Hees, 2006).

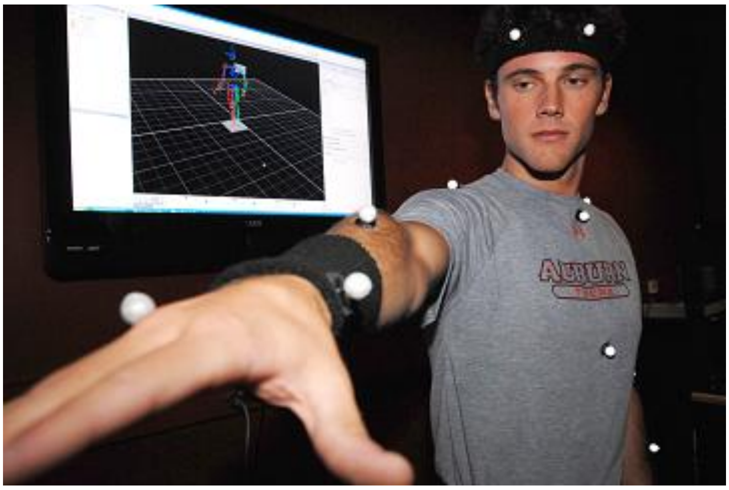

Figure 1.3 Performer wears retro reflective material markers on his hand, wrist, arms, shoulder, head, neck and chest etc, and a full body avatar in 3D space was generated on the screen behind him (taken from Auburn University, 2012).

During motion capture, when the performer is moving with markers on his body, the motion capture camera will find the skeleton dimensions and the markers' positions on the body, and then his movements will be recorded and sampled rapidly by calculating markers' positions from camera images. Then this data is mapped to a 3D model as an avatar by computing joint angles from marker paths, and this avatar will perform the same actions as the performer (see Figure 1.4).

Motion capture systems can also capture movements of a working virtual camera and the performance of actors. With such a technology, a computer is able to generate video images and characters as video camera. The computer directly displays the actors' movements from the desired position of a virtual camera. Retroactively obtaining camera movement data from the captured footage is known as match moving or camera tracking (e.g.Hees, 2006).

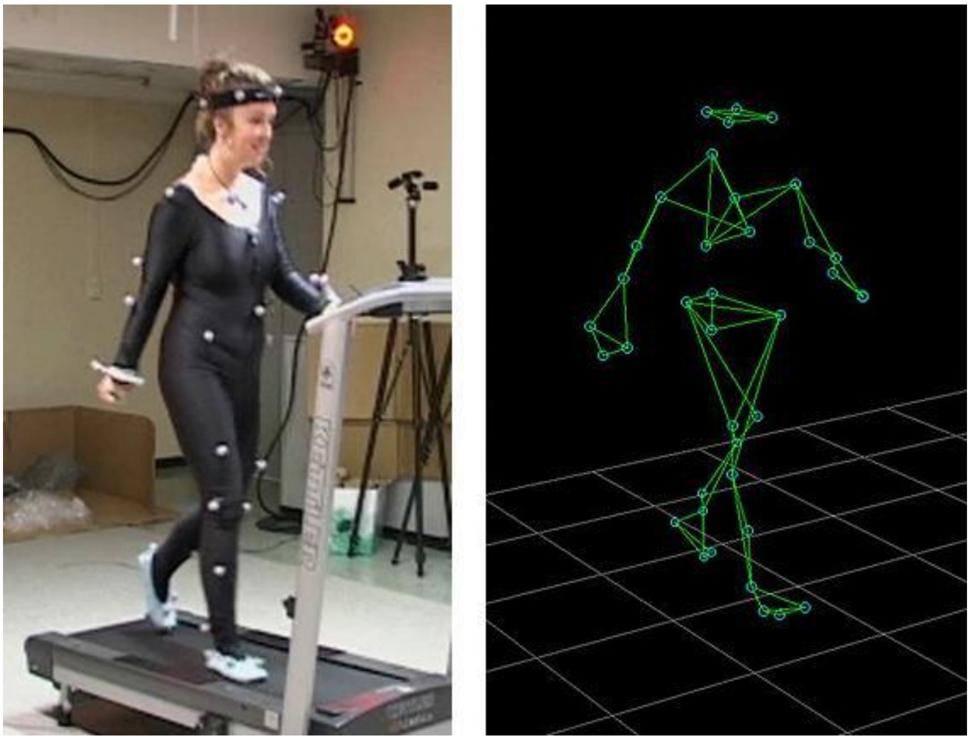

Figure 1.4 When a subject is walking in a locomotion simulator with markers coated (left), the full body avatar is generated in real time with the same locomotion (right) in a 3D space (taken fromWorldPress.com, n.d.).

The current world leading motion capture systems are Vicon, OptiTrack, and also some game consoles such as Wii and Xbox 360 Kinect. Vicon systems are supplied by Vicon, UK. The system includes Vicon MX, Vicon Bonita, ViconMotus and Vicon Cara (Vicon 2008). Richards (1999) reviewed and compared different motion capture systems and stated that Vicon system is better than many other systems due to its small size and less time used for motion tracking and data collection. The OptiTrack system (NaturalPoint, 2012) is made by NaturalPoint, and can be used for whole body motion capture. It mainly uses the motion capture cameras, Arena Express motion capture software and cables. In this thesis, StroMoHab used the OptiTrack system to track subjects' locomotion on the lower-body. In the gaming aspects, Wii and Kinect also use motion capture systems to generate the avatar on the screen and perform the same actions as user. Wii is made by Nintendo, and contains a sensor bar, a console and Wii remoter (Nintendo, 2010). The sensor bar is put close to the TV; it has five infrared LEDs on each left and right side. The remote control contains an embedded optical sensor, and a user holds the remote control during the game, which communicates with the sensor bar to calculate the users distance to the TV, and the users' moving direction, velocity and distance. The Kinect is a totally controller free motion capture gaming system made by Microsoft (XBOX, 2012). It has three cameras, the camera in the middle is a RGB colour camera; the left side is an infrared emitter and the right side is an infrared CMOS camera which is used as a depth sensor to capture video data in 3D.

1.3.2 Gait Analysis

The StroMoHab system uses motion capture technology for the purpose of gait analysis. Gait analysis is also referred to as motion analysis or biomechanics (e.g. Ariel Dynamics, 1999) as applied to walking by using cameras. It is the major application of motion capture in clinical medicine (e.g.Hees, 2006) used to help characterize most gait pathologies (e.g. Bogey, 2009).

In the StroMoHab system, there are a total of six infra-red cameras around a treadmill, and they are linked and can be controlled by a computer. Passive markers are attached to anatomical landmarks, such as on the pelvis, ankle and knee. As a patient walking on the treadmill, markers reflect cameras light and then the data of the positions is sent to the computer to calculate the trajectory of each marker in three dimensions.

The position of joint centres of rotation (e.g. Bogey, 2009) can be estimated by the location of markers on the patient's skin, and the motion about the joint centres is recorded. At least 3 markers are required to capture the motion of each limb, and each marker's position is recorded using multiple cameras.

1.4 Eye Movements

Eye movement is the voluntary or involuntary behaviour of the eyes, which is one of the most important features of locomotion as it is for other human activities. There is a large body of research on eye movements, on using a locomotion simulator and on human locomotion perception, however, eye movement analysis has not been previously applied to the investigation of human perception in virtual locomotion environments (please see Chapter 5 for more information on eye movements and virtual locomotion environments).

1.5 StroMoHab System in this Study

The StroMoHab system is used in this study aimed at improving the motion capture simulation program, and also works as a platform for addressing the central questions of this thesis, which include the conditions and problems with locomotion simulators in how to optimize them to make them comparable to the real environment.

However, when configuring the StroMoHab system, there has usually been no systematic way of defining where the cameras are and showing its capture volume. The only way to setup StroMoHab and similar systems is through a ‘trial & error’ process by using the existing motion capture software. Therefore, the author improved the motion capture simulator in this study by writing a convenient software called StroCap, which stands for the StroMoHab Motion Capture software, and which can represent each camera’s location, capture volume, overlapping and rotation angles when capturing the required space. The design objectives, algorithms and methods that were developed in this study for the above purpose are described in Chapter 2, and the software is also included in the attached DVD.

The overall goal for the simulator is to make the locomotion simulation environment as similar to a real environment as possible; therefore, optimising the StroMoHab simulator system is one of the important aims. Because eye movement are present during natural locomotion, in this thesis we evaluate StroMoHab by comparing locomotion eye movements in a simulated environment and a real environment. Through investigation, we want to understand the differences of eye movement patterns between the two environments, and also aim to find the best conditions to use the StroMoHab simulator.

Chapter 2. StroMoHab Motion Capture Simulation Software

This chapter focuses on the design of a motion capture simulation system for StroMoHab which would help to set up the system for efficient motion capture, called StroCap. We start to look at the actual work in creating applications by showing the design motivation, methods and mathematics, and all of the design ideas and procedures will be talked about in detail.

2.1 Motivation and Application

When configuring the StroMoHab system, it is very important to decide how to place each component in a defined space, such as the treadmill and cameras. It is understood by motion capture practitioners that fewer than three cameras capturing each marker can be unreliable (e.g. Kitagawa & Windsor 2008), therefore in order to make each marker able to be seen by at least three cameras all the time in motion capture, it is necessary to work out each camera's details in space, such as facing direction, capture volume, rotations, overlapping and coordinates. Therefore, we need software which can provide the camera's details when building StroMoHab in a required space; in order to make the setup procedures accurate and efficient, without repeated 'trial & error'. However, the current existing software could not provide all the needed information, which makes the setup work for StroMoHab system very inconvenient. In order to solve this problem, StroCap was written to help with the StroMoHab simulator setup, providing such features as 1) Calculate the volumes of each camera; 2) Calculate and represent the overlap of the volumes of the individual cameras; 3) Present the simulated StroMoHab system in graphical form. Compared with the existing software (Arena Express), StroCap has improved the design and configuration the motion capture in StroMoHab a locomotion simulator.

2.2 Functional Specifications

The purpose of this software was to simulate the motion capture space in the situation that is a new environment for using StroMoHab. This software has two main functions, which are 'Design' function and 'Synchronizing Simulation' function. The 'Design' function is used to design a simulated StroMoHab system in a 3D space, and then is applied to the real system setup according to the design. In this mode, a user can create the virtual room, the treadmill, some cameras and markers (the number of cameras and markers can be selected depending on needed), and then manually place each component based on the design requirement. For any camera's translations (see section 2.3.1) and rotations (see section 2.3.3), the StroCap software will provide each camera's coordinates and rotation angles in real time, and this software also used 'Collision Detection' (see section 2.3.4) methods to measure either the cameras' volume overlaps or the markers captured by motion capture cameras. Also in StroCap, we use different colours to represent the floor or markers that captured by different number of cameras. For example, if the marker is captured by two cameras, the colour of the marker will be presented in yellow (see Figure 2.9) in StroCap, and it will become green in the simulator if the marker was seen by three cameras. Once the design meets the requirements, the StroMoHab system then can be built based on the design.

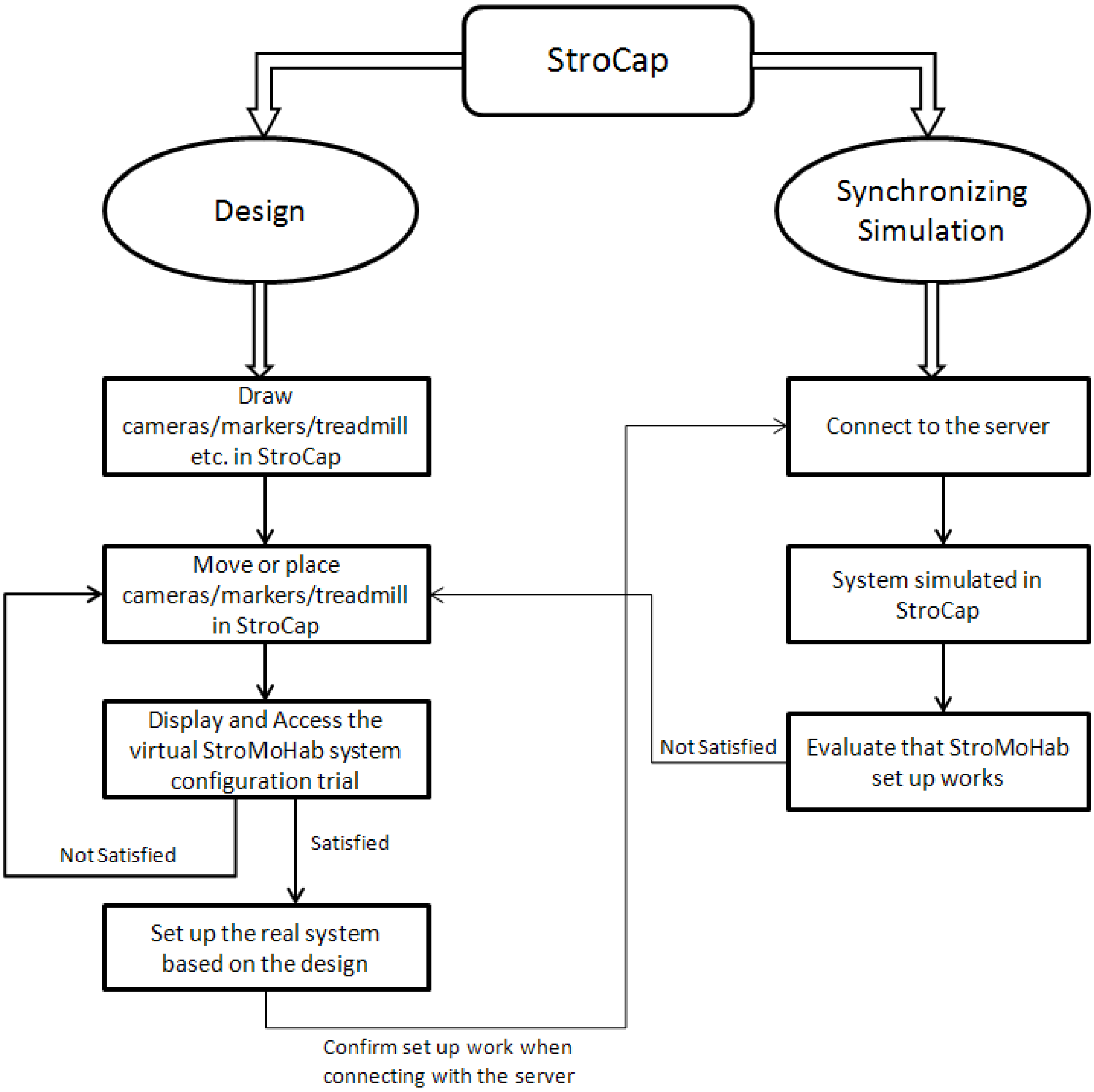

Figure 2.1 A flow-chart of the functionality of the StroCap software. StroCap contains separate 'Design' and 'Synchronizing simulation' functions to help with StroMoHab system configuration. The StroMoHab system can be initially drawn in the StroCap as a trial, without

necessarily relocating the physical cameras, and can then be accessed and tested until design requirements are satisfied.

When a person walks on the StroMoHab simulator with coated markers, the StroCap ‘Synchronizing Simulation’ function can be used to simulate the whole locomotion process in real time via the TCP/IP server, and this is also a method to evaluate if StroMoHab is built properly, such as if the whole treadmill is covered by each camera, or if every marker is captured by at least three cameras at the same time.

StroCap needs to receive and transmit data when simulating the StroMoHab system (see Figure 2.2). The server machine connects to the StroMoHab system via a TCP connection, it receives the cameras’ data and markers’ data and then it sends this data to the client (StroCap) if required. Then, the simulation software will represent the cameras and markers in real time according to this data.

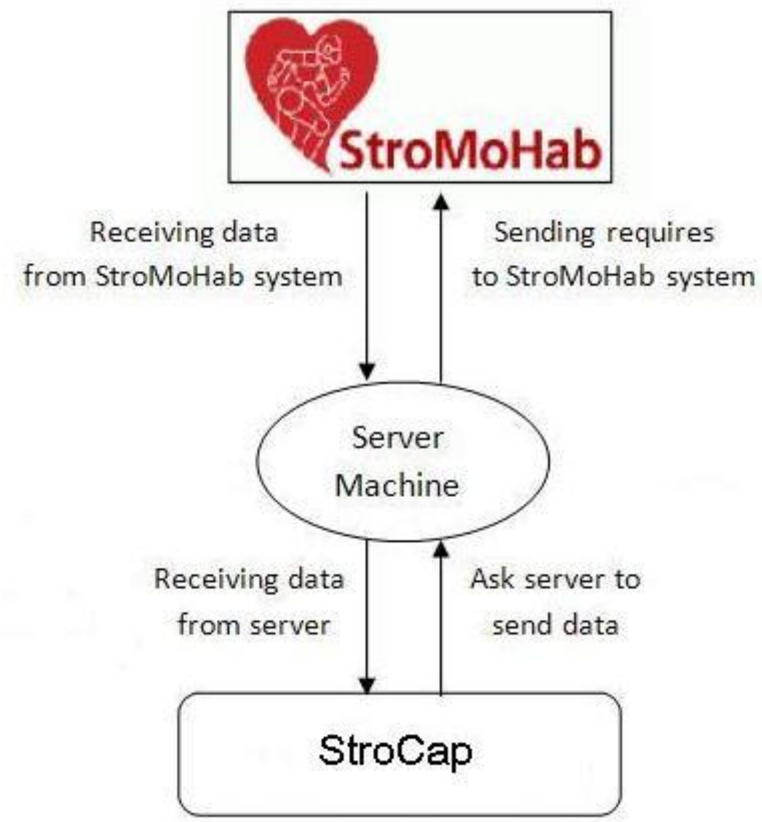

Figure 2.2 Sending and receiving data between the StroMoHab server and StroCap software.

2.3 General Methods and Mathematics

When creating the StroCap program, the most difficult part is to simulate the camera in both 'Design' and 'Synchronizing Simulation' functions, such as to generate a series of cameras, camera's translation and rotation, cameras' overlap and capture space after camera's position has been changed. Here, this section will introduce some main programming and mathematical methods when writing the StroCap software.

2.3.1 Camera Translation

When a camera moves to a new position, we can say the camera is translated. The camera will have new coordinates in the StroCap3D space, and also, the camera's view will be changed and will capture a new space (see Figure 2.3).

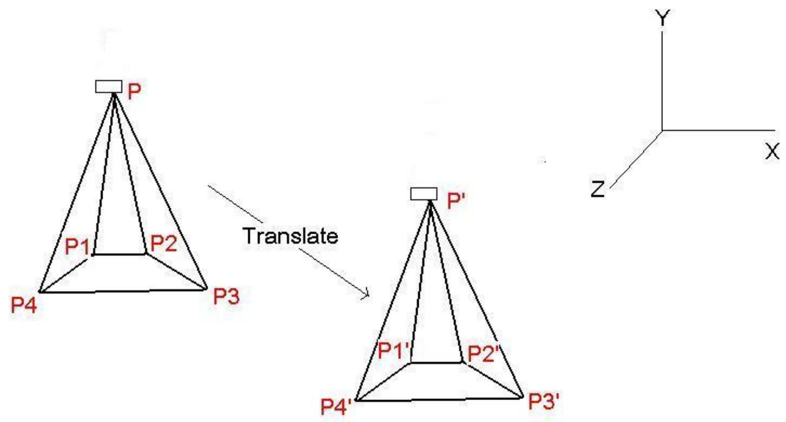

Figure 2.3 Camera translates to a new position in 3D space, and its view will be translated as well. Where P, P_1, P_2, P_3 and P_4 are the coordinates to represent camera and camera view's previous position, P', P_1', P_2' P_3' and $P_4$are the coordinates to show the new position of the camera and the view of the camera after translation in the 3D space.

In Figure 2.3, assuming the camera was initially at point P, the camera's view is at point P_1, P_2, P_3 and P_4. After translating, the camera has now moved to a new position P', and its view is now changed to point P_1', P_2' P_3' and P_4', each point has a (x, y, z) coordinates to represent its position in the 3D space. Once the translation distance is known (set as Tx, Ty, Tz), to find the new position of the camera:

For point P to P'

$$x' = x - Tx$$

$$y' = y - Ty$$

$$z' = z - Tz$$

For point P1 to P1'

$$x_{1'} = x_1 - T_x$$

$$y_{1'} = y_1 - T_y$$

$$z_{1'} = z_1 - T_z$$

For point P2 to P2'

$$x_{2'} = x_2 - T_x$$

$$y_{2'} = y_2 - T_y$$

$$z_{2'} = z_2 - T_z$$

For point P3 to P3'

$$x_{3'} = x_3 - T_x$$

$$y_{3'} = y_3 - T_y$$

$$z_{3'} = z_3 - T_z$$

For point P4 to P4'

$$x_{4'} = x_4 - T_x$$

$$y_{4'} = y_4 - T_y$$

$$z_{4'} = z_4 - T_z$$

Equation 2.1 All the equations above are used to represent how a camera translates from one place to another. By using these equations, the camera's new position can be calculated when cameras are translated.

Algorithm for simple camera translation:

```
if <translation key pressed> then
begin
    get axis (x, y, or z) of translation
    translation distance <= translation distance + 2
    display the camera in the new position
end
```

2.3.2 Convert Rotation Matrix to the Axis Angle

When a camera has been rotated by a known angle, simulating this rotation in the virtual environment is one of goals of this project. To do so, the real camera's current angle in three dimensions is required. However, the motion capture camera in StroMoHab sends its rotation angle data only in a series of matrix numbers (see Equation 2.2); therefore, we need to convert these matrix numbers to the normal axis angles before we can use this data to simulate the camera in the virtual environment.

$$R = \mathrm{r}_{11}\mathrm{r}_{12}\mathrm{r}_{13}\mathrm{r}_{21}\mathrm{r}_{22}\mathrm{r}_{23}\mathrm{r}_{31}\mathrm{r}_{32}\mathrm{r}_{33}$$

Equation 2.2 This is the format of the series of matrix numbers which are sent from each motion capture camera to StroCap, and this series of matrix number has to be re-calculated by using the rotation matrix equations in order to convert to the rotation angle.

A general 3×3 rotation matrix and some math equations (see Equation 2.3) can be applied here to convert the matrix numbers.

$$\cos\theta = \frac{1}{2}(\mathrm{r}11 + \mathrm{r}22 + \mathrm{r}33 - 1)$$

$$\sin\theta = \frac{1}{2}[(\mathrm{r}11 - \mathrm{r}12)^2 + (\mathrm{r}13 - \mathrm{r}31)^2 + (\mathrm{r}32 - \mathrm{r}23)^2]$$

$$\mathrm{R_x} = (\mathrm{r}21 - \mathrm{r}12)/(2 * \sin\theta) = \text{Rotation angle on X axis}$$

$$\mathrm{R_y} = (\mathrm{r}13 - \mathrm{r}31)/(2 * \sin\theta) = \text{Rotation angle on Y axis}$$

$$\mathrm{R_z} = (\mathrm{r}32 - \mathrm{r}23)/(2 * \sin\theta) = \text{Rotation angle on Z axis}$$

Equation 2.3 The equations used to convert the series of matrix numbers to the normal 3×3 rotation matrix.

2.3.3 Camera Rotation

When a camera is rotated in 3D space, the camera view will be rotated as well by the rotation angles. To find the new camera view, we need to use the rotation matrix, in which the rotations can be represented by orthogonal matrices in Euclidean space (EuclideanSpace - building a 3D world, 1998).

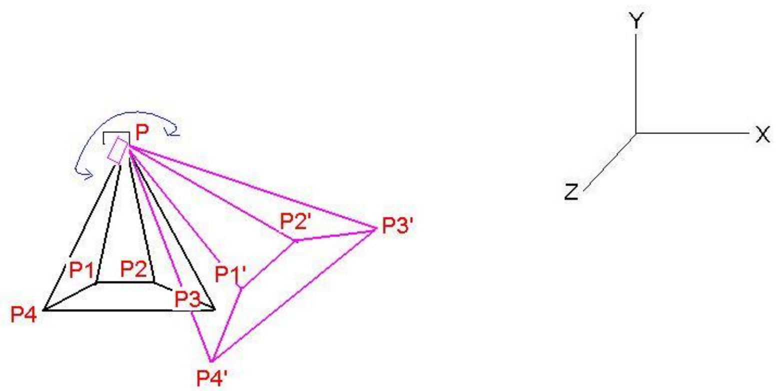

Figure 2.4 Camera view has rotated to a place in 3D space. Where P, P_1, P_2, P_3 and P_4 are the coordinates to represent camera and camera view's previous position; P_1', P_2' P_3' and P_4 are the coordinates to show the new camera view after the camera is rotated in the 3D space.

Rx --- The rotation angle when rotating the *y*-axis towards the *z*-axis

Ry --- The rotation angle when rotating the *z*-axis towards the *x*-axis

Rz --- The rotation angle when rotating the *x*-axis towards the *y*-axis

Assuming the camera has its initial coordinates (x, y, z).

If the camera is rotated on its X axis, the new coordinates of the camera view can be calculated by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cosRx & -sinRx \\ 0 & sinRx & cosRx \end{pmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

If the camera is rotated on its Y axis, the new coordinates of the camera view can be calculated by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} cosRy & 0 & sinRy \\ 0 & 1 & 0 \\ -sinRy & 0 & cosRy \end{pmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

If the camera is rotated on its Zaxis, the new coordinates of camera view can be calculated by:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} cosRz & -sinRz & 0 \\ sinRz & cosRz & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Equation 2.4 A rotation matrix is used to represent how a camera rotates from one place to another. By using these equations, we can get the new coordinates. Therefore, the new camera view after rotation can be calculated.

Camera rotation algorithm:

```
if <rotation key pressed> then

begin

     get axis (x, y or z) of rotation

     get the angle of rotation
```

```
    rotation angle <= rotation angle + 5 degrees
    calculate new position of camera


    Display the camera in new position
End
```

2.3.4 Collision Detection

Collision detection involves algorithms for checking the intersection of two of more objects.

Because we want markers to be captured by at least three cameras at all times in the StroMoHab simulation system, therefore, when designing this simulator in the StroCap software, we need to apply a method to calculate if there are intersection between objects (e.g. cameras, markers and floor). Collision detection is therefore used in StroCap to measure if the marker, floor or treadmill has been captured by a camera, and also use colour to represent when the markers, the floor, or the treadmill is captured by different numbers of cameras (see Figure 2.5 & Figure 2.6).

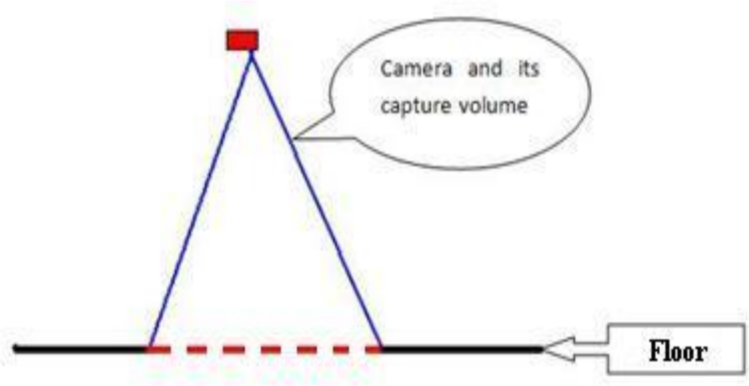

Figure 2.5 Using collision detection to represent the area on ground which captured by one camera. This is a side view of the camera and its capture volume. The red dash lines represent the area on the floor captured by that camera.

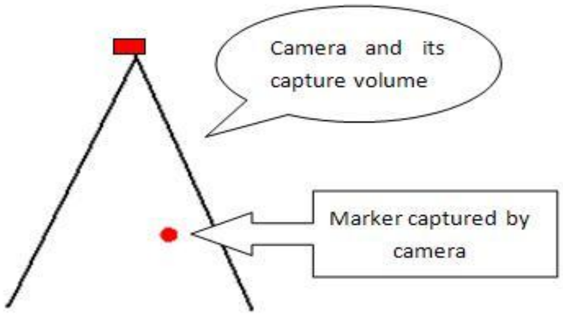

Figure 2.6 Using collision detection to represent the marker (red point) is captured by one camera (side view). When the marker is seen by only one camera, the colour of the marker in StroCap becomes red.

To use collision detection to test if the object has been seen by a camera, the basic principle is to first check the relationship of positions between the object and camera. For example, a camera is directly looking down to the ground:

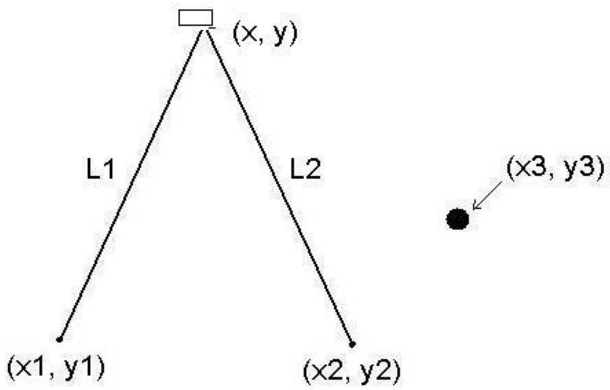

Figure 2.7 Camera view and marker's coordinates in two dimensions view. This is just a simple diagram to help with analysing the position relationship between marker and camera view.

In the collision detection algorithm, to check if the marker is inside of the camera view, it will compare the marker's coordinates and camera's coordinates. Marker's coordinate is (x3, y3), the equation for L1 and L2 are:

$$y_{1'} = k_1 * x_{1'} + C$$

$$y_{2'} = k_2 * x_{2'} + C$$

Where

$$y_1 = y_2$$

$$k_1 = (y - y_1) / (x - x_1)$$

$$k_2 = (y - y_2) / (x - x_2)$$

Equation 2.5 Equations to represent line L1 and line L2.

To test if the marker is inside, horizontally, its X coordinate must be in between the left and right edges of the camera view; vertically, its Y coordinate must be in between the top and bottom of the camera view. Therefore, the ‘Boolean statement’ below must be true in the program:

$$x - (y - y_3) / k_1 < x_3 < x - (y - y_3) / k_2$$

$$\&$$

$$y_1 < y_3 < y$$

Equation 2.6 Equations used to test if the marker is inside of the camera’s view

Collision detection algorithm:

```
if <marker was moved> then
Begin
     get new coordinates of marker
     if <X_camera-min < X_marker < X_camera-max>
     And<Y_camera-min < Y_marker < Y_camera-max>
     And <Z_camera-min < Z_marker < Z_camera-max>
     then
          marker has seen by camera
     display marker in different colour
End
```

Once marker's coordinates reach this condition, it means the marker is inside of the camera view, and also means the marker is captured by that camera, and the marker's colour changes to show the success of the collision detection. By using this method, StroCap software successfully represented the capture space overlap on the ground or marker (see Figure 2.8 & Figure 2.9).

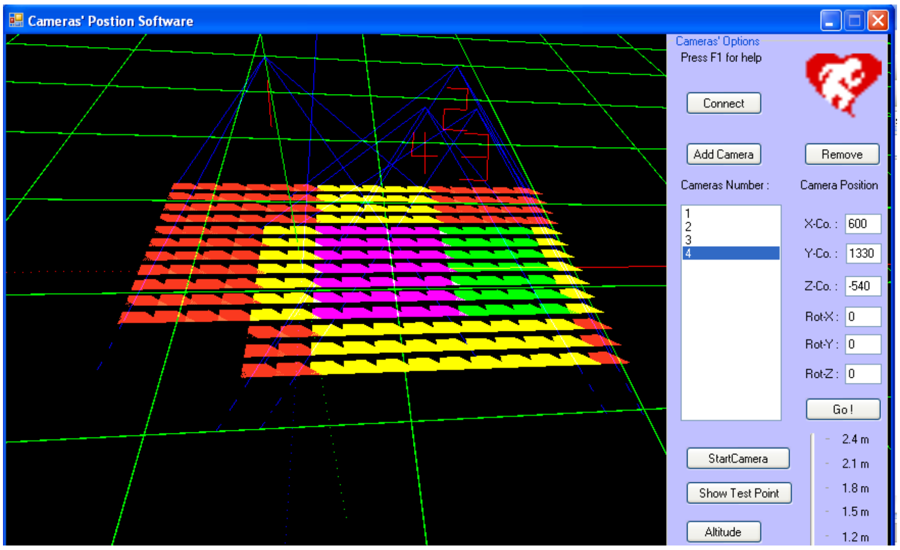

Figure 2.8 The area on the ground which has been captured by four cameras. The red area represents the region that has been captured by just one camera; yellow means that region is seen by two cameras; green means three cameras can see that area; purple represents that the area has been seen by four cameras.

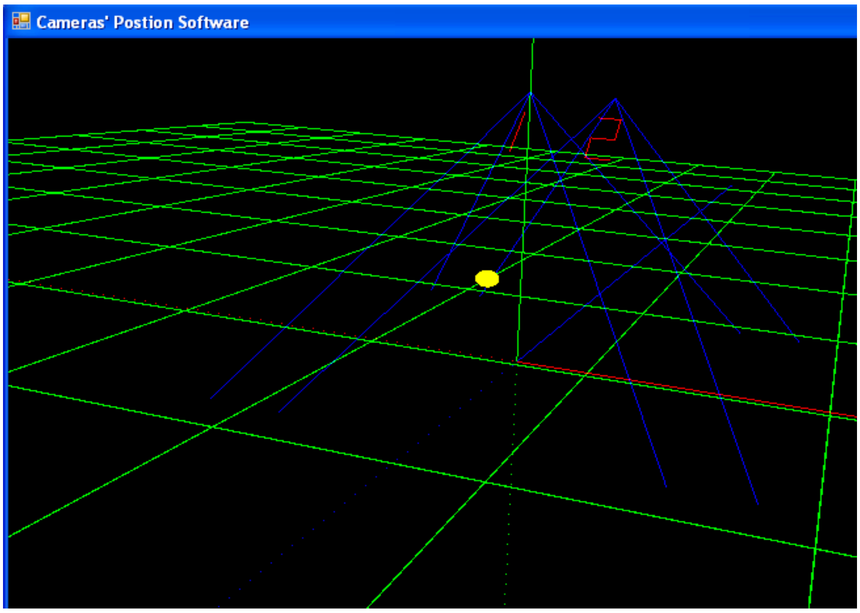

Figure 2.9 Marker is seen by two cameras at the same time in the StroCap, which was represented by using the yellow colour.

The camera capture area at different heights

The motion capture for stroke patients in the StroMoHab system focuses more on the lower body, which means that the patient will put markers on their feet, ankles, knee and waist. Each of these markers has to be seen by at least three cameras when the patient is doing the practice. Therefore, it is very important to find which place on treadmill is the best practice region for a patient, which allows every marker at different heights to be captured by the cameras.

The collision detection method is the best solution to solve this problem. By using this, it will be easy to find the camera's capture region when a marker is at a different height. The range of the height is from the 0m (ground) to 2.4m, because the room space was set as 3m×3m×3m. 2.4m is the reasonable highest distance to the floor. In StroCap, we also created a tracking bar to adjust camera's height, so the user can scroll the tracking bar to find the camera's capture area when putting the marker at different heights (see Figure 2.10).

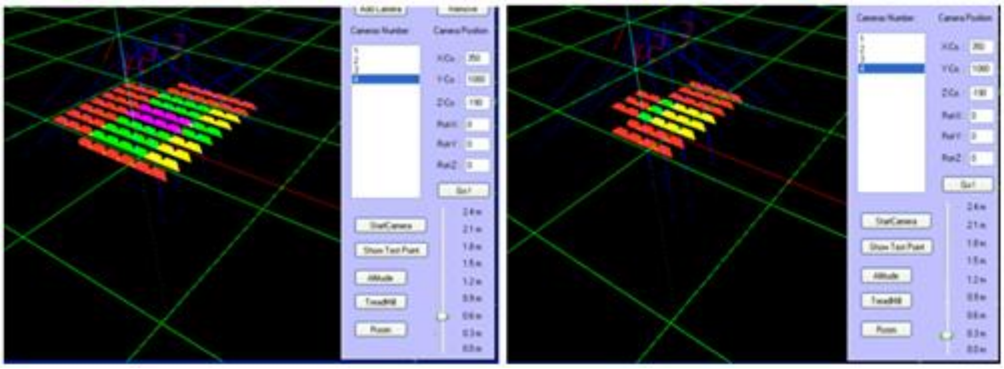

Figure 2.10 Capture space when camera height is 0.3m (right) or 0.6m (left). When a camera is closer to the floor, there is a smaller volume that it can capture, while a higher camera position would capture a larger volume. This figure shows that when cameras at 0.6m to the floor, the floor can be captured by four cameras at the same time, which is represented by four different colours, red, yellow, green and purple; but when the cameras are at 0.3m above the floor, only three cameras intersect on the floor, as shown in red, yellow and green. The colour scheme and graphical representation are intended to make it easier for the user to evaluate camera positions and corresponding capture volumes.

2.4 Camera's Capture Volume Measurement

Before we can create a camera in the StroCap and also represent its capture volume in 3D space, there are two figures we need to know. One is the camera's field of view, and another is the range of view the camera can see.

Theoretically, when you take a shot using a camera, the camera's capture volume is a four - sided pyramid, and its image plane is a perfect rectangle (see Figure 2.11).

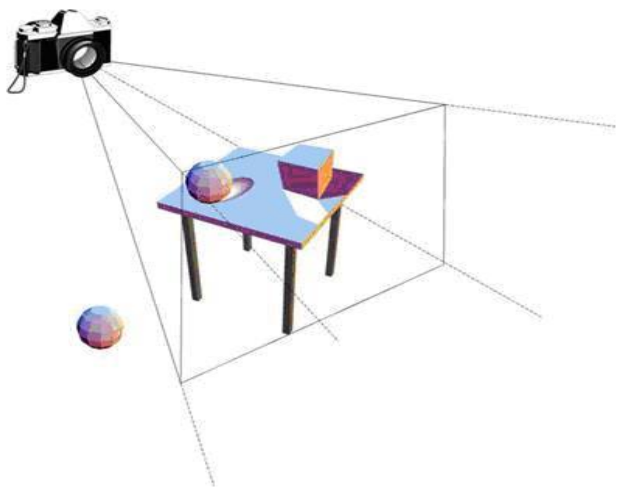

Figure 2.11 Camera's capture volume in theory (taken from PCBookCN, 2000)

However, theoretical data sometimes differs from the real world. Therefore, it is necessary to measure the camera properties when the camera is at various distances from the object, which will exactly define the view and volume a camera can capture, and also we can test if the image plane of camera (the theoretically perfect rectangular shape) is steady the same at different distances.

2.4.1 Camera Measurement Procedures

When initially designing the StroMoHab system, one of its potential purposes is for home use, in which stroke patients can do rehabilitation work at home. Therefore StroMoHab has to fit in a small place, a 3m×3m×3m space is the maximum, and the reasonable distance range of camera to people is between 50cm to 255cm.

In the camera measurement experiment, the camera was measured at five different distances (50cm, 100cm, 150cm, 200cm and 255cm, distance was measured by tape measure) to the wall. At each distance, the camera was fixed in a constant ground height and faced to the wall (see Figure 2.12). The camera was turned on and connected with the Arena Expression software, setting the Arena to the 'Calibration' mode. Then, put the marker on the wall to start measurement.

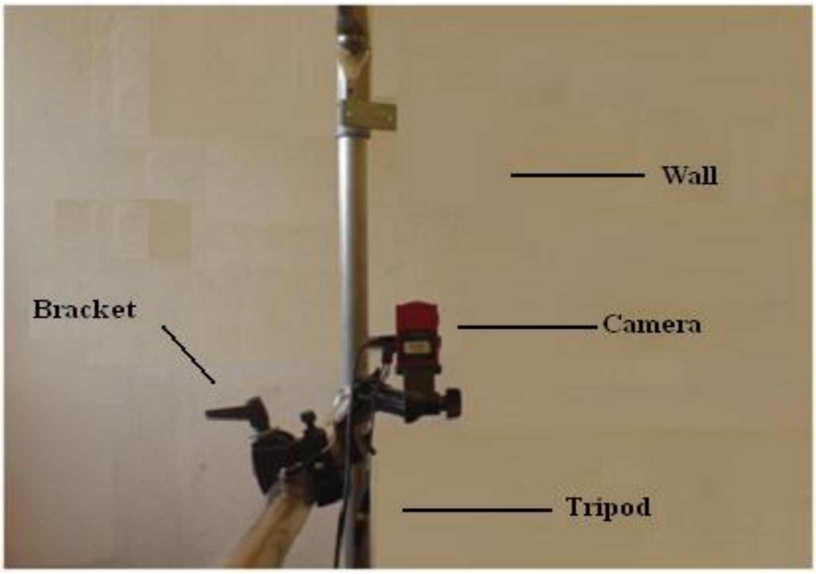

Figure 2.12 Camera was fixed on the bracket and mounted on the tripod. Camera faced to the wall at distance of either 50cm, 100cm, 150cm, 200cm or 255cm.

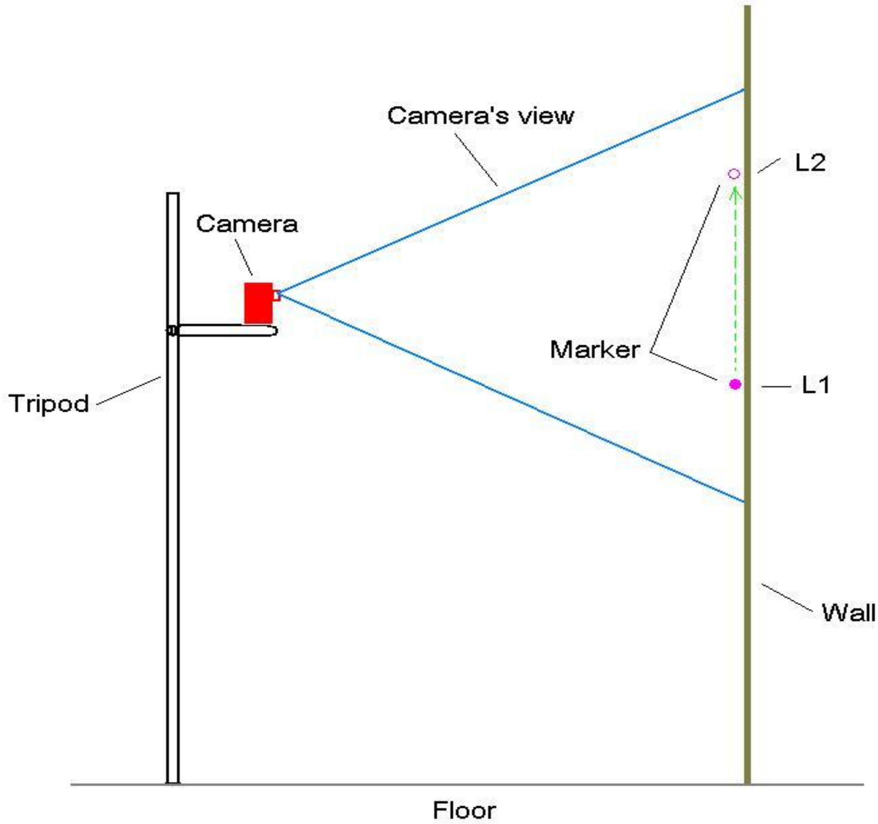

Figure 2.13 A schematic of the side view of the for camera calibration as shown in Figure 2.12. The camera was mounted on a tripod and faced to the wall; the camera's view is shown in blue. The marker was moved on the wall from location 1 (L1) to location 2 (L2) to measure the reliable region of the camera's view.

To measure the top limit, the marker was moved upwards (see Figure 2.12). As the marker was moving on the wall, its movement was simulated in Arena in real time. When moving the marker, the experimenter also needs to watch the Arena on the PC screen to see if the marker could be detected at that point. When the maker was at the top limit point on the wall, the marker in the Arena also reached the top of the window (see Figure 2.14). If the marker is moved upwards from this point, the marker in the Arena window will disappear, so it means that this was the upper limit point, and then mark this point on the wall.

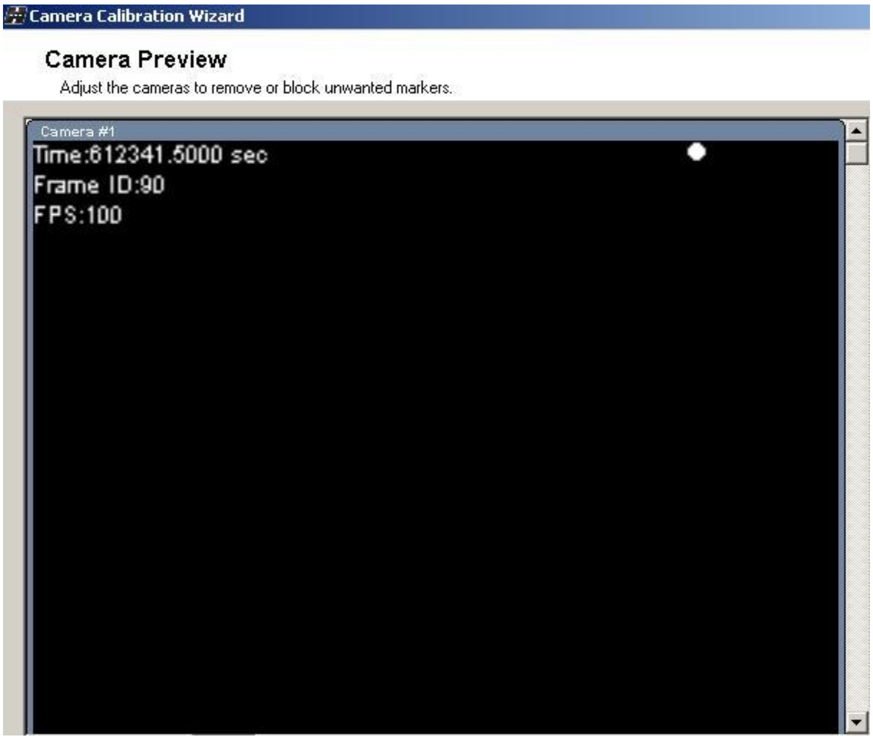

Figure 2.14 Marker reaches to the top limit of the camera view. The white dot represents the marker which is shown in the Arena calibration software.

By using the same method, other limit points can also be defined, such as in the bottom, left direction, right direction, left upper corner, right upper corner, etc.

2.4.2 Draw Camera Volume Shape Diagram

Through the measurements, it has been found that some areas at the edge of camera's view that could not be captured properly. This means those areas are in an unreliable region, markers in that region can be detected sometimes, but sometimes cannot. And also, the further the camera moved away from the wall, the larger the unreliable region occurs. One possible reason might be that when the marker is at the edge of the camera's view, the ability of the camera to receive the reflecting infrared is weaker than in other areas. Another reason might be if the marker is far away from the camera, the infrared reflection signal which transmitting from marker to camera will be distorted.

According to the experimental results, the shapes of camera's view were drawn in diagrams at five different distances and in three view modes.

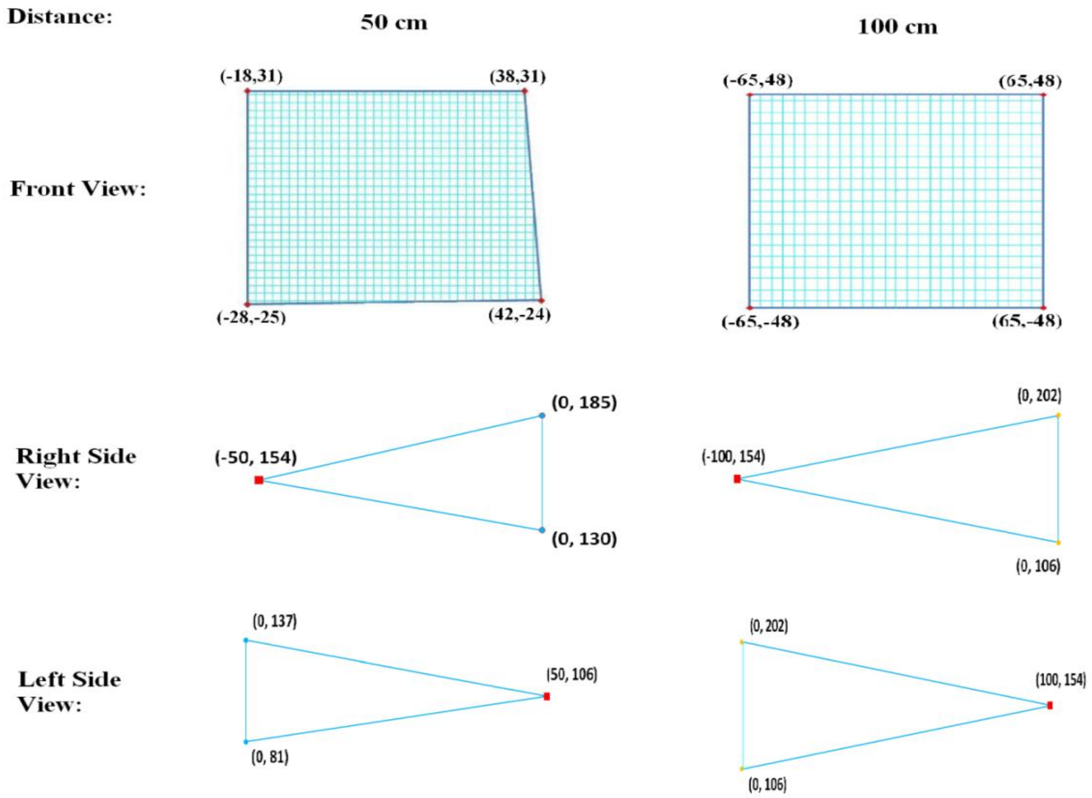

Figure 2.15 Diagrams of cameras' view shape were drawn from three view modes (front, right side and left side view) at different distances (50 and 100cm) to the wall. The coordinates on each diagram were used to show the limit of the camera view. The blue lines represent the reliable capture regions of the camera, which means markers can always been seen in these area.

The red lines are the unreliable regions, where markers in that region cannot be captured sometimes.

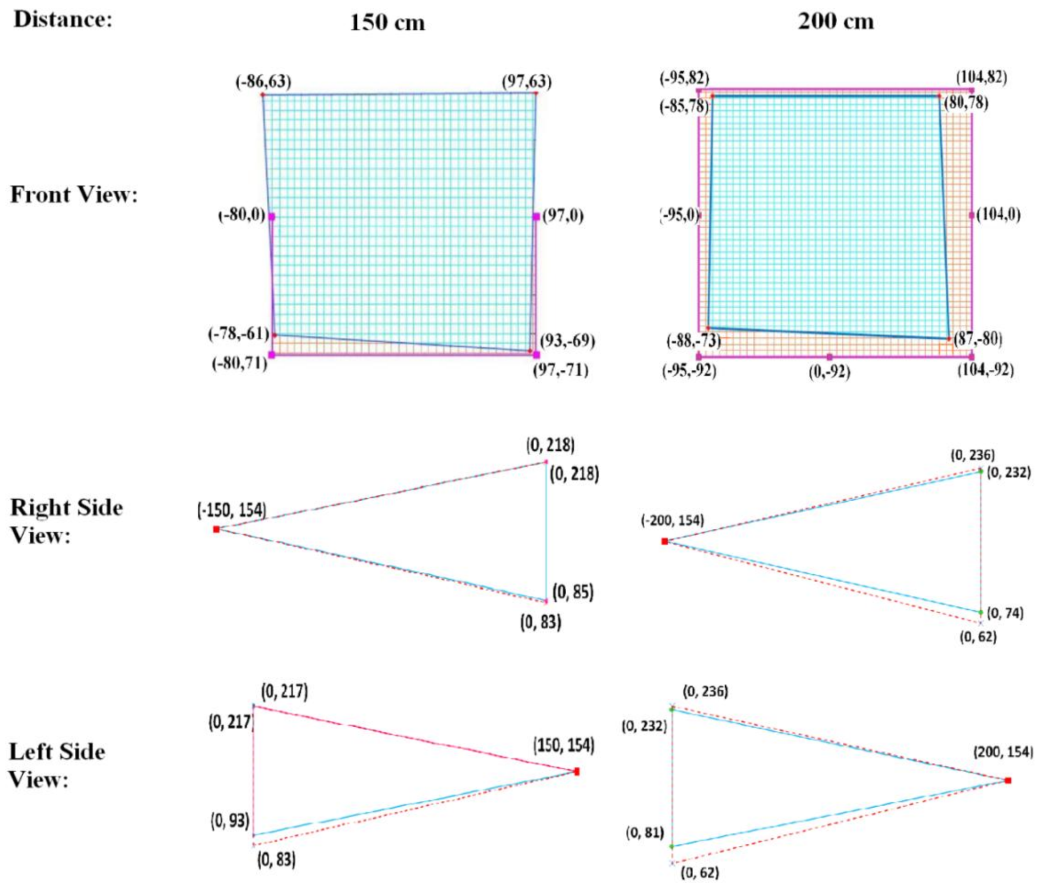

Figure 2.16 Diagrams of cameras' view shape were drawn from three view modes (front, right side and left side view) at different distances (150 and 200cm) to the wall. The coordinates on each diagram were used to show the limit of the camera view. The blue lines represent the reliable capture regions of the camera, which means markers can always been seen in these area. The red lines are the unreliable regions, where markers in that region cannot be captured sometimes.

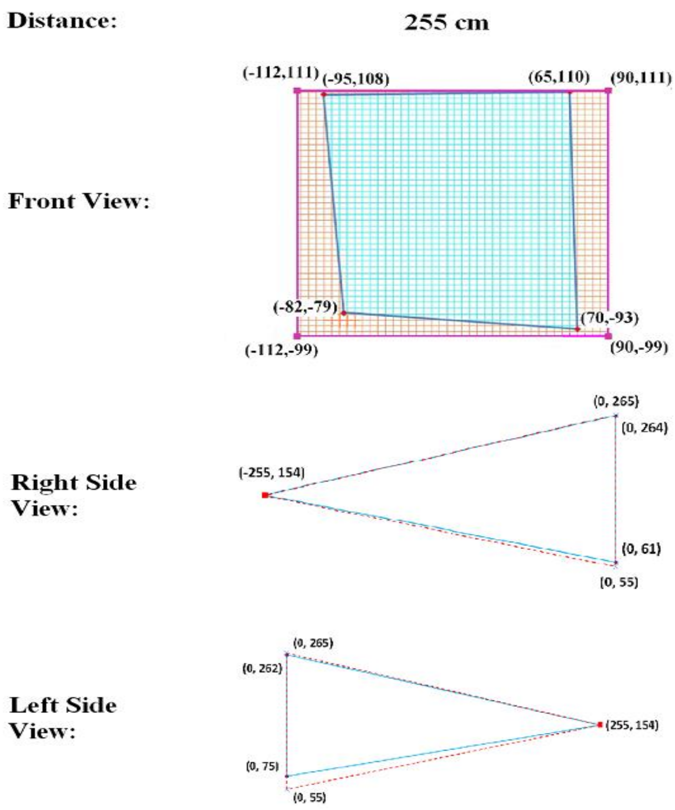

Figure 2.17 Diagrams of cameras' view shape were drawn from three view modes (front, right side and left side view) at distances of 255cm to the wall. The coordinates on each diagram were used to show the limit of the camera view. The blue lines represent the reliable capture regions of the camera, which means markers can always been seen in these area. The red lines are the unreliable regions, where markers in that region cannot be captured sometimes.

The shapes of the camera's view at five different distances were drawn in front view, right hand side view and left hand side view. In the 'Front View' diagrams, the regions in blue colour represent the reliable regions, and the regions in red means the unreliable regions, where markers can't be captured all the time. In the 'Right Side View' and 'Left Side View', the region inside of the blue lines is the reliable region, and the unreliable region exists in between the red dot line and the blue line.

Figure 2.15 - Figure 2.17 clearly show that there is an unreliable region existing in camera's view, and this unreliable region gets bigger as we move the camera backwards towards the wall. These results reveal that the shape of the camera's image plane is not a perfect rectangle all the time, and the reliable region of view will become a quadrilateral when moving the camera far away from its capturing object.

When designing the camera position and shape in the StroCap software, we want the camera view to be very accurate when detecting the marker, therefore, we only used the value of the reliable region to calculate the ratio (R) of the camera's distances and its horizontal or vertical reliable capture scope (see Figure 2.18).

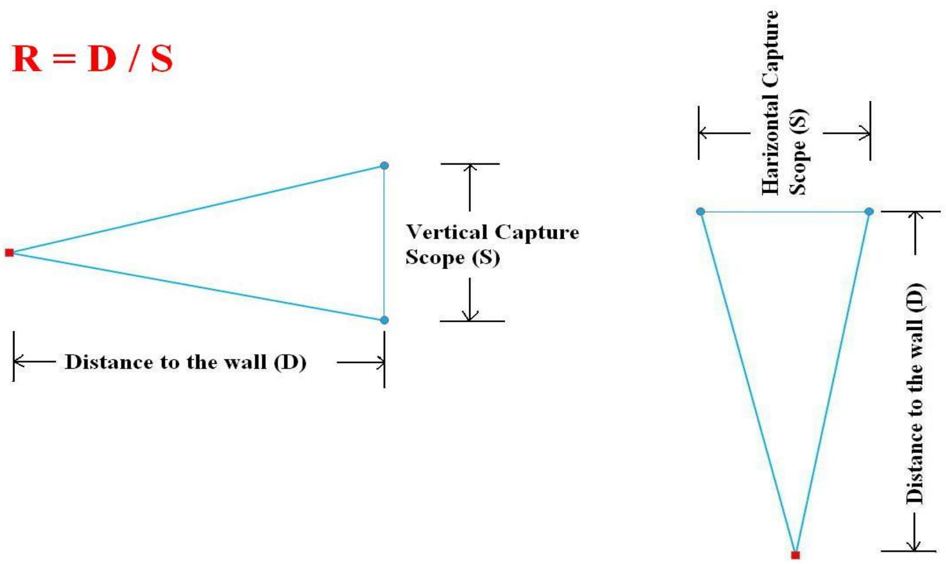

Figure 2.18 The methods for calculating the ratio of distance and capture scope in either vertical direction or horizontal direction.

The ratio of distance and capture scope was used to help with drawing the virtual cameras in StroCap program. By using this ratio value, we can design the shape of the camera volume, and show the camera capture volume at different camera heights. Because the StroMoHab system was suggested to be used in the room with maximum space 3m × 3m × 3m, the reasonable ratio value we selected when camera is at 255cm from the wall. In this case, the distance (D) from the camera to the wall is 255cm, the average reliable capturing scope in horizontal (S_1) is 156cm, and the average reliable capturing scope in vertical (S_2) is 195cm, hence, the horizontal ratio is 1.63, and the vertical ratio is 1.3.

Figure 2.19 NaturalPoint FLEX: V100 camera (taken from NaturalPoint, 2012) was selected to use in the StroMoHab system by its small size and effective image capture ability.

The camera we used in the StroMoHab system is the OptiTrack FLEX: V100 camera (see Figure 2.19), which is produced by the NaturalPoint Company, with image capture and processing integrated (NaturalPoint, 2012) function. The camera captures motion images 100 frames per second by receiving reflected infrared light source in a six meters' range.

The range of camera view is from 0.5cm to 600cm, in StroCap software (see Figure 2.20). When the virtual camera reaches its extreme distance, the shape of this camera is a pyramid, height is 600cm, and the image plane is a 461cm ×368cm rectangular.

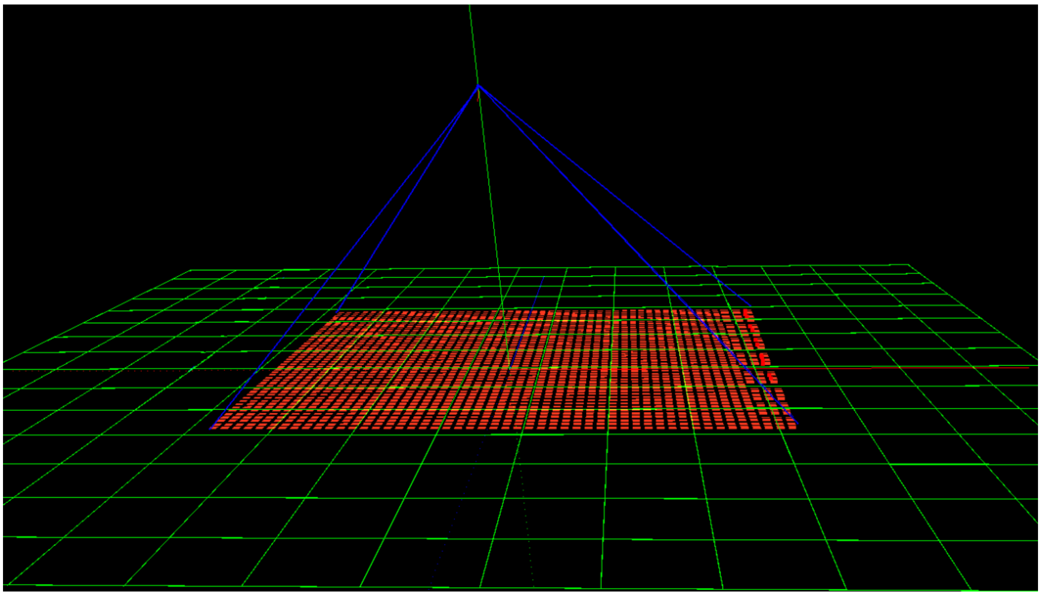

Figure 2.20 Camera's shape and its capture volume in StroCap space. The Green grids represent the floor, the pyramid shape in blue is the camera's view when camera height is 600cm in 3D space, and the red squares are the region (461cm × 368cm) of floor captured by the camera.

Chapter 3. StroCap Software Performance

This chapter describes the testing work for this software. Because this software is used to model the whole StroMoHab system, it is therefore necessary and very important to test its performance after the design and implementation. Some experiments were run and we also applied StroCap to a real case study to assess and evaluate the performance of the software. The results are analysed and discussed in this chapter.

3.1 Experiment Preparation

The experimental equipment was: Lego bricks, a tape measure, a marker, Opti-Track FLEX: V100 cameras and the StroCap software. During the experiment, all external light was blocked from the room to prevent infrared interference.

3.2 Marker Translation Test

The aim was to test how accurately StroCap responds when the marker is moved in the real world. The accuracy was assessed by translating an optical marker through small, medium and large distances. We used Lego to aid in the distance measurements, due to its degree of precision. The tolerance of each Lego brick is smaller than 10 micrometers (see Lego Company Profile, 2010). The size of each Lego brick is 8mm × 8mm × 9.6mm (L × W × H), and the spacing between each stud centre is 8mm (Orion Technologies, 2005). In this experiment, the small movement on the X or Z axis was defined by the distance between each stud centre, which is 8mm (see Figure 3.1); the medium movement is 8cm, which translates to 10 studs (see Figure 3.3). Due to the size of the Lego base plates, the maximum reasonable large movement on X or Z axis was 40cm, which translates to 50 studs. During the translation, the marker was attached to the Lego, (see Figure 3.1) and so for each movement we can just translate the Lego brick from one stud to another.

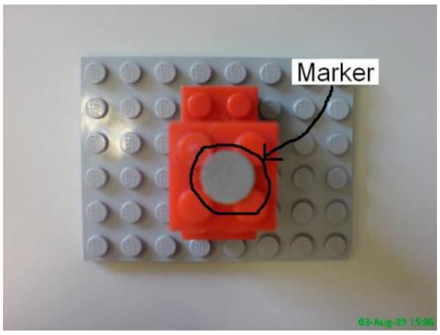

Figure 3.1 Lego with Marker

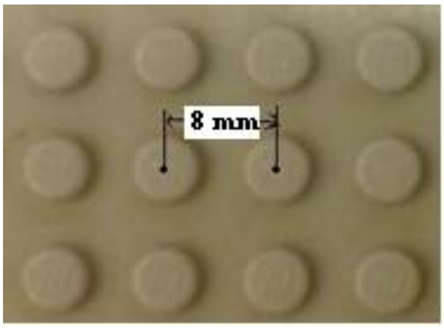

Figure 3.2 8mm between two points

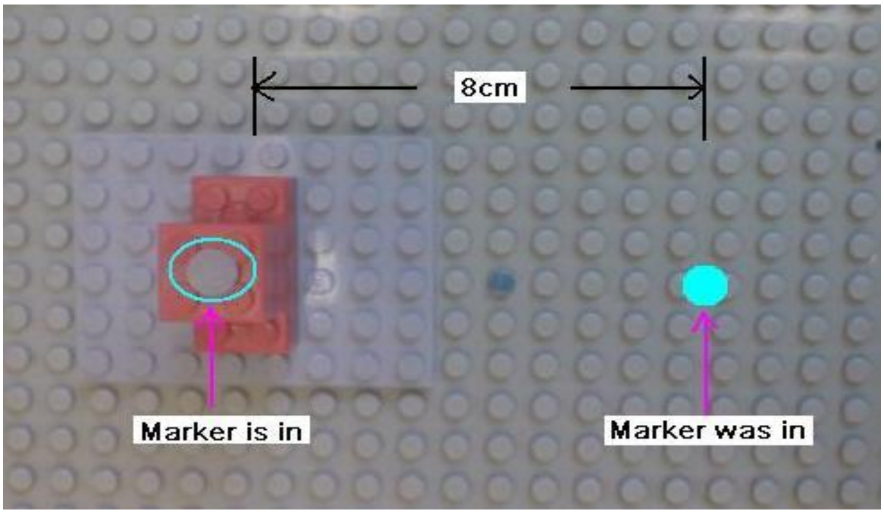

Figure 3.3 The medium distance movement on X, Z axis is 8cm. The blue point represents marker's previous position, and the marker's new location was circled in blue. This diagram shows marker's medium distance movement on X axis which translated the marker by ten studs.

On the Y axis, the small movement is 9.6mm where the marker was translated by just stacking the Lego bricks on top of each other (see Figure 3.4). The large movement was defined by the height of five bricks, which is 4.8cm (see Figure 3.5).

Figure 3.4 The height of one piece of Lego brick is 9.6mm, for each small movement (which is also 9.6mm) on Y axis; we just add one Lego brick onto another (taken from Maroues 2008-2012, edited by author).

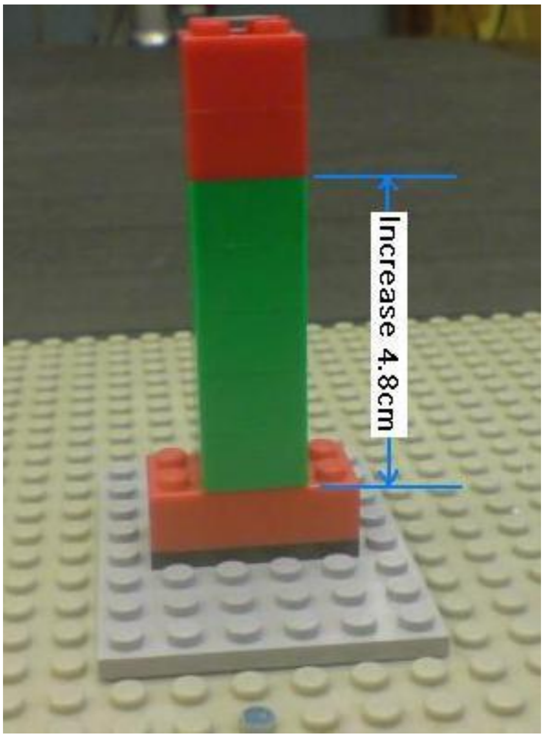

Figure 3.5 The medium movement on Y axis is 4.8cm, which was done by adding five Lego bricks onto the top of the previous brick.

In order to make sure the collected data was reliable we did each movement five times. After each manual translation in each dimension X, Y and Z (see Figure 1.2), the marker's new location data in the real world and StroCap software was recorded, and then we compared the average value of each movement in both environments (see Table 3.1).

	Small Movement (mm)					Medium Movement (mm)				
	Real World		StroCap		Est. Error %	Real World		StroCap		Est. Error %
Axis	Value	s.d.	Value	s.d.		Value	s.d.	Value	s.d.	
+X	8	± 0	7.88	± 0.299	1.50%	80	± 0	79.39	± 1.373	0.76%
-X	8	± 0	7.89	± 0.109	1.38%	80	± 0	79.31	± 0.399	0.86%
+Z	8	± 0	7.98	± 0.677	0.25%	80	± 0	80.19	± 1.799	0.24%
-Z	8	± 0	8.06	± 0.455	0.75%	80	± 0	80.53	± 0.942	0.66%
+Y	9.6	± 0	9.39	± 0.838	2.19%	N/A	N/A	N/A	N/A	N/A
-Y	9.6	± 0	9.72	± 0.725	1.25%	N/A	N/A	N/A	N/A	N/A

	Large Movement (mm)				
	Real World		StroCap		Est. Error %
Axis	Value	s.d.	Value	s.d.	
+X	400	± 0	390.28	± 2.398	2.43%
-X	400	± 0	390.85	± 1.984	2.29%
+Z	400	± 0	390.55	± 0.866	2.36%
-Z	400	± 0	390.25	± 0.420	2.44%
+Y	48	± 0	47.64	± 1.492	0.75%
-Y	48	± 0	47.72	± 2.157	0.58%

Table 3.1 The values of translations in real world and StroCap were shown for every small, medium and large movement, and also the percentage of estimate error was calculated and shown for each movement on each axis.

In Table 3.1, Est. Error % stands for the percentage of estimated error, which is defined by using the estimated error divided by the marker's movement in the real world. The estimated error is represented by the value differences between the real world movement and the StroCap movement.

For the small movements, it has the smallest Est. Error % when the marker is translating on the +Z axis ($F(5,24) = 9.417$, $p < 0.001$), which means that the camera has better performance when capturing the marker's motion on the +Z axis. For the medium movement, the camera has better performance to capture marker's movement also on the +Z axis ($F(3,15) = 1.138$, $p < 0.001$). For the large movement, the camera is more accurate in capturing the movements on -Y axis ($F(5,20) = 0.895$, $p < 0.001$). One reason for the estimated error may be that during the test, the reflected infrared source from the maker was distorted by other noises (e.g. room light), and this makes the camera confused about the marker's exact position.

An Independent-Samples T-Test shows that there is no significant difference between real world movement and the StroCap movement for all size of the movements in three dimensions ($p = 0.966$), and this result shows that the StroCap software is accurate and reliable when simulating the translation movements.

3.3 Camera Rotation Testing

This experiment aims to evaluate how accurate and reliable StroCap responds with the camera rotation in the real world; it was assessed by comparing the rotation angles in StroCap with the real world camera rotations on X axis, Y axis and Z axis. A scale was provided (see Figure 3.6) on each camera's mount, which shows the angle degrees the camera rotated.

Figure 3.6 A scale was provided on every camera's mount to help with recording camera's rotation angle.

When StroCap was connected to the server machine, the camera's initial position and rotation angles in three dimensions would be sent by request (see Figure 3.7).

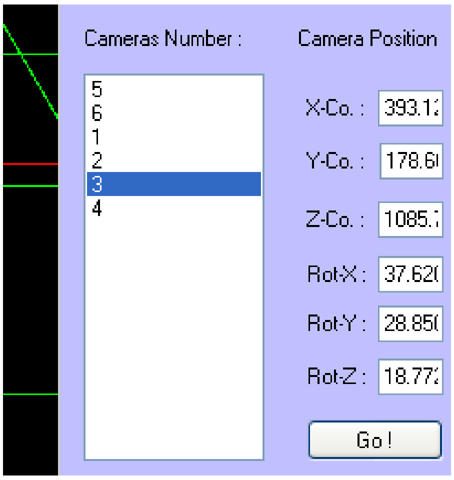

Figure 3.7 Location and rotation angles of camera 3. X-Co., Y-Co. and Z-Co. are the current coordinates of that camera, Rot-X, Rot-Y and Rot-Z are the current rotation angles of that camera.

In order to collect reliable rotation data, camera was rotated four times on each axis. After each manual rotation in each dimension, the marker's new rotation angles in the real world and StroCap software were recorded, and then compared to the average value of each rotation in both environments (see Table 3.2).

Rotation Angles	**Real World**		**StroCap**		**Est. Error %**
Axis	**Value**	**s.d.**	**Value**	**s.d.**	
X	15	± 0	15.84	± 3.686	5.58%
Z	15	± 0	14.59	± 2.268	2.76%
Y	15	± 0	15.45	± 0.805	3.02%

Table 3.2 This table listed the angle of camera rotated in real world and StroCap. The percentage of estimate error was calculated and shown for each rotation on three dimensions.

In Table 3.2, the percentage of estimate errors was defined by using the estimated error divided by the marker's movement in the real world. The estimated error is represented by the value of rotation angle differences between the real world movement and the StroCap movement.

The Est. Errors of camera rotation on each axis are similar ($F(2, 9) = 0.801$, $p = 0.478$). The errors of all rotation values between the real world and StroCap are all less than 1 degree. One possible reason for the error is that the rotation angle is not precise when rotating the camera manually in the real world.

An Independent-Samples T-Test shows that there is no significant difference between real world rotation angle and the StroCap rotation angle in three dimensions ($p > 0.677$). This result reveals that the StroCap software is accurate and reliable when simulating the camera rotations.

3.4 Case Study

StroMoHab system was presented at the Venturefest exhibition in February 2010, which was a good opportunity to show and popularize the system, and was also a good opportunity to test and use this simulation system to help with the StroMoHab system setup. The leader of the StroMoHab team, Dr Adar Pelah said "Venturefest is very important for us; we should think about how to perfectly optimize our system in a restricted space, it is a challenge".

3.4.1 Case Background and Problem

Venturefest is a 1 day expo held in York Racecourse, bringing together the finest science, technology and knowledge entrepreneurs in Yorkshire. Each exhibitor is allocated a certain amount of space as the exhibitor stands.

The host provided a 3m × 3m (Length × Width) square meter space for the StroMoHab system (see Figure 3.8). The whole system consisted of a treadmill (W: 71cm, H: 10cm, L: 175cm), one 50 inch's lifting controllable screen, one PC and six motion capture cameras. Besides the StroMoHab system, it also needed to provide enough room for staff and visitors walking around. When setting up the system, the motion capture cameras are the only component which might require more space. Therefore, it is a difficult problem to find a best place to place those 6 cameras, which then allow each camera to cover the most space, and also make them occupy as little space as possible.

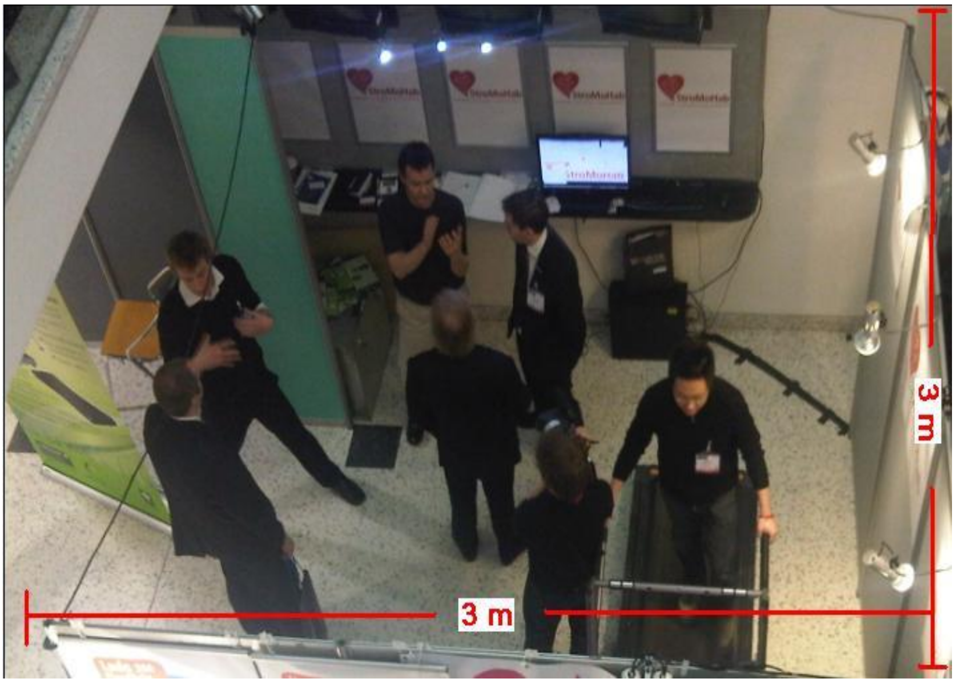

Figure 3.8 The 3m $\times$ 3m square meters exhibition stand for StroMoHab system in Venturefest. The author is seen walking on the treadmill, and the author's colleagues are introducing the StroMoHab system to visitors. This is to illustrate the small space that needs motion capture.

3.4.2 Design of Camera Positions

A team meeting was organized to discuss how to optimize the StroMoHab system in the provided stand, focusing especially on the setup for each camera, in order to provide the best camera view but also use less space. During the meeting, three design ideas were made and discussed as the options to set up the StroMoHab system. Then the StroCap software was used to analyze the three possible designs for the cameras, to see which one was the most effective in terms of capture space over the treadmill. The final design was selected for configuring StroMoHab after detailed analysis with the StroCap program.

Design option 1: Set up cameras on the top of the treadmill with as spider arms.

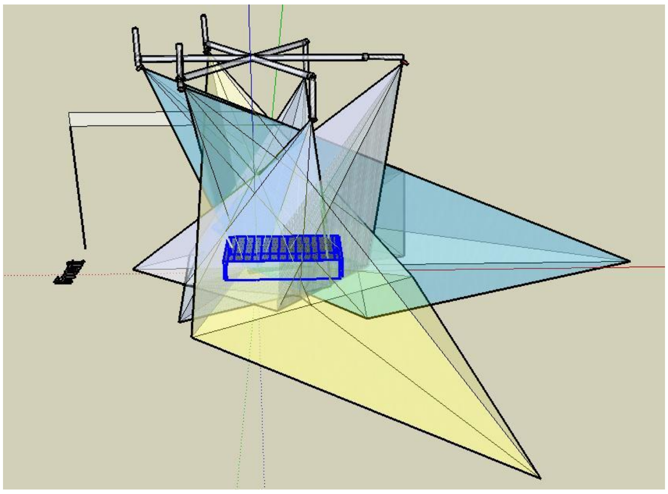

Figure 3.9 Spider arms design. In this design, cameras are fixed on the tubes and mounted on the ceiling, and looking down to the ground.

In this design option, cameras are placed all around on the top of treadmill with aluminium tubes and at 3 meters height (see Figure 3.9). Cameras are looking down to the treadmill with different rotation angles and each of them could capture the whole treadmill area. Also, because cameras are on the top of the system, they will not take any ground space, which will save more space for the exhibitor stands to use.

Design option 2: Set up cameras beside the treadmill as a big ring.

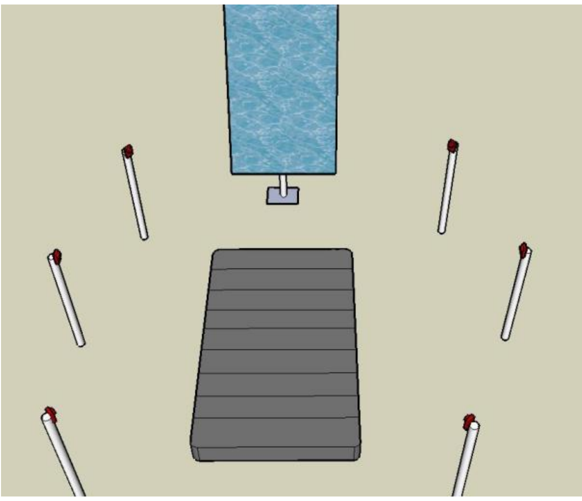

Figure 3.10 Big ring design. Cameras are mounted on the tubes and stand around the treadmill like a big ring.

In this design option, cameras are placed around the treadmill with aluminium stands, in a big ring with a 1.5 meters radius (see Figure 3.10). Cameras are all at 2m height, facing the treadmill with different rotation angles and each camera captures a different amount of space on the treadmill area. As people are walking on the treadmill, cameras will detect markers from the front, two sides and behind, and then people’s whole movement will be covered and captured.

Design option 3: Cameras mounted on the screen

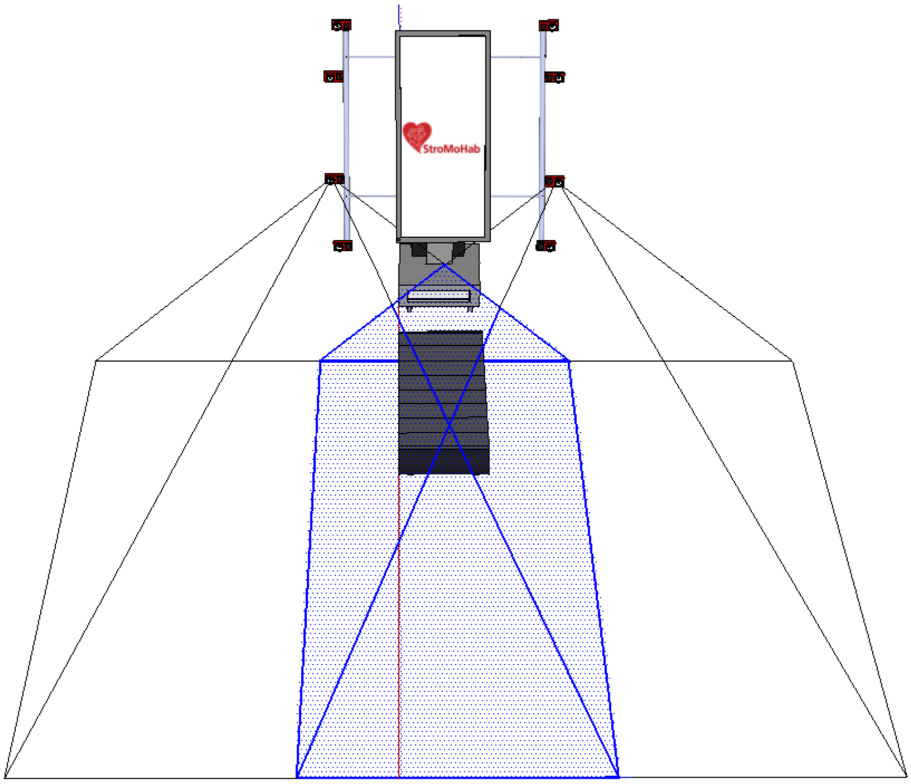

Figure 3.11 Cameras mounted on screen design. In this design, cameras are all mounted around the screen and looking forward to the treadmill. Two cameras' capture volume are shown by the black lines, and their overlap regions on the treadmill and areas of the floor are represented in blue.

Cameras mounted on the screen are placed in front of the treadmill in this design (see Figure 3.11). Cameras that are facing towards the treadmill will make sure that markers' movements will be captured by every camera.

3.4.3 Design Testing on the StroCap and Analysis

The size of the stand is only 3m × 3m, the StroMoHab system therefore must be setup by using as small a space as possible, whilst also allowing every motion capture camera to cover the maximum possible space. All of those design ideas sounds reasonable, but only one of the options can be selected and applied to the system setup.

The traditional method for testing each design option is to use a ‘trial & error’ process by using the existing motion capture software (e.g. Arena Express). In this method, to test each design idea we firstly need to randomly set up the system in the real world, and then use markers to measure the capture volume for each camera. If the calibration fails, we have to manually set up the system again. Therefore, the traditional method is not a systematic way to configure the StroMoHab system. However, if the StroCap software improves the efficiency of the setup, it is useful to use to optimize the system, which can simulate each camera’s location, capture volume, overlapping and rotation angles when capturing the required space, and the actual configuration can be then processed after the design in StroCap.

The first design option is to put cameras above the treadmill, probably to setup all the cameras on the ceiling. However, this design would make the StroMoHab system setup work more complicated and less practical. For example, as the exhibition is in a public place, it is unlikely that we would get permission to drill a hole into the ceiling, and build up a frame with aluminium tubes. This design could be used as an alternative choice in a private and fixed place, such as the experimental lab. The second design option is easier to implement compared to the first option, in that the cameras can be easily set up with aluminium tubes. But the disadvantage for this design is that the ring shape with 1.5 meters radius would occupy almost the whole exhibition space. Therefore option 3 seems to be the best and most suitable design for the Venturefest exhibition, because all cameras are mounted on the screen in this design. It would not take any extra space on the ground, and would leave enough room for other needs.

When simulating the third option in StroCap, six cameras were added in the front of the treadmill with different rotation angles. Camera positions can be adjusted to ensure that each camera could capture the most space on the treadmill (see Figure 3.12).

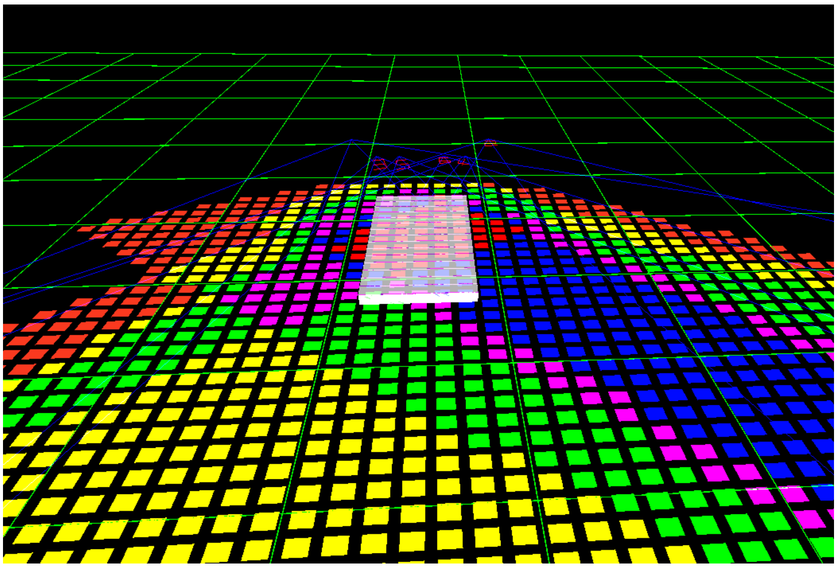

Figure 3.12 Cameras capture areas on the treadmill and floor in design option 3. Some of the areas on the treadmill were captured by all six cameras.

Figure 3.12 shows the area around whole treadmill (in white colour) is captured by 6 cameras at the same time. The red region shows areas captured by one camera; yellow colour means two cameras could see this place at same time; green represents areas captured by three cameras; areas captured by four cameras are coloured blue; purple regions are captured by five cameras and dark red places can be seen by all six cameras.

After discussing the result, the third design option is the most suitable and was used at the Venturefest Exhibition. Figure 3.13 shows the StroMoHab system setup picture according to the design.

Figure 3.13 StroMoHab was configured based on the design option 3. Cameras were all mounted on the virtual reality display and looking forward to the treadmill, markers were put at the front of the treadmill.

After the StroMoHab setup, we randomly put two markers at the front, two sides and the bottom of the treadmill in order to test cameras' performances in this design. The Arena Express software shows markers are successfully captured by all the cameras (see

Figure 3.14) when makers were put anywhere on the treadmill, which reveals that using StroCap is a good systematic method for positioning motion capture cameras to capture the highest amount of marker movement.

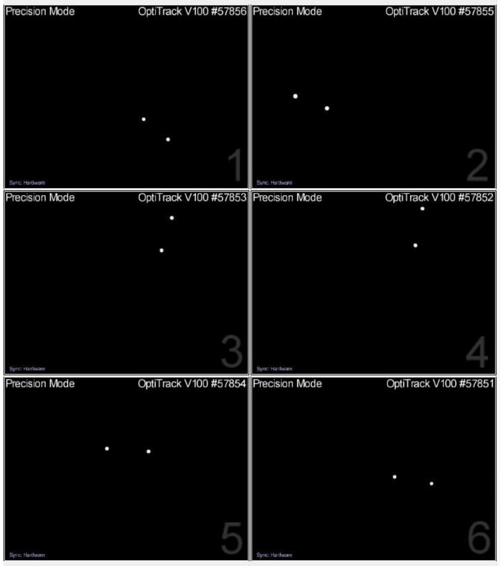

Figure 3.14 Two markers are all captured by each camera when put on the front of the treadmill and their positions are shown in the Arena Express software. The black panel show the view of each camera, which is numbered from one to six to represent the six cameras. The two white dots are the markers which were captured by cameras.

3.4.4 Case Study Conclusion

In this real case study, StroCap successfully provided a platform for setting up the StroMoHab system, and accurately provided cameras' coordinates, facing direction, capture volume, rotations and overlapping area. These have helped with the design and configuration of the StroMoHab system, and make the setup procedures more accurate and efficient. StroCap improved the existing motion capture simulator, which perfectly optimized the system in the defined space in the Venturefest exhibition booth without requiring a prolonged ‘trial & error’ process.

Chapter 4. Background of Eye Movements in Virtual Locomotion Simulators

Previous chapters described the procedures for designing and implementing the StroCap, and the results have proved to be a success for this software after it was tested. The next part of this thesis will introduce the study of eye movements while walking in a real and a StroMoHab locomotion simulator. This will be used as a method to evaluate the locomotion simulator by comparing the eye movement performance in the real world and using the StroMoHab simulator. We are seeking a way of making the StroMoHab simulator, and locomotion simulators generally, more realistic, i.e. more like the real world, by comparing the differences in the eye movements between those in the simulator and those in the real world under controlled conditions; the environments of both are matched as closely as possible. Prior research and an exploration of the theoretical background to this topic will also be presented in this chapter.

4.1 Eye Movement Studies

The eye is a sensory organ (see Figure 4.1) that reacts to light and converts it into electro-chemical impulses by the photoreceptor cells of the retina. The impulse signal goes through the ganglion cell layer and optic nerve fiber, and finally reaches the visual cortex, which itself contains many feedback networks.

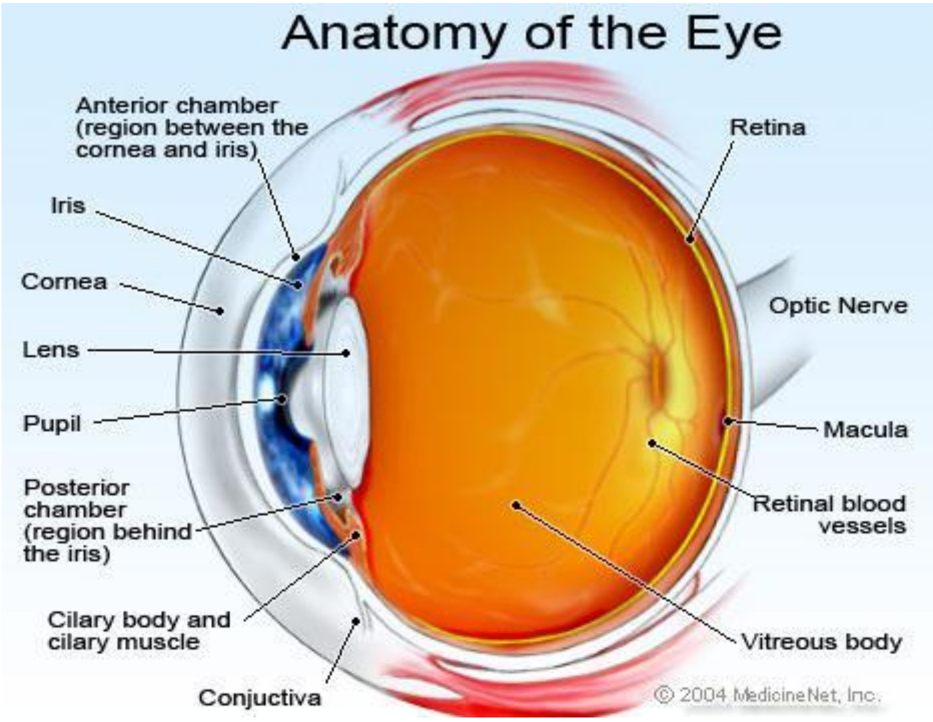

Figure 4.1 The anatomy of the human eye. The eye has two sets of muscles, which change the diameter of the pupil to adjust the amount of light, and the retina converts the light to an electrochemical signal that propagates action potentials along the optic nerve to the brain. (taken from MedicineNet 1996-2012).

Eye movement can be a voluntary or involuntary behaviour of the eyes. A small cortical region in the brain's frontal lobe initiates eye movements by moving the frontal lobe (e.g. Heinen & Liu 1997; Tehovnik et al., 2000), delivering the reflex movements without voluntary control (e.g. Encyclopaedia Britannica, 2012). Eye movement is one of the most important features of human activity during locomotion. For example, a human hunter will need to move their eyes to find predators during the chase, or to avoid tripping on obstacles as they run along a rough terrain.

By a variety of methods, eye movements can be classified as three voluntary movements, which are vergence movements, saccades and pursuit movements. Vergence movements, also known as convergence, occur when both the eyes are moving to make sure the target image is captured by both retinas (e.g. Carlson & Heth, 2010). Saccades represent short and rapid eye movements between target objects or visual scenes (e.g. Murray, 2003). Pursuit movements, also called smooth pursuit, in which the eyes track the movements of an object, so that a moving object can remain static on the fovea (e.g. Carlson & Heth, 2010). Physiologically, there are three main basic types of eye movements: ductions, versions, or vergences (e.g. Kanski, 1989; Awwad, n.d.). A duction is an eye movement which only uses one eye; a version involves both eyes and each eye moves in the same direction (e.g. saccades); in vergence movements, both eyes are involved with each eye moving in an opposite direction.

There are a large number of studies around eye movement, such as eye movements and visual attention (e.g. Corbetta et al., 1998), eye movement while in locomotion (e.g. McDonald, Bahill & Friedman, 1983), and visual searching in a virtual environment (e.g. Lessels & Ruddle, 2005), most of which investigate topics outside the scope of this thesis. McDonald, Bahill & Friedman (1983) found that in normal everyday activities, such as walking, and head movement, these can make people have less precision on fixating objects than other behaviours such as standing still. Lessels & Ruddle (2005) investigated the ability of subjects to target while walking. Two experiments were carried out in two field of view conditions (full vs. restricted) in both real and virtual environments (VE). The searching performances were different between all the conditions, but were getting closer after the subjects practiced. The best performance was found in the VE group, with results approaching those who participated in real world experiment. Those results indicated that it is important to use high fidelity scenes in VE, and the study also included using a simple control interface for spatial orientation during searching.

Although the research on eye movement has been done for different purposes, eye movements have not previously been applied to locomotion simulators, and this is the purpose which will be addressed in this thesis. We will study eye movement during normal locomotion, and will do investigations by comparing eye movement parameters during locomotion as a performance measure in simulated environments versus real environments.

4.2 Virtual Locomotion Environments

A virtual locomotion environment (VLE) or locomotion simulator is a device or system that allows humans to experience walking as naturally as possible in a controlled environment by presenting the illusion of motion activities while remaining stationary in a room or laboratory. A locomotion simulator is the highly controlled, relatively low cost system (although very expensive ones exist also), and any errors made by users are arranged that they will not have dangerous consequences, unlike potential situations in the real world. (e.g. Grechkin et al., 2010). These benefits make virtual locomotion environments potentially applicable to many areas, including rehabilitation, training, and behavioural research. An experiment was conducted by Mohler, Creem-Regehr & Thompson (2006) which investigated the effect of virtual feedback on distance estimation in a HMD (head mounted display) locomotion simulator. Before the experiment started, subjects viewed a rendered target on the floor in HMD VLE, and were instructed to remember the location of the target. The blind walking was conducted in a pre-test, in which subjects' eyes were closed during walking in the VLE, and ceased walking and opened their eyes until they believed they were standing on the target location. The feedback was given in the adaptation session, such as verbal feedback in which participants were walking with eyes closed to a previously viewed target until the experimenter told them that they were standing at the target location. After this session, subjects were instructed to take the post-test, which is a similar process as the pre-test. The results showed that the feedback significantly improved the accuracy of blind walking (e.g. 13% increase by verbal feedback) in the estimated distance of the task in VLE.

To date, many of the applications of virtual environments are to investigate and understand the differences of people's perception and action (e.g. Creem-Regehr & Kunz, 2010). For further background, see a special issue on walking in real and virtual environments (e.g. Pelah & Koenderink, 2007).

4.3 VLE Display Technology and Walking Interfaces

Two common technologies are used for display in virtual locomotion environments: One is a head-mounted display (HMD) (see Figure 4.2), where the user wears a helmet with attached displays to view and move within a computer generated environment, but this equipment normally provides a relatively narrow field of view (FOV) and the weight of the helmet and cables can encumber the user (e.g. Grechkin et al. 2010). Another is a large-screen immersive display (LSID) system (see Figure 4.3), which has a large screen to provide the user with a wider FOV and less encumbrance (e.g. Grechkin et al. 2010). In the locomotion simulator studies, an HMD has often been used for blind walking experiments for distance perception (e.g. Kuhl, Thompson & Creem-Regehr, 2006; 2008; Creem-Regehr et al. 2004). The timed imagined walking experiments normally involved LSIDs, in which subjects start a stopwatch when they begin to imagine walking toward a target and stop the stopwatch when they believe they have reached the target; subjects are not allowed to look at the stopwatch during the task (e.g. Klein et al., 2009). A comparison showed that the distance perception performance was better in the LSID than in the HMD in virtual locomotion environment (e.g. Plumert et al., 2005).

Figure 4.2 User wears Head-mounted display (HMD) (taken from VRarchitect 2004).

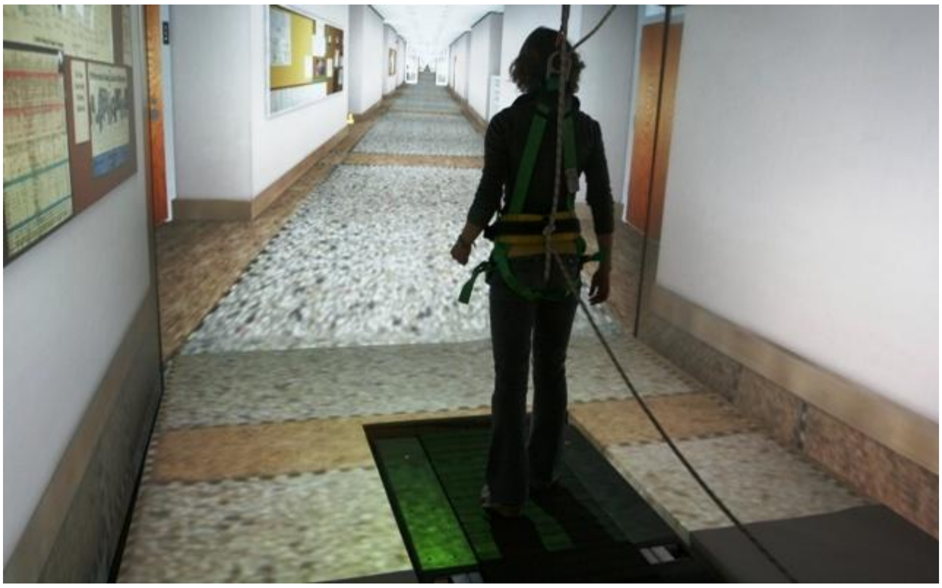

Figure 4.3 Participant walking in the Large-screen immersive display (LSID) system (taken from Hank Virtual Environments Lab, 2010).

The walking interfaces normally used in VLE studies are the conventional treadmill and, more recently, the omni-directional treadmill. The conventional treadmill only supports forward walking, as users make distance progress in the locomotion simulator. Omni-directional treadmills (see Figure 4.4) are a potentially revolutionary device (e.g. Darken, Cockaybe & Carmein, 1997) which allows a person to perform locomotion in any direction in a large-scale virtual locomotion environment (Ruddle & Lessels, 2009). Other locomotion interfaces can also be used in the VLE, such as the Walking-Pad (e.g. Bouguila, Florian & Courant, 2005) (see Figure 4.5), which could provide the user with a life-like walking experience in VLE by stepping in place; this device is able to simulate a user's locomotion activities such as walking directions, walking speed, jumping etc. It can be controlled by connecting it to any computer via a USB cable; the advantage for this device is less weight and size, and it is also more portable compared with treadmills; the disadvantage is that it requires walking in place as opposed to forward walking as in natural locomotion.

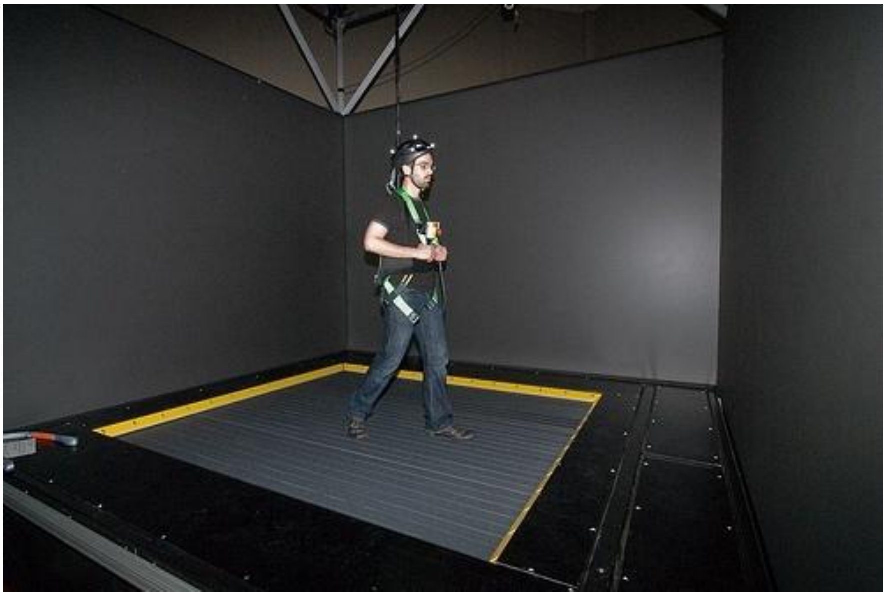

Figure 4.4 User walking on the Omni-directional treadmill which allows user to walk in any direction (taken from Kuntz, 2007).

Figure 4.5 Participant walking on the 'Walking-Pad' (taken from Bouguila, Florian & Courant, 2005)

We use the StroMoHab locomotion simulator in this thesis. StroMoHab has a 50 inch virtual reality display, which is similar to the LSID but not as large. The locomotion interface in StroMoHab is a conventional linear treadmill, which allows the user to walk linearly whilst facing the display. The rotation (i.e. turning) movements are not considered in this thesis.

4.4 Human Perception in VLEs

There are several research studies that measure people's locomotion perception in the VLEs and comparing subjects' behaviours between the real world and the virtual locomotion environment. In Geuss et al., (2010), subjects completed three perceptual judgement tasks (size estimates, affordance judgments, and blind walking estimates) in order to compare subjects' performance in terms of locomotion and perceptive distance between real and virtual locomotion environments. The results show that in the affordance judgments and size estimates tasks, subjects had comparable performances in the locomotion simulator and the real world; however, in the blind walking task, distance estimates were different between the two environments. Potential reasons for this may be that in the virtual locomotion environments, the depth intervals are compressed in the virtual scene, leading to perceptual differences in egocentric distance estimation. However, in affordance judgments and size estimate tasks, subjects were asked to judge the exocentric distance. Findings showed that participants have similar perceptions in both the real environment and the VLE. This is consistent with another finding by Kenyon et al. (2007), that size perception in virtual locomotion environments is similar to the size perception in the real world. Sahm et al. (2005) did a study on measuring subjects' locomotion versus distance perception in the HMD VLE. In this study, participants made blind walking and blind throwing in the hallway in both real and HMD virtual locomotion environments, and aimed to assess the performance in terms of accuracy across the two environments. In design, targets were placed on the floor at a distance of 3, 4, 5 and 6m in both environments, with each distance presented three times during the experiment. Subjects were instructed to either blind walk or blind throw beanbags to targets on the floor in both environments during the experiment, and the results indicate that the performance in terms of blind walking accuracy and blind throwing accuracy in both real and virtual locomotion environment is similar. This finding differs with Geuss et al. (2010) and Kenyon et al. (2007), which said that blind walking in real world and VLE was different. Sahm et al. (2005) also found that the absolute egocentric distance was compressed; this is similar with the finding by Durgin et al. (2002), Thompson et al. (2004), and Willemsen & Gooch (2002), who showed that the egocentric distance was compressed by nearly 70% in VLE. Sahm et al. (2005) suggested that performance differences between real and locomotion simulator may be because of an underlying perceptual error with respect to the virtual locomotion environment. The cause of that error remains unknown.

4.5 About this Study

While there have been many studies about the eye movements themselves, eye movements have not been used to study human perception in virtual locomotion environments. The present study uses eye movements as a measure of performance in both real and simulated environments. Eye movement experiments were run on a treadmill-based locomotion simulator, provided with a video stimulus which was presented on a 50” Sony plasma display oriented in portrait format (see Figure 1.2). A video stimulus was chosen to make the simulated locomotion environment as comparable as possible to the real environment. Video platforms in VLE were also used in other applications, such as in rehabilitation (e.g. Rand et al., 2005; Weiss et al., 2006), which was considered to increase the interaction between simulator and patients (e.g. Weiss et al., 2004).

The overall goal of the study is to make the locomotion simulator as similar as possible to the real environment, and to optimize the StroMoHab simulator system by comparing locomotion eye movements in simulated locomotion environment and the real environment. Through the investigation, we want to learn the differences of eye movement parameters between the two environments, aiming to find the best conditions to use the StroMoHab simulator. Also, we will use a natural realistic yet complex task, to activate several brain mechanisms, in order to understand locomotion simulators and eye movements in a more general way.

Chapter 5. Eye Movement in Locomotion: Experimental Methods

This chapter describes the experimental methods in detail that will be used in all the experiments. Methods include the design for each experiment, status of involved participants, experimental equipment, description of experiment environment, experiment tasks design, details of performance measurement and analysis methods.

5.1 Experiment Design

Totally six experiments were conducted, abbreviated Exp-1 to Exp-6, and three independent variables were varied: speed of walking, motion speed of the video and a no-video walking condition. Table 5.1 outlines the walking and video speed combinations, numbers of subjects and conditions used in the six experiments.

Experiment	Exp-1	Exp-2	Exp-3	Exp-4	Exp-5	Exp-6
No. of Subjects	30/Env	10	5	15/Env	5	5
Environment (Env)	SE & RE	SE	SE	SE & RE	SE	SE
SE Video Speed (km/h)	1.3	——	0.7/2.0/3.3/4.6	2.1/2.8/3.8	0.7/2.0/3.3/4.6	3.3
SE Walking Speed (km/h)	0/1.3/2.5	1.3/2.6/5.2	Vary	0	0.7/2.0/3.3/4.6	3.3
RE Walking Speed (km/h)	1.3	——	——	2.1/2.8/3.8	——	——

Table 5.1 The different combinations of video and walking speed, numbers of participating subjects and the conditions in the experiments conducted in the real (RE) and simulated (SE) environment.

5.2 Subjects

A total of 70 subjects took part in the study (55 male, 15 female), all of whom were university students between the ages of 21 and 32, with normal or corrected-to-normal vision. Each subject signed an informed consent form according to departmental standards on ethics and following appropriate risk assessment procedures.

5.3 Eye Tracker

All subjects wore portable eye-tracking equipment in the experiments. The eye-tracker used was a 'WearCam' (Figure 5.1), developed by the Learning Algorithms and System Laboratory, in the Department of Microengineering at EPFL, Luasanne, Switzerland. The eye-tracker has two cameras (top and bottom) mounted on a headband (Figure 5.2), the first being used to record the scene viewed by the subject. The second camera captures the user's eyeball movements. The video signals are captured and analysed by a notebook PC carried in a small rucksack by the subject. Following the experiments, the provided 'Eye Detection SVM' software is used to first calibrate the eye recorded video and then to analyse the eye movement data.

Figure 5.1 The Wear Cam eye-tracking equipment is shown, including rechargeable batteries (left), connector box with USB cables (middle) for PC input, and cameras mounted on the headband worn by the subject.

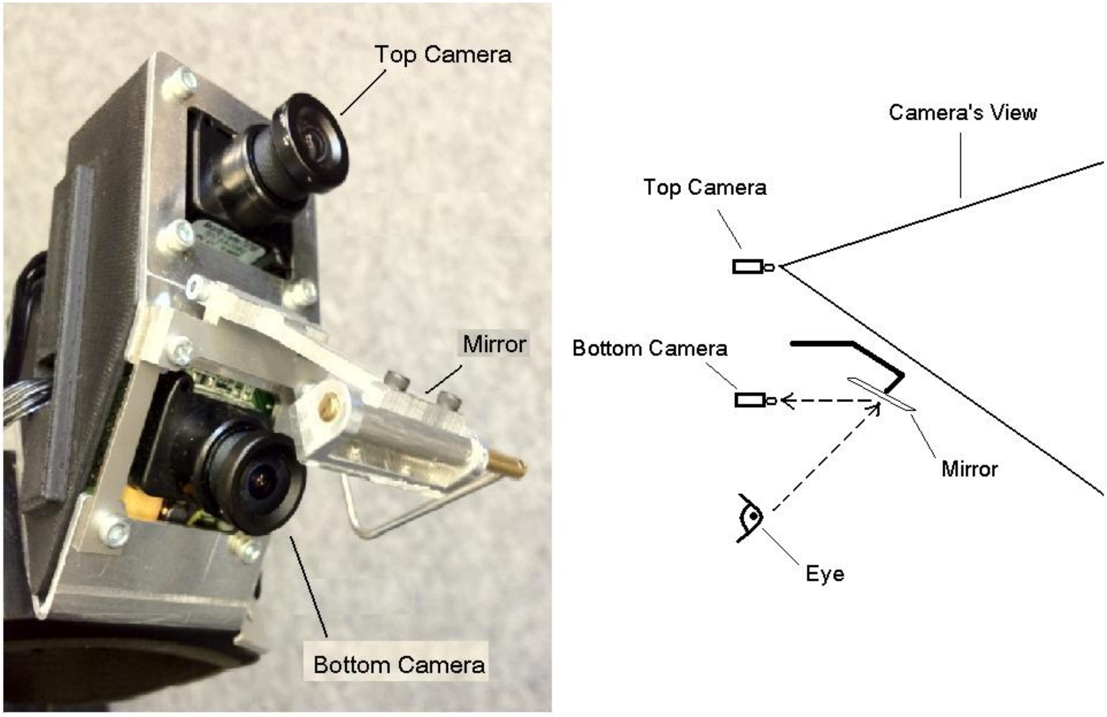

Figure 5.2 The eye tracker device diagram and the schematic of the optics of this device. The eye tracker contains two cameras mounted on the headband on subject's head. The camera on the top was used to record the scene viewed by the subject. The bottom camera captures the user's movements of their eyeballs reflected from the mirror.

5.4 Real Environment Setup

The real world condition (RE) in this study was run on a 38m long university department corridor (Figure 5.3), with between 20 and 30 coloured tennis balls (red and yellow) and ping pong balls (orange, yellow and blue) as the target objects distributed on either side of the subject's walking path. There were also are five buffers on the sides of the path placed there in order to occlude a full view of all targets from any one position, and to prevent subjects from reviewing the full path prior to the start of the trials. The numbers of target objects were varied and their positions semi-randomly applied for each experimental run.

Subjects were asked to carry a metronome with them and their steps were following the metronome to make sure every subject walked at the same frequency; the metronome frequency was varied in different experiments. Before starting every subject was trained several times to help them walk with the metronome speed. Then, in the experiment, each subject was instructed to walk the path once at a steady pace from a specified starting point, while wearing the WearCam eye tracker and a small rucksack and to stop upon reaching the end of the corridor. They had to also keep their body and head steady during walking.

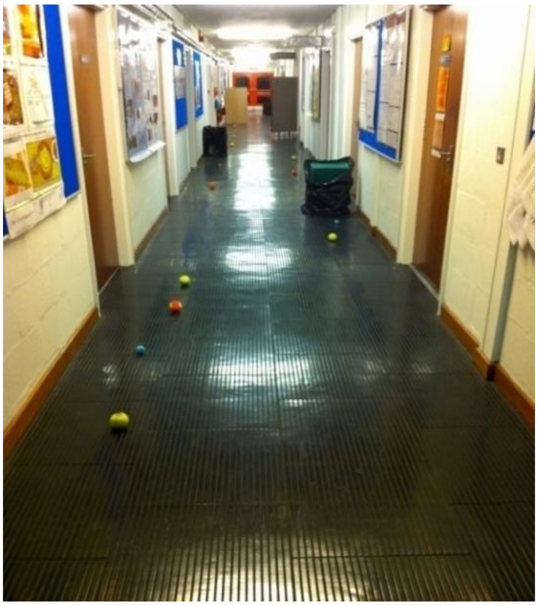

Figure 5.3 The real environment in this study, which is a walking corridor in the department. Objects were randomly placed on the corridor, and also some buffers on the side of the corridor. In this real environment, subjects were provided with normal indoor lighting conditions.

5.5 Visual Stimulus and the Locomotion Simulator

When creating the simulated video environment, a cameraman was asked to stand on a trolley and hold the video camera in the portrait form at a fixed height (170cm) and also make sure all the scene objects are viewed in the lens. The video camera in this study is Canon HV20 with 3.1Mega pixels, and set to HDV mode when recording. Another person pushed the trolley steady and slowly through the corridor, and a steady video was captured from the cameraman's viewpoint while the trolley was moving in the corridor path.

After recording, the video was then stabilized using 'SynthEyes' software and played back at different speeds as required on the StroMoHab system. The video speeds were matched precisely to the walking speeds in the locomotion simulator by computing the required video run time from the known length of the RE path for each walking speed. The stimulus was presented on a 50" Sony FWD-50PX2 plasma display positions oriented in portrait format, chosen to present a wider vertical angular view of the walking path to better represent the subject's view during the forward walking path in the real environment (see Figure 5.4). The video stimulus was chosen to make the simulated environment as comparable as possible to the RE, as in our judgment observers would probably regard a video of a previously recorded (somewhat cluttered) real world scene as a more faithful representation of a familiar real world environment than a rendered virtual scene. The treadmill handlebars were used by subjects in the SE for better safety.

Figure 5.4 Subject walks on the StroMoHab locomotion simulator while wearing the eye-tracker, as the simulated video environment is played in the screen. In this simulated environment, subjects were provided with normal indoor lighting conditions, while avoiding reflections on the screen.

5.6 Matching Visual Angles in Real and Simulated Video Environment

In order to enable comparisons between the RE and SE, the viewed scenes (real and virtual locomotion) needed to be made as similar as possible (inherent differences are discussed later in the thesis), and particularly in equalizing the visual angles given in which eye movements were measured (see Figure 5.5). We first computed the RE visual angle from the viewpoint at the start of the 38m long corridor of the floor-to-ceiling height of 2.5m. We then analyzed the video frame taken from the same starting location, measuring the floor-to-ceiling height on the image at 0.07m, and adjusted the viewing distance from the screen at 1.06m thus ensuring the same angle (3.8 deg) was subtended at the nodal point of the eye in the SE. In SE, the field of view of subjects are 61.3 deg vertically and 39.1 deg horizontally. Unlike the SE, the visual field in the RE was of course not limited to a physical display device. Target objects were visible only within the shared field sizes of the two environments.

Figure 5.5 A photograph to illustrate the matched visual angles in the RE (left) and SE (right) within the scene of the corridor path. The visual angles of target objects were equal in the two environments, although the full field of view in SE was narrower in both height and width in comparison to the RE. Regions outside the SE view did not contain target objects for the SAC task. Note that any geometric distortions in this figure are due to capturing the image and were not present in experiments for this figure.

5.7 Performed Task

A 'Search and Count' task (SAC) was used to measure subjects' performance (both in SE and RE), in which subjects were instructed, for example, to count the number of red tennis balls, orange ping pong balls and blue ping pong balls during the walk. Three ball colours from the two possible types of balls were selected for the counting tasks, with variations made in these (and in target positions) between the experiments and the captured videos of the path shown in the SE.

We invented the SAC task in the eye tracking experiment because 1) it requires eye movements while walking in the task, 2) the SAC task can provide subjects with visual conditions that can be made similar between RE and SE in the experiments., and 3) this is a complex task that is similar to a number of natural tasks, involving colour detection, eye movements, visual attention and cognition etc.

The SAC task was chosen to illicit the desired behaviour and neural and cognitive activity in subjects and not as a measure of performance. Difficulty levels were adjusted such that it was sufficiently challenging but not to the extent that subjects performed poorly at it. After a few pilot trials, participants felt that it was too easy for them to count only two ball colours, but four ball colours were too difficult for them to count and remember. Therefore, subjects were asked to search and count three different ball colours in every experiment in this study. Subjects' usually performed correctly in counting balls, as verbally reported at the end of each run, and this was recorded to ensure subjects complied with instructions.

5.8 Performance Measures

In order to quantitatively compare how well or differently subjects performed in RE and SE, we developed a new measure of performance that is based on their eye movements. The total number of task-relevant saccades, S, was calculated from the number of eye movements made during an experimental run between balls or regions of interest (ROI), and defined as the overlap of the tracking software's eye fixation marker (see Figure 5.6) with the pixels in the analysed or captured videos that form the images of the target objects (balls). The definition assumes reasonably that fixating at or very near a target object is needed in order to count it over a relatively short time and within a fairly cluttered environment. S therefore excludes saccades outside the regions that are unlikely to be contributing to task performance, and this to an extent reduces the influence of possible variations in strategy between subjects. When a SAC task trial begins, the subject starts to search the target ball and count it. Subjects first move their eyes to search for the balls on the corridor, and then 'capture' the ball by fixating their eye on the ball (see Figure 5.6); subjects next have to define if it is the target ball which needs to be added to the mental count, and then move their eyes to search for the next ball. In this procedure, when subject moves his eyes onto a ball, we will count this movement as one saccade. However, if a subject moves their eyes for other proposes, such as to look elsewhere at the floor or wall, this eye movement was not defined as an eye saccade to be added to the metric S.

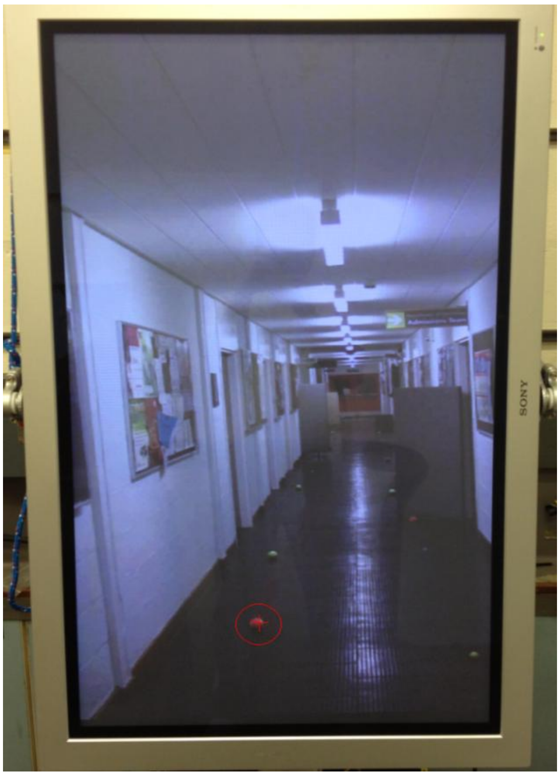

Figure 5.6 Subject's eyes move onto a target in the SAC task in SE, and this movement is counted as an eye saccade. The red circle and cross represent the eye fixation marker which was analysed and computed by the 'Eye Detection SVM' software.

When considering S, a low number of saccades correspond to better performance than a high number since, in general, fewer eye movements would correlate with lower energy expenditure to plan and generate the movement (e.g. Schultz, 1997). We combine this measure with the total fixation durations, F, the total time during which fixation rested within ROIs, to produce an *energy cost function* (ECF), a hypothetical scalar metric based on the assumption that the energy expended by a series of eye movements during a search and count task is proportional to the product of the number of discrete movements (saccades) and their total duration. We define ECF using the Michelson formulation for both duration and number of saccades as in Equation 5.1. Michelson formulation is often used for luminance contrast, with possible values ranging from 0 to less than 1, such that values closer to 0 indicate a lower energy cost in the present case.

$$\mathrm{ECF} = \left(\frac{\mathrm{F} - \mathrm{F_{ideal}}}{\mathrm{F} + \mathrm{F_{ideal}}}\right) \times \left(\frac{\mathrm{S} - \mathrm{S_{ideal}}}{\mathrm{S} + \mathrm{S_{ideal}}}\right)$$

$$\mathrm{F_{ideal}} = (\mathrm{t} \times \mathrm{N_{ROI}})$$

$$\mathrm{S_{ideal}} = (1 \times \mathrm{N_{ROI}})$$

Equation 5.1 The *energy cost function* (ECF) defines the measure used to express performance based on eye movements and fixation durations for real (RE) and simulated (SE) environments used in the experiments. See text for explanation.

In Equation 5.1, $\mathrm{N_{ROI}}$ is the total number of regions of interest (ROI), and $\mathrm{F_{ideal}}$ describes optimal performance based on t, the minimal usable fixation duration (taken arbitrarily in the analysis as 1 video frame, or 0.04 seconds, because there are 25 frames in 1 second). $\mathrm{S_{ideal}}$ is defined as the minimal number of saccades needed to count the target objects in the task, which assumes that each target optimally requires only one saccade. Therefore the $\mathrm{S_{ideal}}$ equals to the $\mathrm{N_{ROI}}$, or the total number of targets times 1 saccade on each of them. Note that when $\mathrm{F} = \mathrm{F_{ideal}}$ or $\mathrm{S} = \mathrm{S_{ideal}}$, ECF assumes its lowest (best) cost value of 0.

Based on the ECF, we also define *efficiency* as the reciprocal of ECF, and for convenience chose efficiency as the principal metric we use to compare performance between the different environments and conditions of the experiments.

$$\text{Efficiency} = \frac{1}{\text{ECF}}$$

Equation 5.2 Efficiency is the principal metric devised for the study to compare eye movement performance in the SAC task across conditions and environments, defined as the reciprocal of the *energy cost function* (ECF). Efficiency grows exponentially approaching infinity under ideal conditions and is not defined when ECF=0.

5.9 Statistical Analysis

All statistical analyses were calculated with the Statistical Products and Services Solution (SPSS). The chosen level of significance was p= 0.05 and lower. We conducted analysis of variance (ANOVA) on the number of saccades, S, total fixation time, F, counting accuracy and efficiency for all six experiments, either One-Way ANOVA (Exp-1 and Exp-4) or repeated measures ANOVA (Exp-2, Exp-3, Exp-5 and Exp-6). In Exp-4, we also conducted the Independent-Sample T-Test to compare the F (total fixation time) between real world and simulator groups (the data in each group fits the normal distribution).

In the One–Way ANOVA and Independent-Sample T-Test, Levene’s test was provided and used to assess the Variances of Homogeneity. If the significance value is less than or equal to 0.05 in Levene’s test, it means the variances are significantly different, therefore in the Independent-Sample T-Test, should use the data which in the row of ‘Equal variances not assumed’ as the result; in the One-Way ANOVA, it should use other methods such as Tamhane's T2, which is used when Equal Variance Not Assumed. Otherwise, if the significant value is greater than 0.05, it shows that equal variances assumed, and the One-Way ANOVA can be used, and in Independent-Sample T-Test, we can use the data which in the row of 'Equal variances assumed'. In this study, the Levene's test for all of the experiments are not significantly different ($p > 0.05$), hence, One-Way ANOVA and Independent-Sample T-Test (Equal variances assumed) can be used in all of these experiments to evaluate the statistical comparison.

In the One-Way Repeated Measures ANOVA, Mauchly's Test of Sphericity is used to validate the assumption of Sphericity, which if Mauchly's test value p is greater than 0.05, then Sphericity is assumed and can be used for the One-Way Repeated Measures ANOVA. Otherwise, if Mauchly's test is significant ($p < 0.05$), it therefore means that the assumption of Sphericity has not been met, and in this case, Huynh-Feldt (used when Epsilon value is greater than 0.75) or Greenhouse-Geisser correction (used when Epsilon value is less than 0.75) should then be used. Epsilon value indicates different ways to calculate an appropriate adjustment to the degrees of freedom of the F-test. In this thesis, Exp-2 and Exp-6 met the assumption of Sphericity; in Exp-5, the Mauchly's test was significant, and the Epsilon value was less that 0.75, therefore Exp-5 used Greenhouse-Geisser correction for the statistic comparison.

Least Significant Difference (LSD) method was used as the Post Hoc Test to find the significance between the levels of dependent variable after the One-way ANOVA or repeated measurement ANOVA in all the experiments.

5.10 Experiment Procedures

The procedures for each experiment is stated and shown as below:

5.10.1 Exp-1 Design and Procedures

--- Comparing Eye Movements during Walking in RE & SE

Experiment 1 was designed to compare the eye movement performance in terms of the number of eye saccades and total time of eye fixation between real and simulated video environments while doing an equivalent SAC task during walking. A total of thirty subjects joined this experiment and were randomly assigned to four groups for these four conditions. Each group contained ten random subjects. To avoid subjects feeling familiar with the task, this experiment was run on different days for different groups. The first group of subjects only walked once in the real environment at 1.3 km/h, group two, three and four were all walking on the treadmill in the simulated video environment, at speed of 0 km/h (stand on the stationary treadmill), 1.3 km/h and 2.5 km/h. Among them, the second group of subjects in SE have the same walking speed (1.3 km/h) with the RE walking group. In this experiment, subjects were instructed to search and count the number of red tennis balls (8 balls in total), yellow ping pong balls (3 balls in total) and blue ping pong balls (4 balls in total) from totally 26 objects in both RE and SE, and reported their answers after each run. In both environments, subjects were provided with an equivalent room illumination lighting conditions.

How to measure and define the walking and video speed

In Exp-1, subjects did the SAC task in the real environment first. When they were walking, subjects were asked to carry a metronome with them to make sure every subject was walking at the same frequency. The metronome can be set to different frequencies which range from 40bpm (bits per minute) to 110bpm.

Subject walked in RE by following the metronome at a random selected frequency which is 70bpm, after the RE group (10 subjects) completed the SAC task, we calculated that on average, it took them 60 seconds to walk through the 38m corridor, and hence their mean walking speed is approximately 1.3km/h.

In SE, the video needed to play back at the same speed as the real world walking, so the length of video was edited to 60 seconds by using the 'Ulead video studio' software. Because the aim of Exp-1 is to compare subject's total number of saccades in RE and SE at different walking speed, therefore in SE, the treadmill speed was set to 0km/h (0 speed), 1.3km/h (slower) and 2.5km/h (faster speed) for those three groups' subjects to walk. The 0km/h and 2.5km/h walking speed were randomly selected.

5.10.2 Exp-2 Design and Procedures

--- Eye Movement while Walking in SE with Static Visual Stimulus

Exp-2 measured subjects' number of saccades versus different walking speed when walking in a non-video input condition in SE, in order to investigate the relationship between eye movement and walking at different speed in SE. There were totally 10 participants in this experiment. Non-video input means that there are no video playing, and also no SAC task. Therefore, instead of the video in Exp-1, this time the screen only displays a static image. Subjects were all wearing the eye-tracking equipment and were required to look at the screen all the time during the walking.

Subjects were instructed to focus looking at the static screen while walking on the treadmill at three different speeds: 1.3km/h (slow speed), 2.6 km/h (medium speed) and 5.2 km/h (fast speed), all of these speeds were randomly chosen. Each subject walked on the treadmill for totally 1 minute. It started with a random speed (one of those set speeds), after the subject walked 20 seconds, the experiment was paused, and the current subject was asked to move their eyes away from the screen. The Experimenter then changed the treadmill to another speed and asked the subject back to the experiment to walk for another 20 second, and then we changed the treadmill speed again for the last 20 seconds.

How to create the static image

The stimulus in Exp-2 is a static image which consisted of nine rectangular boxes (Figure 5.7). This image was rendered by the C# program in Microsoft Visual Studio.

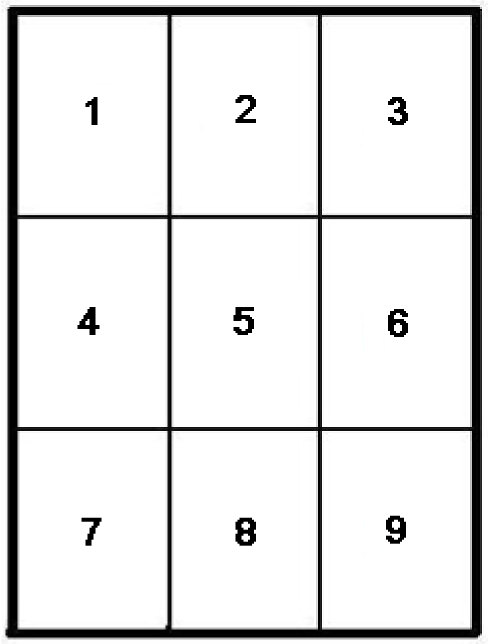

Figure 5.7 The display presented white outlines on a dark background of nine rectangular boxes according to the format shown in the figure. The numbers shown are included in the figure to describe the data analysis of fixations in the boxed regions and did not appear in the displayed stimulus.

5.10.3 Exp-3 Design and Procedures

--- Matching Locomotion Speed to Optic Flow Speed in SE

In this experiment, subjects were instructed to walk on the treadmill (without wearing the eye-tracking equipment) by following a video on the SE, and they were asked to perceive if their current walking speed matched the video playing speed. If they feel that their locomotion speed is not matched to the video clip speed, they either increase or decrease their walking speed by adjusting the treadmill speed until they feel it is matched. There were a total of four different tasks in this experiment which consisted of four different video clips, all of these which captured from the real environment and played back at four different speeds at 0.7, 2.0, 3.3 or 4.6 km/h. There were five subjects participating in this experiment, each of them had five attempts for each task. The aim of this experiment was to find if subjects' perception of locomotion speed could match with the video's actual speed from RE; if they don't match, we are looking for evidence to see if a learning effect is taking place between each attempt.

In the beginning of the task, the subject is provided with an initial treadmill speed (e.g. 0.32 km/h), and the clip in this task was randomly chosen and played from those four different video clips. After it starts, the task was paused every five seconds, subjects were then asked if they needed to either increase or decrease the treadmill speed, if they feel their walking speed is not matching to the video speed, then they can verbally report they need to either increase or decrease the treadmill speed, and the treadmill speed then will be adjusted by ±0.165 km/h (±0.1 mph) each time. The speed adjustment was allowed to be repeated as many times as necessary, until the subject was happy with the setting. When subjects feel their walking speed is matched with the video speed, they will report “Matched”, and the current task will end.

After one task is finished, the subject will be given another task to do. In this new task, the subject will be provided with a new clip with a different playing speed. Same as the last task, the subject will be asked to walk by following the new clip, and adjust the treadmill speed until they feel that the walking speed is matched with the video playing speed.

When this subject finishes the first attempt for all the four tasks, he will be given a one minute break, and then he starts again for the second round of the experiment. After the first subject has completed all the five attempts, the next subject will be asked to start. The experiment will be ended after all five attempts are completed for every video clip by five subjects.

5.10.4 Exp-4 Design and Procedures

--- Eye Movements Measured while Walking in RE and Standing in SE

Experiment 4 removed walking in the SE, the purpose of which was to investigate if waking affects subjects’ performance in SE. And also, in order to achieve optimal tuning of SEs, Exp-4 (and also Exp-5) invented the efficiency metric to compare subjects’ efficiency in SAC task when they were walking in RE at 2.1, 2.8 or 3.8 km/h and standing stationary while watching different speed of video (at 2.1, 2.8 or 3.8 km/h) in

SE. The SAC task in Exp-4 (same SAC task in both RE and SE) instructed subjects to search and count the amount of yellow tennis balls (7 balls in total), orange ping pong balls (3 balls in total) and blue ping pong balls (5 balls in total) from a total of 23 objects, and report their answers after each run. In both RE and SE, subjects were provided with sufficient lighting condition.

How to measure the locomotion speed in RE and video speed in SE

Same as Exp-1, subjects in this experiment also did the SAC task in the real environment first. Totally fifteen subjects joined this experiment in RE and were randomly assigned to three groups, each group contained five random subjects.

When they were walking, subjects were asked to carry a metronome with them to make sure every subject was walking at the same frequency. The metronome can be set to different frequencies which has a range from 40bpm (bits per minute) to 110bpm.

The first group of subjects walked in RE by following the metronome at a random selected frequency which is 70bpm, after this group completed the SAC task, we found on average, it took them 63 seconds to walk through the 38 meter’s corridor, and hence their walking speed was calculated as approximately 2.1 km/h. Subjects in the second walking group were also asked to hold a metronome which had a frequency set randomly at 90bpm, and their walking speed was found at 2.8km/h in average. The metronome’s frequency was 110bpm in the third walking group, and those subjects’ walking speed was calculated as 3.8km/h through calculating.

The second part of Exp-4 was held in SE, and with the same fifteen subjects. To avoid subjects feeling familiar with the task, the experiment in SE was run on a different day to the RE task. The video in SE was captured from the real environment, and was played back at the same speed as the real world walking. According to different speeds, the length of video was then edited to 63 seconds (70bpm, 2.1km/h), 48 seconds (90bpm, 2.8km/h) and 37 seconds (110bpm, 3.8km/h) by using the 'Ulead video studio' software. Because the one of the aims of Exp-4 is to measure and compare subject's efficiency versus different video speed when standing in SE, therefore the treadmill speed was set to 0 km/h (0 speeds) only.

5.10.5 Exp-5 Design and Procedures

--- Eye Movement Performance at Matched Locomotion and Optic Flow Speed in SE

In Exp-5, the efficiency of search and count was measured in the same task scene but at four matched speeds (walking speed matched video playing speed) at 0.7, 2.0, 3.3 or 4.6 km/h in the SE, in order to find that if there is an optimal speed in which efficiency is highest, to tune SE displays so that they can reproduce as well as possible the properties of RE.

SAC task and matched speed measurements in SE

The video in Exp-5 SE was captured from the real environment. This experiment measured the efficiency in four different matched speeds for five new subjects; therefore the original video scene was edited and then generated four videos at 0.7, 2.0, 3.3 or 4.6 km/h playing speed. According to these speeds, the length of video was then edited to 56 seconds (0.7km/h), 20 seconds (2.0km/h), 13 seconds (3.3km/h) and 11 seconds (4.6km/h) by using the 'Ulead video studio' software. Because this is a matched speed, so after measuring the video speed, the next step is to set the treadmill speed to its corresponding video speed.

The SAC task in these four videos are the same, in which subjects were instructed to search and count the amount of yellow tennis balls (3 balls in total), red tennis balls (3 balls in total) and blue ping pong balls (2 balls in total), and report their answers after each run. To avoid subjects feeling familiar with the task, the experiment was run on a few different days. In this simulated environment, subjects were provided with sufficient lighting conditions.

5.10.6 Exp-6 Design and Procedures

--- The Learning Effect Investigation of SAC Task in SE

Exp-6 was designed to discover if the familiarity with the environment could affect the subjects' performance in SE. In this experiment, five new subjects were at the same matched walking and video speed combinations at 3.3km/h, and SAC task in here was to search and count the amount of yellow tennis balls (3 balls in total), red tennis balls (3 balls in total) and blue ping pong balls (2 balls in total), and report their answers after each run. In order to increase their familiarity, subject was asked to repeat each task five times in total. In this simulated environment, subjects were provided with sufficient lighting condition.

During each period of walking, we accessed the subject's performance by measuring subject's total number of saccades, total fixation time and efficiency. We are assuming that if the familiarity to an environment affects subjects, their task performance should be getting better after each task is repeated; otherwise, if there is no improvement on their performance, it will indicate that familiarity has no effect on subjects' task performance in SE.

Chapter 6. Eye Movement Experiments Results

Exp-1 and Exp-2 obtained subjects’ eye movement patterns by measuring their total number of eye saccades in RE and SE. Exp-3 was run only in the SE, which provided subjects with four different matched walking and video speed combinations. Instead of eye movements, it measured subjects’ perception of their locomotion speed for different video playing speeds. It is advantageous to consider how well subjects performed in the SE during different conditions, in order to better understand the underlying mechanisms and factors that influence this performance, and to potentially tune the parameters of locomotion simulators to enhance this performance. In Exp-4, Exp-5 and Exp-6, we applied the efficiency metric (see Equation 5.2) to the analysis of the results for this purpose. Other interesting results in these experiments are also shown and discussed in this chapter.

AS shown in Figure 6.1, analysis was carried out for all experiments (except Exp-2 and Exp-3) to consider the proportions of time for subjects’ fixation on targets and non-targets. Targets include coloured balls that were required to be counted by the subject in the task. Non-targets included other balls and all other parts of the scene, such as sections of floor, wall and the other balls which were not required for counting.

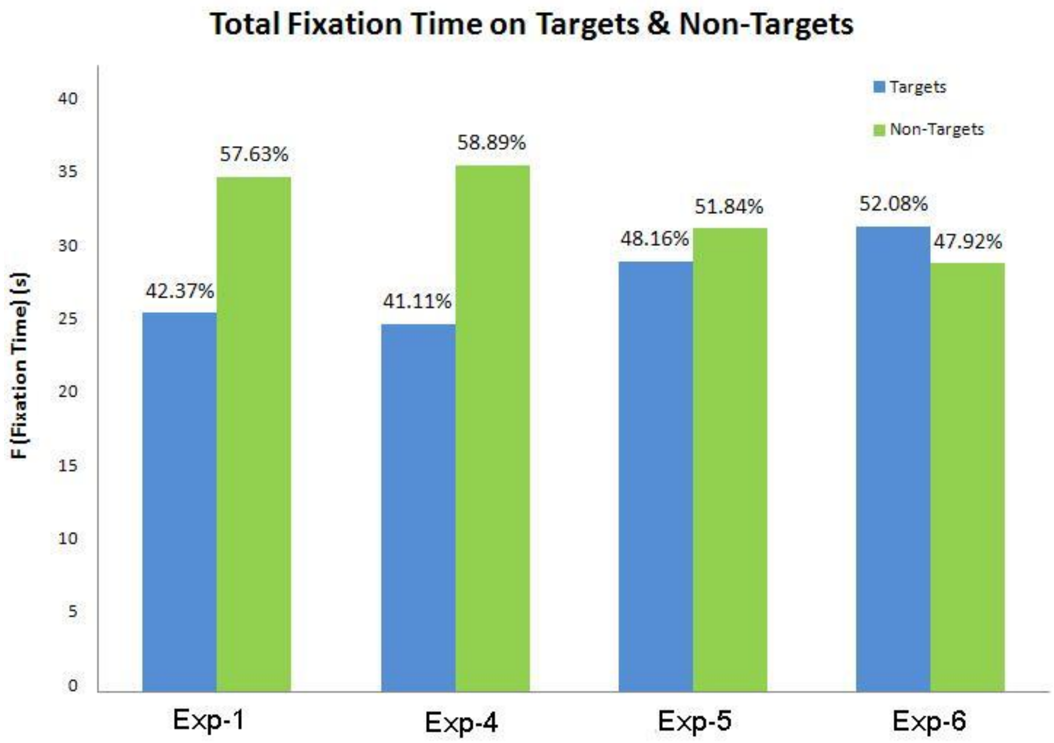

Figure 6.1 The average total fixation time for all subjects, and relative percentages, as directed to targets and non-targets in Exp-1, Exp-4, Exp-5 and Exp-6 (see Text).

As shown in Figure 6.1, in general subjects on average spent approximately 40% - 53% of their time looking at required targets. The reminder of the analysis will concentrate entirely on these required target times and corresponding saccades counts.

6.1 Exp-1 Results

--- Comparing Eye Movements during Walking in RE & SE

The aim of Exp-1 is to compare eye movement performance between real and simulated environments in terms of the numbers of task-relevant saccades (S) and time fixation on ROI or target balls (F) while doing an SAC task during walking.

Total Number of Saccades in Exp -1

The total number of saccades (S) between the ROI in the different conditions and environments are as shown in Figure 6.2.

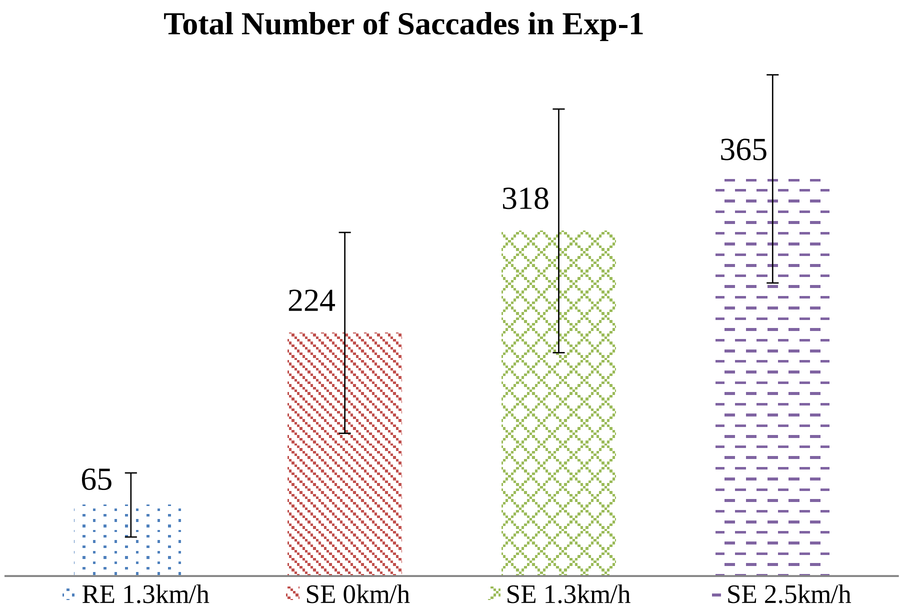

Figure 6.2 Total number of saccades (S) in each condition in the simulated environment (SE) while performing the search and count (SAC) task with a constant speed video, compared to the real environment (RE) with natural viewing. Error bars represent ± 1 standard deviation of the mean.30 different subjects were participated in each environment.

Results revealed that, firstly, that despite equivalent visual angles and the presence of walking in both environments, the numbers of saccades in SE are approximately five times higher than in RE while walking at the same speed. When walking at 0 km/h (i.e. standing) in SE, the numbers of saccades are approximately four times higher than in RE; when the walking speed increased to 2.5 km/h, the numbers of saccades are about six times higher than in RE. Secondly, within the SE, the number of saccades applied to perform the SAC task appears to be higher while walking even through the video speed is constant. One way ANOVA shows that S has significant differences between all of the groups in the SE ($F(2, 27)=7.137$, $p = 0.003$). Then use the Post Hoc Test to find where the differences are between these SE groups. It has been found that S increases significantly between 0 and 1.3 km/h ($p = 0.036$) and 0 and 2.5 km/h ($p = 0.01$), despite no change in video speed; no significant difference is seen between 1.3 and 2.5 km/h (p = 0.312). This is the case even though the visual task is also equivalent, as the video is played at a constant speed at all treadmill speeds. This experiment demonstrates that subjects have much higher number of saccades when performing the SAC task in SE than in RE; also, in SE, as the walking speed increasing, the number of saccades will increase with it.

Time Spent on each ROI & Total Time of Fixation in Exp-1

In Exp-1, one interesting result is the measurement for the time spent on each region of interest (ROI) in each condition. The walking corridor was divided to 22 interest regions in this task; each region of interest has 1-2 objects in it, as shown in Figure 6.3

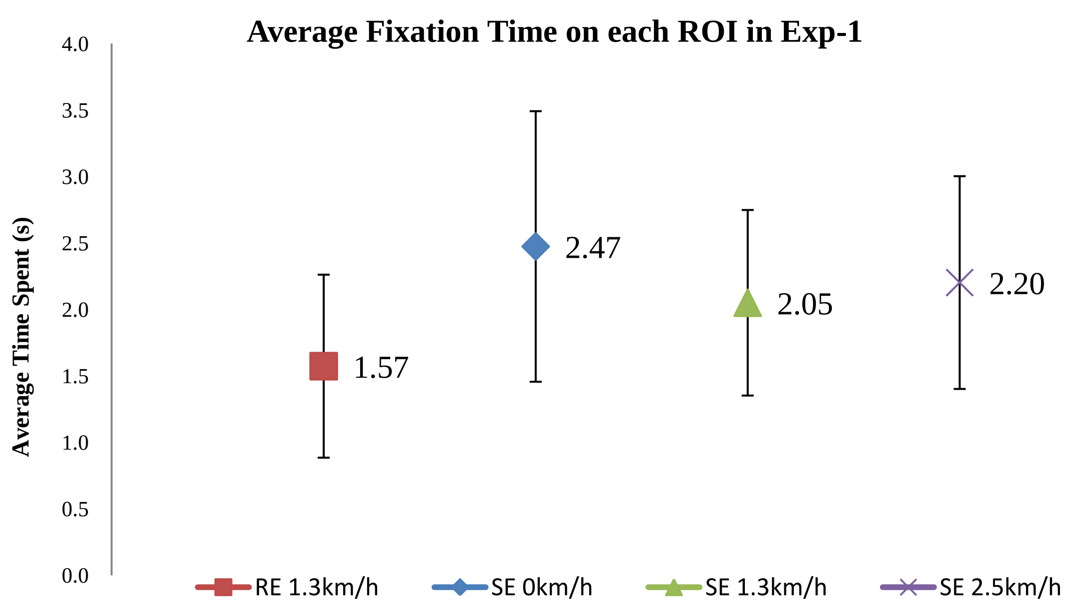

Figure 6.3 Average time spent on each ROI in each condition. The red square line is in the real world walking condition; the blue diamond line is the time spent on ROI at 0 km/h in SE; the green triangle line at the 1.3 km/h walking speed condition in SE and the purple cross line represents the time spent on each ROI at 2.5 km/h in SE. Error bars represent ± 1 standard deviation. 30 different subjects participated in each environment condition.

The average time spent on the ROI in the real world environment is 1.57s; in SE, the average fixation time on targets at 0 km/h is 2.47, the fixation time at 1.3 km/h is 2.05, and when at 2.5 km/h, the time spent is 2.2. Statistical analysis using One-way ANOVA, shows the difference between each condition is significant ($F(3, 84) = 4.753$, $p = 0.004$). But from the Post Hoc Test analyse, it shows no significant differences were observed between 2.5 km/h and either 0 km/h ($p = 0.271$) or 1.3 km/h ($p = 0.537$) in SE. And also there is no statistical difference between SE 1.3km/h and either RE 1.3km/h ($p = 0.055$) or SE 0 km/h ($p = 0.087$). But it has found there are statistical differences between 1.3 km/h in RE and either 0 km/h ($p < 0.001$) or 2.5 km/h ($p < 0.012$) in SE. This shows that subjects have averagely less fixation time on the ROI in RE than when walking in SE, which also means when subjects were doing the SAC task in the real environment, they spent much less time on searching for the targets. These results show that in both environments, even though subjects have been provided the same angle of view, the same length of walking and the same task, subjects in the RE walking group can still complete the task quicker and easier than those subjects who walked in the SE.

Total Fixation Time in Exp-1

Another interesting result is measuring subjects' fixation time on all of the targets in each walking condition. Total time of fixation is the total time which subjects spent exclusively on the target when they were doing the SAC task.

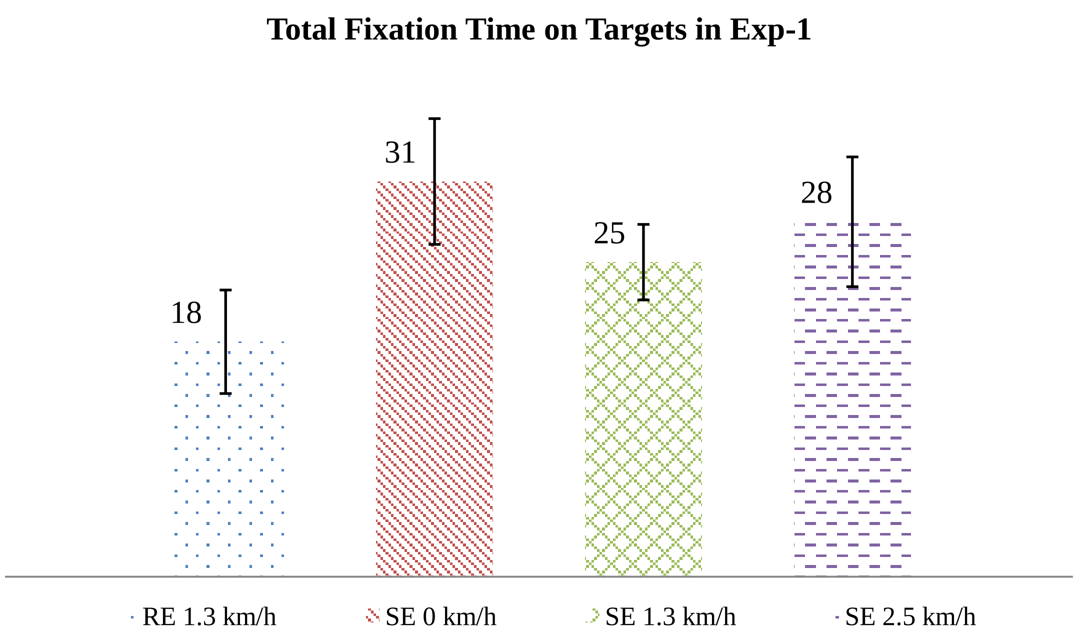

Figure 6.4 The total fixation time (F) for all the targets in different conditions. The blue bar represents the time spent in the real world walking at 1.3 km/h; the red bar is the time spent when walking speed is 0 km/h in SE; green bar is at the 1.3 km/h walking speed condition and the purple represents the time spent at the 2.5 km/h speed. Error bars represent ± 1 standard deviation. 30 different subjects participated in each environment condition.

The total time spent on the targets in the real world environment is 18s; in SE, the fixation time on targets at 0 km/h is 31s, the fixation time at 1.3 km/h is 25s, and when at 2.5 km/h, the time spent is 28s. The statistical analysis using One-way ANOVA, shows the difference between each condition is significant ($F (3, 36) = 15.999$, $p< 0.001$). But from the Post Hoc Test analysis, it shows no significant differences were observed between 2.5 km/h and either 0 km/h ($p = 0.077$) or 1.3 km/h ($p = 0.126$). But it has found there are statistical difference between 1.3 km/h in RE and 0 km/h ($p < 0.001$) or 1.3 km/h ($p = 0.006$) or 2.5 km/h ($p <0.001$) in SE, and also between 0 km/h and 1.3 km/h ($p = 0.003$) in SE. These findings show that subjects have less fixation time in RE than they had in walking in SE, which also means when subjects were doing the SAC task in the real environment, they spent much less time on searching for targets. These results lead to the conclusion that in both environments, even though subjects have been provided with the same angle of view, the same length of walking and the same task, subjects in the RE walking group can still complete the task quicker and easier than those subjects who walked in the SE.

The measure for the average fixation time on each ROI is very similar to the measurement of total time of fixation on the target, and also the measurement of fixation time on each target is more precise than just measuring the ROI, therefore, we will only measure the total time of fixation on each target from now on.

Target Counting Accuracy in Exp -1

In this SAC task, all the subjects were asked to count the numbers of red tennis balls (8 balls in total), yellow ping pong balls (3 balls in total) and blue ping pong balls (4 balls in total). The results of counting accuracy when counting the targets were different between each condition (see Figure 6.5).

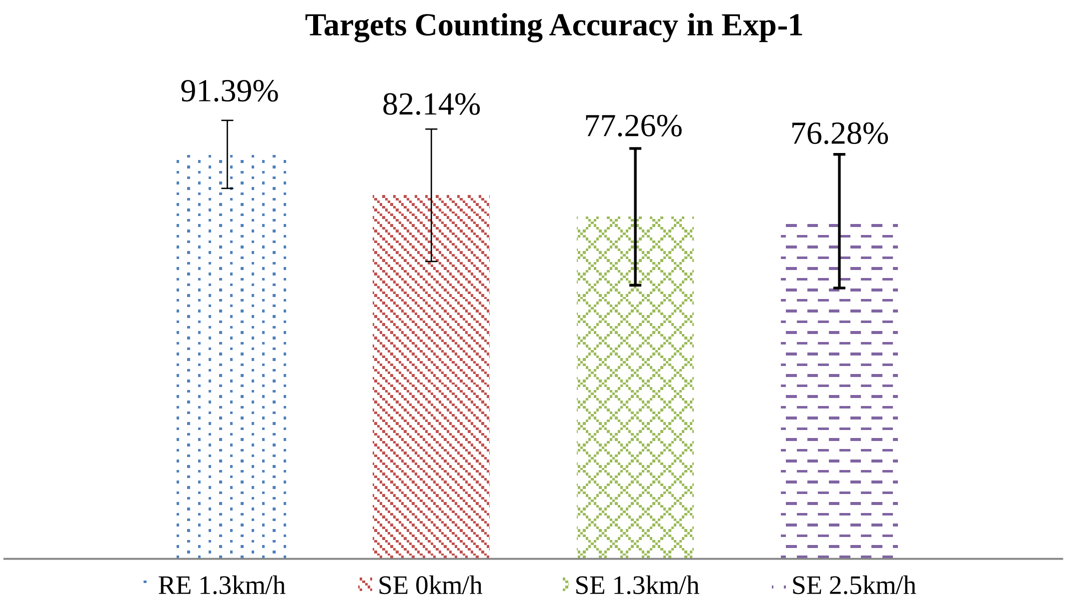

Figure 6.5 Counting accuracy in different environments and conditions. The blue bar represents the counting accuracy in real world walking; the red bar is the counting accuracy when walking speed is 0 km/h in SE; green bar is at the 1.3 km/h walking speed condition and the purple bar represents the counting accuracy at the 2.5 km/h speed. Error bars represent ± standard deviation. 30 different subjects participated in each environment condition.

Subjects in the real world have the most accuracy for counting targets, at 91.39%. Subjects in the video environments achieved less counting accuracy than in the real world. The group at 0 km/h walking speed scored 82.14% of accuracy; groups at 1.3 km/h and 2.5 km/h speed got even less accuracy, these scores were 77.26% and 76.28%. Statistically, counting accuracy in the real world is only similar with the stationary condition (0 km/h walking speed) in SE ($F(3,36) = 2.224$, $p = 0.175$). The counting of accuracy in three different conditions in SE are comparable with each other ($F(3,36) = 2.224$, $p > 0.05$).

These results show that within these four conditions, natural (RE) counting accuracy is better than all SE walking conditions (except 0km/h walking condition).

6.2 Exp-2 Results

--- Eye Movements while Walking in SE with Static Visual Stimulus

Exp-1 demonstrated a higher numbers of saccades while performing the SAC task in the SE compared with in the natural environment, even though the task, the viewing angles and the speeds of walking were equivalent. In addition, however, a change in walking speed also resulted in a change of total saccades performed in SE. In Exp-2 therefore, we assessed whether there are non-task dependent effects on eye movements that might be due to the action of locomotion itself. We measured the number of saccades when subjects viewed a static screen (see Figure 5.7) and instructed simply to look at the display while walking at the appropriate speeds. The SAC task was therefore removed, and any requirements associated with it should therefore not contribute to eye movements.

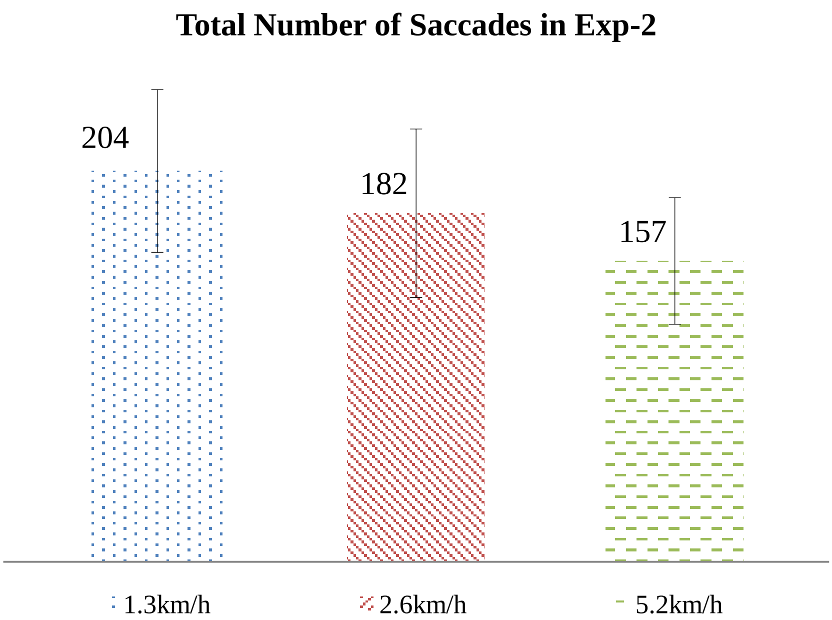

Figure 6.6 Number of saccades at each walking speed in the no-video condition in SE. Subjects viewed a static scene with no instructions to perform a specific visual task. Error bars represent ± 1 standard deviation of the mean. A total of 10 subjects participated in this experiment.

Repeated Measurement ANOVA shows that S decreases significantly between these three groups ($F(2, 18)=5.541$, $p = 0.013$). Furthermore, by using LSD methods for the Post Hoc Test to compare each group, there are no significant differences observed between 2.6 km/h and either 1.3 km/h ($p = 0.124$) or 5.2 km/h ($p = 0.74$); the total number of saccades only decreased significantly between 1.3 km/h and 5.2 km/h ($p = 0.02$). These results reveal that when the visual tasks and indeed all video motion are removed, the number of saccades is nevertheless still affected by walking speed. But, in contrast to the results of Exp-1, saccades appear to *decrease* with increasing walking speed. Had saccade counts increased with walking speed in this experiment, the conclusion would have been that locomotion itself was implicated in perhaps adding additional eye movements regardless of the visual task. However, a decrease in saccade numbers with locomotion speed; as was indeed found, would indicate that a mechanical jitter or a possible vestibular influence on saccades, this effect cannot be explained - walking speed in itself does not apparently add additional saccades.

Where subjects look most when they walk

When subjects stand on the treadmill, their eyes were approximately at the same height level as the centre of the top region of the screen, the visual gaze direction angle on each region of screen (top, middle and bottom) is $\Theta_{1,}$ Θ_2 and Θ_3 and which is shown as in Figure 6.7.

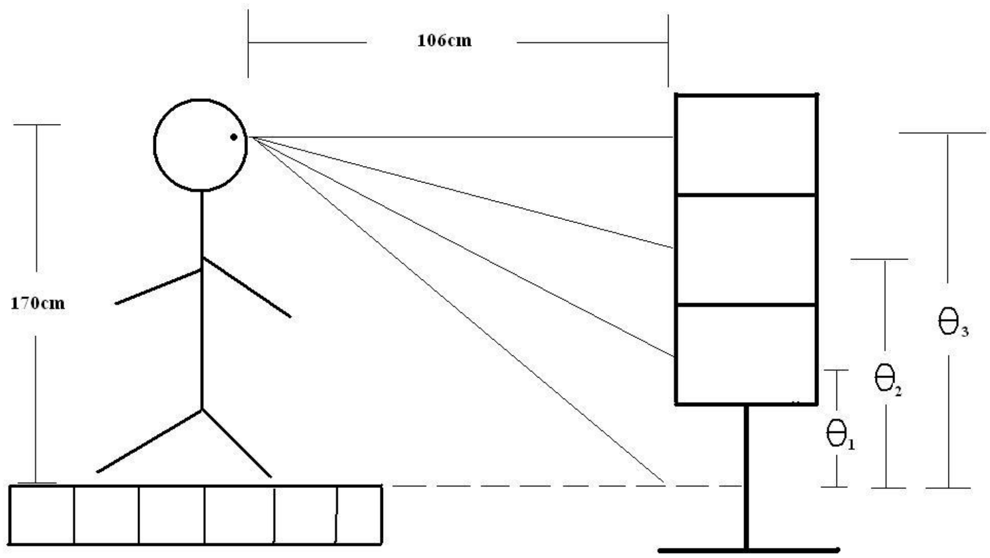

Figure 6.7 A subject's visual gaze direction angles when looking on the static screen while walking in Exp-2. Θ_1 represents the vertical visual gaze direction angle when subject is looking at the top region of screen, measured from the centre of the top region of the screen to subject's feet level; Θ_2 and are the visual gaze direction angle when subject is looking at the middle region of screen, measured from the centre of the middle region of the screen to the subject's

feet level; Θ_3 is the visual angle which from the centre of the bottom region of the screen to the subject's feet level when subject is looking at the bottom region of the screen. A total of 10 subjects participated in this experiment.

The visual gaze direction angle was calculated by using trigonometric methods. The height of the display is 125cm (49.4 inch), so the distance between the centre of each region is 41.7cm; the distance between subject to screen is 106cm (see chapter5.6); and the distance from the centre of the top region of the screen to the subject's feet level is approximately 170cm. Hence, subject's visual gaze direction angles are shown in Equation 6.1.

$$\Theta_3 = \tan^{-1}(170 \ / \ 106) = 58.1°$$

$$\Theta_2 = \Theta_3 - \tan^{-1}(41.7 \ / \ 106) = 36.6°$$

$$\Theta_1 = \Theta_3 - \tan^{-1}(41.7 \times 2 \ / \ 106) = 19.9°$$

Equation 6.1 Subject's visual gaze direction angles were calculated by using trigonometry.

An interesting additional result for Exp-2 is the spatial distribution of total viewing time across the static image on the display.

0.1% 0.1% 0%	0.08% 0.18% 0%	0.16% 0.18% 0%
10.41% 14.13% 17.17%	16.17% 14.15% 11.84%	6.84% 4.52% 6.19%
12.91% 14.63% 8.5%	36.12% 35.4% 33.63%	17.21% 16.71% 22.67%

Table 6.1 Percentage of total viewing time in the regions of the static screen image while walking at speeds of 1.3 km/h (on the top of each region), 2.6 km/h (middle of each region) and 5.2 km/h (bottom of each region), averaged across all subjects participating in Exp-2. A strong bias is clearly observed towards the lower and vertically central regions. The grey level (colour) value of each region is selected as approximately proportional to the average of the total viewing times for the three speeds for that region, with white representing the highest total duration.

As shown in Table 5.1, it is clear that there is a significant tendency for subjects to spend the largest proportion of their time looking at the lower regions of the display. There is no apparent bias, on average across all subjects, towards either the left or the right sides of the display but instead, of the nine possible regions, subjects remain looking at the vertically central and lower regions, with an average 35% in region 8, the lower central region, alone (see Figure 5.7 for region numbers). There were no significant differences ($F(2, 16)=0.000$, $p = 1$) in the above distribution of looking times between the different walking speeds. This finding demonstrates that when walking in the SE without visual input, subjects tend to look down and with a visual angle of 11.5 degrees on the screen.

6.3 Exp-3 Results

--- Matching Locomotion Speed to Optic Flow Speed in SE

The aim of this experiment is to observe whether there are differences between subjects' locomotion perception and video playing speed in SE, and if the differences exist, then to establish if the differences could be reduced following several episodes of training.

This experiment was run at four different video speeds at 0.7, 2.0, 3.3 or 4.6 km/h in the SE. Every subject had five attempts to set their walking speed for each of the four clips. The differences between their perception of locomotion speed and video speed were measured during the task.

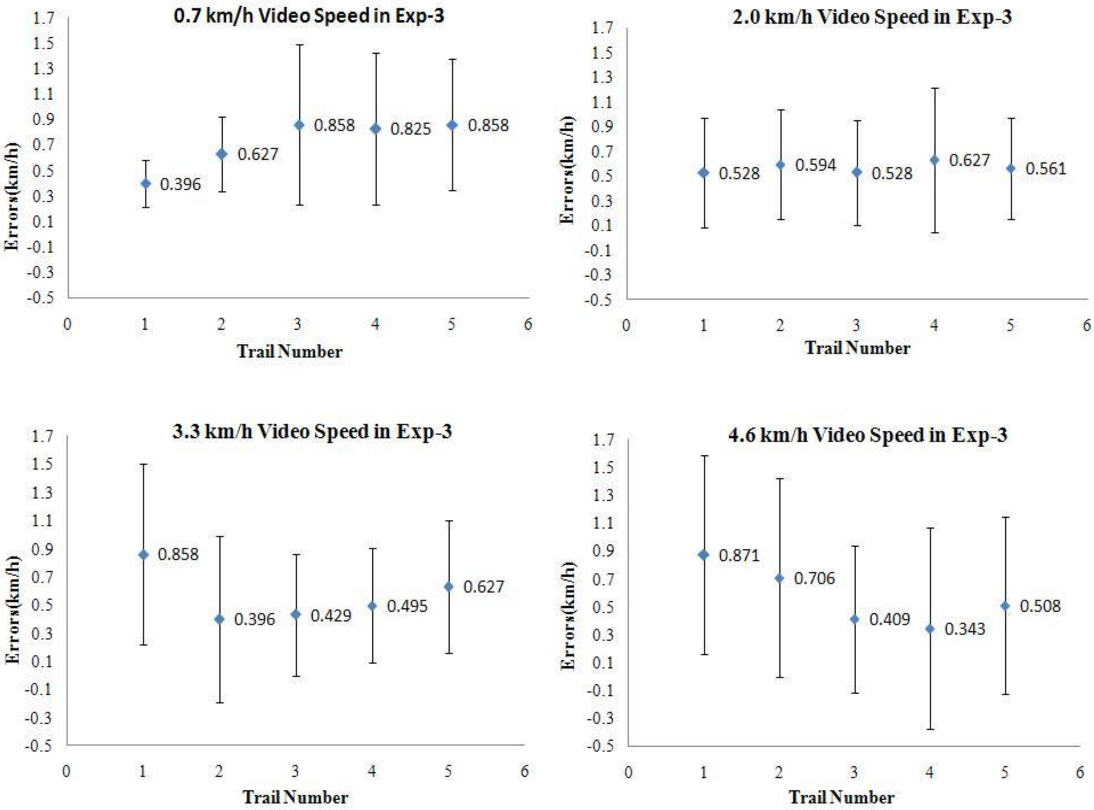

Figure 6.8 Errors between locomotion perception and video speed at 0.7, 2.0, 3.3 and 4.6 km/h in SE. Error bars represent ± 1 standard deviation of the mean. 5 subjects participated in this experiment.

From Figure 6.8, results show the errors between locomotion perception and video speed. In One-Way Repeated Measure ANOVA, it was shown that the error wasn't eliminated or reduced with the more times of training when the video speeds were at 0.7km/h ($F(4, 16)=1.826$, $p = 0.173$), 2.0km/h ($F(4, 16)=0.65$, $p = 0.991$) and 3.3km/h ($F(4, 16)=1.193$, $p = 0.352$). However, the errors were significantly reduced after a few times training when the video speed was 4.6km/h, which ($F(4, 16)=4.617$, $p = 0.11$). These findings reveal that there is no apparent learning effect on subjects' locomotion perception when subjects aimed to match their walking speed to video speed at 0.7, 2.0 and 3.3km/h, and especially in the slowest video speed (0.7 km/h), subjects even have poorest performance after more times training. When video speed was 4.6km/h, subjects' performance of matching walking speed to video speed was getting better with training.

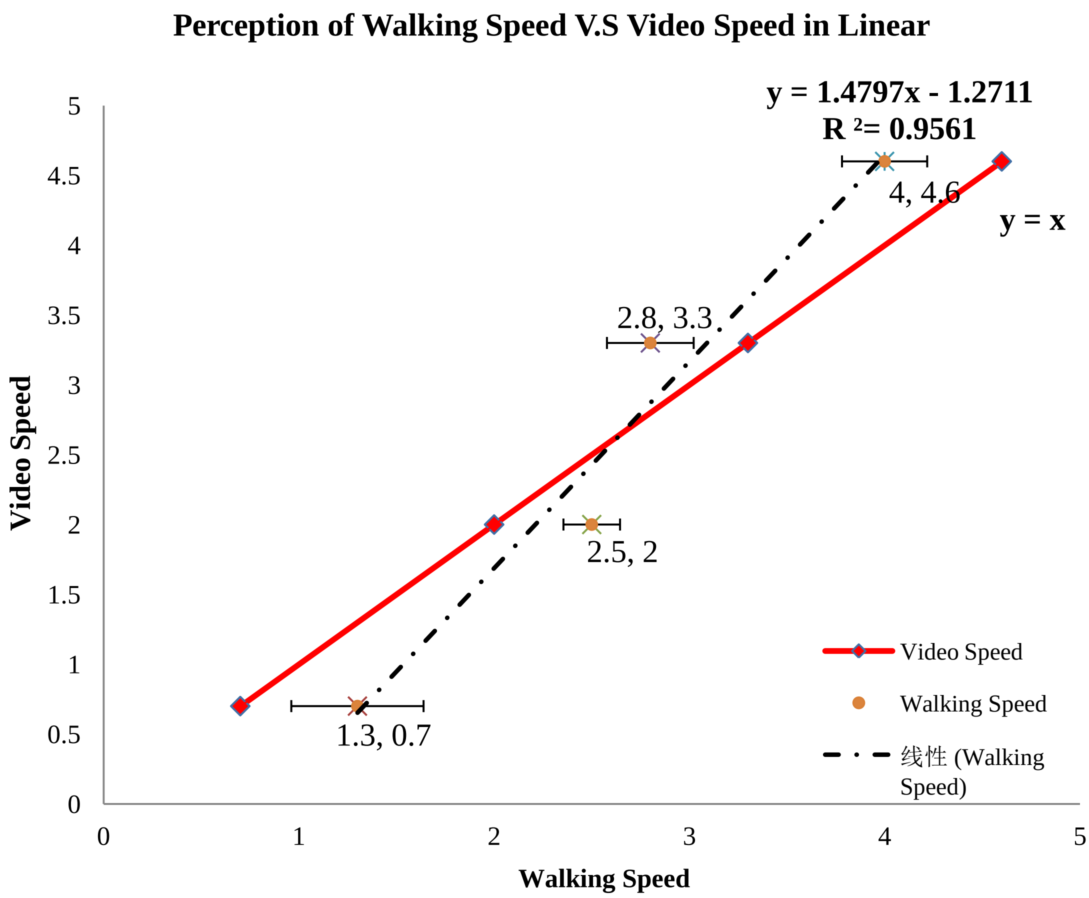

Figure 6.9 Comparing the perception of locomotion speed with video speed in SE. The red line represents the four video playing speeds; the orange dots represent subjects' actual walking speed; the black dashed line is the linear trend line of subjects' walking speed. Error bars represent ±1 standard deviation of the mean. Presently the dependant variable on the axis is less common but perhaps more intuitive to understand in this case. 5 subjects participated in this experiment.

If subjects matched the video speed perfectly with their walking, then all the data points would lie on the line of video speed (the red linear line). But in Figure 6.9, when video playing speed is at 0.7 km/h or 2.0 km/h, the corresponding locomotion speed is 1.3 or 2.5 km/h, which is faster than the video speed. When video speed is at 3.3 or 4.6 km/h, the locomotion speed of subjects is 2.8 or 4.0 km/h, which is slower than video speed. The black dashed linear trend line for locomotion speed is y = 1.4797x – 1.2711, which intersects with the line of video speed (y = x) when both video speed and walking speed are approximately 2.65 km/h. The R square value for the linear trend line is 0.9561, which means that the trend line is fitted to the data very closely. By analysing the linear trend line and the video speed line, it shows that in SE, when locomotion speed is less than 2.65 km/h, subjects overestimate the video speed, which means the video speed in subjects' perception is faster than its real speed; therefore, this mis-perception leads subjects to walk faster than the video playing speed. When the locomotion speed is higher than 2.65 km/h, subjects underestimate the video playing speed; this mis-perception makes subjects walk slower than the video speed.

6.4 Exp-4 Results

--- Eye Movements Measure while Walking in RE and Standing in SE

A group of subjects first walked in RE at speeds of 2.1, 2.8 or 3.8 km/h; in the SE, subjects merely stood while viewing the videos at the equivalent video speeds to the exposure conditions experienced by the RE subjects.

Total Number of Saccades in Exp-4

The results in Figure 6.10 reveal that S for three different video speeds for standing subjects in SE are approximately two and half times higher than walking in the three different speeds in RE. In RE, no significant differences were seen between walking speed 2.1, 2.8 and 3.8 km/h ($F(2, 12)=0.928$, $p = 0.422$), as shown in Figure 6.10.

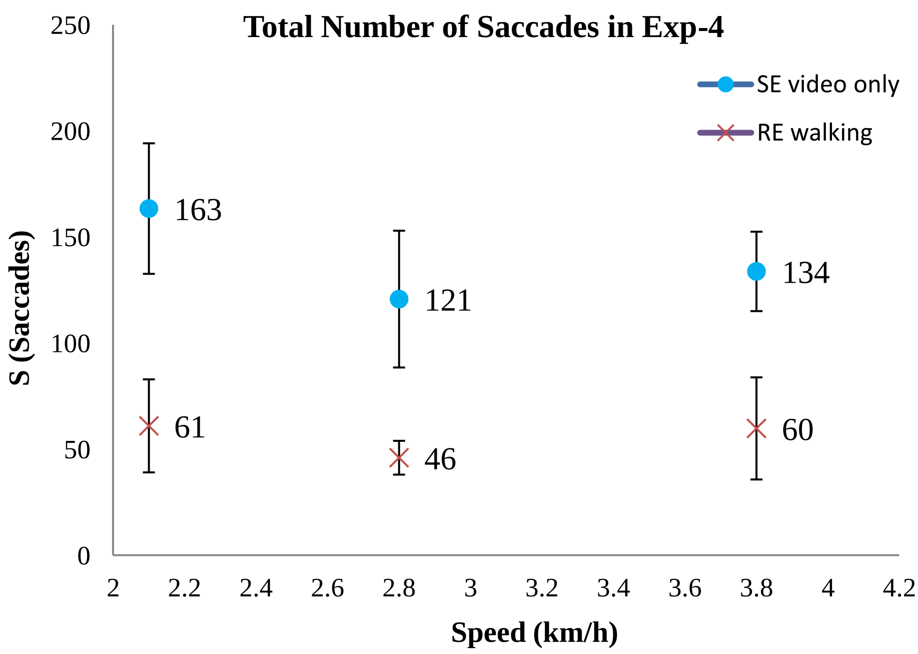

Figure 6.10 Number of saccades shown for three different video speeds for standing subjects in SE and walking at three different speeds in RE in Exp-4. Error bars represent ± 1 standard deviation of the mean. 15 subjects participated in each environment in this experiment.

In SE condition, One-Way ANOVA shows that there was also no significant difference when the video was played between 2.1, 2.8 and 3.8 km/h speeds ($F(2, 12)=2.919$, p =

0.096). This result reveals that when subjects stand in SE, S does not change as the changing of video speed. It also means that the video speed does not affect subjects' saccades in this condition.

Total Fixation Time in Exp-4

After comparing the S, we also measured the total time of fixation on all targets when subjects doing the SAC task in both RE and SE. Figure 6.11 shows that at each different condition, the total time of fixation in RE are slightly lower than in SE; and as speed (both RE walking speed and SE video playing speed) increase, the fixation time in SE appears to approach level of RE. Statistically, Independent-Sample T-Test results shows there are no significant difference between RE and SE groups when both groups of subjects at same speed at 2.1 km/h (p = 0.177), 2.8 km/h (p = 0.108) and 3.8 km/h (p =0.613).

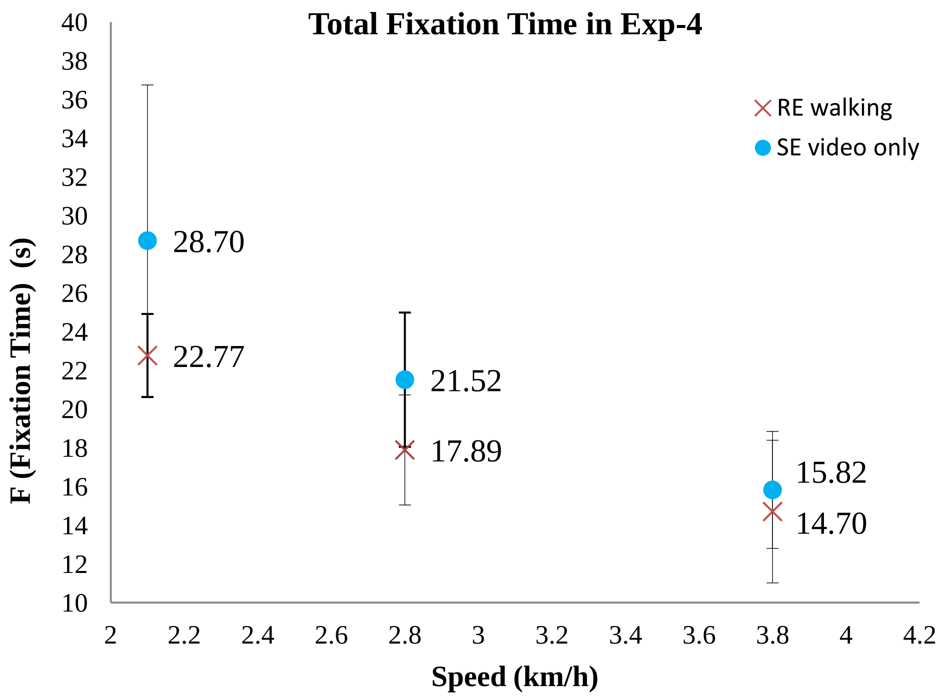

Figure 6.11 The total fixation time, F, for all the targets in the different conditions. The blue round dots represent the time spent at different video speed when subjects were standing in the SE; the red cross represents the fixation time at different walking speed in RE. Error bars represent ±1 standard deviation. 15 subjects participated in each environment in this experiment.

In the real environment, when walking speed was 2.1 km/h, the total time of fixation is 22.77s, but it decreased to 14.7s as speed was increased to 3.8 km/h. One-Way ANOVA shows the total fixation time between these groups has significant difference ($F(2,12) = 9.417$, $p = 0.003$). Through the Post Hoc Test for these groups, the differences occur between 2.1 km/h and either 2.8 km/h ($p = 0.023$) or 3.8 km/h ($p = 0.001$), but no significant difference between 2.8 km/h and 3.8 km/h walking speed ($p = 0.115$).

In the simulated environment, the total fixation time was 28.7s when the video speed at 2.1 km/h, and decreased to 15.82s when video speed increased to 3.8 km/h. In One-Way ANOVA, it shows the total time of fixation has significantly decreased between these speeds ($F(2,12) = 7.28$, $p = 0.009$). Further, the Post Hoc Test shows, the differences were occurred between 2.1 km/h and 3.8 km/h ($p = 0.002$), but no significant difference between 2.8 km/h and either 2.1 km/h ($p = 0.55$) or 3.8 km/h video playing speed ($p = 0.118$).

These results reveal that in both RE and SE, while the number of saccades does not change, the total time of fixation is affected by the video (and also locomotion speed in RE), which decreases as the video speed increase.

Target Counting Accuracy in Exp -4

In this SAC task, all the subjects were asked to count the numbers of yellow tennis balls (7 balls in total), orange ping pong balls (3 balls in total) and blue ping pong balls (5 balls in total). The results of counting accuracy when counting the targets are different between each condition (see Figure 6.12).

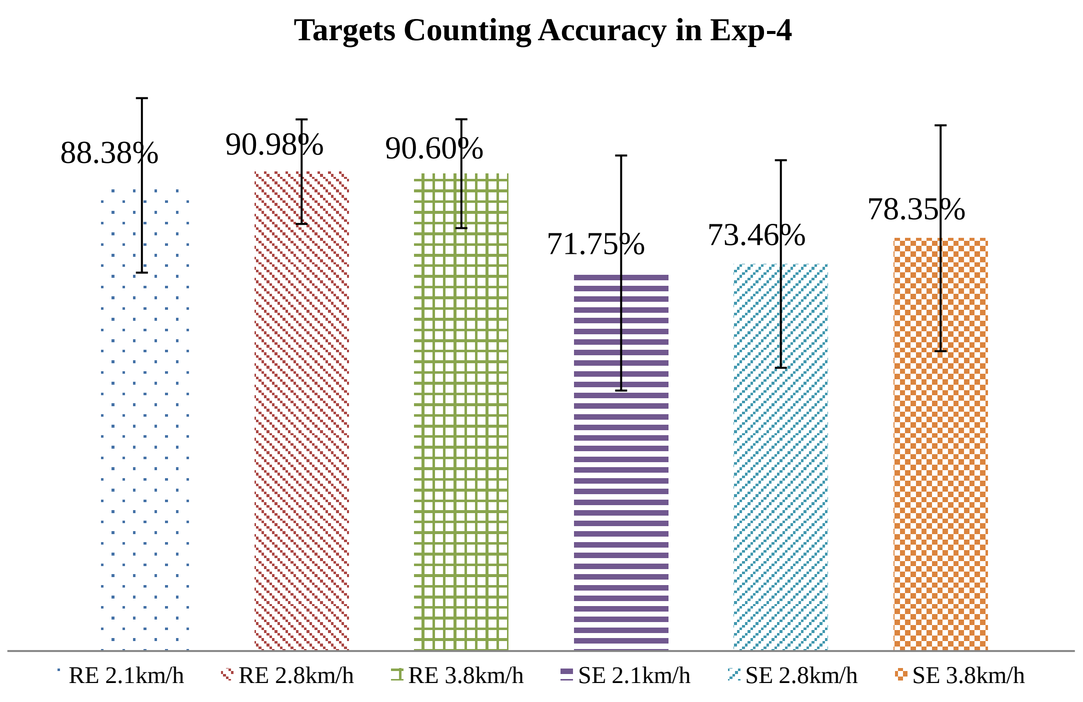

Figure 6.12 Counting accuracy in different environments and conditions. The dark blue bar represents the counting accuracy at 2.1 km/h in the real world walking; the red bar is at 0 km/h in RE; green bar is at the 3.8 km/h walking speed in RE. The purple bar is video speed at 2.1 km/h SE; the light blue bar is at 2.8 km/h in SE; orange bar is at the 3.8 km/h video speed in SE. Error bars represent ± standard deviation.15 subjects participated in each environment in this experiment.

Subjects in the real world achieved more accuracy for counting targets (p = 0.017), which on average is 89.99%, while only 74.14% on SE on average. In RE, statistically, counting accuracy is similar between these three ($F(2,12) = 0.067$, $p = 0.936$). Also, the counting accuracy in three different speed in SE are comparable with each other ($F(2,12) = 0.124$, $p = 0.884$).

These results reveal that within these two environments, natural counting accuracy is better than counting in the simulator; but the counting results have no significant differences between groups in either RE or SE. Therefore, we found generally, walking doesn't impair the counting accuracy performance in the SAC task in RE; and also, when the walking speed is 0 km/h in SE, the counting accuracy isn't affected by the video speed.

6.5 Exp-5 Results

--- Eye Movement Performance at Matched Locomotion and Optic Flow Speed in SE

Total Number of Saccades in Exp-5

Figure 6.13 shows the Total number of saccades, and this is a main factor for calculating the efficiency. In Exp-5, S were measured from new runs using four matched walking and video speed combinations at 0.7, 2.0, 3.3 or 4.6 km/h in the SE.

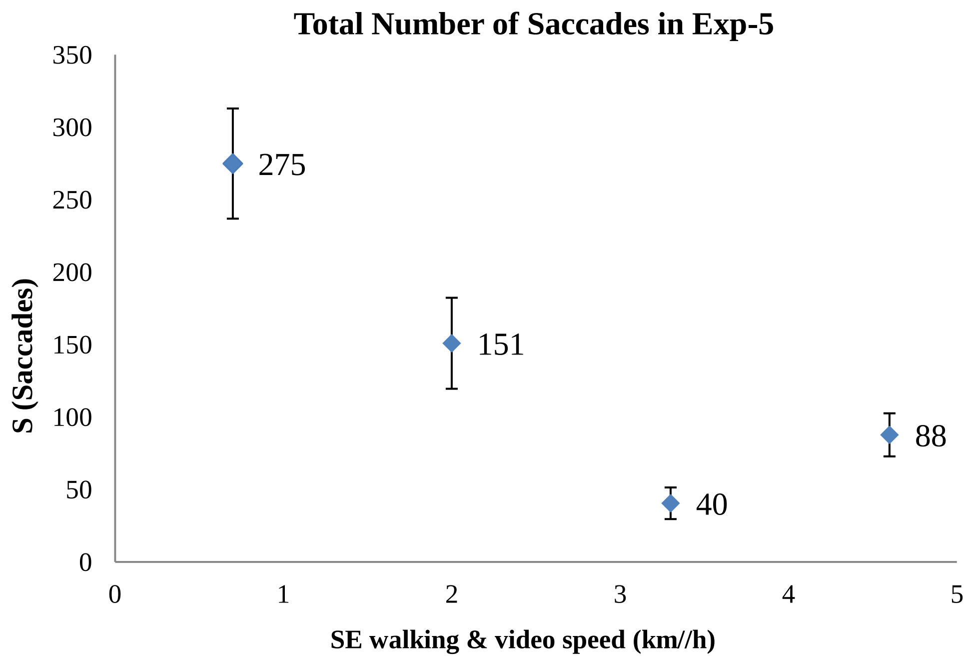

Figure 6.13 Total number of saccades (S) shown at four matched speeds in SE in Exp-4. Error bars represent ± 1 standard deviation of the mean. A total of 5 subjects participated in this experiment.

One-Way Repeated Measures ANOVA shows that there were significant differences between these groups at different speeds ($F(1.182, 4.73)=77.799$, $p < 0.001$). Also, LSD Post Hoc Test shows that the S decreased significantly when walking and video playing speed were increasing from 0.7 to 2.0 km/h ($p = 0.008$), and then from 2.0 to 3.3 km/h ($p < 0.001$). But when this matched speed increased from 3.3 to 4.6 km/h, the S were also increased with it ($p < 0.001$).

Results found that the S was affected by the matched walking and video speed, and subjects had the lowest number of S when at the intermediate speeds of matched walking and video of 2 – 4 km/h.

Total Fixation Time in Exp-5

After comparing the S, the total time of fixation on all targets when subjects doing the SAC task in SE were also measured. Figure 6.14 shows that subjects spent the maximum time of fixation (36.392s) on average on all the targets when at 0.7 km/h walking and video playing speed. This is approximately five times longer than at 3.3 km/h speed (7.784s).

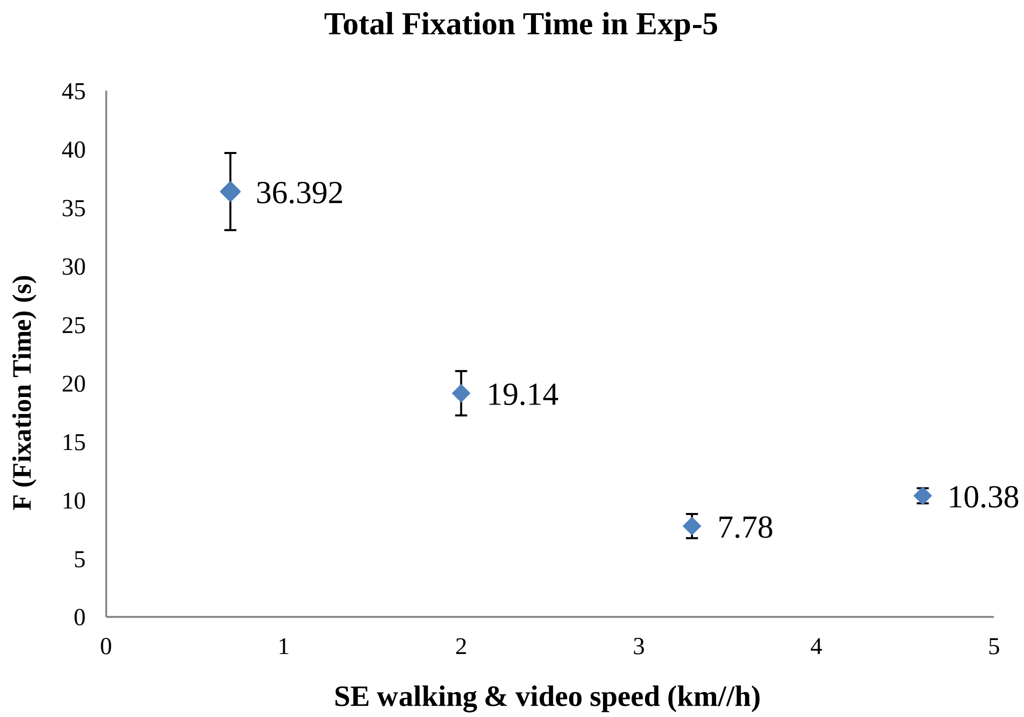

Figure 6.14 The total fixation time (F) at four matched walking and video speed combinations at 0.7, 2.0, 3.3 or 4.6 km/h in the SE. Error bars represent ± 1 standard deviation. A total of 5 subjects participated in this experiment.

One-Way Repeated Measures ANOVA shows that there were significant differences between these groups at different speeds (F(1.401, 5.603)=161.236, p < 0.001). Also, LSD Post Hoc Test shows that the total fixation time decreased significantly when walking and video playing speed were increasing from 0.7 to 2.0 km/h ($p = 0.001$), and then from 2.0 to 3.3 km/h ($p < 0.001$). But when this matched speed increased from 3.3 to 4.6 km/h, the total fixation time were also increased with it ($p = 0.002$).

Results found that the total fixation time was affected by the matched walking and video speed, and subjects have the least fixation time when at the intermediate speeds of matched walking and video of 2 – 4 km/h.

Targets Counting Accuracy in Exp -5

In this SAC task, all the subjects were asked to count the numbers of yellow tennis balls (3 balls in total), red tennis balls (3 balls in total) and blue ping pong balls (2 balls in total). The results of counting accuracy when counting the targets are different between each condition (see Figure 6.15).

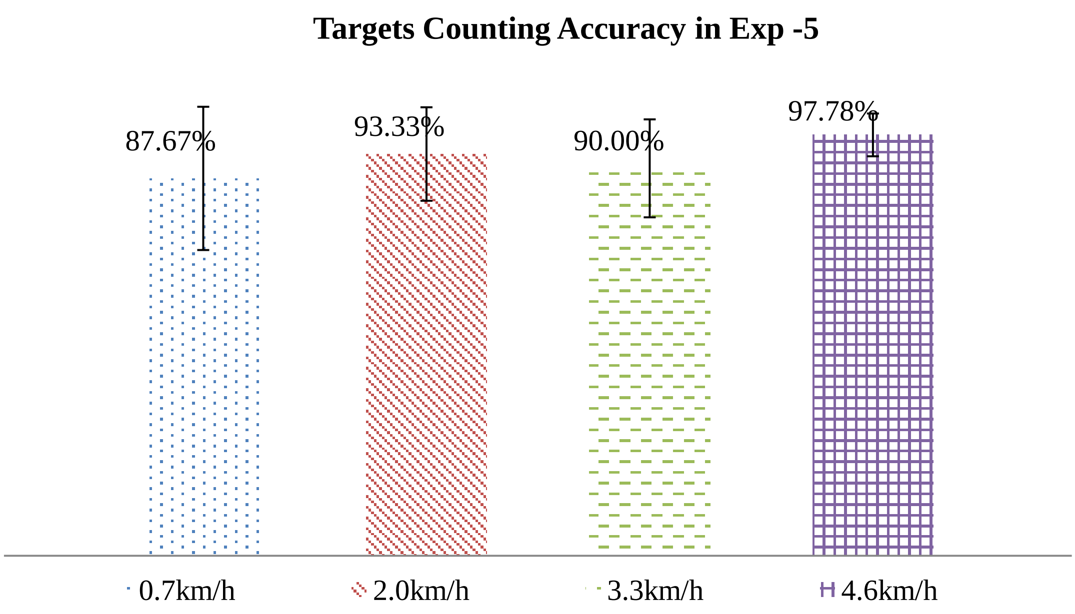

Figure 6.15 Counting accuracy at four matched walking and video speed combinations in SE. The blue bar represents the counting accuracy at 0.7 km/h; the red bar is at 2.0 km/h; green bar is at the 3.3 km/h walking speed and the purple bar is at 4.6 km/h SE. Error bars represent ± 1 standard deviation. A total of 5 subjects participated in this experiment.

Subjects at 4.6 km/h speed have the most accuracy for target counting, which is 97.78%, and it only has 87.67% of accuracy at 0.7 km/h speed. But statistically, counting accuracy is similar between all the groups ($F(3,12) = 0.702$, $p = 0.569$). This result reveals that the counting accuracy always remains at a similar level when at these four matched walking and video speed combinations in SE.

In addition, at these matched speed combinations, the average counting accuracy is higher in SE than in the video only (i.e. without walking) as in Exp-4 (see Figure 6.12).

Similar to previous experiments, the counting accuracy is very dependent on the quality of the video, the performance of the subjects ability to recognize and separate the two types of balls and the contrast of image etc., also the results for each group are very easily and strongly affected by every subjects' performance. All of these reasons make the results unreliable, therefore, target counting accuracy is just an interesting but minor result, therefore will not be used much to compare or further discuss in the results.

6.6 Exp-6 Results

--- The Learning Effect Investigation of SAC Task in SE

The results in Exp-1 and Exp-4 clearly show that performance in RE is much better than in SE. Experience in real life results in people being more familiar, more confident and more comfortable with their activities in the real environment than in the simulated environment. Therefore, one possible explanation for the previous experimental results may be that subjects have undergone a familiarity or learning in RE which we do not have in SE. In order to answer this question, this experiment is designed to discover the familiarity or learning effect by measuring the subjects' total number of saccades and total fixation time when they are walking in a familiar SE. In Exp-5, the result shows when walking at 3.3 km/h in SE, subjects have the best performance in terms of S (40) and T (7.78s). Therefore, this walking speed was chosen for this experiment.

Total Number of Saccades in Exp-6

Besides the efficiency, some other useful data can also be looked at in Exp-6 (see Figure

6.16), firstly the total number of saccades, S, as this is the main factor for calculating the efficiency. In Exp-6, subjects were asked to perform a SAC task five times at the same matched walking and video speed combinations, and the S was measured from each time of task completion in SE.

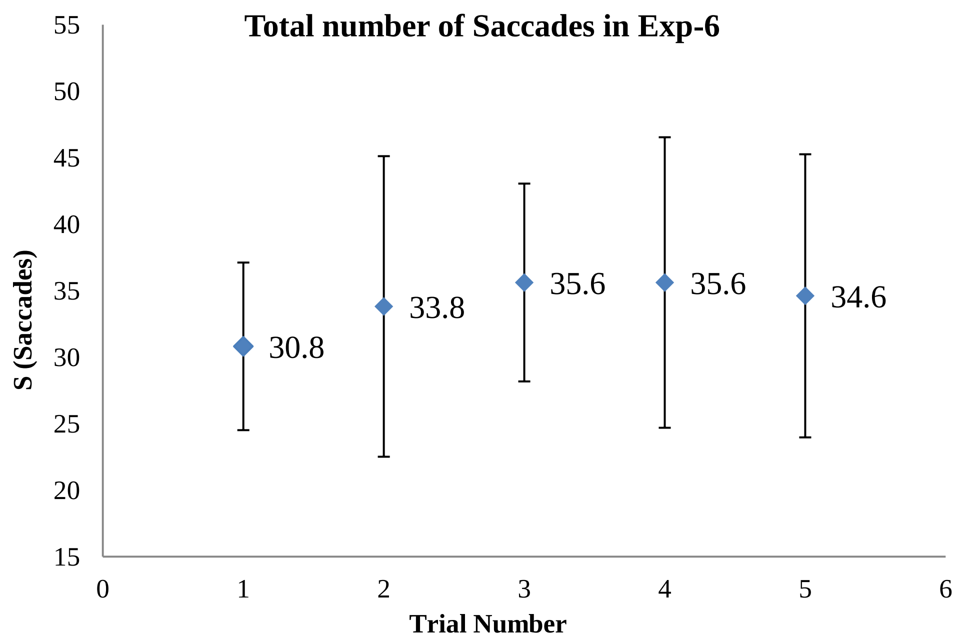

Figure 6.16 Total numbers of saccades (S) for walking at 3.3km/h matched speeds five times in SE. Error bars represent ± 1 standard deviation of the mean. 5 subjects participated in this experiment.

One-Way Repeated Measures ANOVA shows that there are no significant differences of S between each time of walking at 3.3km/h matched speeds in SE ($F(4, 16)=0.271$, $p = 0.892$). This reveals that in SE, even when subjects were trained 5 times in the same condition to increase their familiarity, but there was still no decrease for their S in this SAC task. We then can say that when doing the SAC task, subjects have much fewer saccades in RE than in SE, and this is not changed with increasing familiarity in that environment.

Total Fixation Time in Exp-6

After comparing the S, another result is measuring subjects' fixation time in SE. Total time of fixation is the total time subjects eyes were on the targets when they were doing the SAC task. Figure 6.17 shows the total fixation time on all the targets measured from the time each task was completed.

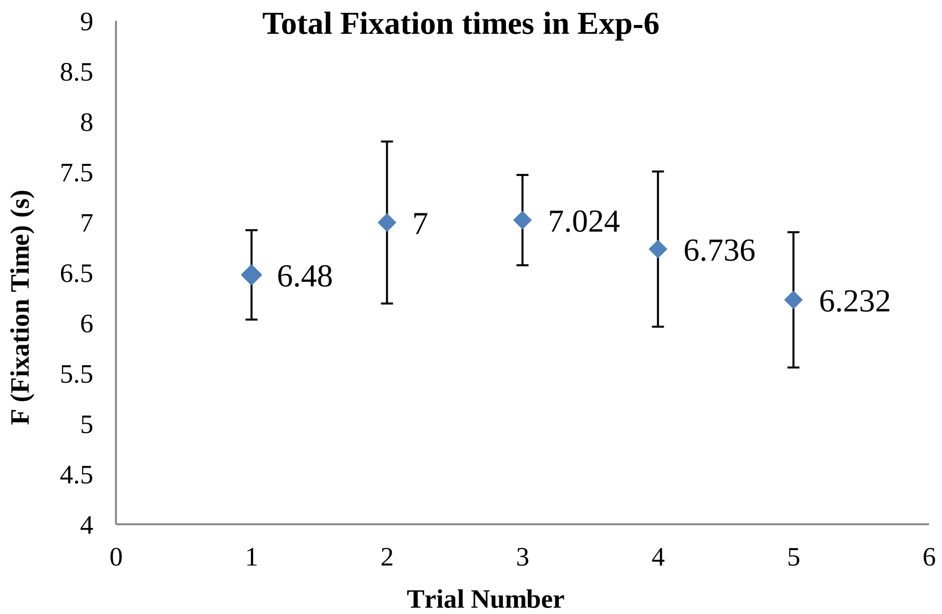

Figure 6.17 Total fixation time (F) for walking at 3.3km/h matched speeds five times in SE. Error bars represent ±1 standard deviation of the mean. A total of 5 subjects participated in this experiment.

One-Way Repeated Measures ANOVA shows that there were no significant differences between each task ($F(4, 16)=2.263$, $p = 0.108$). The results found that, even though subjects were trained 5 times in the same condition to increase their familiarity, there were still no decreases for their total fixation time in this SAC task. We then can say that when doing the SAC task in SE, the total fixation time is not affected by the familiarity level in that environment.

Targets Counting Accuracy in Exp -6

In this SAC task, all the subjects were asked to count the numbers of yellow tennis balls (3 balls in total), red tennis balls (3 balls in total) and blue ping pong balls (2 balls in total). The results of counting accuracy when counting the targets are different between each condition (see Figure 6.18).

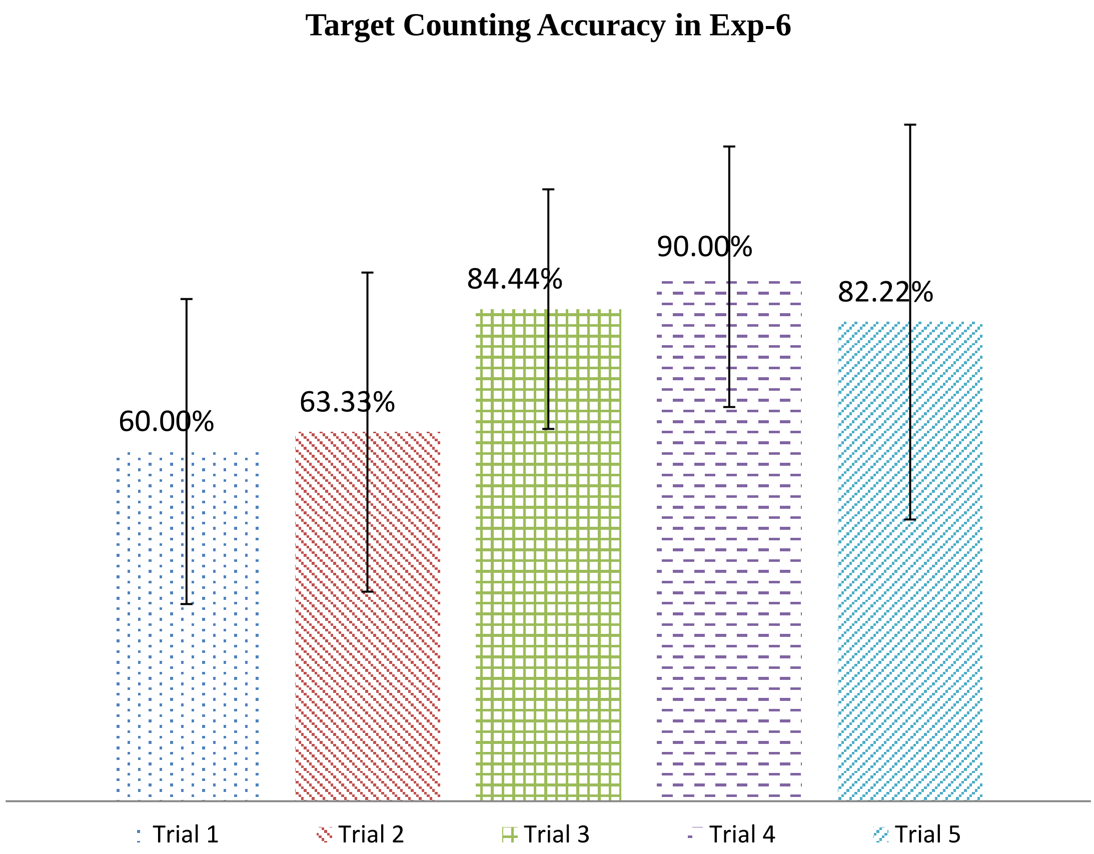

Figure 6.18 Counting accuracy for walking at 3.3km/h matched speeds five times in SE. The dark blue bar represents the counting accuracy at the first Trial of walking; the red bar is in second Trial; green bar is in Trial 3; the purple bar is in Trial 4 and the light blue bar in the walking in the fifth Trial. Error bars represent ± 1 standard deviation. A total of 5 subjects participated in this experiment.

Subjects who did the task at the fourth Trial have the most accuracy for targets counting, which is 90.00%, and it was only 60.00% of accuracy at the first Trial. As subjects walk more often, the accuracy is increasing from Trial 1 to Trial 5. But statistically, the counting accuracy did not increase very significantly between all the groups ($F(4,16) = 2.818$, $p = 0.06$).

Chapter 7. Discussion

This thesis measures eye movements to compare the real environment with a simulated environment during locomotion while doing a basic ‘Search and Count’ task (SAC) that uses several natural visual mechanisms. The SAC represents a realistic-world task which engages many brain mechanisms, and should thus help to explore and understand locomotion simulators in a more general and natural way. In addition to eye movements, the task also activates mechanisms for attention, cognition, search, colour and contrast. The task combines these mechanisms in the experiments with the aim of achieving the main objectives of the study 1) to develop and test methods for improving the design and configuration of motion capture in locomotion simulators (this objective includes the work reported in Chapter 3 as well as the approach taken in Chapters 4-5); 2) to improve the understanding of human interaction within locomotion simulator environments; 3) understand the differences and find the best conditions under which a locomotion simulator environment perform as similarly as possible to a real world environment; 4) Improve the understanding of the mismatch in speed between locomotion and video optic flow, and find the method which makes locomotion and optic flow speeds comparable.

When analyzing recorded eye movements, we look at S (total number of saccades in a trial), F (total fixation time in a trial) and efficiency (see Equation 5.1 and Equation 5.2).

Also, we used a video of the real environment (RE) for the simulated environment (SE), and although differences remain (e.g. spatial and temporal resolution, field of view size, focal distances, binocularity/stereopsis), we matched the visual angle, distances and speeds of locomotion so as to make the SE and RE conditions as similar as possible.

7.1 Eye Movement performance in Real Environment Versus Simulated Environment

The main result in this thesis is that, in general, the eye movement performance in the real environment is better than in the simulated environment. This is found in the total number of saccades (S), total fixation time (F) as well as in the efficiency (described in this chapter).

In Exp-1 and Exp-4, we compared subjects' task performance (measured by the total saccades, S, total fixation time, F and the counting accuracy) and found that there are differences between real and simulated environments. In Exp-1, we firstly found that S increases significantly with walking speed even though the video is set at the same speed (see Figure 6.2). The results show that, the numbers of saccades in SE are approximately five times higher than in RE while walking at the same speed. When subjects were standing in SE, the numbers of saccades are approximately four times higher than in RE; when the walking speed increased to the fastest speed (2.5 km/h), the numbers of saccades are about six times higher than in RE. In Exp-4, results showed that the average S for all three different video speeds for standing subjects in SE are approximately two and half times higher than all the walking in the three different speeds in RE (see Figure 6.10).Next, we found that subjects performed better in RE than SE in terms of the total time of fixation, F, which is always lower in RE. Results in Exp-1 (see Figure 6.4) and Exp-4 (see Figure 6.11) reveal that subjects have less fixation time in RE than when walking in SE, which also means when subjects were doing the SAC task in the real environment, they spent less time on searching for targets. These results lead to the conclusion that in both environments, even though subjects have been provided with the same angle of view, the same length of walking and the same task, subjects in the RE in both Exp-1 and Exp-4 can still perform the task quicker (less fixation time on each target) and easier (less saccades between targets) than those subjects who were in the SE.

In addition to the above, subjects' accurate response for the target counting in RE was better than in the SE. In Exp-1, subjects in the real world had the highest accuracy for counting targets (see Figure 6.5), but had less counting accuracy in SE when at the same walking speed as RE ($F(3,36) = 2.224$, $p = 0.037$), and also less counting accuracy at the fastest walking speed in SE ($F(3,36) = 2.224$, $p = 0.013$). Similarly in Exp-4, subjects in the real world have greater accuracy in target counting ($p = 0.017$) (see Figure 6.12).

In Exp-1 and Exp-4, we compared subjects' task performance (measured by the S, counting accuracy and efficiency, as later described) and found that there are differences between real and simulated environment that favour the real environment. These results echo the findings of Lessels & Ruddle (2005) who investigated subjects' target searching ability in both real and virtual environments, and whose results showed that the searching performance in the real environment were better than in the virtual environment, and that the participants in the real-world had fewer numbers of visits on the targets than most groups in the virtual environment. The results in both Exp-1 and Exp-4 show subjects in the SE had more total eye saccades than subjects in the RE. As in Lessels & Ruddle (2005), the results of Exp-1 (real-world experiment) shows the searching performance in terms of time taken and total number of visits to targets between each group were similar, consistent with Exp-4 results in this study, which finds that the total number of eye saccades are similar between all the RE groups.

However, our findings are different from Sahm et al. (2005). In the Sahm et al study, participants compared blind walking and blind throwing in a hallway in both a real and an HMD virtual environment, aiming to assess the performance in terms of accuracy across the two environments. Results found that the performance in terms of blind walking accuracy and blind throwing accuracy in both real and virtual environment is similar. Our findings are also different from Mohler et al. (2004), which hold that the perception/action couplings involving locomotion are similar in the real and virtual environments. Both Sahm et al. (2005) and Mohler et al. (2004) focused on measuring and comparing subjects' perception between SE and RE; however, in this thesis, we measured subjects' performance in RE and SE in terms of eye movements, a different measure than those used by Sahm et al. and Mohler et al. studies, which may therefore be one reason to explain the different findings.

7.2 Interactions between Optic Flow and Locomotion

We investigated the interactions between optic flow (as expressed by video speed) and locomotion, discovering that the relationship between visual optic flow and walking is complex, in that it does not depend in a simple way on either walking, speed of walking or the visual task.

In Exp-1, when we measured S in the SE condition, the video speed was held constant, yet S increased with walking speed. Because optic flow was constant, this finding suggests that subjects' eye movement might be affected by the action of walking (see Figure 6.2). In other words, the faster the subject walks, the more saccades are produced, possibly as a result of the effect of mechanical action that occurs with increased speed of locomotion (e.g. McDonald, Bahill & Friedman, 1983).

7.2.1 SE Walking Only

To test the above hypothesis on mechanical interference with eye movements, in Exp-2 we tested how walking by itself affected eye saccades by comparing S in the simulated environment with no task present and no changing visual input. If walking itself was responsible for the increase in S then we would expect this increase to take place even in the absence of the SAC task. However, results show that, rather than increasing, S significantly *decreases* as walking speed increases when the video task is removed (see Figure 6.6). Therefore, a decrease in saccades with locomotion speed is not consistent with walking speed itself adding additional saccades through mechanical interference or by other means.

7.2.2 SE Video Input Only

As walking itself does not increase S (it actually decreases it), another possible reason is that it might be related to the SAC task. Although the video speed does not change, it may be that the increasing walking speed is biasing the perception and judgement of targets motion speeds during locomotion and thus increasing S in response, related perhaps to known effects of walking on optic flow speed perception (e.g. Pelah, 1996; Thurrell & Pelah, 2005; Durgin, Gigone & Scott, 2005). In order to answer this question, in Exp-4 we measured subjects' saccades (see Figure 6.10) in SE with increasing video speed in this case when the walking was removed (i.e. while standing). If SAC task is a reason for the effect then when walking is removed, the S between those different video speeds would be expected to be different as well. However, the statistical result showed that there was no significant difference in S when the video was played at different speeds. This indicates that the video speed does not affect subjects' total saccades in SE while doing the SAC task, within the range of conditions tested.

7.2.3 Video and Walking Are Not Matched in SE

If it is not the task alone (i.e. no walking), nor is it the walking alone (i.e. no task) that is increasing S, we may then speculate that perhaps it is the combination of the two together. As discussed previously the known interactions between vision and locomotion (e.g. Pelah, 1996; Thurrell & Pelah, 2005; Durgin, Gigone & Scott, 2005; Durgin, 2009), offer an hypothesis that the reason for increasing S with walking speed at constant video may be the mismatch between optic flow and walking speeds as compared to what takes place during natural locomotion, where optic flow and walking speed are coupled and are thus matched. We test this hypothesis by examining performance while matching the video speeds and walking speeds in the locomotion simulator. In Exp-5, S was measured using four matched walking and video speed combinations in the SE. It was found that S now decreases significantly ($F(1.182, 4.73) = 77.799$, $p < 0.001$) when matched walking and video playing speed were increasing, just as it does without the task or video present (see Figure 6.6 in Exp-2). In addition, when the matched speed is at the medium-fast (3.3 km/h), S reduces to its lowest value; indeed, this total number of eye saccades is similar to the RE level in Exp- 4 (see Figure 6.10). The total fixation time, F, was also decreased to its minimum at medium-fast speed (3.3km/h), and is also at comparable levels to the RE performance (see Figure 6.4 and Figure 6.11). Thus, in terms of S and F, it is not only decreasing instead of increasing, but matching video and walking speed also makes the eye movement performance comparable to that in the real world.

In summary, the experiment results so far suggest that the mismatched video and locomotion speed in combination is probably the reason of production of the additional saccades. This is argued because 1) when we took away the video stimulus, S decreased when walking speed increased, showing video speed itself does not have the same increasing effect on S; 2) when we increased the video speed without walking at all, S did not change significantly; and 3) when we match video and walking speed, S no longer increases with walking speed but (along with total fixation times F) is instead at levels comparable to real-world performance. The level of S at the medium-fast matched speed (3.3km/h) was reduced (40), which is lower and closer to the real world levels in Exp-1 (65) and Exp-4 (56 on average).

It may be considered that, although conditions were matched as closely as possible, display differences between the RE and SE may be responsible for the differences in performance, for example due to the differences in resolution, frame rate and field-of-view size, as well as other differences (e.g. binocularity/stereopsis, accommodation). However, the fact that performance in SE *at its best* is comparable to RE (i.e. counting accuracy, S, F and efficiency as later described), argues strongly against such display artefacts or differences explaining the results.

To conclude, all the discussed results lead to the finding that the mismatched optic flow and locomotion is likely to be the reason for the increase of saccades with walking speed at constant video speed. In addition, subjects in the matched SE condition have better performance than in the mismatched SE in terms of counting accuracy, eye saccades and total fixation time.

7.3 Unmatched Perception of Locomotion Speed and Video Speed in SE

In Exp-3, it was found that the perception of locomotion speed on the treadmill is generally not well matched by subjects with video playing speed. This result conforms with Kong et al. (2012), which found that once perception of speed was influenced by the treadmill, subjects were unable to match their corresponding self-induced motion, and also consistent with the findings of Pelah & Barlow (1996). The reason for this effect is likely to be due to the distortion to normal visual and locomotor inputs resulting from the discrepancy between observed and expected optic flow. Figure 6.9 shows that the possible range for good matching between subjects' perception of locomotion speed and video speed is when both speeds are around 2.65km/h. By looking at the two trend lines we can say that when the locomotion speed is under 2.65km/h, subjects overestimated the video speed; this made them walk faster than the video speed. In contrast, when subjects walked faster than 2.65km/h, they underestimated the video speed, in that their chosen locomotion speed is slower than the video speed. Thus, from Figure 6.9, we can predict that in general if locomotion speed is faster or slower than 2.65km/h, the errors of mismatching speed will be increased with the locomotion speed. This finding also supports the view of Thurrell & Pelah (2005), which holds that the increasing speed will bring greater distortion to people's visual perception when walking on the treadmill.

In addition, we also found that although we provided five attempts for each subject in each task, the differences between the perception of locomotion speed and video playing speed still remain and are not reduced with training (see Figure 6.8). This reveals that when walking on the treadmill in SE, no matter in which speed, there is no apparent learning effect to help people with reducing the distortion. It therefore appears to be the case that, although the visual angles of the moving (as a result of self-motion) targets and objects are correctly matched in the experimental setup to their real-world sizes, the information is insufficient or overly distorted to produce a correct walking to video speed match. Overall, these findings indicate that eye movements are linked to locomotion in other ways than only mechanically (e.g. McDonald, Bahill & Friedman 1983). Total numbers of saccades and fixation durations (as well as efficiency, see section 7.4), during the SAC task seem to also depend on the relationship between optical flow speed and locomotion speed.

7.4 The Efficiency

In this study, we defined the efficiency as the model metric with which we compare overall eye movement performance between the different environments and conditions of the experiments. It is defined as the reciprocal of the energy cost function (ECF) (see Equation 5.2), and is useful in that it combines both S (total eye fixations) and F (total fixation time) and can be used for tuning the conditions towards achieving optimal performance in the use of locomotion simulators such as that used in the present study.

In order to achieve optimal tuning of SEs, using the results of Exp-4 (and also Exp-5) we apply the measure to compare subjects' efficiency in the SAC task when they were walking in RE as compared to while they remained stationary and viewed different speeds of video in SE.

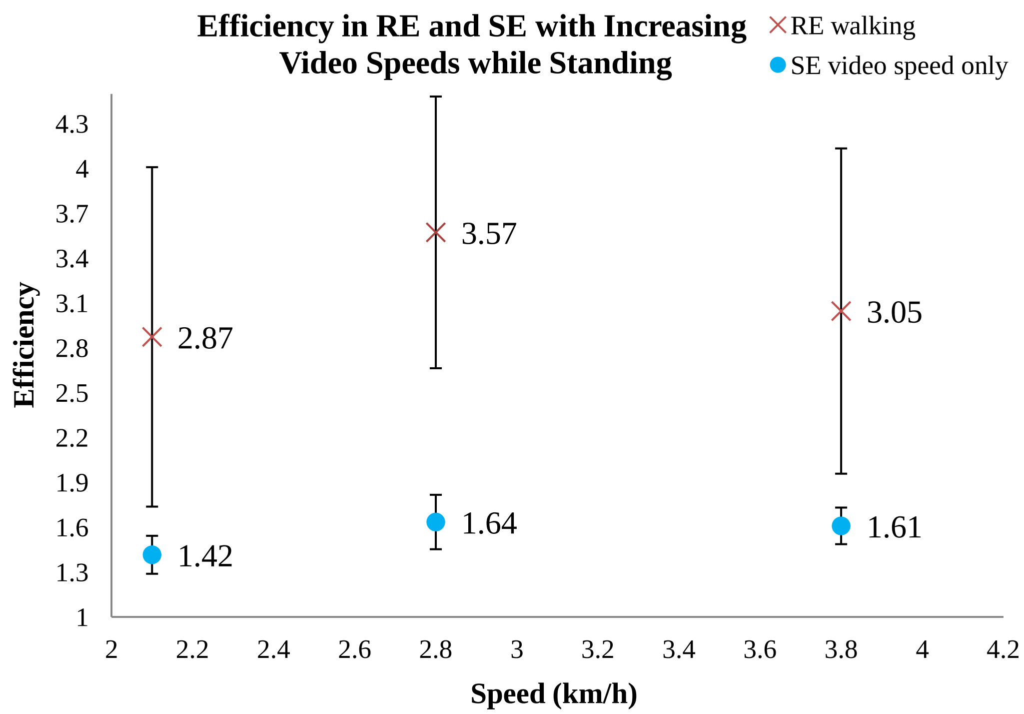

Figure 7.1 Efficiency is shown for three different video speeds (at 2.1, 2.8 or 3.8 km/h) for standing subjects in SE and walking in three different speeds in RE (at 2.1, 2.8 or 3.8 km/h) in Exp-4. Error bars represent ± 1 standard deviation of the mean. 15 subjects participated in each environment in this experiment. The LSD Post Hoc Test (see Section 6.4) found that the efficiency increases significantly for SE between 2.1 km/h and 2.8 km/h (p = 0.035); no significant differences were observed between 3.8 km/h and either 2.1 (p = 0.059) km/h or 2.8 km/h (p = 0.784). Therefore following an initial increase in efficiency from the slower video speed of 2.1 km/h no further improvement is seen by 3.8 km/h for stationary (standing) subjects.

As shown in Figure 7.1, examining the efficiency of performing the SAC task at different walking speeds in RE, no significant differences were found (F(2, 12 = 0.603, p = 0.563). There was therefore no loss or gain in efficiency with walking speed, and the associated visual speed, under the natural RE conditions. However, efficiency in SE is lower than it is in the RE, with efficiency in the real environment being approximately twice that obtained in the simulator, even in the absence of walking in SE. Since subjects in Exp-4 were standing in SE, the lower efficiency found cannot be attributed to treadmill walking. Within the SE, there were also no significant differences for all the different video speeds (F(2, 12 = 3.351, p = 0.07). The LSD Post Hoc Test (see Section 6.4) found that the efficiency increases significantly for SE between 2.1 km/h and 2.8 km/h (p = 0.035); no significant differences were observed between 3.8 km/h and either 2.1 (p = 0.059) km/h or 2.8 km/h (p = 0.784). Therefore following an initial increase in efficiency from the slower video speed of 2.1 km/h no further improvement is seen by 3.8 km/h for stationary (standing) subjects.

Exp-4 has shown that efficiency in the SE can vary depending on the speed of the video stimulus for subjects that are performing a SAC task while standing (though efficiency does not vary in the case of RE). The result raises the possibility that, once locomotion is again added to the SE, an efficient set of parameters could be found for walking and video speeds to optimally run a locomotion simulator for a given configuration and setup.

In Exp-5, efficiency of subjects performing a SAC task was computed for four matched walking and video speed combinations in the SE.

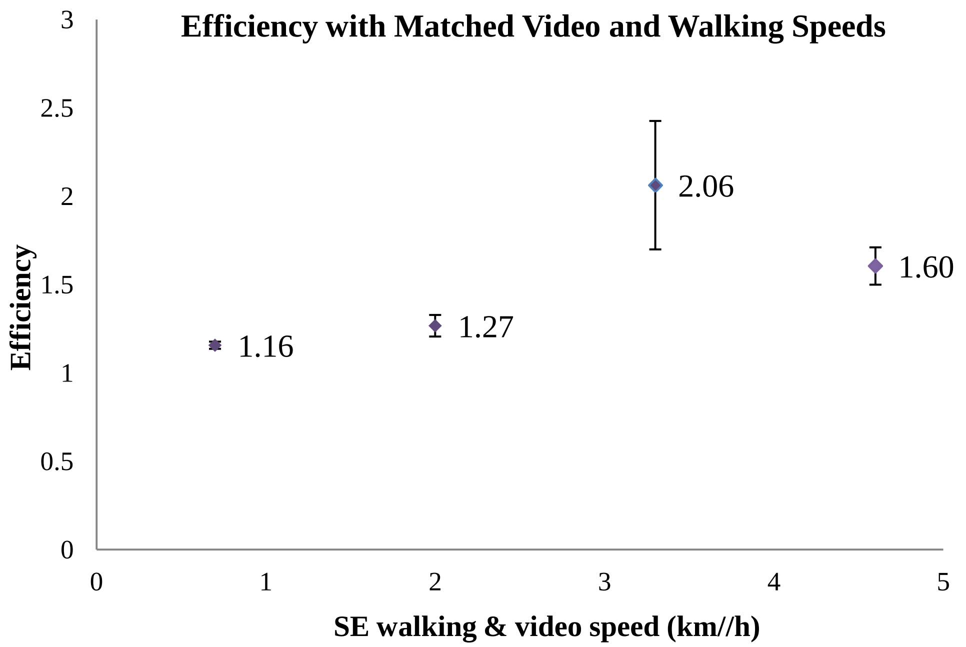

Figure 7.2 Efficiency in SE at four matched walking and video speed combinations at 0.7, 2.0, 3.3 or 4.6 km/h. The LSD Post Hoc Test shows that the efficiency increases significantly between 0.7 and 2.0 km/h (p = 0.027), also between 2.0 and 3.3 km/h (p = 0.006). Efficiency then decreases significantly between 3.3 km/h and 4.6 km/h (p = 0.017). Error bars represent ± 1 standard deviation of the mean. Thus, good efficiency may depend on selecting a range of matched parameters for walking and video speeds.

The statistics reveal that efficiency was significantly different between matched video and walking groups ($F(1.036, 4.143)=29.977$, $p < 0.001$). It appears in addition that intermediate speeds of matched walking and video of 2 – 4 km/h offer a better simulation of the RE for the SAC task and under the conditions tested.

By taking the mean efficiency in the RE from Figure 7.1 (i.e. 3.16), a comparative increase in efficiency can be computed that would result from selecting the efficiencies from the worst (37% of RE at 1.16) to the best (65% of RE at 2.06) walking and video speed parameters. Therefore, in terms of the metrics and methods applied, and under the conditions and task used in the experiments, the tuning of walking and matched video optic flow parameters in a locomotion simulator can lead to significant improvements of between 37% - 65% in efficiency. Although not considered in this study, it is suggested as well that, by reducing the numbers of additional saccades and/or the duration of fixations from these eye movements (i.e. increasing efficiency) the levels of usability and comfort of locomotion simulators will probably also be improved. Further experiments, however, would be needed to confirm any improvements from matched tuning by these other measures.

7.4.1 Learning Effect on the Efficiency

In Exp-6, we investigated learning effects to test if familiarity or learning that would naturally be higher in RE (i.e. due, for example, to greater experience with everyday living in the real world) may have caused or contributed to the difference in SAC task performance between RE and SE, and also to measure if subject performance in SE can be improved after training. Using the efficiency model (see Equation 5.2), results of Exp-6 were assessed for learning in SE by comparing subjects' performance in terms of efficiency between each successive trial.

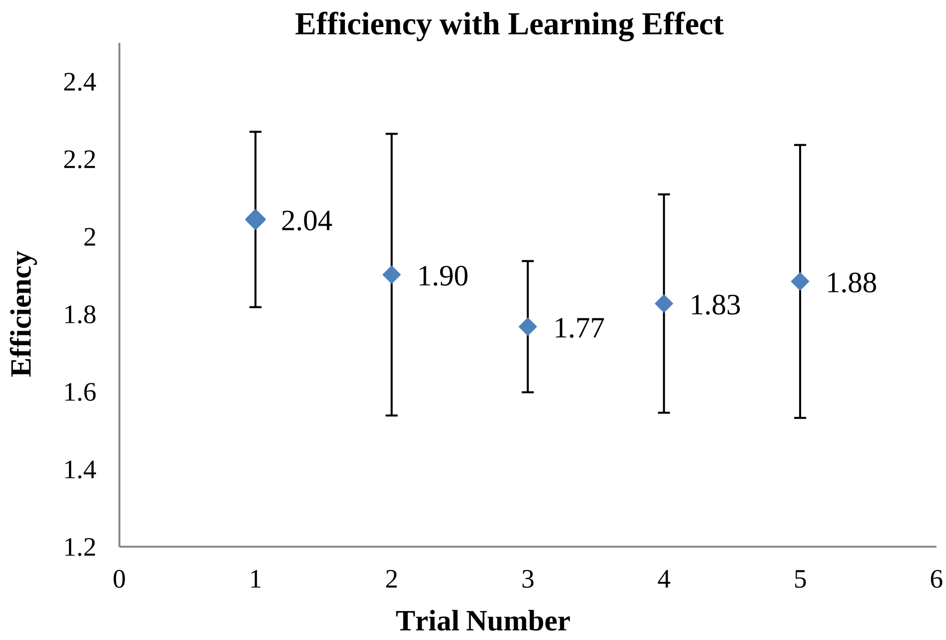

Figure 7.3 Efficiency for walking at 3.3km/h matched speeds five times in SE. Error bars represent ± 1 standard deviation of the mean. In total, five subjects participated in this experiment (Exp-6). Statistical results shows that for efficiency no significant differences are found ($F(4, 16)=0.732$, $p = 0.584$).

Figure 7.3 shows that efficiency is similar at each successive trial of walking in SE. Thus in SE, even when subjects were trained five consecutive times in the same condition (i.e. same matched walking and video speed combinations, same SAC task) in order to increase their familiarity or to stimulate learning, there were still no improvements in their efficiency. Conditions between our experiments and Lessels and Ruddle (2005), who did in fact find improvements with training, were not comparable, because they had used the keyboard cursor key for movement in an HMD instead of the real walking in a large-screen simulator as we had. In Exp-4, when doing the SAC task, subjects have much better efficiency in RE than in SE, in agreement with our other findings (e.g. Figure 6.10 and Figure 6.11), and Figure 7.3 shows that the efficiency in

SE is not improved with familiarity or learning of that environment. Thus, it cannot be said that improved performance in RE can be accounted for by the greater everyday experience of walking and seeing in the real world.

Chapter 8. Conclusion

This thesis reported research and experiments on motion capture and locomotion simulators. We successfully developed and tested methods for improving the design and configuration of motion capture in a locomotion simulator; improved understanding of human interaction within locomotion simulator environments. Next, we increased understanding of the differences between real world walking and locomotion in simulators, and found ways of tuning the best conditions under which locomotion simulator environments (SE) perform as similarly as possible to the real world environment (RE); also, we found that the link between eye movements and locomotion is not only mechanical. Finally, we found that in order to best match RE in the locomotion simulator, it requires matching video and walking speeds at a walking speed of approximately 2-4km/h. We propose that the approach represents a new method for tuning and assessing SE locomotion simulators.

In Chapters 2 & 3, a new motion capture design tool (StroCap) was developed and tested. Testing results showed that StroCap successfully simulated the StroMoHab system, and accurately provided camera coordinates, facing direction, capture volume, rotations and overlapping capture area. These helped with the design and configuration of the StroMoHab system, and made the setup procedures more accurate and efficient. StroCap improved the existing motion capture simulator, which assisted in arranging the system in a defined space without using a tedious ‘Trial & error’ process.

Chapters 4 & 5 provided background and methods for experimental results reported in Chapter 6, comparing eye movements performance in the real world and simulated environments in terms of S (total number of eye saccades), F (total fixation time) and counting accuracy for a natural Search-and-Count task (SAC) that involves several brain mechanisms that may be active during locomotion. As discussed in Chapter 7, results found that 1) generally the eye movement performance in the simulated environment is not as good as in the real environment as expressed by the lower values of S and F, and reduced accuracy in target counting. 2) there are interactions between optic flow and locomotion which affected subjects’ SE performance in terms of eye movements, and the relationships do not depend in a simple way on either walking, speed of walking or the visual task; it was suggested that the mismatched video optic flow and locomotion speed is the main the reason for production of poorer performance in SE. 3) the differences in performance between real and simulated environments could not be improved by increasing the familiarity and learning in SE through training.

In addition, we devised efficiency in Chapter 7 as a model metric to quantify performance using the methods proposed in this thesis to analyse and optimize locomotion simulators in order to match as closely as possible to the real world. By matching video and walking speeds, we found that the efficiency, as well as S and F, during the SAC task seem to also depend on the relationship between optical flow speed and locomotion speed, and that there is an optimal speed range of approximately 2-4km/h in which efficiency is highest in the simulated locomotion environment used in this study. The methods and approach reported in this thesis may prove useful for new designs and comparisons of similar virtual simulation environments.

Appendices

Appendix 1: Marker's movements with Errors

- Set marker's initial position, small movement on -X

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	132.41	28.086	40.7
	132.92	24.818	37.89
	132.75	28.67	41.487
	129.31	27.87	40.952
	132.51	27.64	38.02
Initial value chose:	**136.38**	**27.83**	**40.98**

Move on -X	X(program)	Z(program)	Y(program)	X(By hand)	Errors
Small Move	128.61	27.42	40.626	128.38	0.23
	120.55	26.279	39.576	120.38	0.17
0.8cm each time	112.658	27.962	40.793	112.38	0.278
	104.719	27.484	40.028	104.38	0.339
	96.882	26.335	39.183	96.38	0.502

- Reset marker's initial position, small movement on +X

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	131.75	24.59	39.44
	135.89	27.34	39.07
	131.62	28.887	42.25
	134.52	27.72	40.17
Initial value chose:	**136.463**	**27.481**	**40.521**

Move on +X	X(program)	Z(program)	Y(program)	X(By hand)	Errors
Small Move	144.612	27.585	41.159	144.463	0.149
	152.217	27.036	40.223	152.463	0.246
0.8cm each time	159.82	27.895	41.214	160.463	0.643
	168.065	27.89	41.49	168.463	0.398
	175.904	27.151	41.185	176.463	0.559

- Reset marker's initial position, small movement on +Z

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	136.59	191.936	33.38
	138.29	194.83	41.15
	133.67	193.51	40.29
	136.34	195.33	40.63
Initial value chose:	**138.509**	**195.142**	**40.392**

Move on +Z	X(program)	Z(program)	Y(program)	Z(By hand)	Errors
Small Move	138.825	203.526	40.127	203.142	0.384
	138.868	211.261	40.713	211.142	0.119
0.8cm each time	138.916	219.545	40.262	219.142	0.403
	139.277	226.974	39.855	227.142	0.168
	139.398	235.446	40.488	235.142	0.304

- Reset marker's initial position, small movement on -Z

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	138.51	195.59	36.45
	136.31	191.07	42.98
	135.75	195.17	39.04
	134.67	193.72	40.818
Initial value chose:	**138.87**	**195.083**	**40.649**

Move on -Z	X(program)	Z(program)	Y(program)	Z(By hand)	Errors
Small Move	138.825	203.526	40.127	203.142	0.384
	138.868	211.261	40.713	211.142	0.119
0.8cm each time	138.916	219.545	40.262	219.142	0.403
	139.277	226.974	39.855	227.142	0.168
	139.398	235.446	40.488	235.142	0.304

- Reset marker's initial position, small movement on +Y

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	138.12	192.159	77.88
	135.90	194.134	79.34
	139.07	196.49	78.51
	138.47	193.07	72.25
Initial value chose:	**138.917**	**193.01**	**78.96**

Move on +Y	X(program)	Z(program)	Y(program)	Y(By hand)	Errors
Small Move	138.03	193.406	87.031	87.133	0.102
	138.367	193.194	96.687	96.133	0.554
0.9cm each time	138.013	194.053	106.312	105.133	1.179
	138.101	193.255	115.54	114.133	1.407
	137.457	196.302	125.886	123.133	2.753

- Reset marker's initial position, small movement on -Y

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	138.03	194.131	79.92
	134.131	196.53	78.66
	137.58	192.11	78.13
	136.29	193.62	76.83
Initial value chose:	**137.032**	**193.755**	**77.413**

Move on -Y	X(program)	Z(program)	Y(program)	Y(By hand)	Errors
Small Move	134.731	193.364	67.596	68.413	0.817
	135.86	193.256	57.837	59.413	1.576
0.9cm each time	136.719	193.03	48.355	50.413	2.058
	136.906	194.198	39.592	41.413	1.821
	137.18	193.418	28.815	32.413	3.598

- Reset marker's initial position, large movement on +X

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	-50.401	27.53	40.93
	-50.47	25.19	38. 47
	-43.25	26.598	41.32
	-51.18	27.34	38.93
Initial value chose:	**-50.268**	**27.365**	**38.495**

Move on +X	X(program)	Z(program)	Y(program)	X(By hand)	Errors
Large Move	31.02	25.437	37.384	31.541	0.521
	109.027	24.261	36.892	111.541	2.514
8cm each time	188.084	21.321	33.671	191.541	3.457
	267.312	26.852	40.812	271.541	4.229
	31.02	25.437	37.384	31.541	0.521

- Reset marker's initial position, large movement on -X

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	311.366	23.88	35.18
	322.213	31.47	37. 23
	314.56	25.26	38.46
	313.922	24.19	41.22
Initial value chose:	**314.683**	**25.979**	**37.759**

Move on -X	X(program)	Z(program)	Y(program)	X(By hand)	Errors
Large Move	235.121	26.09	38.218	234.121	1
	156.06	24.852	36.556	154.121	1.939
8cm each time	76.722	23.028	34.037	74.121	2.601
	-2.074	27.619	38.739	-5.879	3.805
	-91.883	30.065	38.073	-85.879	6.004

- Reset marker's initial position, large movement on +Z

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	147.891	-3.82	40. 313
	142.332	-3.195	38. 17
	147.81	-3.41	41. 46
	145.57	-2.177	38.551
Initial value chose:	**147.272**	**-3.339**	**40.013**

Move on +Z	X(program)	Z(program)	Y(program)	Z(By hand)	Errors
Large Move	147.647	74.312	36.779	77.895	3.583
	148.148	156.011	40.434	157.895	1.884
8cm each time	148.273	236.51	40.97	237.895	1.385
	149.218	315.706	40.057	317.895	2.189
	149.201	397.631	40.099	397.895	0.264

- Reset marker's initial position, large movement on -Z

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	148.31	405.056	39. 74
	144.72	397.001	40. 32
	149.330	404.98	39. 109
	149.28	405.22	38.492
Initial value chose:	**149.664**	**405.725**	**39.874**

Move on -Z	X(program)	Z(program)	Y(program)	Z(By hand)	Errors
Large Move	148.579	324.328	40.732	325.725	1.397
	147.934	244.36	40.044	245.725	1.365
8cm each time	147.473	164.305	40.907	165.725	1.42
	146.81	82.631	39.247	85.725	3.094
	146.401	3.073	38.483	5.725	2.652

- Reset marker's initial position, large movement on +Y

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	151.12	190.21	34. 92
	148.322	191.73	40. 177
	147.08	196.58	37. 322
	148.49	194.31	40.031
Initial value chose:	**148.634**	**196.753**	**40.087**

Move on +Y	X(program)	Z(program)	Y(program)	Y(By hand)	Errors
Large Move	148.088	195.904	87.918	87.048	0.87
	148.934	191.919	133.18	134.048	0.868
4.7cm each time	147.837	191.908	180.687	181.048	0.361
	153.972	191.776	228.952	228.048	0.904
	148.181	190.642	278.265	275.048	3.217

- Reset marker's initial position, large movement on -Y

Marker initial position	X Coordinate	Z Coordinate	Y Coordinate
	145.93	194.157	280. 52
	141.14	190.53	231. 89
	146.27	195.149	288. 115
	141.722	191.22	287.64
Initial value chose:	**141.549**	**191.162**	**287.234**

Move on -Y	X(program)	Z(program)	Y(program)	Y(By hand)	Errors
Large Move	148.067	194.307	238.681	240.234	1.553
	150.876	194.187	189.845	193.234	3.389
4.7cm each time	148.451	197.581	145.939	146.234	0.295
	146.219	195.546	97.742	99.234	1.492
	145.969	194.217	48.647	52.234	3.587

Glossary of Acronyms

RE – Real Environment

SE – Simulated Environment

VR – Virtual Reality

StroMoHab – Stroke Mobility Rehabilitation

StroCap – StroMoHab Motion Capture Simulation Software

VLE –Virtual Locomotion Environment

HMD – Head Mounted Display

LSID – Large-Screen Immersive Display

FOV – Field of View

S – Total Number of Saccades

F – Total Fixation Time

ECF – Energy Cost Function

SAC – Search and Count

List of References

Ariel Dynamics. (1999). *Principle of Gait Analysis.* Retrieved 10 2008, from http://www.arielnet.com/Main/adw-39a.html

Auburn University. (2012). *Vicon Motion Capture System.* Retrieved 9 2012, from http://www.eng.auburn.edu/research/centers/oseipp/research/labs/biomechanics_lab/biomechanics-lab.html

Awwad, S. (n.d.). *Eyeweb.org.* Retrieved 10 2008, from http://www.eyeweb.org/motility.htm

Banton, T., Stefanucci, J., Durgin, F., Fass, A., & Proffitt, D. (2005, August). The perception of walking speed in a virtual environment. *Presence: Teleoperators and Virtual Environments , 14* (4), pp. Volume 14(4), 394-406.

Bogey, R. (2009). *Medscape:Gait Analysis.* Retrieved 2 2010, from http://emedicine.medscape.com/article/320160-overview#aw2aab6b2

Bouguila, ,. L., Florian, E., & Courant, M. (22-27 July 2005). Active walking interface for human-scale virtual environment. *11th International Conference on Human-*

Computer Interaction, HCII'05. Las Vegas, Nevada, USA.

Carlson, N. R., & Heth, D. C. (2010). *Psychology: The Science of Behaviour* (4,illustrated ed.). Pearson Education Canada, ISBN:0205645240, 9780205645244, pp 140, .

Carranza, J., Theobalt, C., Magnor, M., & Seidel, H.-.. (2003). Free-viewpoint video of human actors. *ACM SIGGRAPH*, (pp. 565—577).

CodaMotion. (2012). *Technology Overview.* Retrieved 9 2012, from http://www.codamotion.com/systems/technology-overview.html

Computer Stroies. (2011). *Motion capture explained.* Retrieved 9 2012, from http://computerstories.net/motion-capture-explained/

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., & Drury, H. A. (1998). A common network of functional areas for attentional and eye movements. *Neuron* (21), pp. Volume 2(1), 761–773.

Creem-Regehr, S. H., & Kunz, B. R. (2010, November/December). Perception and action. *Wiley Interdisciplinary Reviews: Cognitive Science , 1* (6), pp. Volume 1(6), 800-810.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., & Thompson, W. B. (2004). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception , 34*, pp. Volume 3(16), 191-204.

Darken, R. P., Cockaybe, W. R., & Carmein, D. (October 14-17, 1997). The Omni-Directional Treadmill:A Locomotion Device for Virtual Worlds. *Proc. of UIST'97* (pp. 213-221). Banff, Canada: ACM Press.

Durgin, F. H. (2009, February). When walking makes perception better. *Current Directions in Psychological Science , 18*, pp. Issue 18, 43-47.

Durgin, F. H., Fox, L. F., Lewis, J., & Walley, K. A. (2002). Perceptuomotor adaptation: More than meets the eye. *Poster presented at the 43rd Annual Meeting of the Psychonomic Society.* Kansas City, MO.

Durgin, F. H., Gigone, K., & Scott, R. (2005, April). Perception of Visual Speed While Moving. *Journal of Experimental Psychology: Human Perception and Performance* , pp. Volume 31(2), 339-353.

Encyclopaedia Britannica. (2012). *Human Eye - MOVEMENTS OF THE EYES.* Retrieved 9 2012, from http://www.britannica.com/EBchecked/topic/1688997/human-eye/64889/Movements-of-the-eyes

EuclideanSpace - building a 3D world. (1998). *Maths - Orthogonal Matrices.* Retrieved

5 2009, from
http://www.euclideanspace.com/maths/algebra/matrix/orthogonal/index.htm

Fajen, B. R., & Warren, W. H. (2005). Behavioral dynamics of steering, obstacle avoidance,and route selection. *Journal of Experimental Psychology: Human Perception and Performance* , pp. Issue 29, 343–362.

Fink, P. W., & Warren, W. (2002, November 20). Velocity dependence of optic flow strategy for steering and obstacle avoidance. *Journal of Vision* , p. Volume 2(7).

Fink, P. W., Foo, P. S., & Warren, W. H. (2007, January). Obstacle avoidance during walking in real and virtual environments. *ACM Transactions on Applied Perception (TAP) , 4*, p. Volume 4.

Frenz, H., & Lappe, M. (2005). Absolute travel distances from optic flow. *Vision* , pp. Issue 45, 1679--1692.

Fung, J., Richards, C. L., Malouin, F., McFadyen, B. J., & Lamontagne, A. (2006). A Treadmill and Motion Coupled Virtual Reality System for Gait Training Post-Stroke. *Cyberpsychology and Behavior* , pp. Volume 2(9), 33-39.

Geuss, M., Stefanucci, J., Creem-Regehr, S., & Thompson, W. B. (July 23-24, 2010). Can I pass?: using affordances to measure perceived size in virtual environments. *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization.*

Los Angeles, California.

Grechkin, T. Y., Nguyen, T. D., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2010, July). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception (TAP) , 7* (26), pp. Volume 7(26), 1-18.

Guerra-Filho, G. (2005). Optical Motion Capture:Theory and Implementation. *Brazilian Computing Society Revista de Informática Teórica e Aplicada* , p. Volume 12.

Hank Virtual Environments Lab. (2010). *Perception & Adaptation Research.* Retrieved 6 2011, from http://homepage.cs.uiowa.edu/~hank/research/page2.html

Hees, H. (2006). *3D Computer Graphics.* Retrieved October 16, 2008, from http://zh.scribd.com/doc/75892491/3D-Computer-Graphics

Heinen, S. J., & Liu, M. (1997, September). Single-neuron activity in the dorsomedial frontal cortex during smooth-pursuit eye movements to predictable target motion. *Visual Neuroscience , 14* (05), pp. Volume 14(5), 853-865.

Jaffe, D. L., Brown, D. A., Pierson-Carey, C. D., Buckley, E. L., & Lew, H. L. (2004, May/June). Stepping over obstacles to improve walking in individuals with poststroke. *Journal of Rehabilitation Research & Development , 41* (3A), pp. Volume 41, 283-292.

Kanski, J. J. (1989). *Clinical Ophthalmology: A Systematic Approach* (Vol. 111). (V. 1.-1. ISBN 0080449697, Ed.) Boston: Butterworths.

Kenyon, R. V., Sandin, D., Smith, R. C., Pawlicki, R., & Defanti, T. (2007, April). Size-Constancy in the CAVE. *Presence: Teleoperators and Virtual Environments , 16* (2), pp. Volume 16(2), 172-187.

Kitagawa, M., & Windsor, B. (2008). *MoCap for Artists: Workflow and Techniques for Motion Capture.* Focal Press.

Klein, E., Swan, E. J., Schmidt, G. S., Livingston, M. A., & Staadt, O. G. (March 14-18, 2009). Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays. *Proceedings of the 2009 IEEE Virtual Reality Conference* , (pp. 107-113).

Kong, P. W., Koh, T. M., Tan, W. C., & Wang, Y. S. (2012, May). Unmatched perception of speed when running overground and on a treadmill. *Gait & Posture , 36* (1), pp. Volume 36(1), 46-48.

Kuhl, S. A., Thompson, W. B., & Creem-Regehr, S. H. (August 09-10, 2008). HMD calibration and its effects on distance judgments. *Proceedings of the 5th symposium on Applied perception in graphics and visualization.* Los Angeles, California.

Kuhl, S. A., Thompson, W. B., & Creem-Regehr, S. H. (July 28-29, 2006). Minification influences spatial judgments in virtual environments. *Proceedings of the 3rd symposium on Applied perception in graphics and visualization.* Boston, Massachusetts.

Kuntz, S. (2007). *VR Geek.* Retrieved 7 2009, from http://cb.nowan.net/blog/2007/06/28/omni-directional-treadmill/

Lego Company Profile. (2010). *An introduction to the LEGO Group.*

Lessels, S., & Ruddle, R. A. (2005). Movement around real and virtual cluttered environments. *Presence: Teleoperators and Virtual Environments* (14), pp. Volume 14(5), 580-596.

Maroues. (2008-2012). *Class 46.* Retrieved 9 2012, from http://www.marques.org/class46/default.asp?D_A=20120820

McDonald, J. D., Bahill, A. T., & Friedman, M. B. (1983). An adaptive control model for human head and eye movements. *IEEE Transactions on Systems, Man, & Cybernetics* (SMC-13), pp. Volume 13(2), 167–174.

MedicineNet. (1996-2012). *Black Eye (cont.).* Retrieved 9 2012, from http://www.medicinenet.com/black_eye/page4.htm

Metamotion. (n.d.). *Magnetic Motion Capture.* Retrieved 2011, from http://www.metamotion.com/motion-capture/magnetic-motion-capture-1.htm

Moeslund, T. B., & Granum, E. (2001, March). A Survey of Computer Vision-Based Human MotionCapture. *Computer Vision and Image Understanding , 81* (3), pp. Volume 81(3), 231-268.

Mohler, B. J., Creem-Regehr, S. H., & Thompson, W. B. (2006). The influence of feedback on egocentric distance judgments in real and virtual environments. *Proceedings of the 3rd symposium on Applied perception in graphics and visualization* . Boston, Massachusetts.

Mohler, B. J., Thompson, W. B., Creem-Regehr, S., Pick, H. L., Warren, W., Rieser, J. J., et al. (August 07-08,2004). Visual motion influences locomotion in a treadmill virtual environment. *Proceedings of the 1st Symposium on Applied perception in graphics and visualization.* Los Angeles, California.

Murray, R. F., Beutter, B. R., Eckstein, M. P., & Stone, L. S. (2003). Saccadic and perceptual performance in visual search tasks. II. Letter discrimination. *JOSA A , 20* (7), pp. Volume 20(7), 1356-1370.

Murray, W. S. (2003). Behavioral and Brain Sciences. (26), pp. Volume 2(6), 446.

NaturalPoint. (2012). *OptiTrack.* Retrieved 9 2012, from http://www.naturalpoint.com/optitrack/

NHS Greater Glasgow and Clyde. (2002). *Mangement of Patients with Stroke.*

Retrieved October 2008, from http://www.nhsggc.org.uk/content/mediaassets/pdf/HSD/sign64.pdf

Nintendo. (2010). *Wii*. Retrieved 8 2012, from http://www.nintendo.co.uk/NOE/en_GB/wii_54.html

Orion Technologies. (2005). *Lego Specifications*. Retrieved June 2009, from http://orionrobots.co.uk/Lego+Specifications

PCBookCN. (2000). *Camera's capture volume in theory*. Retrieved February 2009, from http://www.pcbookcn.com/img_article/20041123162518.jpg

Pelah, A. (1997). The vision of natural and complex images. *Vision Research , 37* (3201), pp. Volume 37 (23), 3201-3202.

Pelah, A., & Barlow, H. B. (1996). Visual illusion from running. *Nature* , pp. Volume 381, 283.

Pelah, A., & Koenderink, J. J. (2007, January). Editorial: Walking in real and virtual environments. *ACM Transactions on Applied Perception (TAP)* , p. Volume 4(1).

Pelah, A., Secker, B., Bishop, A., & Askham, C. (1998). A wide-field simulator for studying visuo-motor interactions in locomotion. *Journal of Physiology* , pp. 506p, 13-14.

Plumert, J. M., Kearney, J. K., Cremer, J. F., & Recker, K. (2005, July). Distance perception in real and virtual environments. *ACM Transactions on Applied Perception (TAP) , 2* (3), pp. Volume 2(3), 216-233.

Qualisys. (2011). *Biomechanics.* Retrieved 9 2012, from http://www.qualisys.com/applications/biomechanics/sport-science/

Rand, D., Kizony, R., Feintuch, U., Katz, N., Josman, N., Rizzo, A. A., et al. (2005). Comparison of two VR platforms for rehabilitation: video capture versus HMD. *Presence, Teleop Virt Environ , 14*, p. Volume 1(4).

Richards, J. G. (1999). The measurement of human motion: a comparison of commercially available systems. *Human Movement Science* , pp. Volume 5 (18), 589-602.

Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E. (1995, June). Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance* , pp. Volume 21(3), 480-497.

Rudd, A. G., Irwin, P., Rutledge, Z., Lowe, D., Wade, D., Morris, R., et al. (1999). The national sentinel audit for stroke: a tool for raising standards of care. *Journal of the Royal College of Physicians of London* , pp. Volume 33(5), 460-464.

Ruddle, R. A., & Lessels, S. (2009, April). The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI) , 16* (1), pp. Volume 16(1), 1-18.

Saboune, J., & Charpillet, F. (2005, cs/0510063, October). Markerless Human Motion Capture for Gait Analysis. *ArXiv Computer Science e-prints* .

Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., & Willemsen, P. (2005, January). Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Transactions on Applied Perception (TAP) , 2* (1), pp. Volume 2(1), 35-45.

Schultz, W. (1997, December). Pointing with focussing devices. *Behavioral and Brain Sciences , 20* (04), pp. Volume 20 (4), 755-756.

Simon, R. (2004). Quantification of human motion: gait analysis benefits and limitations to its application to clinical problems. *Journal of Biomechanics* , pp. Volume 37, 1869-1880.

Stroke Association. (2005). *About stroke.* Retrieved October 2008, from http://www.stroke.org.uk

Sturman, D. J. (1994). A Brief History of Motion Capture for Computer Character

Animation. *SIGGRAP*, (p. Course Notes 9). Los Angeles.

Tehovnik, E. J., Sommer, M. A., Chou, I. H., Slocum, W. M., & Schiller, P. H. (2000). Eye fields in the frontal lobes of primates. *Brain Research Reviews* (32), pp. Volume 3(2), 413-448.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments , 13* (5), pp. Volume 13(5), 560-571.

Thurrell, A. E. (1998). *The influences of non-visual signals of walking on the perceived speed of optic flow.* Physiology Project, University of Cambridge, Physiological Laboratory, Cambridge.

Thurrell, A., & Pelah, A. (2005). Matching visual and nonvisual signals: evidence for a mechanism to discount optic flow during locomotion. *In Proceedings of the SPIE-IS&T Electronic Imaging*, (pp. Volume 5666, 434-448).

U.S. Army Research Institute for the Behavioral and Social Sciences. (2005). *Introduction to and Review of Simulator Sickness Research.* Retrieved June 2012, from http://www.hqda.army.mil/ari/pdf/RR%201832.pdf

Vicon. (2008). *Vicon Products.* Retrieved 9 2012, from http://www.vicon.com/products/

VRarchitect. (2004). *Head-Mounted Display (HMD).* Retrieved 6 2010, from http://www.vrarchitect.net/anu/cg/Display/HMD.en.html

Weiss, P. L., Kizony, R., Feintuch, U., & Katz, N. (2006). Virtual reality in neurorehabilitation. (M. E. Selzer, L. Cohen, F. H. Gage, S. Clark, & P. W. Duncan, Eds.) *Textbook of neural repair and rehabilitation* , pp. Volume 2(13), 182-197.

Weiss, P. L., Rand, D., Katz, N., & Kizony, R. (2004). Video capture virtual reality as a flexible and effective rehabilitation tool. *NeuroEng Rehabilitaion , 1* (12), p. Volume 1(12).

Willemsen, P., & Gooch, A. A. (March 24-28, 2002). Perceived Egocentric Distances in Real, Image-Based and Traditional Virtual Environments. *Proceedings of the IEEE Virtual Reality Conference*, (p. 275).

WorldPress.com. (n.d.). *Motion Capture.* Retrieved 2012, from World Press: http://lukebeech.wordpress.com/motion-capture/

XBOX. (2012). *Xbox 360 + Kinect.* Retrieved 10 2012, from http://www.xbox.com/en-GB/Xbox360?xr=shellnav