NTNU – Trondheim
Norwegian University of
Science and Technology

Telepresence Quality

Haijiao Zhang

Master of Telematics - Communication Networks and Networked Services (2
Submission date: June 2013
Supervisor: Leif Arne Rønningen, ITEM
Co-supervisor: Otto Wittner, Uninett

Norwegian University of Science and Technology
Department of Telematics

Problem Description

Name of the student: Haijiao Zhang

Unlike a general video conferencing system, a telepresence system makes users feel more natural and gives a more "accurate" sensation of nearness to the people who are interacting with other people through the system. In this thesis, a telepresence system model which is a simple lab model of the DMP (Distributed Multimedia Plays) system will be proposed.

There are many factors that influence users' perceptions of telepresence systems, such as the quality of sound and video. However, one important factor is delay, which will be dealt deeply in this thesis. In addition, a series of stimulus-response tests will be conducted to discover other factors that influence the perceptions of users. In the end, a simple 3D system will be proposed and a series of pilot tests will be performed to see how 3D can influence the perceptions of users.

Assignment given: February 5, 2013

Supervisor: Leif Arne Rønningen, ITEM

Abstract

One aim of a telepresence system is to make users' collaborations become more natural and gives a more "accurate" sensation of nearness to the people who are interacting with other people through the system. There are many factors that can directly influence people's perceptions about telepresence systems. This thesis helps to find out what typical factors are and how they can influence the perceptions of telepresence systems.

One of the most important factors that influences the quality a telepresence system is delay. Hence, in order to have a control of the delay, a telepresence experimental model was established, and a serous of tests had been carried out with the model. Based on the measurement results, we can conclude that the optical delay is directly related to the frame rate of the camera and the refreshing rate of the screen.

In addition, a lot of stimulus-reaction time measurements had been done. Based on the test results, we got indications that the perception and reaction time is directly involved with age, tiredness, and concentration of the people using the system. The perception and reaction time does not seem to correlate with the education of the people using the system.

Finally, in order to find out whether the space perception can influence the perception and reaction time, two stereoscopic system models had been constructed. A lot of pilot tests had been done with the model. According to the measurements, we can conclude that with the same contents and same frame rates of videos, the perception and reaction time with 3D videos seem to be shorter than that with 2D videos.

Preface

This master thesis is written at the department of Telematics in the Norwegian University of Science and Technology (NTNU) during the Spring Semester, 2013. The responsible supervisor of the subject has been Leif Arne Rønningen.

I would like to take this opportunity to thank my professor – Leif Arne Rønningen for his helpful supports, invaluable advices and quick feedbacks on my work. From the initial brainstorming to the completion of my thesis, there was no single moment when my questions were not answered. Whenever I met tough questions, he always tried to help me and kept my spirits up, which provided me with a huge motivation. In addition, I must be grateful to Pål Sturla Sæther who helped us get all the equipment needed. Also, I would like to express my gratitude to Otto J Wittner for lending us the equipment from Uninett. Overall, I strongly believe that the quality of this master thesis has been significantly improved with their helps.

I am thankful to Fenglin Han, Wenjing Sun, Shunwen Duan, Leif Arne Rønningen for participating in different pilot tests.

Finally, I would like to thank my family and all of my friends for their supports and encouragements during my study in Norway.

Contents

List of Figures

List of Tables

1 Introduction

Nowadays, many companies have set up a system to have remote meetings in order to reduce cost and save time. This kind of systems enable two or more people from different locations to have real-time communications and let people transfer video and audio files. However, it might be questions as whether the system being used is a telepresence system or a general video conferencing system and what are the differences between these two systems? In this chapter, all the answers are clarified. In addition, the meaning of telepresence quality is specified in this chapter.

Telepresence Quality Telepresence quality means the quality of a telepresence system. Telepresence quality involves many aspects. First of all, a system with good telepresence quality should be able to provide people who are interacting with the system with a sensation of nearness. In addition, a good telepresence quality system should be very convenient to use for end-users. Also, a system with good telepresence quality should be able to provide good reliability. Reliability, in this context, means the ability of system equipment working functionally. Finally, a good telepresence quality system should be able to provide real-time video imageries and stereophonic audio sounds.

Differences between a Telepresence System and a Video Conferencing System
A telepresence system with good quality contains at least the following features:

1. High-definition images
2. Stereophonic sounds
3. Real sizes of objects

This type of systems are designed to deliver an illusion that meeting participants in completely different locations are in the same room. In contrary, many video conferencing systems, most of time, provide people with poor quality and unreliable experience. It has the problems of inconsistent sizes of objects, fuzzy images, and even worse, unsynchronized pictures and sounds.

Distributed Multimedia Play (DMP) A DMP system presents a system architecture that can handle collaboration space distributed services. The goals of DMP systems are to guarantee optical end-to-end delay and graceful degradation of quality when traffic overloads the network or errors reduce the total network capacities, to provide a near-natural virtual collaboration, to provide a system that is easy to use, and to protect end-users' privacy. In other words, DMP systems are very intelligent and responsible in that the systems can track and interpret end-users' behaviors, and then take appropriate actions. Also, a DMP system can fulfill the functions of providing near-natural imageries and multi-channel sounds.

The reason why a general telepresence system such as Cisco telepresence system cannot be called a DMP system is that the collaboration space proposed by DMP systems makes end-users "believe" that they are in the same space by having high-quality multi-view videos and multi-channel sounds.

The quality requirements of DMP system can be summarized as follows: [1]

- Near-Natural Virtual Collaboration
- Real 3D sound
- Multi-view vision, adaptive
- Guaranteed maximum end-to-end time-delay
- Quality allowed to varies with time
- Quality variation guarantee
- Graceful degradation of quality when traffic overloads the network
- Graceful degradation of quality when system components fail (reduced capacity as well)
- Defined security level

Due to the facts that implementing a sturdy DMP system is out of the scope of this thesis and that implementing a whole DMP system requires a huge amount of work, starting from lower layer design, modeling, all the way to the higher layer implementation, the model of the DMP system realized in this thesis is a simple lab model of the DMP system. The proposed system is constructed by a PC, a display screen, two high-definition cameras, and a shutter glasses, which realizes some basic functions of a DMP system. This proposed system does not provide the following features:

1. Quality allowed to varies with time
2. Quality variation guarantee
3. Graceful degradation of quality when traffic overloads the network
4. Graceful degradation of quality when system components fail.
5. Multi-view, adaptive
6. Defined security level

Work The emphasis of this thesis lies in practical work. The goal of the thesis is to get a running lab model of the DMP system and then do various tests to investigate what are typical factors that can influence the users' perceptions.

The first step started from studying and using the available equipment in the laboratory. Then, it was necessary to choose and order other necessary equipment to do the desired tests. Obviously, choosing equipment, configuring the equipment properly and using the equipment to judge results imply a thorough understanding of the equipment. After that, proper tests and appropriate research methodologies were selected to find out the factors that influence the perceptions of the users. Finally, limitations and improvements of the thesis were mentioned.

Thesis Layout There are seven chapters that make up this whole thesis. Chapter 1 provides the definitions of the telepresence quality and the DMP. The differences between a telepresence system and a general video-conferencing system are specified. Chapter 2 introduces the methodologies that are used in the thesis. Chapter 3 specifies parameters that can influence the quality of a telepresence system. Chapter 4 illustrates how frame rates of a camera and refreshing rates of a screen can influence the optical delay of a system. Chapter 5 presents the stimulus-reaction time experiment and then concludes the factors that can influence the perception and reaction time. Chapter 6 introduces the concept of the space perception. A simple 3D system has been proposed and various pilot tests have been done to test how 3D can influence the perceptions of users. Chapter 7 summarizes the whole thesis, and future work is illustrated.

2 Methodology

In order to measure the optical delay of telepresence systems and investigate people's perceptions of telepresence systems, some methodologies are needed. This chapter illustrates the concepts of the chosen methodologies, and the advantages and disadvantages of the methodologies are specified. Some empirical investigation researchers distinguish between two approaches: quantitative approach and qualitative approach. However, empirical investigation can also be divided into three groups: descriptive investigations, relational investigations, and experimental investigations. The second way of dividing empirical investigation is preferred in this thesis, because it is more refined for our purpose and easier to understand. Answers of what these three terms mean and what the differences between these groups are can be found in the following sections.

2.1 Descriptive Investigations

As the name suggests, descriptive investigations focus on forming an accurate description of what is happening. Two typical examples of this type of investigation are survey and interview, which are further specified in the following.

2.1.1 Survey

A survey is a set of predefined questions to which a range of individual are asked to respond.

Target of a Survey The target of a survey is usually obvious, because the target of a survey is always a population of interest.

Survey Questions The goal of writing a survey is to get proper answers from the interest group. Hence, the key of achieving this goal is to ask proper questions. In general, the survey questions are close-ended, which provides an ordered response that requests the respondents to choose from.

Data Analysis After having all the necessary data, it is essential to analyze these data. By analyzing the data, it is easy to understand the relationship between variables and how could they influence each other.

Advantages and Disadvantages This investigation method does have its own advantages. It is easy to collect data from a large number of people at a relatively low cost. However, because questions of surveys are usually prepared in advance on paper or email and questions are usually fixed, it is impossible to ask follow-up questions if the answerers give very interesting and offbeat answers.

2.1.2 Interview

Because of the disadvantages of survey mentioned above, it is important to introduce the concept of interview. As everyone must know, an interview is usually a conversation between two people where questions are asked by the interviewer. Then, interviewees, based on their experience or knowledge, come up with answers.

Target When talking about interview, it is always a question that who is the person that should be interviewed. Based on the opinion from [2], "For investigations of broader concerns, such as system requirements or overall evaluation of system operation, a broader pool of interviewees drawn from all categories of stakeholder might be more informative. A stakeholder is anyone who is affected by the use of a system."

Structure It is necessary to consider about how to ask questions for the interview. One type of interview is fully structured interview, which the interviewer uses a script to present questions in a defined order. This type of interview makes it easy to analyze data, but if the interviewer comes up with some unanticipated questions during the interview, it is not possible to follow. Another type of the interview is semi-structured which enables the interviewer to let the conversation go wherever it may.

2.2 Relational Investigations

Relational investigations are used to discover the relationships between multiple different factors. For instance, in the equation $x+1=y$, we can tell that the value of y changes as the value of x changes. However, this type of investigation can seldom determine the casual relationship between several factors. For instance, it cannot determine whether there is a relationship between the lifespan of a person and the mood of that person

2.3 Experimental Investigations

Experimental investigations can be used to determine the causal relationships among factors. That is, it tells how it happens and why it happens.

Steps of Experimental Investigations Based on the suggestion from[2], the experimental investigation can be divided into the following steps:

1. Identify a research hypothesis
2. Specify the design of the study
3. Run a pilot study to test the design, the system, and the study instruments.
4. recruit participants
5. run the actual data collection sessions.
6. analyze the data
7. report the results

Based on the descriptions mentioned above, it is not hard to find out that these three types of investigations are not totally independent in that experimental investigations can compensate the limitations of relational investigations. Therefore, it is a common approach to include a combination of two or even three kinds of investigations to do a research project.

3 Telepresence Quality

Compared with a general video conferencing system, a DMP system makes a user feel that the collaborative interface is more natural. A DMP system gives a more "accurate" sensation of nearness to the people who are interacting with other people through the system. There are many factors that can influence users' perceptions of a DMP system. It is totally up to users to set which degree of tolerance is acceptable. Typical factors that influence the sensation of nearness of a general system are user equipment parameters and network parameters which are specified in the following sections. Another factor - delay which plays an important part in influencing the sensation of nearness is clarified in the next chapter.

3.1 Equipment Parameters

3.1.1 2D Video

The quality of a 2D video is typically determined by two parameters, the spatial resolution and the frame rate. Frame rate means the number of frames a camera can capture in a second, and it defines the temporal resolution. Typical modern cameras can capture 120 or more frames per second.To achieve an illusion of a moving image, the minimum frame rate required is 15 frames per second. Hence, for fast movements, in order to avoid the 'ghost effect', it is necessary to have a high frame rate.

The spatial resolution is another factor that influences the quality of a video. It is often represented by n*m, where the n represents the number of pixel columns (width), and the m is the number of pixel rows (height). This resolution is indeed the physical resolution, and should be higher than the perceived resolution. The higher the resolution, the better the details of an image can be shown. In general, two techniques are commonly used to project a video image on a display screen: interlaced scan and progressive scan. An image usually consists of a certain number of lines of pixels and a certain number of columns of pixels. With progressive scan, an image is captured, transmitted, and displayed line by line from the top to the bottom. In contrast, for interlaced scan, only a half of the total number of lines of an image is sent at a time, alternating between odd and even lines. Hence, this technique efficiently reduces the use of bandwidth by half. However, the problems of this scanning technique are that it can introduce flicker, that it reduces the spatial resolution and that it introduces quality issues.

Figure 1: Car Driving at 20km/h

As can be seen from the Figure 1[3], at left, a JPEG image is captured by an analog camera using interlaced scan technique. At right, a JPEG image is captured by a network camera using progressive scan technique. Both cameras used the same type of lens, and the car was driving at the same speed of 20km/h. It is obviously that the background is clear in both images, but the driver is only clearly visible in the image using progressive scan technique.

3.1.2 3D Video Collaboration Spaces

One of the biggest problems of current video-conferencing systems and general telepresence systems is that they lack space perception. In other words, they do not provide the function of telling the depth differences for different objects. A 3D video collaboration space can effectively solve this problem. It can provide accurate information about the depth relationships of objects in a scene. Detailed specification can be found in the chapter 6.

3.1.3 Audio

A major factor that affects the quality of the audio is bit rate, measured in kilobits per second (kbps). Normally, the higher the bit rate, the better the quality of the audio. Another important term has to be mentioned here, and that is audio channel. A channel is something through which a data signal travels from a source to a destination. Physically, an audio channel can be a wire through which an audio signal is propagated. In the case of stereo audio, it is essential to have multiple channels. The coordination between these channels directly influences the feeling of nearness of people. A typical case with multiple audio channels can be found at [4].

3.2 Network Parameters

Typical network parameters that influence the quality of a general telepresence system are packet loss, jitter, and bandwidth.

3.2.1 Packet Loss

Packet loss is the phenomenon that packets of data travelling across the network fail to reach their destinations. There are several factors that could result in this situation, such as network traffic congestions, the periods of changing network protocol for maintenance, and failures of network equipment. Typical problems caused by packet loss are blur, image freezing, and blockiness of video.

3.2.2 Bandwidth

Bandwidth is also called throughput, which is used to describe the amount of data that is allowed to transmit. It is measured in bit per second (bit/s). The bigger the bandwidth is, the larger the amount of data can be transmitted. In order to have a better coordination of audio channels and video channels, it is essential to have enough bandwidth.

3.2.3 Jitter

Jitter refers to the variation of delay, which is usually generated by the network congestion. Sound is very sensitive to network jitter in that jitter can cause a packet misses its playing-time slot. Hence, in order to overcome the problems of jitter in real-time sound transmission, it is a common practice to let the receiving system wait a sufficient period of time. Within this period of time, the incoming packets are saved in a terminal called a buffer. However, this technique does introduce one obvious disadvantage – delay.

4 Optical Delay Measurement

Optical delay is one of the most important factors that sets the level of satisfaction of the users in DMP systems. Hence, this chapter starts by illustrating the concept of optical delay. Then, the reasons of doing this measurement are specified. After that, the specification of the chosen methodology is mentioned. Finally, different scenarios of optical delay measurements have been done followed by the steps of the chosen methodology.

4.1 Optical Delay

Optical delay is the time spanning from an optical event occurs in the physical scene till it is actually shown on a screen. Optical delay commonly consists of two parts: network delay and equipment delay.

4.1.1 Network Delay

Network delay means the time a packet travels from its access node to its destination node. This delay includes the time used for transmitting the packet, the time of queuing when it meets network congestion, and the time used for the packet retransmitting when the packet is lost. It is worthy to mention that the time of network delay is not a constant value even when you use the same laptop and stay at the same place, because the routers or switches choose the path they “think” is the most appropriate based on the network status. There are several parameters that can influence the paths the gateway chooses, such as the network congestion, the number of jumps from source to destination, and network topology.

4.1.2 Equipment Delay

Equipment delay means the time equipment needs to fulfill its functionalities. Typical equipment in our DMP system that introduce delay are a camera and a display screen. For a camera, because of the fact that the shutter of a camera cannot be synchronized with the movement of an object, this asynchronous start of the camera introduces some delay. For example, let us assume that a camera shoots at two frames per second. As can be seen from the Figure 2, consider one line as a frame.

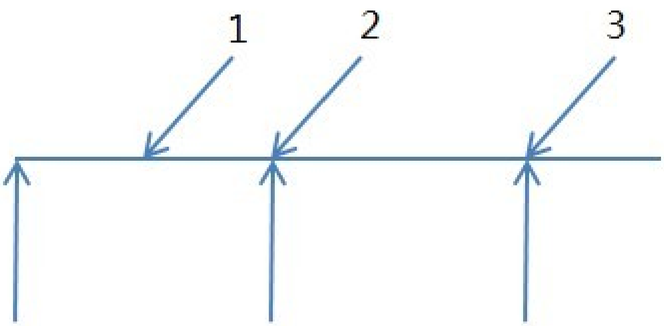

Figure 2: Frame Rate of a Camera

When an optical event happens at the point 1, the optical event will be processed immediately and then sent out at point 2, because it happens within the duration of the frame. However, if an optical event happens at the point 2 which is the end of the first frame, it will be sent out at point 3, because the camera needs certain time to process and store the current frame in the RAM of the camera. Then, the second frame of the camera can process the optical event. The time that is used to process and store optical events in the RAM of camera and the time caused by the asynchronous start of camera constitute the variable equipment delay of a camera.

In order to let a display screen present the optical event, the video board first needs a constant time to process the optical event. Then, because the refreshing rate of the screen cannot be synchronized with the optical event, which is of the same reason as the camera specified above, the optical event will be processed either by the current frame or the next frame of the screen and then be presented on the screen. Therefore, the time that is used to process the optical event and the time caused by the asynchronous start of the screen compose the variable equipment delay of the display screen.

Hence, for different equipment, it takes a certain time to fulfill its own functionalities, and the total time used by all the equipment in an end-to-end path of our DMP system contributes to the equipment delay.

4.2 Reasons of Doing the Measurement

A preliminary measurement of the optical delay with camera shooting at 60Hz can be found in [5]. The limitations of the measurements are that they were done with a camcorder only shooting at 60Hz, and the relationship between frame rates of the camera and the optical delay cannot be concluded. In real life, with a camera shooting at a higher rate, it enables people to catch fast movement and can provide people with a better quality image. Therefore, it is a common practice to let cameras shoot at a frame rate of 60Hz or more. This gives an acceptable spatial resolution in most cases, but for a high-quality telepresence system, it is of great interest to know the relationship between frame rates of cameras and the optical delay.

4.3 Methodology

In order to measure the optical delay, the methodology of experimental investigation is adopted and carried out in the following sections. Detailed definition and steps of experimental investigation can be found in the chapter 2. The purpose of this investigation is to measure the optical delay caused by equipment delay, so it does not include human test persons. Hence, the step 4 - recruit participants of the experimental investigation does not suit for this experiment, but all the other steps are suitable for the investigation and are therefore adopted.

4.4 Experimental Investigation

4.4.1 Research Hypothesis

Since people prefer cameras to shoot at higher frequency if it is possible because of the ability to get a better quality image, the research hypothesis of the experimental investigation is that with camera shooting at a higher frequency, it introduces more optical delay.

4.4.2 Design of the Experiment

Because the main purpose of this thesis is to investigate the realtionship the optical delay of a DMP system and frame rates of cameras, it is essential to propose an experimental model of the DMP system in order to do the measurements. Hence, for the simple DMP system presented in the following, we simplify the connection and use direct cables to connect all the equipment. Therefore, the measurement of the optical delay in the following sections includes only the equipment delay. However, when using the DMP system to do a remote communication, for example, from Trondheim to Oslo, the network delay must also be taken into account.

The experiment consists of using a flashing LED, a camera capable of shooting at a high frequency, an oscilloscope, a PC and a display screen supporting to display at a high frequency, which is shown in the Figure 3.

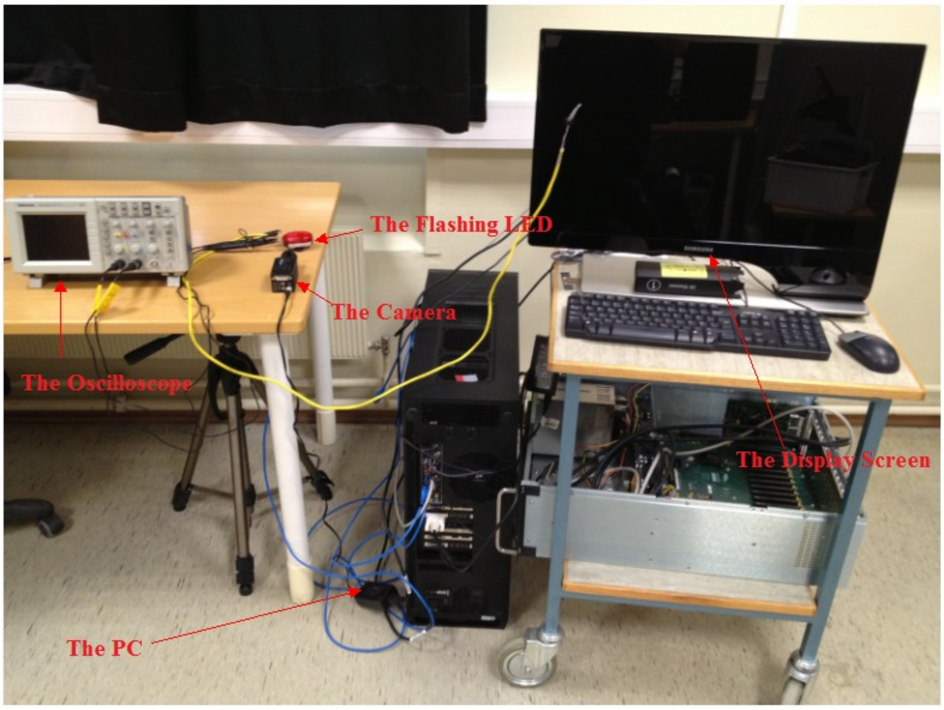

Figure 3: The Measurement System

The flashing LED is connected with two ends, one end is a sensor and another end contains three optical cables which contain the signal for the ground, the synchronized signal for the LED and the synchronized signal for the sensor separately. When the flashing LED flashes, the released light signal is captured by the camera in an asynchronous way in that it is impossible to synchronize the shutter of the camera with the pace of movement of an object. The camera is connected to the screen through the PC. The optical signal captured by the camera is transmitted to the screen through the PC and then displayed on the screen. The sensor of the LED is attached to the display screen. The oscilloscope contains two channels connected to the three lines of the flashing LED. One channel is connected to the cables with the synchronized signal for LED and the ground signal, which monitors the light signal when the flashing LED flashes. Another channel is connected to the cable with synchronized sensor signal and ground signal, which monitors the same light signal shown on the screen. The time used between the LED flashes and the same light signal being displayed on the screen is the optical delay.

4.4.3 The Instruments

4.4.3.1 The Flashing LED

Figure 4: The Flashing LED

As can be seen from the figure 4, the flashing LED connects with two ends, one end is a sensor as shown in the left and another end is the three optical cables which represent three types of signals mentioned above. The flashing LED can be built with different methods, such as using type 555 timer as mentioned in [6].

4.4.3.2 The Oscilloscope

Figure 5: Tektronic TDS 2012 Oscilloscope

As shown in the Figure 5, the Tektronic TDS 2012 is used. It is a two channel digital real-time oscilloscope. It contains two channels which are used to connect to the three cables of the flashing LED. It also possesses a screen which enables people to read the desired output. Most importantly, it contains two time sensors which enable people to read the time difference between time sensors.

4.4.3.3 The Camera

Figure 6: FL3-U3 USB 3.0 Digital Camera

Originally, the idea was to use the Toshiba camera IK-HR2D which is the same camera used by [5] to measure the optical delay in 60Hz. Then, by adjusting the shooting frequency to 120Hz, we can measure the optical delay. After that, based on the results from this measurement and the results from the [5], we can get a conclusion about whether the camera with a higher frame rate introduces more delay. However, because the Toshiba camera cannot support 120Hz, the FL3-U3 USB 3.0 digital camera has been selected. This camera provides a 32megabyte frame buffer, on-camera power, and status monitoring. Most importantly, it supports shooting frequency as high as 150Hz. The disadvantage with this camera is that it is a black-white camera, but because the main objective is to measure the optical delay, it is assumed that it does not influence the results of the measurement. In order to study the relationship between the frames rate of camera and the optical delay, the optical delay measurements with camera shooting at 60Hz, 120Hz, and 150Hz are selected for the following optical delay measurements.

4.4.3.4 The Display Screen

The display screen used is a SAMSUNG SyncMaster S23A950D. The maximum resolution of the screen can reach to 1920 * 1080. The detailed properties of this screen can be found in [7]. The reasons of choosing this screen are that it can support 3D, that it can provide high-quality pictures, and that it provides with refreshing rate as high as 120Hz. In order to find out whether the refreshing rates of the display screen can influence the optical delay, the refreshing rates of 60Hz, 100Hz and 120Hz are chosen for the following optical delay measurements.

4.4.3.5 The PC

The PC used to connect the display screen to the camera is not an ordinary PC. The CPU of the PC is Intel Core I7 – 3820 @3.6GHz. The RAM of the PC is 32GB. This PC is built specifically for the experiment in order to have the shortest processing time, and it can provide the best processing speed in the market. However, due to the fact that the equipment of measuring the processing speed of the PC is not available, we do not know how much delay has been introduced by the PC.

4.4.4 Actual Data Collection

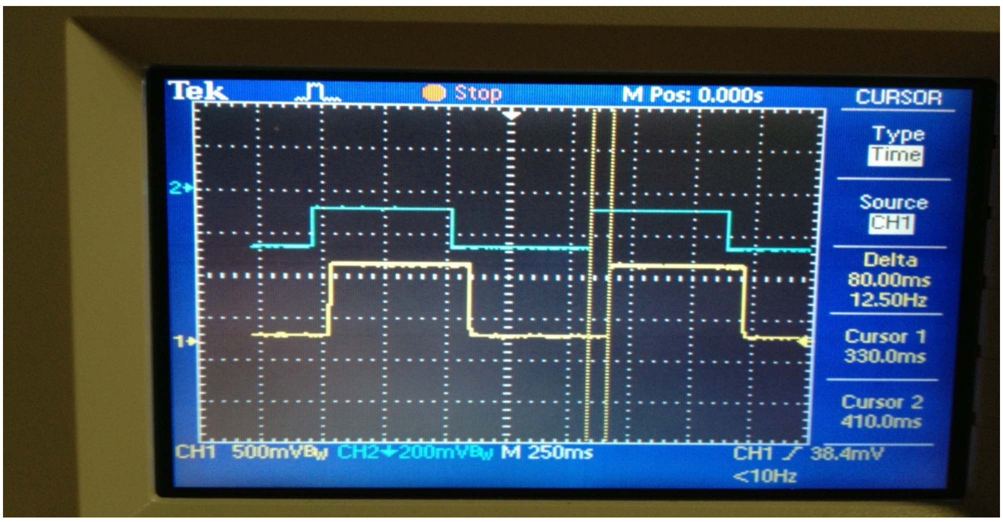

Figure 7: The Output From the Oscilloscope

As can be seen from the Figure 7, two waves are generated. The upper wave is generated when the LED flashes. The lower wave is generated when the sensor senses the light signal displayed on the screen. The flashing LED flashes in a constant interval, so the wave appears to be regular. There are two vertical lines shown on the screen that represent two time cursors. When putting the two time cursors at the same position of the separate waves, such as the peak of the waves, the time difference between these two cursors is the optical delay. The readout of Delta value is this time difference between two time cursors.

This optical delay measured with the configuration mentioned above consists of the equipment delay of the camera, the processing time of the PC and the equipment delay of the screen. We can assume the equipment delay of the camera is Tcam, the processing time of the PC is Tpc, and the equipment delay of the screen is Tsc. The Tcam contains two parts, the processing time of the camera and the time from the optical event happened until start of next frame. The processing time of each frame of the camera is a constant value, and we can regard it as K1. The asynchronous start of the next frame is a variable value and can be denoted as V1.The Tsc also includes two parts, the image processing time and the time for the asynchronous arrival of the image and the start of the refreshing frame of the screen. Since the processing time of the video board is constant, we can assume that it equals to K2. Because the time from the arrival of the image until the asynchronous start of the refreshing frame of the screen is a variable value, we can denote it as V2. Therefore, we can make a conclusion that:

$$Optical Delay = K1 + V1 + Tpc + K2 + V2$$

Because there are many interruptions with the PC, Tpc should be a variable value. Therefore, there are totally three variable values, and they are V1, V2 and Tpc.

Since this experiment is a about statistics, the adviser suggests that each measurement should be done twenty times in order to have comparatively accurate results. Hence, all the following measurements have been done twenty times. In order to make a comparison between different scenarios, the average value is calculated. Besides that, in order to calculate how far the data spreads out, the variance is also calculated. An unbiased estimator for the average value is:

$$Average\,M = (x1 + x2 + x3 + \ldots + xn)/n$$

In the equation, x represents the measurement result for each time, and n is the total number of measurements. An estimator for the variance is:

$$Variance\,v = \left[(M - x_1)^2 + (M - x_2)^2 + (M - x_3)^2 + ... + (M - x_n)^2\right]/n$$

In the equation, M is the average value calculated above. X is the measurement result for each time. V1, V2, and Tpc are assumed to be uniformly distributed. The average value that is calculated in each scenario should be:

$$K1 + average\,of\,V1s + average\,of\,Tpcs + K2 + average\,of\,V2s$$

Here, V1s means the sum of V1 of twenty times, and average of V1s means the mean value of the sum of V1. V2s means the sum of V2 of twenty times, and average of V2s means the mean value of the sum of V2. However, detailed study of how values distributed for each variable is out of the scope of the thesis. Calculating the maximum and minimum values is to make the results become straight forward.

4.4.4.1 Scenario 1

Camera (Hz)	Screen (Hz)
60	60

Table 1: The Camera With 60 Hz and the Screen With 60Hz

This is the first scenario where the camera is shooting at 60Hz and the refreshing rate of the screen is also 60Hz. The detailed measurement results for each scenario can be found in the Appendix A. The Figure 8 is the plot of measurement results. It shows that the optical delay is not a constant value, and most of the time, it varies between 70ms and 80ms.

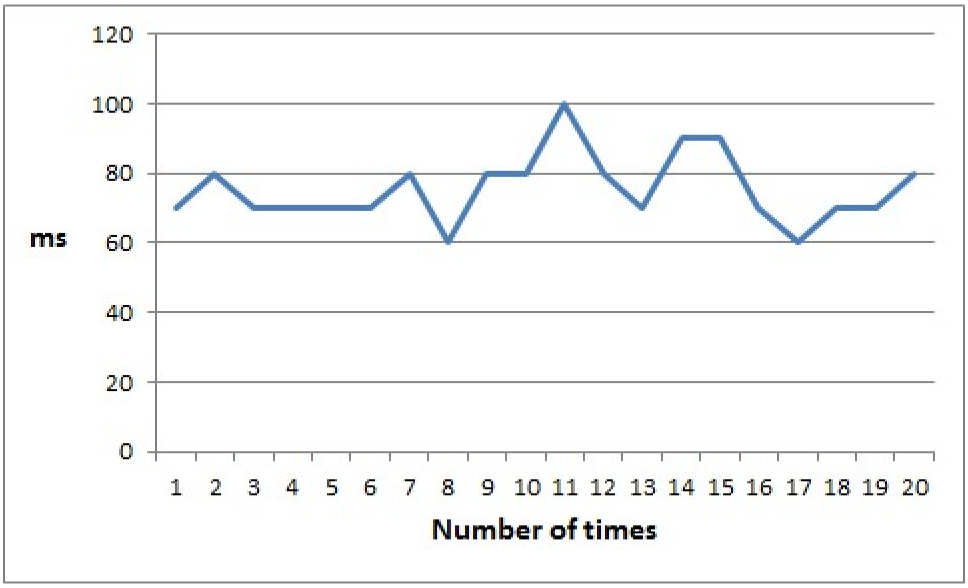

Figure 8: The Plot With the Camera Shooting Rate of 60Hz and the Screen Refreshing Rate of 60Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
60	100	75	95	40

Table 2: Optical Delay Measurement of Scenario 1

As mentioned previously that this measured optical delay consists of the time for asynchronous start of the frame of the camera, the processing time caused by the camera, the processing time of the pc for transmission, and the processing time of video board, and the asynchronous start of refreshing frame of the screen. Assuming the measurement results are normally distributed, the min value of 60ms is supposed to happen at the time that the current frame of the camera captures the optical event as soon as the optical event happens and the current frame of the screen catches the optical event as soon as the video board finishes processing the optical event.

The max value of 100ms should happen at the time that the optical event happens at the end of current frame of the camera and therefore it has to wait certain time till next frame of the camera to process it, and that the optical event has been sent to the end of the current frame of the pc screen and therefore it has to wait certain time till next time of refreshing. This analysis applies for the whole measurements for the following scenarios.

4.4.4.2 Scenario 2

Camera (Hz)	Screen (Hz)
120	60

Table 3: The Camera With 120Hz and the Screen With 60Hz

This is the second scenario where the camera is shooting at 120Hz and the refreshing rate of the screen is 60Hz. The Figure 9 is the plot of the measurement results. The Figure 9 shows that most of the time, the optical delay is between 60 and 70ms.

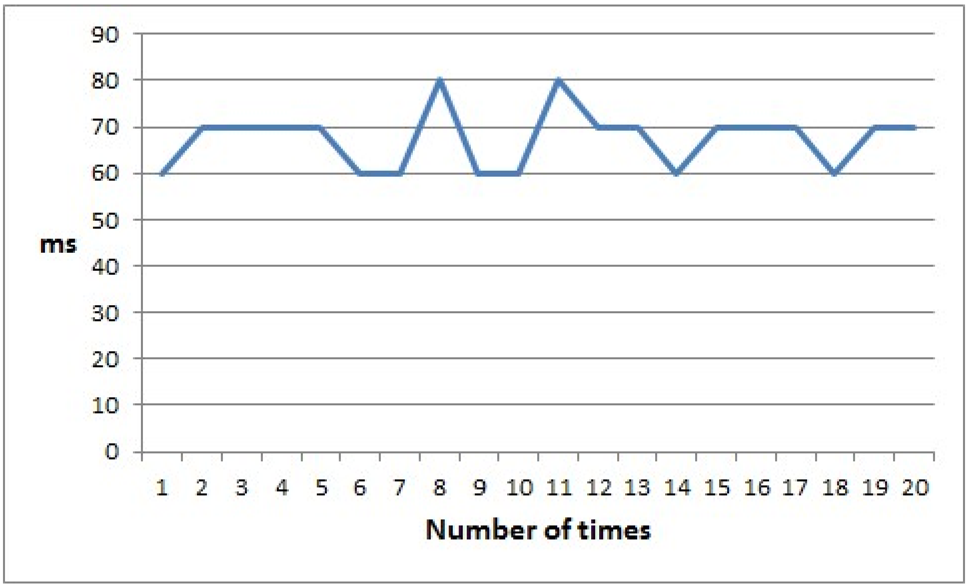

Figure 9: The Plot With the Camera Shooting Rate 120Hz and the Screen Refreshing Rate 60Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
60	80	67	39	20

Table 4: Optical Delay Measurement of Scenario 2

4.4.4.3 Scenario 3

Camera (Hz)	**Screen (Hz)**
150	60

Table 5: The Camera With 150Hz and the Screen With 60Hz

This is the third scenario where the camera is shooting at 150Hz and the refreshing rate of the screen is 60Hz. The Figure 10 is the plot of the measurement results. The Figure 10 shows that most of the time, the optical delay is between 60 and 70ms.

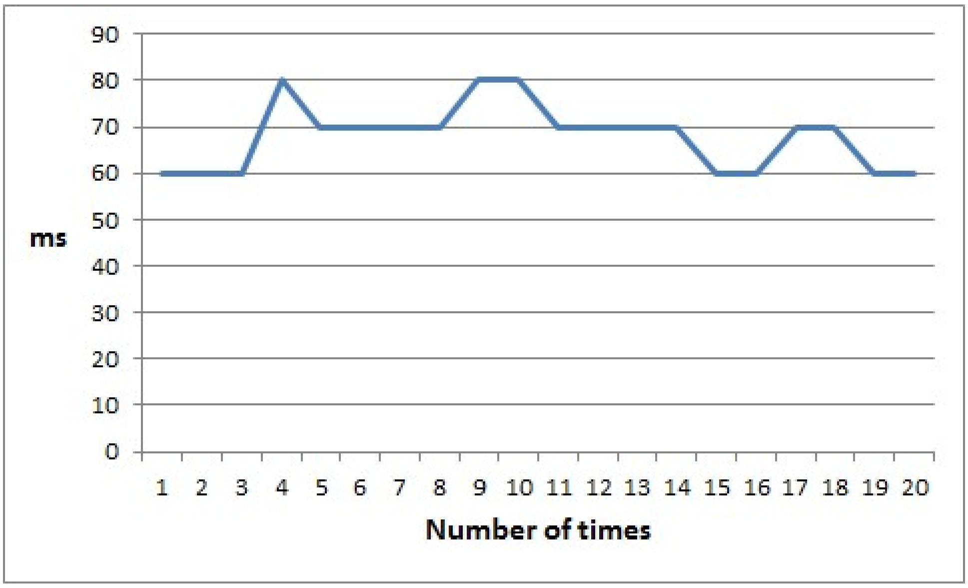

Figure 10: The Plot with the Camera Shooting Rate 150Hz and the Screen Refreshing Rate 60Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
60	80	68	46	20

Table 6: Optical Delay Measurement of Scenario 2

4.4.4.4 Scenario 4

Camera (Hz)	**Screen (Hz)**
60	100

Table 7: The Camera With 60Hz and the Screen With 100Hz

This is the fourth scenario where the camera is shooting at 60Hz and the refreshing rate of the screen is 100Hz. The Figure 11 is the plot of the measurement results. The Figure 11 shows that most of the time, the optical delay is between 60 and 70ms, and the optical delay can seldom be between 50ms.

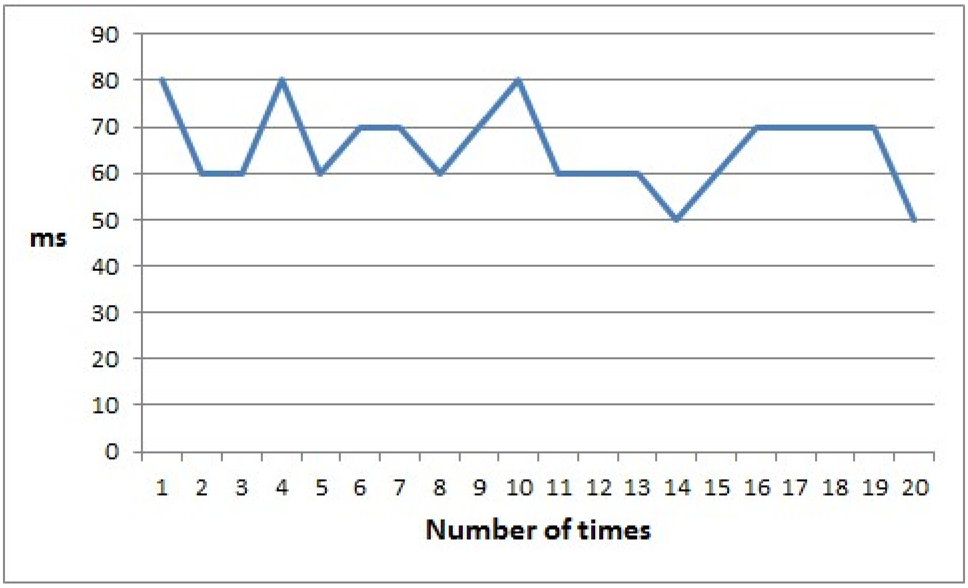

Figure 11: The Plot With the Camera Shooting Rate 60Hz and the Screen Refreshing Rate 100Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
50	80	65	75	30

Table 8: Optical Delay Measurement of Scenario 4

4.4.4.5 Scenario 5

Camera (Hz)	**Screen (Hz)**
120	100

Table 9: The Camera With 120Hz and the Screen With 100Hz

This is the fifth scenario where the camera is shooting at 120Hz and the refreshing rate of the screen is 100Hz. The Figure 12 is the plot of the measurement results. The Figure 12 shows that the optical delay is always between 50 and 70ms.

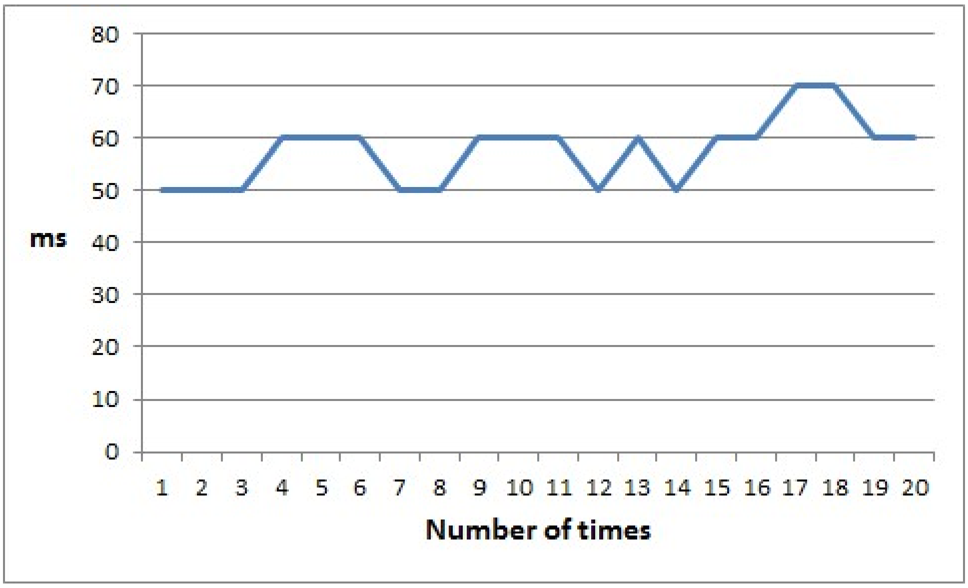

Figure 12: The Plot With the Camera Shooting Rate 120Hz and the Screen Refreshing Rate 100Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
50	70	57	39	20

Table 10: Optical Delay Measurement of Scenario 5

4.4.4.6 Scenario 6

Camera (Hz)	**Screen (Hz)**
150	100

Table 11: The Camera With 150Hz and the Screen With 100Hz

This is the sixth scenario where the camera is shooting at 150Hz and the refreshing rate of the screen is 100Hz. The Figure 13 is the plot of the measurement results. The Figure 13 shows that most of the time, the optical delay is between 50 and 60ms, and it can seldom be 70ms.

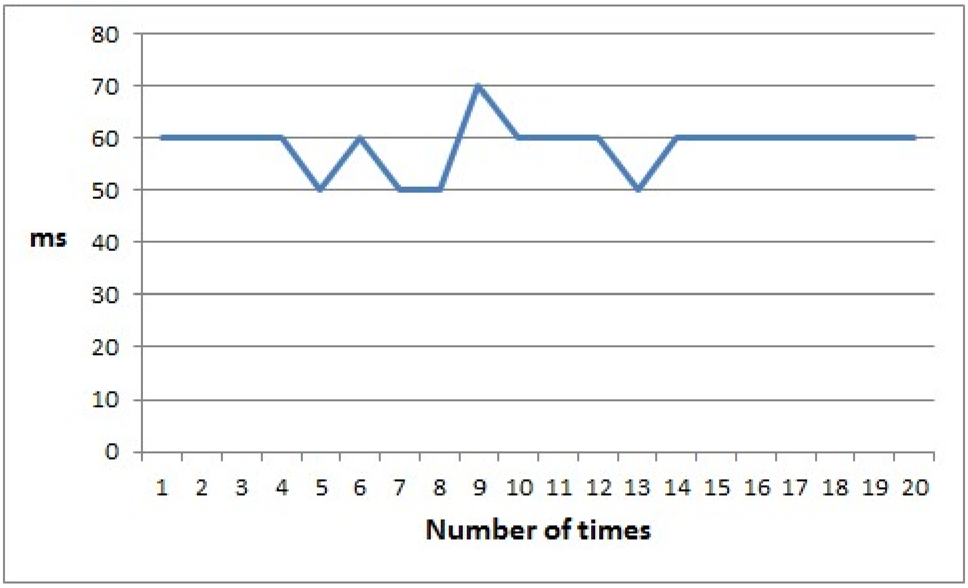

Figure 13: The Plot With the Camera Shooting Rate 150Hz and the Screen Refreshing Rate 100Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
50	70	58	23	20

Table 12: Optical Delay Measurement of Scenario 6

4.4.4.7 Scenario 7

Camera (Hz)	**Screen (Hz)**
60	120

Table 13: The Camera With 60Hz and the Screen With 120Hz

This is the seventh scenario where the camera is shooting at 60Hz and the refreshing rate of the screen is 120Hz. The Figure 14 is the plot of the measurement results. The Figure 14 shows that most of the time, the optical delay is between 50 and 70ms, and it can seldom be 80ms.

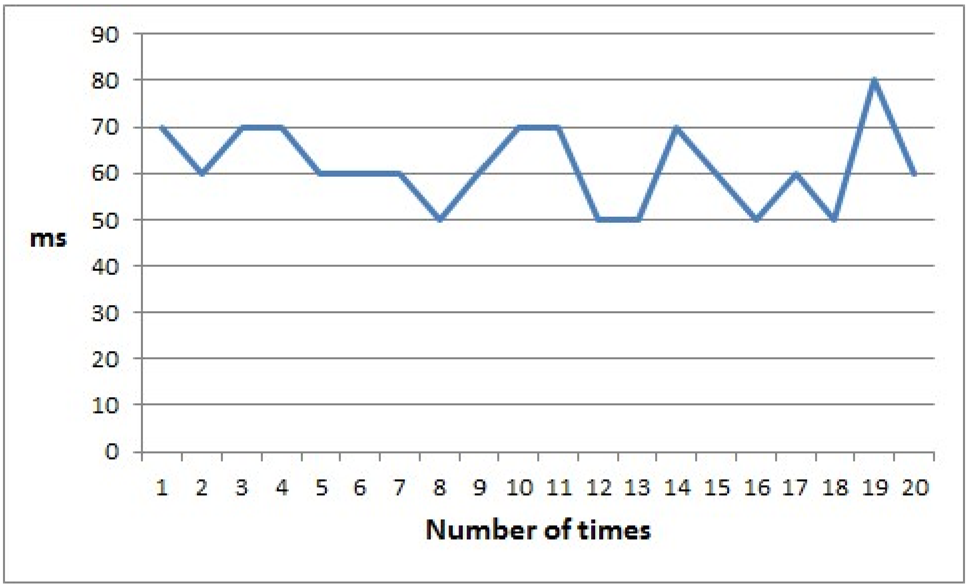

Figure 14: The Plot With the Camera Shooting Rate 60Hz and the Screen Refreshing Rate 120Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
50	80	61	73	30

Table 14: Optical Delay Measurement of Scenario 7

4.4.4.8 Scenario 8

Camera (Hz)	**Screen (Hz)**
120	120

Table 15: The Camera With 120Hz and the Screen With 120Hz

This is the eighth scenario where the camera is shooting at 120Hz and the refreshing rate of the screen is 120Hz. The Figure 15 is the plot of the measurement results. The Figure 15 shows that most of the time, the optical delay is between 50 and 60ms, and it can seldom be 40ms or 70ms.

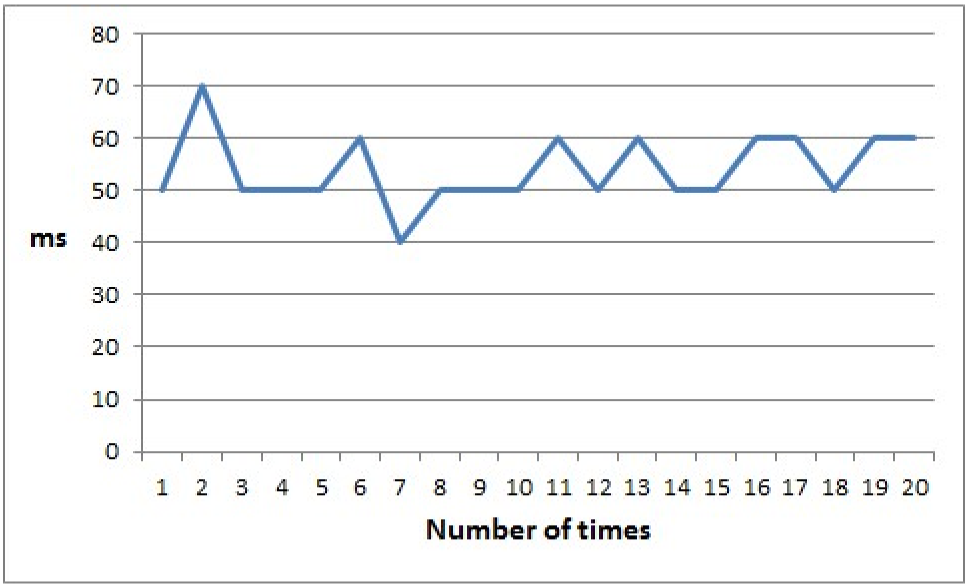

Figure 15: The Plot With the Camera Shooting Rate 120Hz and the Screen Refreshing Rate 120Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
40	70	54	44	30

Table 16: Optical Delay Measurement of Scenario 8

4.4.4.9 Scenario 9

Camera (Hz)	**Screen (Hz)**
150	120

Table 17: The Camera With 150Hz and the Screen With 120Hz

This is the ninth scenario where the camera is shooting at 150Hz and the refreshing rate of the screen is 120Hz. The Figure 16 is the plot of the measurement results. The Figure 16 shows that the optical delay is always between 50 and 60ms.

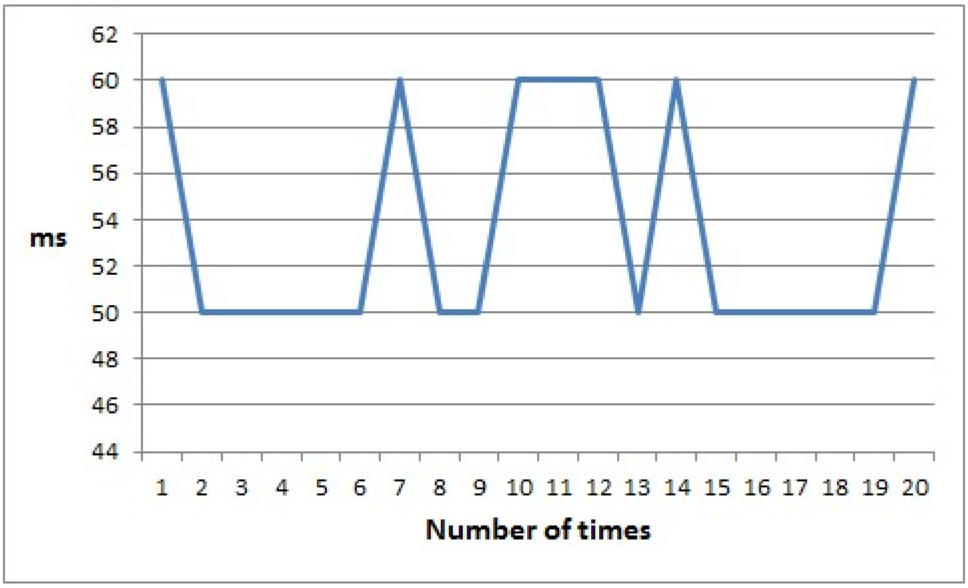

Figure 16: The Plot With the Camera Shooting Rate 150Hz and the Screen Refreshing Rate 120Hz

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
50	60	53	23	10

Table 18: Optical Delay Measurement of Scenario 9

4.4.5 Data Analysis

According to the measurement results above, several conclusions can be drawn. All the following conclusions are based on the average values of each scenario.

First of all, contrary to the hypothesis, with the same refreshing rate of the screen, the camera shooting at the higher frequency introduces less optical delay than the camera shooting at the lower frequency. This is proved by the fact that even though with the same refreshing rate of the screen, the optical delay with camera shooting at 120Hz is similar with the optical delay caused by the camera shooting at 150Hz, the optical delay with camera shooting at 120Hz and 150Hz are obviously shorter than the optical delay with the camera shooting at 60Hz.

In addition, the relationship between the refreshing rate of the screen and the optical delay cannot be ignored. Evidently, with the camera shooting at the same frequency, the higher the refreshing rate of the screen, the shorter the optical delay is.

One disadvantage has to be admitted that the positions of the two time cursors of the oscilloscope are modified and justified by human beings instead of machine, so it probably introduces small human errors with the final readouts of the measurement results. Hence, in order to increase the accuracy, it is better to automate the process of measurements.

4.4.6 Results

All the detailed data results can be found in the Appendix A .

5 Stimulus-reaction Time Measurement

One goal of DMP systems is to make users feel more natural and give a more "accurate" sensation of nearness to people who are interacting with other people through the system, thus getting the same perceptions and reactions that a person experiences in an ordinary environment. Hence, it is of interest to know users' perception and reaction time and what factors can influence users' perception and reaction time. This chapter firstly introduces the concept of stimulus-reaction time. Then, the specification of the chosen methodologies is mentioned. Finally, stimulus-reaction time measurements are carried out followed by the steps of the chosen methodology.

5.1 Concept of Stimulus-reaction Time

Stimulus-reaction time is the time from an optical event being shown on the screen until a test person has given a response to it. In our case, the response is to make a sound – 'ah'. This stimulus-reaction time includes both the perception and the reaction time.

5.2 Methodology

In order to find out the factors that influence people's perception and reaction time, the methodologies of experimental investigation and a survey, one type of descriptive investigation, are adopted and carried out. Detailed definition and steps of experimental investigation can be found in the chapter 2. Specification of descriptive investigation and the analysis of survey including both advantages and disadvantages can also be found in the chapter 2. One thing has to admit that this is a pilot test and with limited test persons, the results can only provide as a theory, but it cannot be generalized until a lot of more testers involve in the tests to prove it.

5.3 Experimental Investigation

5.3.1 Research Hypothesis

When talking about the perception and reaction time, it is a natural practice to think of it having relationships with age, gender and education. Hence, based on these common senses, the research hypothesis is that the perception and reaction time is directly involved with gender, age and education.

5.3.2 The Instruments

The experiment consists of the flashing LED, the camera, the PC, the oscilloscope and the display screen, which are specified in the previous chapter. In addition to these instruments, in order to measure people's reaction time, another equipment, Sennheiser ew 152 G3, has been chosen and been purchased. The main reason of choosing this equipment is that the equipment can provide with a high audio sensitivity and a high quality sound. The equipment consists of a receiver, a wireless headset, and a transmitter. Detailed property descriptions can be found at [8].

5.3.3 Design of the Measurement

Based on the study of the previous chapter, the configuration of the camera shooting at 120Hz and the screen with the refreshing rate of 120Hz introduces 54ms delay in average, which is shorter than most of the other scenarios. Hence, these parameters have been chosen and configured with the screen and the camera. Based on the definition of stimulus-reaction time specified previously, stimulus-reaction time in our system is the time between a tester seeing an optical event happening on the screen and the tester reacting to the event. This time equals the time between the sensor senses the optical signal being shown on the screen and the time the transmitter connected to the headset outputs the signal. The Figure 17 shows how the measurement is done.

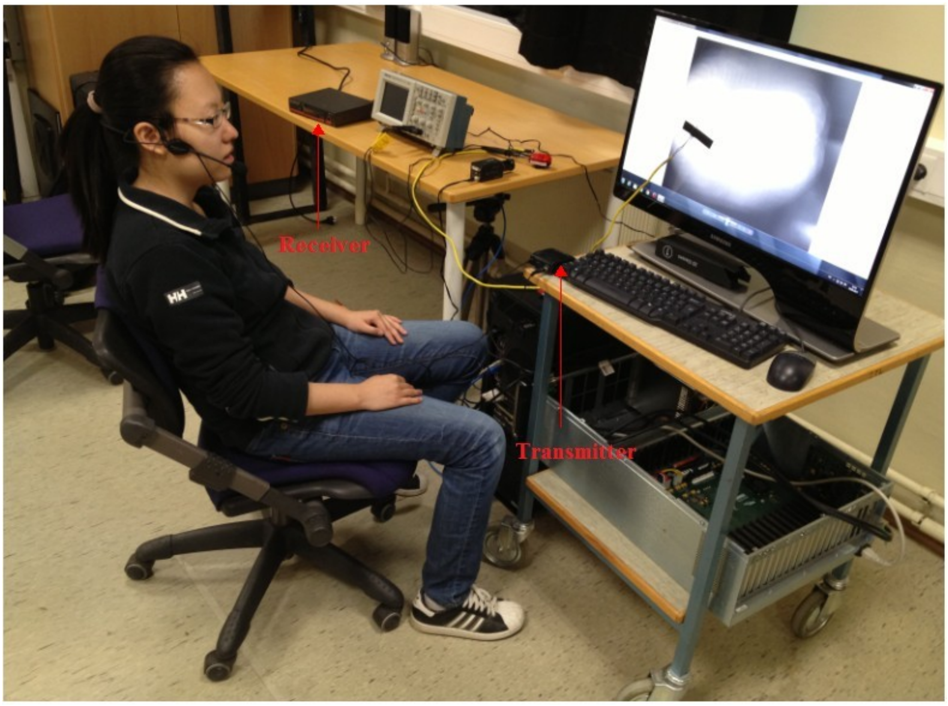

Figure 17: Reaction Time Measurement

For the measurement, one channel of the oscilloscope is connected to the synchronized signal of the sensor and the ground signal. Another channel of the oscilloscope is connected to the output signal of Sennheiser ew 152 G3's receiver and the ground signal. There are two types of the optical events that are carried out for each test person. One type of the optical events happen in a fixed interval. Another type of the optical events appear in a variable interval. The reason of choosing the fixed interval optical event is that with the measurements being done continuously, it is easy for testers to predict the time of the next optical event happening on the screen, which is similar to some cases in real life. For example, when a conductor conducts a concert with a choir, the people of the choir can, based on their knowledge, predict the gesture that the conductor will do, thus reducing the perception and reaction time. Another situation with optical event happening in a variable interval has also been selected, because this also reveals the real situations in life. When the testers see the optical events, they are required to utter the sound of 'ah'.

5.3.4 Recruit Participants

Because the measurement needs a qualitative result and the fact that the time is really limited, five testers are chosen and each of them is tested twenty times for each scenario. That is, a total number of forty times of the measurement has been done for each tester. Permission contract has been signed with testers, which can be found in the Appendix B.

5.3.5 Actual Data Collection

For the fixed interval optical events, when the LED does not flash, the display screen looks like the Figure 18. When the LED flashes, the display screen is shown as the Figure 19.

Figure 18: The Screen With No Optical Event

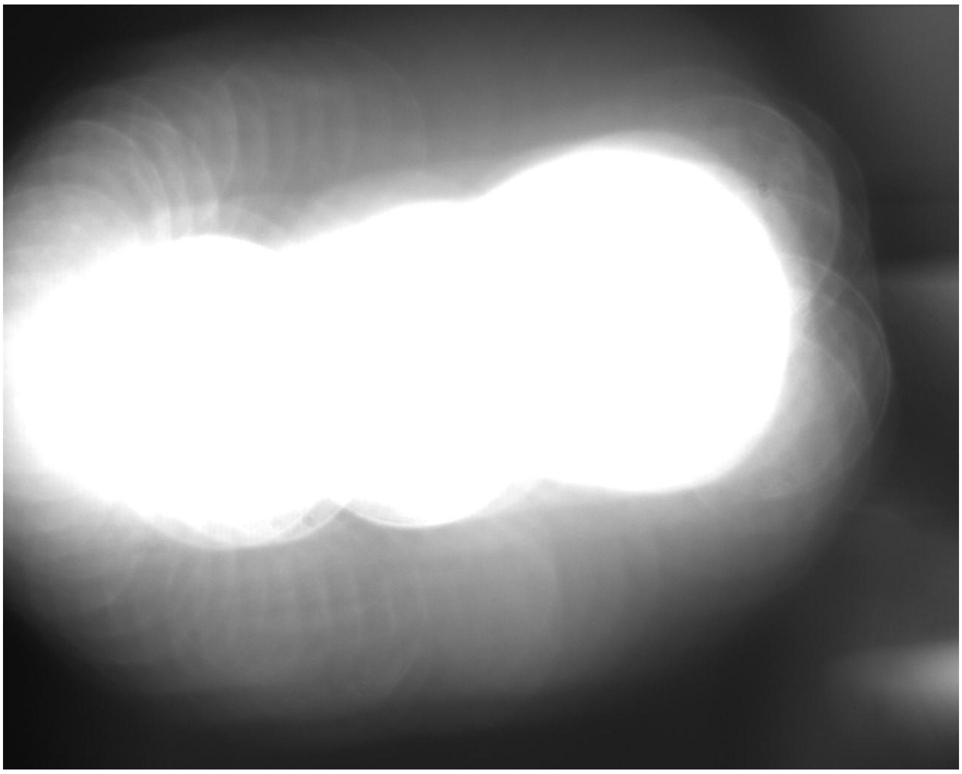

Figure 19: The Screen With an Optical Event

For the variable interval optical events, when no optical event happens, the screen looks like the Figure 18. When the optical event happens, the screen is shown as the Figure 20.

Figure 20: Optical Event Happened in a Variable Interval

Tester 1

Gender	**Age**	**Education**
Female	26	Master

Table 19: Tester 1's Information

Scenario 1: optical event happens in a fixed interval The plot of the measurement results is shown in the Figure 21. As can be seen from the Figure 21, most of the time, the perception and reaction time is between 150ms and 200ms. However, because of the prediction, the shortest time can be less than 100ms.

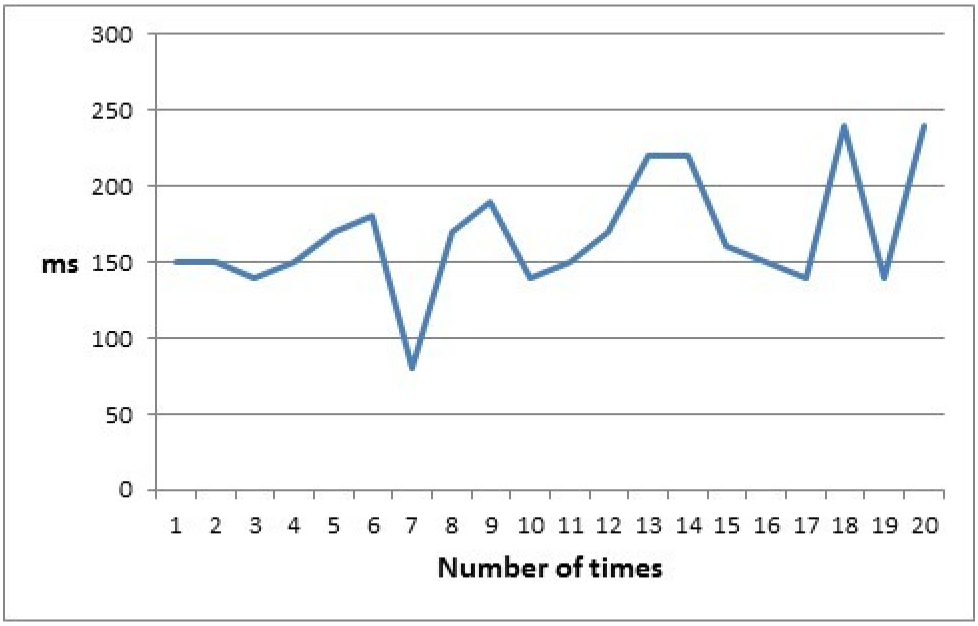

Figure 21: Tester 1: the Optical Events With a Fixed Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
80	240	167	1449	160

Table 20: Tester 1 With a Fixed Interval

One thing has to be emphasized here that the optical delay measured from the previous chapter does not influence the reaction time, because testers start to react to the optical event as soon as it happens on the screen.

Scenario 2: optical event happens in a variable interval The plot of the results is shown in the Figure 22. As can be seen from the Figure 22, since the tester cannot predict the optical event, the distribution is comparatively more regular and the perception and reaction time is always between 200ms and 300ms.

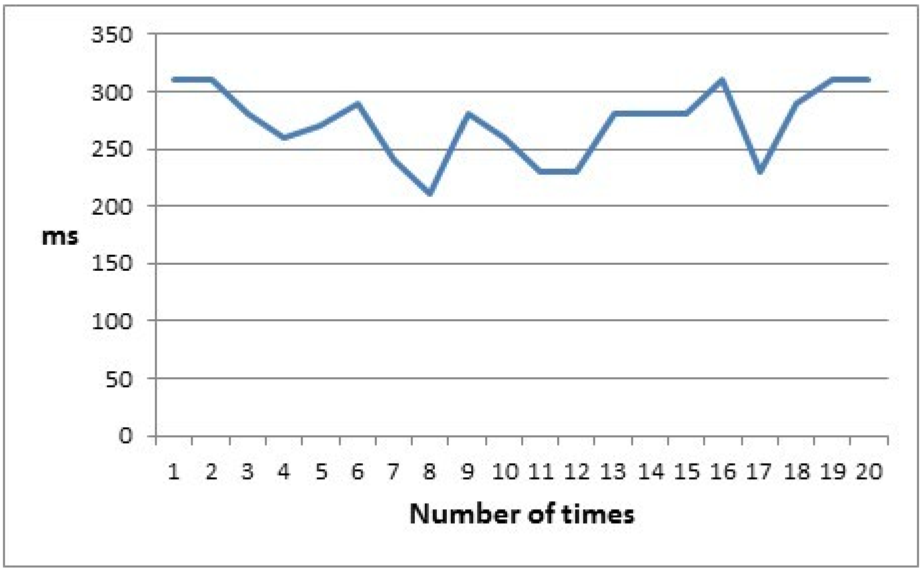

Figure 22: Tester 1:The Optical Event With a Variable Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
210	310	272	931	100

Table 21: Tester 1 With a Variable Interval

Tester 2

Gender	**Age**	**Education**
Male	29	Phd Student

Table 22: Tester 2's Information

Scenario 1: optical event happens in a fixed interval The plot of the results is illustrated in the Figure 23. As can be seen from the Figure 23, most of the time, the perception and reaction time is between 200ms and 300ms. However, because of the prediction, the shortest time can be less than 100ms.

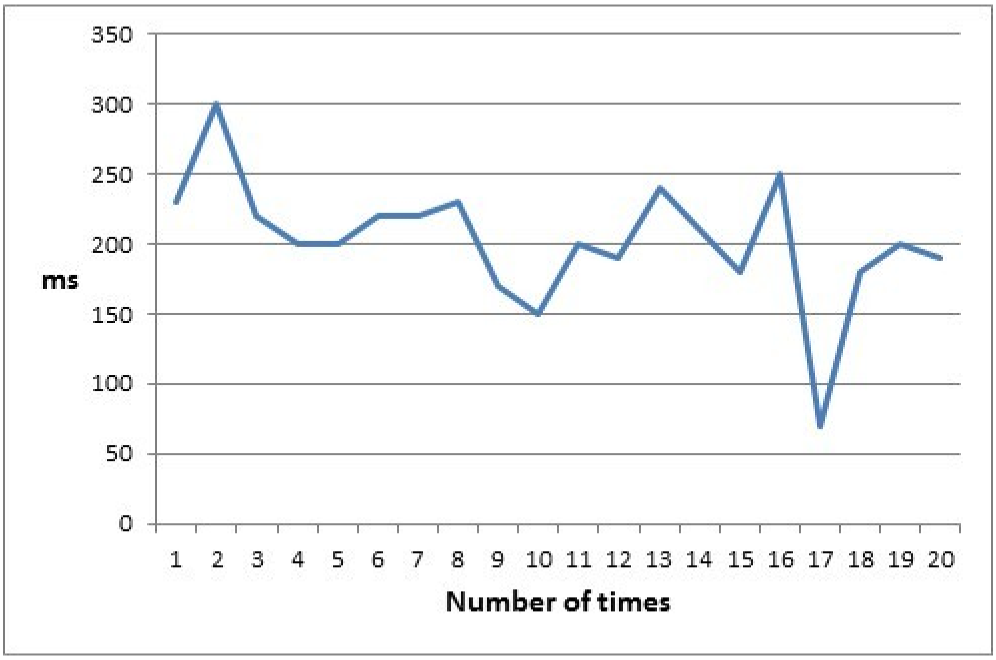

Figure 23: Tester 2: The Optical Event With a Fixed Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
870	300	201	1919	230

Table 23: Tester 2 With a Fixed Interval

Scenario 2: optical event happens in a variable interval As it shown from the Figure 24, for this scenario, the distribution is much more regular compared the plot of optical event happening in a fixed interval, and the perception and reaction time is mostly between 250ms and 350ms.

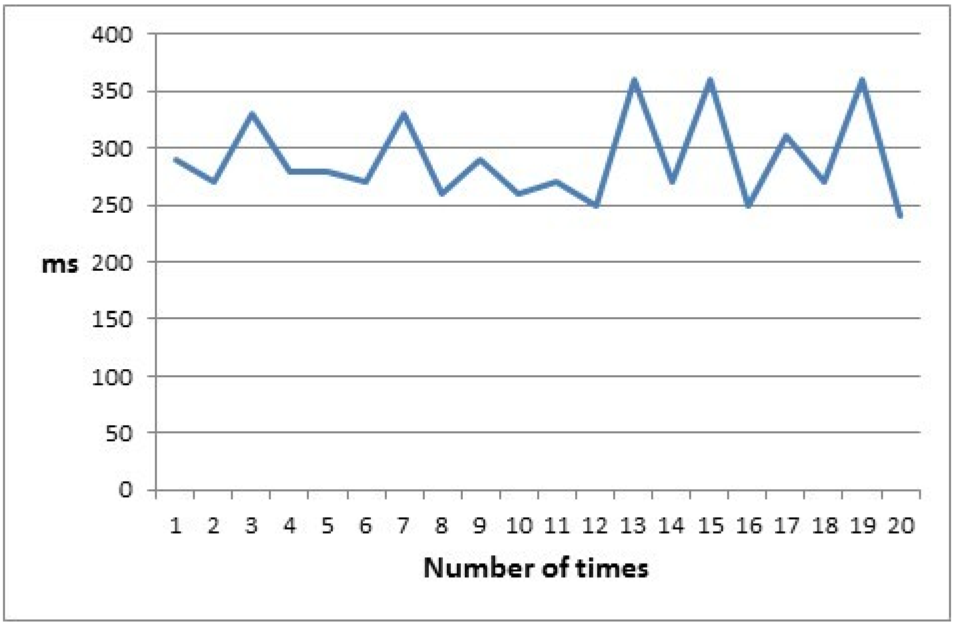

Figure 24: Tester 2: The Optical Event With a Variable Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
240	360	291	1400	120

Table 24: Tester 2 With a Variable Interval

Tester 3

Gender	**Age**	**Education**
Male	65	PhD

Table 25: Tester 3's Information

Scenario 1: optical event happens in a fixed interval The plot of the results is shown in the Figure 25. As can be seen from the Figure 25, most of the time, the perception and reaction time is between 200ms and 250ms. In addition, the distribution is much more regular compared with other testers. However, because of the indispensable prediction, the shortest reaction time is close to 150ms.

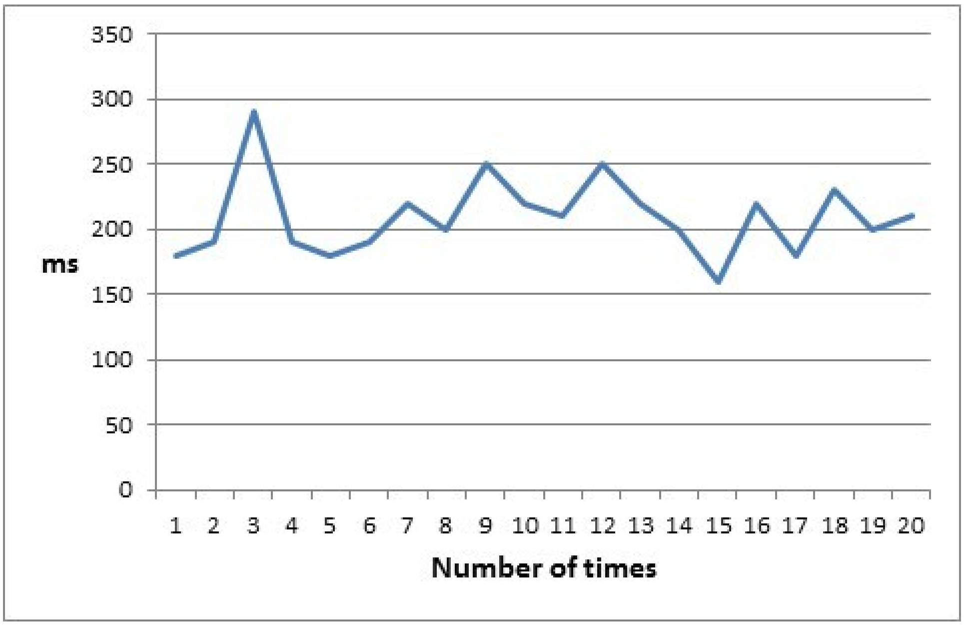

Figure 25: Tester 3: The Optical Event With Fixed Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
160	290	211	855	130

Table 26: Tester 3 With a Fixed Interval

Scenario 2: optical event happens in a variable interval The plot of the results with optical event happening in a variable interval is shown in the Figure 26. As it can be seen from the Figure 26, the distribution is much more regular compared with other testers, and the perception and reaction time is mostly between 250ms and 350ms.

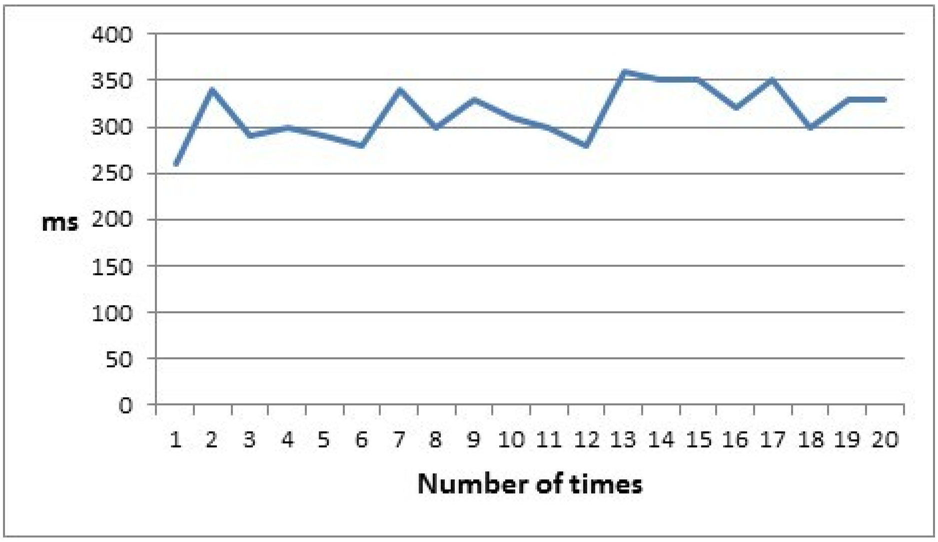

Figure 26: Tester 3: The Optical Event With Variable Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
260	360	315	765	300

Table 27: Tester 3 With a Variable Interval

Tester 4

Gender	**Age**	**Education**
Male	26	Master

Table 28: Tester 4's Information

Scenario 1: optical event happens in a fixed interval The plot of the results with fixed interval is shown in the Figure 27. As can be seen from the Figure 27, because of the prediction, the distribution is not regular and the shortest time can almost be 100ms.

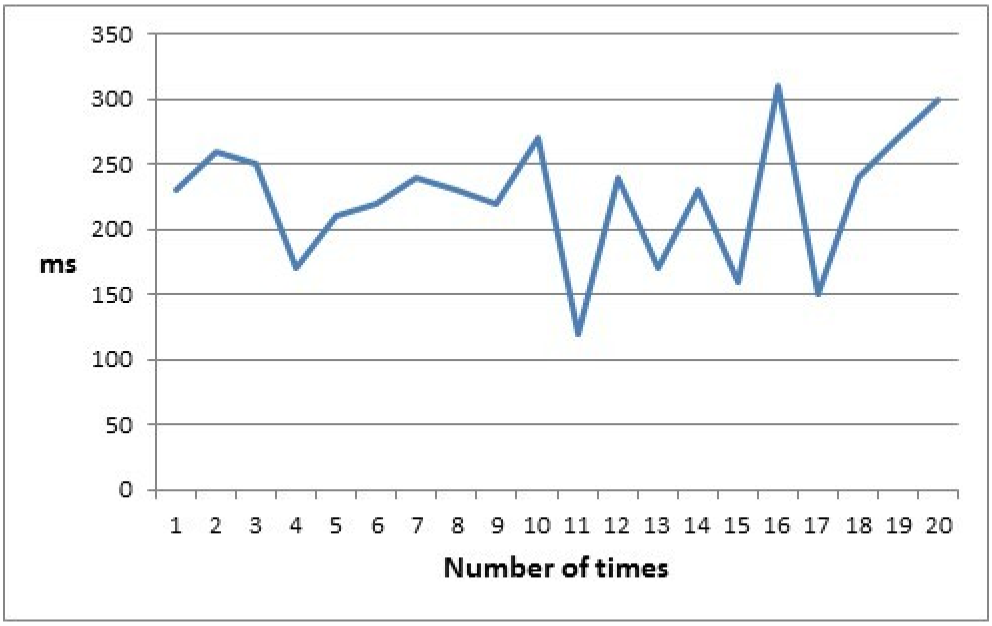

Figure 27: Tester 4: The Optical Event With a Fixed Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
120	310	224	2335	190

Table 29: Tester 4 With a Fixed Interval

Scenario 2: optical event happens in a variable interval The plot of the results with optical event happens in a variable interval is shown in the Figure 28. As it shown from the Figure 28, the distribution is much more regular, and the time is mostly between 200ms and 300ms.

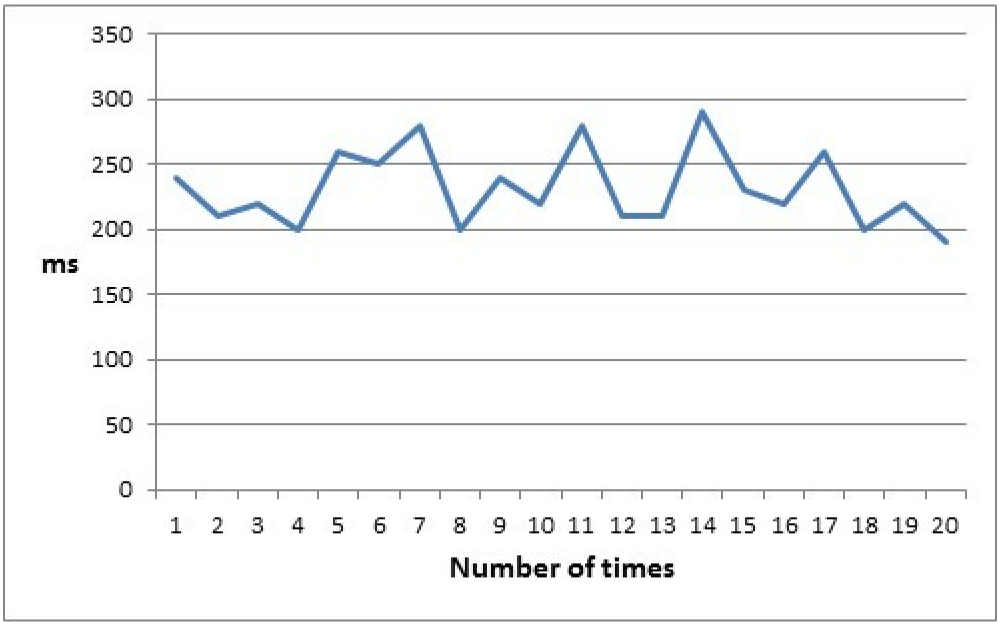

Figure 28: Tester 4: The Optical Event With a Variable Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
190	290	232	843	100

Table 30: Tester 4 With a Variable Interval

Tester 5

Gender	**Age**	**Education**
Female	26	Master

Table 31: Tester 5's Information

Scenario 1: optical event happens in a fixed interval The plot of the results with optical event happening in a fixed interval is shown in the Figure 29. As can be seen from the Figure 29, the result is comparatively more regular than other testers and the time is mostly between 150ms and 200ms. The phenomenon of prediction seldom happens.

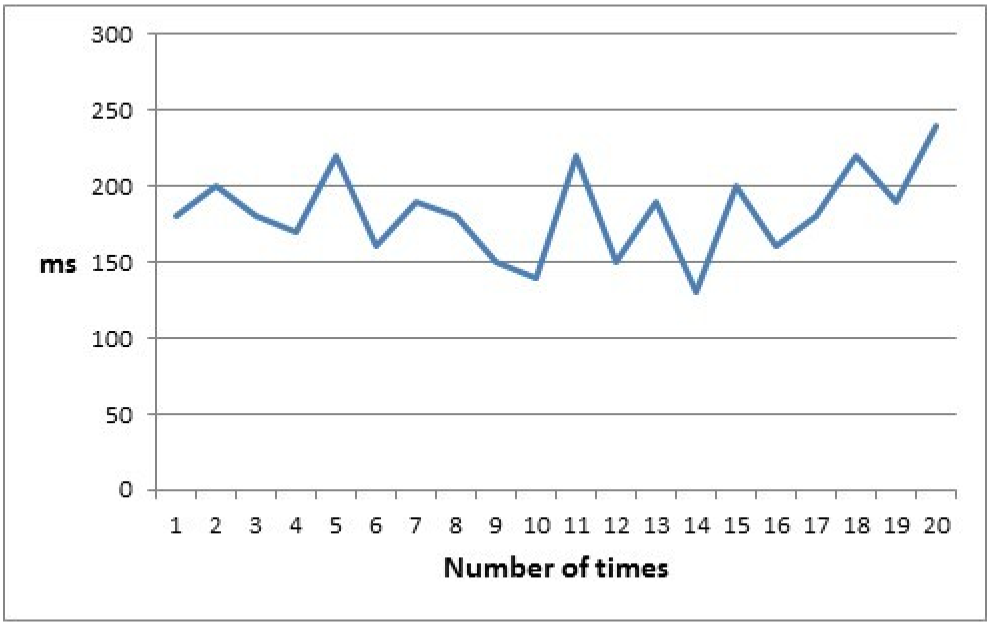

Figure 29: Tester 5: The Optical Event With a Fixed Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
130	240	183	809	110

Table 32: Tester 5 With a Fixed Interval

Scenario 2: optical event happens in a variable interval The plot of the results with optical event happening in a variable interval is shown in the Figure 30. As it shown from the Figure 30, the perception and reaction time is mostly between 200ms and 300ms.

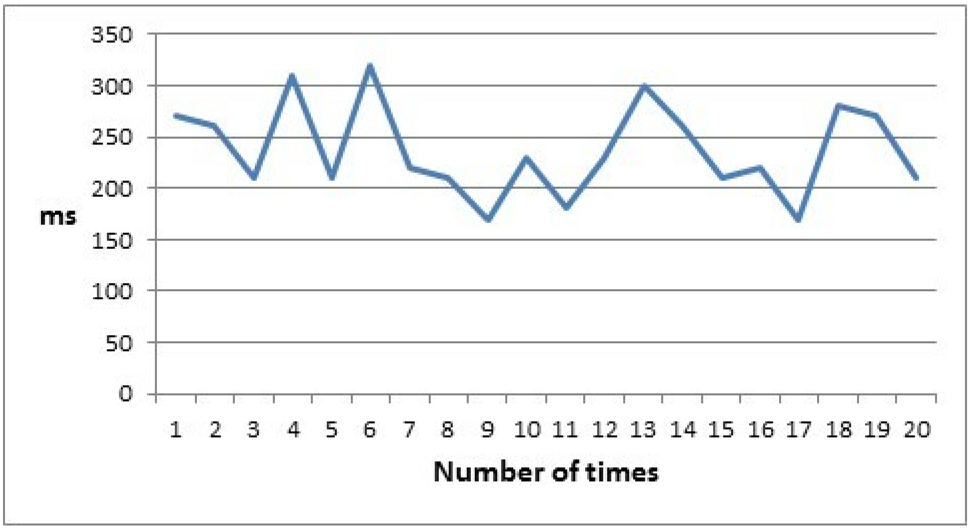

Figure 30: Tester 5: The Optical Event With a Variable Interval

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
170	320	238	1891	150

Table 33: Tester 5 With a Variable Interval

5.3.6 Data Analysis

5.3.6.1 Delay of Wireless Transmission

Because of the fact that in order to measure the reaction time, the wireless equipment of Sennheiser ew 152 G3 has been used. It contains a transmitter and a receiver, and the answer to the question of whether the wireless transmission introduces delay is unknown. Hence, a measurement has been done to measure the delay introduced by wireless transmission. The measurement has been done with two microphones, one ordinary microphone with the wire and another one is the Sennheiser ew 152 G3. Both output signals from the microphones are connected to the oscilloscope, and the time difference between two microphones shown on the oscilloscope is the delay introduced by the wireless transmission. As can be seen from the Figure 31, the time difference between two microphones is quite small and it is always between 3ms and 8ms.

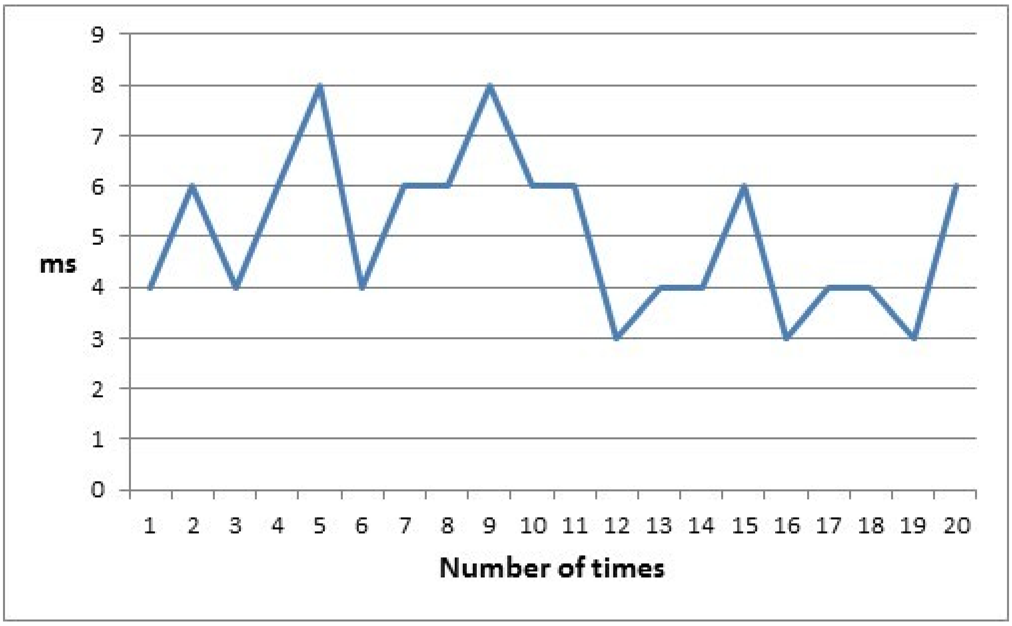

Figure 31: The Delay Caused by Wireless Transmission

Result:

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
3	8	5	2	5

Table 34: The Delay of Wireless Transmission

5.3.6.2 Data Analysis

Because the number of testers is limited, the conclusion cannot be generalized. Hence, the analysis is only valid with the proposed measurement results. One thing can be suggested that the perception and reaction time seems to relate to age of the testers. As can be seen from the data, the younger the person is, the shorter the perception and reaction time is. However, from the results of the measurements, it seems that the gender does not influence too much of the perception and reaction time. In addition, the education seems to have no relationship with the perception and reaction time. What is more, from the plots of the results, it indicates that the female intends to predict more than the male, even though they have been told not to predict.

As can be seen from the measurement results, the smallest difference of perception and reaction time for a tester between the optical event happening in a fixed interval and the optical event happening in a variable interval is 90ms. In other words, if the user of a telepresence system can provide a constant hint or stimulus to the chat partners, it can drastically reduce the perception and reaction time.

5.3.7 Results

All the results can be found in the Appendix C.

5.4 Descriptive Investigation

The detailed information about how to make a survey has been explained in the chapter 2. The survey being conducted can be found in the Appendix D.

Analysis of the Survey It is worthy to mention that the following conclusions are made based on the survey of five testers, and it cannot be generalized. From the survey, it is easy to make a conclusion that the perception and reaction time is directly involved with the tiredness of the people. In addition, concentration is the most important influencer of the perception and reaction time. Next, practice and life habit can also, to some extent, influence the perception and reaction time. However, education is certainly has nothing to do with the perception and reaction time.

6 Space Perception

Space perception is one of the most important factors that distinguish a general telepresence system and a DMP system. It can help to give a more "accurate" sensation of nearness to the people who are interacting with other people through a DMP system. However, the answer to the question about whether the space perception can influence the perception and reaction time is unknown. The detailed study of the influences of the space perception to the perception and reaction time is illustrated in this chapter. This chapter starts by introducing the related knowledge about the space perception. Then, the specification of the chosen methodologies which are used to find how space perception can influence perception and reaction time is mentioned. In the end, the steps of the chosen methodologies are carried out.

6.1 Space Perception

6.1.1 Concept of Space Perception

Space perception is to tell the depth difference between objects. Most of the time, 2D videos eliminate the space perception. However, because of the monocular depth cues, it can help to create space perception in 2D videos. Monocular depth cues help to create space perception by enabling observers to learn the physical significance of different retinal images. This is achieved by interposition, linear perspective, light and shade, relative size, texture gradient, and aerial perspective. Detailed definitions and explanations of these concepts can be found in the [9]. In contrast, binocular vision enables people to derive from the small differences in the location of homologous points in the two images incident on the retina of the eyes [9].

6.1.2 3D

A stereoscopic 3D can effectively provide space perception. The most commonly used approach of 3D video collaboration spaces is stereoscopy. Stereoscopy creates the illusion of depth of an image by means of stereopsis for binocular vision.

It presents two offset of a 2D image to the left and right eye of the viewer separately, and then these two 2d images are combined in the brain to give the depth perception.

6.1.3 3D Format

There are mainly three types of 3D formats, and they are side-by-side 3D, top-and-bottom 3D and frame sequential 3D.

Side-by-side 3D: For side-by-side 3D, a frame that is intended to be seen by both eyes at the same time is divided into two halves. Hence, the entire frame scaled down horizontally to fit for the left eye and right eye respectively.

Top-and-bottom 3D: For top-and-down 3D, a frame that is intended to be seen by both eyes at the same time is divided into two halves, and the entire frame is vertically scaled down to fit for the left eye and right eye respectively.

Frame Sequential 3D: For frame sequential 3D, all the frames are sent one by one. One frame is transmitted to an eye first. Then, another frame is transferred to the other eye next.

6.2 Methodology

In order to determine the relationship between the space perception and the perception and reaction time, the methodologies of experimental investigation and an interview, one type of descriptive investigation, are preferred and carried out. Detailed definition and steps of experimental investigation can be found in the chapter 2. Specifications of descriptive investigation and interview including both advantages and disadvantages can also be found in the chapter 2.

6.3 Experimental Investigation

6.3.1 Research Hypothesis

It is obviously that 3D produces space perception, thus giving a more “accurate” sensation of nearness to people who are interacting with each other through a

DMP system. Because of this sensation, we assume that it can also shorten the perception and reaction time.

6.3.2 Design of the Measurement

In order to study the relationship between the space perception and the perception and reaction time, a 3D system has been achieved. There are two types of 3D systems that can be achieved. One is a real-time 3D system, and another is an off-line 3D system. To achieve a real-time 3D system, we use two Toshiba 1K-HR1D cameras to shoot the same image with a slightly different angle, which is to create two images for left eye and right eye separately. Then, we use an EZblend video combiner to merge two images. After that, a F35 AS3D projector is used to project the image on a screen. The person who is watching the image is required to wear the SyncMaster Samsung shutter glasses. Detailed property of Toshiba 1K-HR1D can be found at [10], and the specification of EZblend video combiner is available at [11]. The description of F35 AS3D can be found at [12], and the features of SyncMaster Samsung shutter glasses can be found at [13].

However, since the purpose is to find out the relationship between the space perception and the perception and reaction time, we do not focus on real-time, interactive 3D, but rather on making comparisons between 2D and 3D displays, and finding out how different frame rates of a video can influence human's perception and reaction time to optical stimuli. It is then important to be able to repeat the sequence of optical events shown to testers and to have full control over the experiments. An off-line solution with a PC for playing out stored videos to a display screen and software capable of editing the videos is preferred. The Figure 32 shows how the system is constructed.

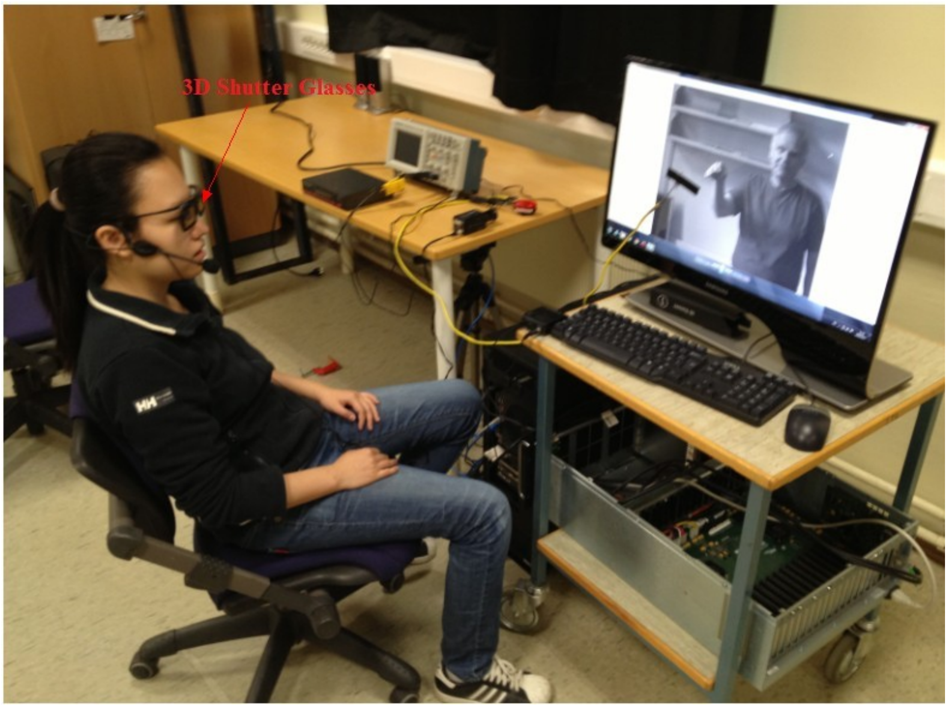

Figure 32: System Structure

By having the adviser mimicking the hand movements of conducting a concert, we use the FL3-U3 USB 3.0 digital camera to capture four groups of pictures. Four groups of pictures are the hand moving at a constant speed with a wooden frame around, the hand moving at a variable speed with a wooden frame around, the hand moving at a constant speed without a wooden frame around, and the hand moving at a variable speed without a wooden frame around. Instead of shooting four videos, we shoot four groups of pictures, because we want the same sequence of optical events. Then, for each group of pictures, we convert it to 2D video with 25fps, 2D video with 50fps, 2D video with 100fps, and 3D video with 25fps. Because the sequence of the optical events of the video are the same, the only thing that influences the testers' perception and reaction time is the frame rate of 2D videos or the space perceptions introduced by 3D videos. To create 2D videos from pictures with different frame rates, we use the software - Videomach. To create 3D videos, we use the software – 3DCombine. For the situation of having the wooden frame around, the tester is asked to pronounce 'ah' sound when they see the adviser's hand hitting the wooden frame. For the situation of not having the wooden frame around, the tester is asked to pronounce ‘ah’ sound whey they see the adviser’s hand reaching to the most far point.

6.3.3 The Instruments

To test how space perception influences perception and reaction time of people, same set of equipment used in the chapter 5 is adopted. All the videos are played on the screen. For 3D tests, all the testers are required to wear the SyncMaster Samsung shutter glasses.

6.3.4 Recruit Participants

In order to make a comparison, we make four groups of pictures, and each group of pictures are converted to four videos. Hence, there are totally sixteen videos. For each video, 20 times of measurements are required to get a comparatively accurate result. Therefore, 320 times of measurements have to be done for each tester. Because the time is really limited, we decide to invite one tester to do all the pilot tests.

6.3.5 Actual Data Collection

Tester 1

Gender	Age	Education
Female	26	Master

Table 35: Tester’s Information

Scenario 1 2D Video with 25fps - Hand Moving at a Constant Speed with a Frame This is the first scenario where the adviser’s hand moves at a constant speed with a wooden frame around. The 2D video is with 25fps. As can be seen from the Figure 33, the perception and reaction time varies between 100ms and 300ms, and most of the time, the time is between 150ms and 200ms.

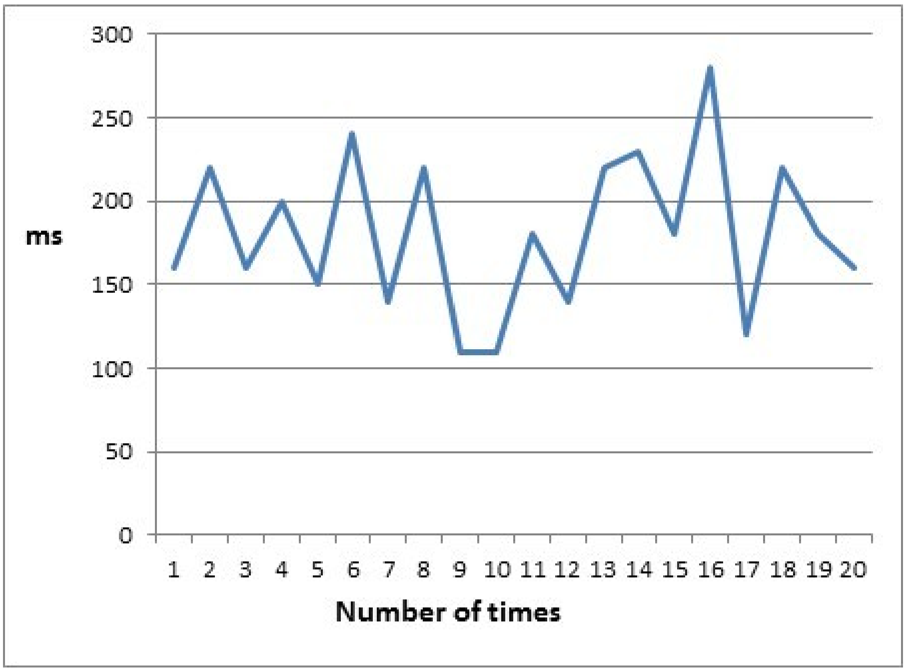

Figure 33: Scenario 1

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
110	280	181	2079	170

Table 36: Results of Scenario 1

Scenario 2 2D Video with 50fps - Hand Moving at a Constant Speed with a Frame This is the second scenario where the adviser's hand moves at a constant speed with a wooden frame around. The 2D video is with 50fps. As can be seen from the Figure 34, the perception and reaction time varies between 40ms and 140ms.

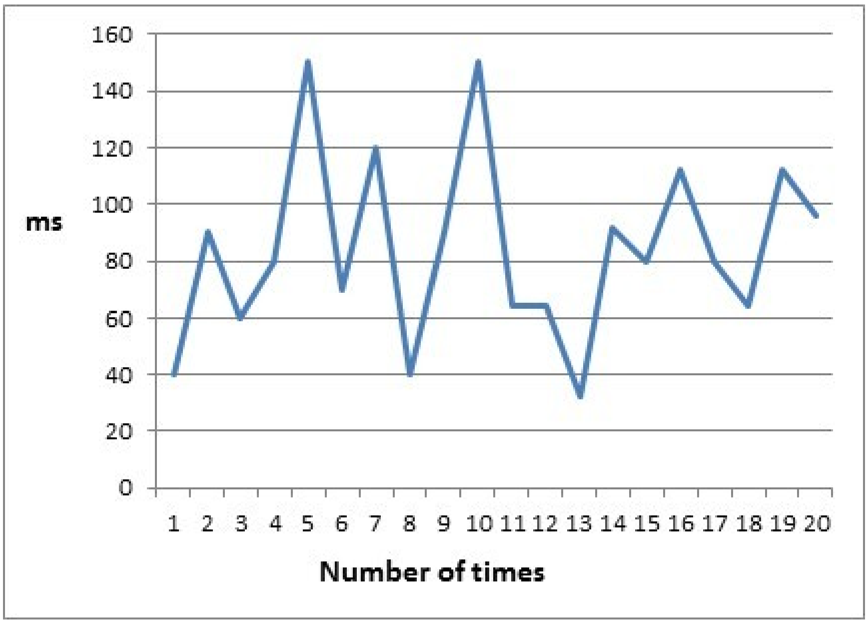

Figure 34: Scenario 2

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
32	150	84	1023	118

Table 37: Results of Scenario 2

Scenario 3 2D Video with 100fps - Hand Moving at a Constant Speed with a Frame This is the third scenario where the adviser's hand moves at a constant speed with a wooden frame around. The 2D video is with 100fps. As can be seen from the Figure 35, the perception and reaction time varies between 30ms and 150ms.

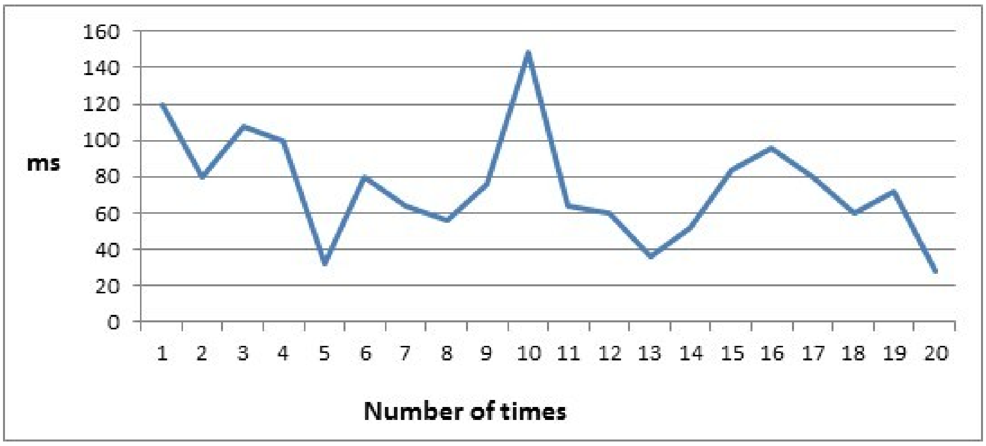

Figure 35: Scenario 3

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
28	148	75	842	120

Table 38: Results of Scenario 3

Scenario 4 3D Video with 25fps - Hand Moving at a Constant Speed with a Frame This is the fourth scenario where the adviser's hand moves at a constant speed with a wooden frame around. The 3D video is with 25fps. As can be seen from the Figure 36, the perception and reaction time varies between 50ms and 200ms.

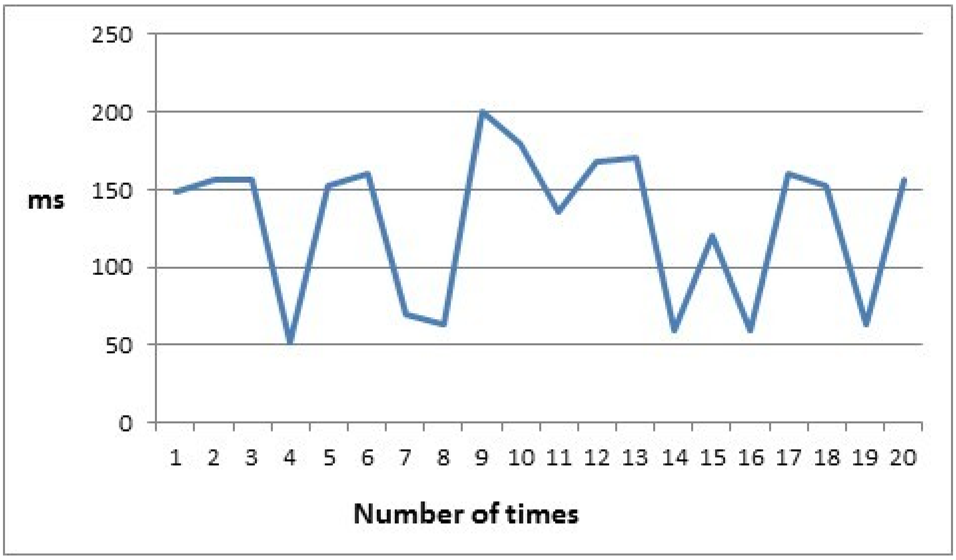

Figure 36: Scenario 4

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
52	200	129	2194	148

Table 39: Results of Scenario 4

Scenario 5 2D Video with 25fps - Hand Moving at a Variable Speed with a Frame This is the fifth scenario where the adviser's hand moves at a variable speed with a wooden frame around. The 2D video is with 25fps. As can be seen from the Figure 37, the perception and reaction time varies dramatically, and the time is between 10ms and 180ms.

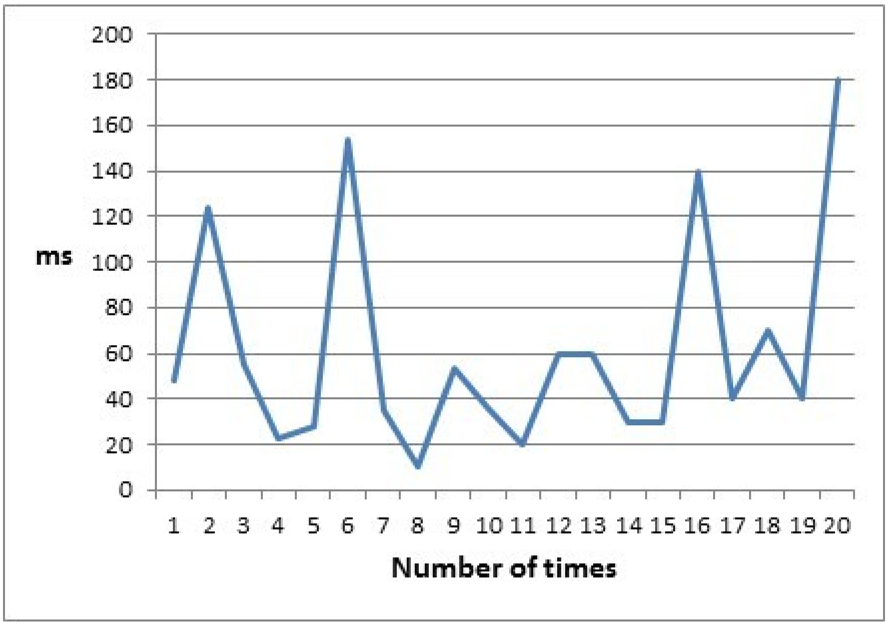

Figure 37: Scenario 5

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
10	180	62	2214	170

Table 40: Results of Scenario 5

Scenario 6 2D Video with 50fps - Hand Moving at a Variable Speed with a Frame This is the sixth scenario where the adviser's hand moves at a variable speed with a wooden frame around. The 2D video is with 50fps. As can be seen from the Figure 38, the perception and reaction time varies very dramatically. The longest time reaches to over 200ms. The shortest time can be around 5ms.

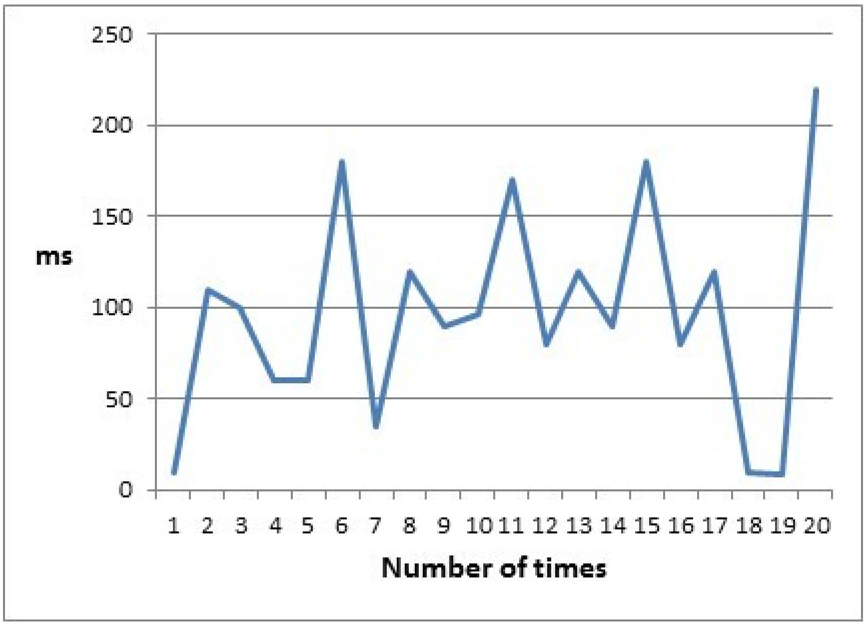

Figure 38: Scenario 6

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
9	220	97	3307	201

Table 41: Results of Scenario 6

Scenario 7 2D Video with 100fps - Hand Moving at a Variable Speed with a Frame This is the seventh scenario where the adviser's hand moves at a variable speed with a wooden frame around. The 2D video is with 100fps. As can be seen from the Figure 39, the perception and reaction time varies very dramatically. The longest time reaches to over 250ms. The shortest time can be around 5ms.

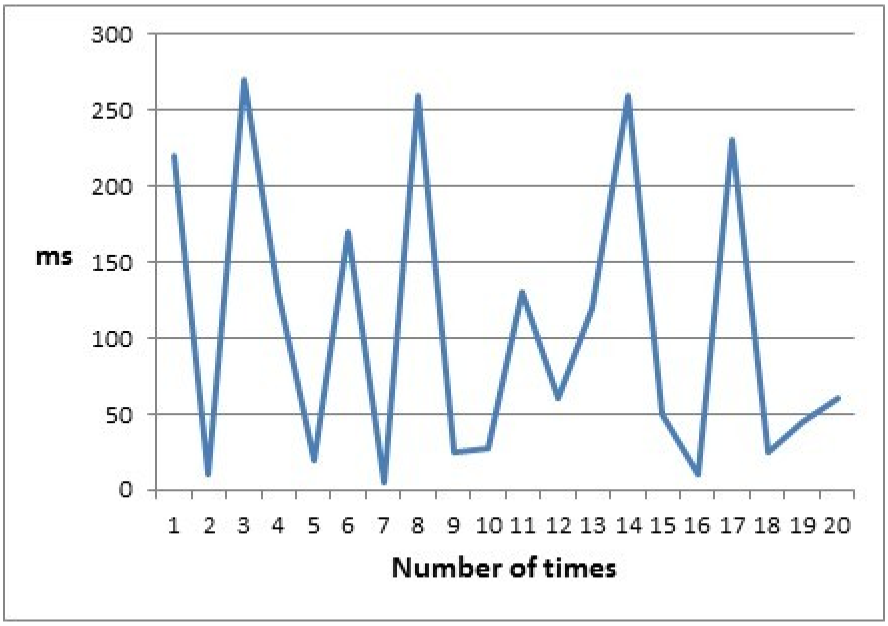

Figure 39: Scenario 7

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
5	270	106	8723	265

Table 42: Results of Scenario 7

Scenario 8 3D Video with 25fps - Hand Moving at a Variable Speed with a Frame This is the eighth scenario where the adviser's hand moves at a variable speed with a wooden frame around. The 3D video is with 25fps. As can be seen from the Figure 40, the perception and reaction time varies very dramatically. The longest time reaches to over 200ms. The shortest time can be around 5ms.

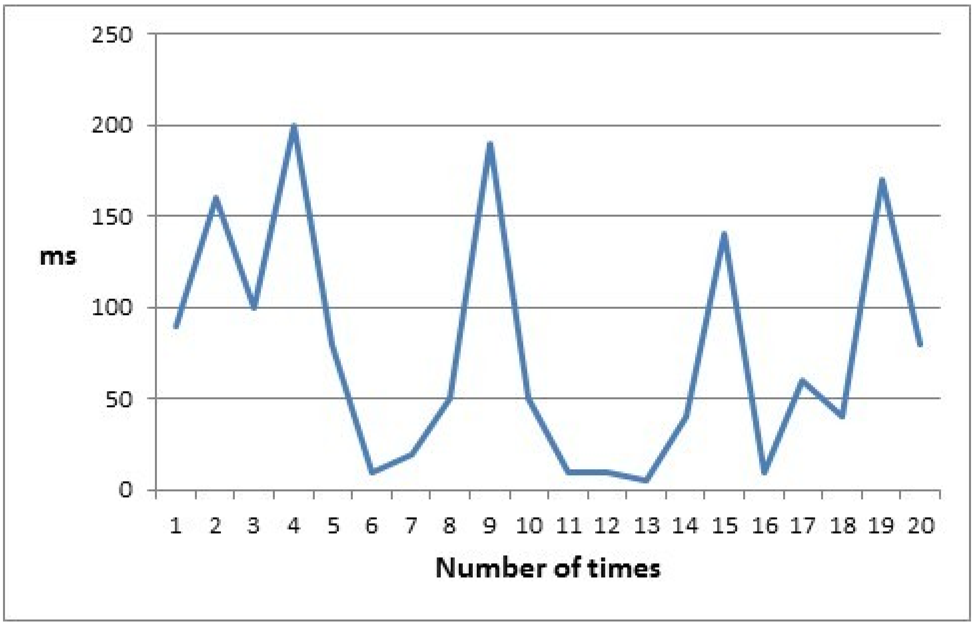

Figure 40: Scenario 8

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
5	200	76	3498	195

Table 43: Results of Scenario 8

Scenario 9 2D Video with 25fps - Hand Moving at a Constant Speed without a Frame This is the ninth scenario where the adviser's hand moves at a constant speed without a wooden frame around. The 2D video is with 25fps. As can be seen from the Figure 41, the perception and reaction time can have negative value. This means, because of the prediction, people react before the hand reaching the most far point.

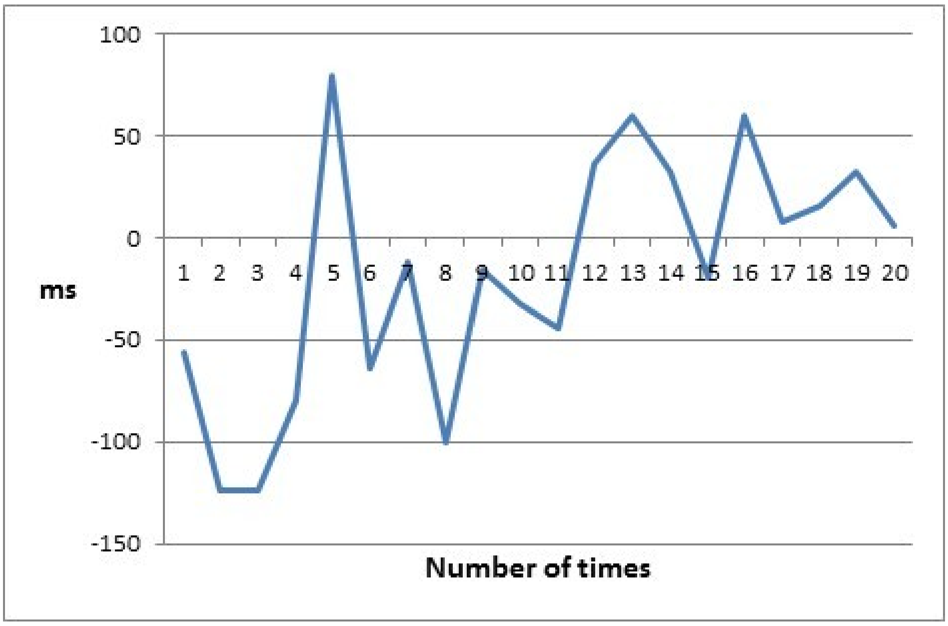

Figure 41: Scenario 9

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-124	80	-17	3480	224

Table 44: Results of Scenario 9

Scenario 10 2D Video with 50fps - Hand Moving at a Constant Speed without a Frame This is the tenth scenario where the adviser's hand moves at a constant speed without a wooden frame around. The 2D video is with 50fps. As can be seen from the Figure 42, the perception and reaction time distributes not regular, and it also has negative values.

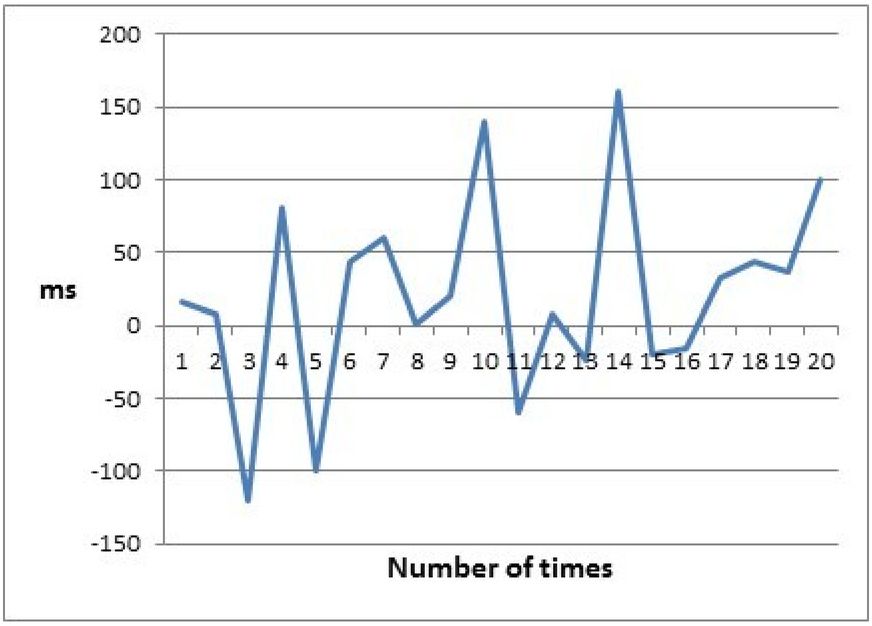

Figure 42: Scenario 10

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-120	160	20	4652	280

Table 45: Results of Scenario 10

Scenario 11 2D Video with 100fps - Hand Moving at a Constant Speed without a Frame This is the eleventh scenario where the adviser's hand moves at a constant speed without a wooden frame around. The 2D video is with 100fps. As can be seen from the Figure 43, the perception and reaction time can also be negative value. However, the number of times that the perception and reaction time is negative is less than previous two situations.

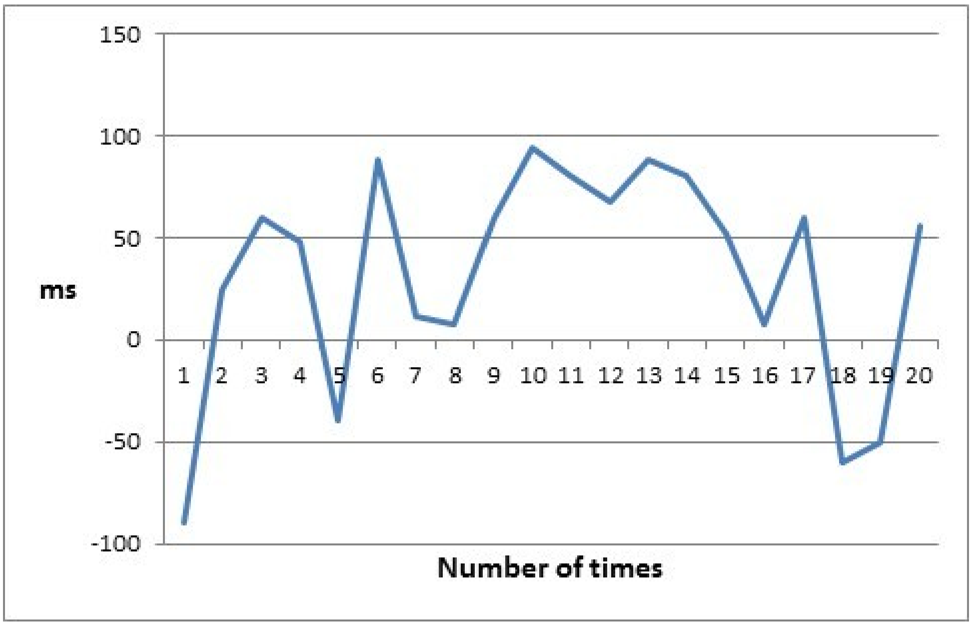

Figure 43: Scenario 11

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-90	94	32	2824	184

Table 46: Results of Scenario 11

Scenario 12 3D Video with 25fps - Hand Moving at a Constant Speed without a Frame This is the twelfth scenario where the adviser's hand moves at a constant speed without a wooden frame around. The 3D video is with 25fps. As can be seen from the Figure 44, the distribution of perception and reaction time is more regular than previous three scenarios, and it contains only one time that the perception and reaction time is negative.

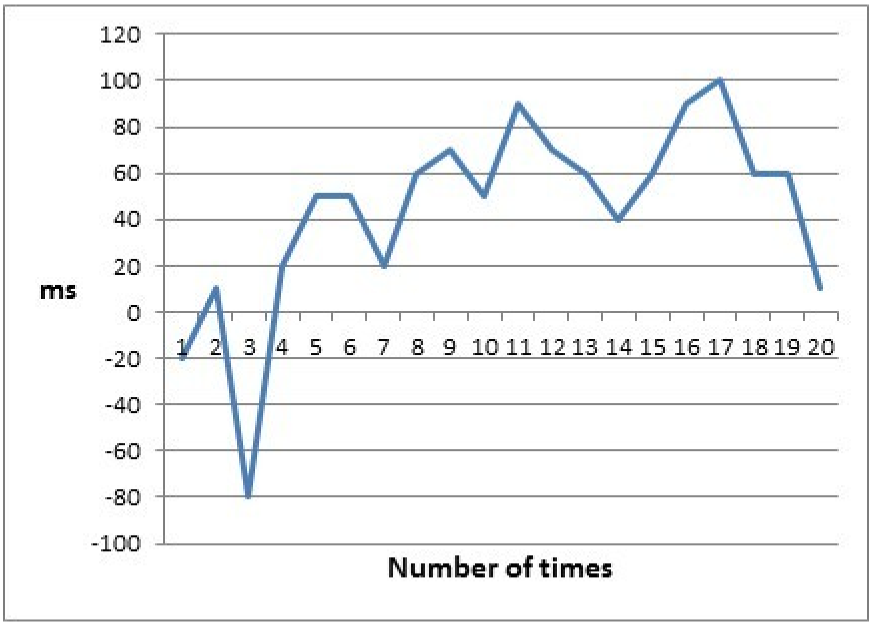

Figure 44: Scenario 12

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-80	100	43	1653	1840

Table 47: Results of Scenario 12

Scenario 13 2D Video with 25fps - Hand Moving at a Variable Speed without a Frame This is the thirteenth scenario where the adviser's hand moves at a variable speed without a wooden frame around. The 2D video is with 25fps. As can be seen from the Figure 45, the perception and reaction time can also be negative value.

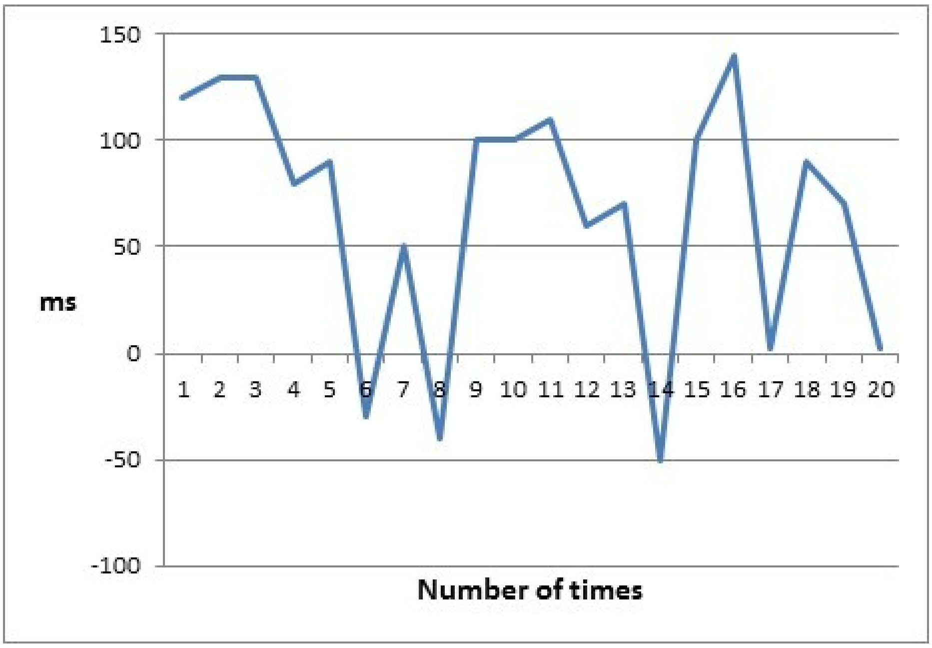

Figure 45: Scenario 13

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-50	140	66	3288	190

Table 48: Results of Scenario 13

Scenario 14 2D Video with 50fps - Hand Moving at a Variable Speed without a Frame This is the fourteenth scenario where the adviser's hand moves at a variable speed without a wooden frame around. The 2D video is with 50fps.

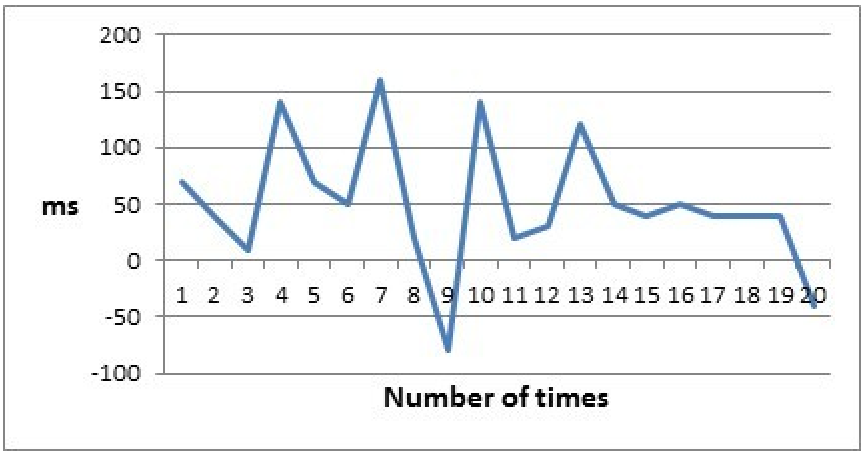

Figure 46: Scenario 14

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-80	160	50	3165	240

Table 49: Results of Scenario 14

Scenario 15 2D Video with 100fps - Hand Moving at a Variable Speed without a Frame This is the fifteenth scenario where the adviser's hand moves at a variable speed without a wooden frame around. The 2D video is with 100fps. As can be seen from the Figure 47, the perception and reaction time varies a lot, and it can seldom be 0.

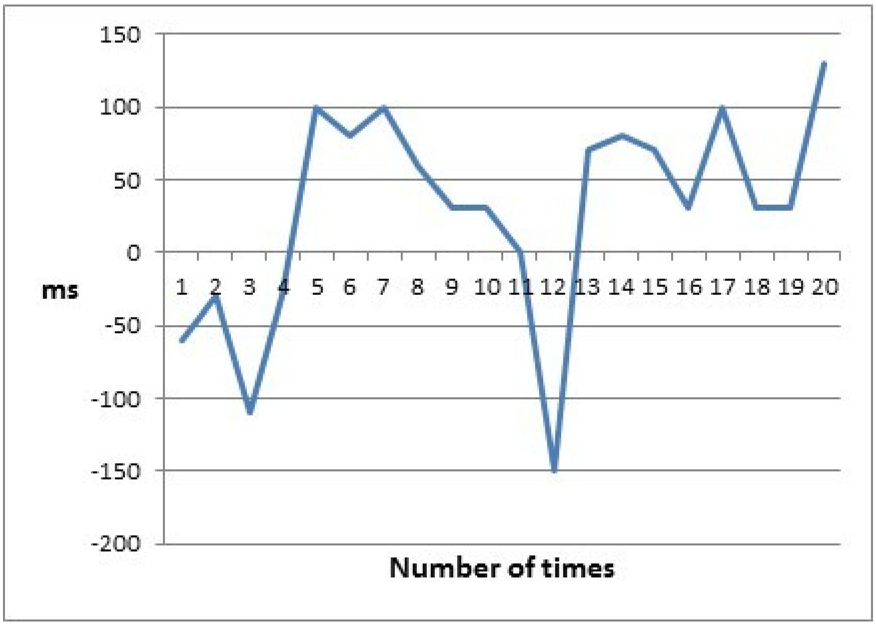

Figure 47: Scenario 15

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-150	130	28	5096	280

Table 50: Results of Scenario 15

Scenario 16 3D Video with 25fps - Hand Moving at a Variable Speed without a Frame This is the sixteenth scenario where the adviser's hand moves at a variable speed without a wooden frame around. The 3D video is with 25fps. As can be seen from the Figure 48, the perception and reaction time changes a lot and it can be negative.

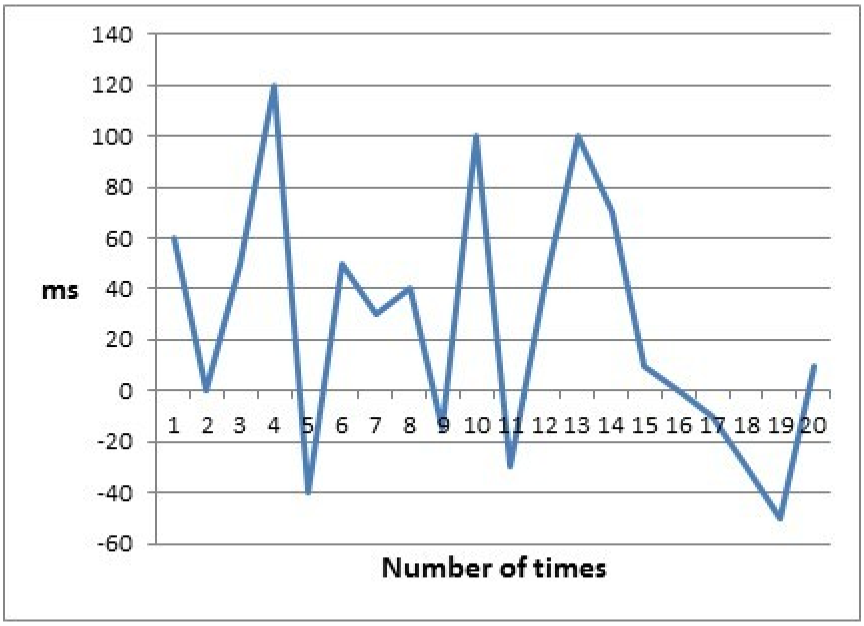

Figure 48: Scenario 16

Min Value (ms)	Max Value (ms)	Average Value (ms)	Variance (ms)	Range Observed (ms)
-50	120	25	2284	170

Table 51: Results of Scenario 16

6.3.6 Data Analysis

Based on all the plots and measurement results presented above, there are several conclusions can be drawn. First of all, based on the results of first four scenarios, one conclusion can be drawn that the frame rate can directly influence the perception and reaction time of the tester. The bigger the frame rate, the shorter the perception and reaction time is.

In addition, based on the results of the scenarios with a wooden frame around and the scenarios without a wooden frame around, we can conclude that with a wooden frame, the tester knows when exactly she needs to respond to an optical stimuli. When without a reference, like a wooden frame, the tester intends to predict, thus reacting earlier than it is supposed to be. Hence, for some cases, the perception and reaction time can be negative values, and in extreme case, it can be zero. This reveals the real situations in real life. For instance, when a conductor conducts a concert with a choir, it is indispensable to have some asynchronization sound among people of a choir, because some people intend to predict the conductor's hand movements.

Next, when compare a 2D video with 25fps with a 3D video with 25fps, the perception and reaction time of the 3D video is obviously shorter than that of the 2D video. However, in comparison, the perception and reaction time of the 3D video with 25fps is always longer than the perception and reaction time of the 3D video with 50fps.

Because of the fact that only one tester participating all the pilot tests, the conclusions cannot be generalized. However, these conclusions can be a great hypothesis for further work.

6.3.7 Results

All the results can be found in the Appendix E.

6.4 Interview

In order to find out how space perception influences the perception and reaction time, it is also important to get the feeling from the tester. Hence, an interview has been conducted. The questions and the results of the interview can be found in the Appendix F.

7 Conclusion

Several aspects related to the perceived quality of a telepresence system have been discussed. A lab model of the DMP system has been presented. Taking advantage of the model, a method of measuring the optical delay is proposed. Based on the measurement results, the conclusion has been drawn that the optical delay is directly related to the frame rate of the camera and the refreshing rate of the display screen. The higher the frame rate of the camera, the shorter the optical delay is. The higher the refreshing rate of the display screen, the shorter the optical delay is.

In addition, by using the presented DMP system model, an approach of measuring stimulus-reaction time is suggested. Based on a series of tests, we get the indication that the perception and reaction time is directly related to age, tiredness, and concentration of the testers. However, the perception and reaction seems to have no connection with the education and gender.

Finally, two ways of achieving 3D systems have been proposed. In order to find out the relationship between space perception and perception and reaction time, several pilot tests have been done. Based on the measurement results, we suggest that the frame rate of the video can directly influence the perception and reaction time. The higher the frame rate, the shorter the perception and reaction time is. Moreover, for 3D videos, it makes people feel more real, and the measurements indicate that the perception and reaction time of 3D videos is shorter than the 2D videos with the same frame rate and same content of the video.

Limitation and Future Work: Because the time is really limited, the stimulus-reaction time measurements and the final 3D measurements have been done with limited amount of testers. Hence, the results cannot be generalized. For future work, in order to generalize it, a large amount of testers have to be invited. In addition, the two time cursors of the oscilloscope are adjusted and modified by human beings instead of machine, so it may introduce some errors. Hence, for accuracy, it is better to automate the process of adjusting time cursors in the future.

References

[1] Rønningen, L. A. (2011). The Distributed Multimedia Plays Architecture. Norwegian University of Science and Technology, Department of Telematics.

[2] Lazar, J., Feng, J., and Hochheiser, H. (2010). Research Methods. United Kingdom: A John Wiley and Sons, Ltd.

[3] AXIS COMMUNICATIONS (2013). Image scanning techniques. Retrived on 1st April, 2013 from www.axis.com/products/video/camera/progressive_scan.htm.

[4] Auten et al. System and method for generating and controlling a simulated musical concert experience. United States Patent. 5990405. November 23, 1999.

[5] Conca, D.P. (2012) Telepresence Quality. Master's thesis. Trondheim. NTNU.

[6] 555 timer circuits. Flashing LED Circuit. Retrived on 20th April, 2013 from www.555-timer-circuits.com/flashing-led.html.

[7] SAMSUNG. (2013). 23" LED Monitor S23A950D. Retrived on 20th April, 2013 from http://www.samsung.com/sg/consumer/pc-peripherals-printer/monitor/led-monitor/LS23A950DS/XS-features.

[8] SENNHEISER. (2013). ew 152 G3. Retrieved on 28th April, 2013 from http://en-us.sennheiser.com/wireless-headworn-cardoid-microphone-headmic-singers-speakers-presenters-ew-152-g3.

[9] Holliman, N. (2005). 3D Display Systems. Department of Computer Science, University of Durham.

[10] TOSHIBA. (2013). IK-HR1D. Retrieved on 1st May, 2013 from http://www.toshibacameras.com/products/prod_detail_ikhr1d.jsp.

[11] WESTAR. (2013). EZblend Video Combiner. Retrieved on 5th May, 2013 from www.westardisplaytechnologies.com/products/video-combiner-ezblend.

[12] projectiondesign. (2013). F35 AS3D. Retrieved on 20th May, 2013 from www.projectiondesign.com/products/discontinued-products/f35-as3d.

[13] SAMSUNG. (2013). 3D Active Glassess SSG-3500CR. Retrieved on 1st June, 2013 from www.samsung.com/us/video/tvs-accessories/SSG-3500CR/ZA.

Appendix A

Results of Optical Delay Measurements

The measurements have been done twenty times, so the total number of times is 20. The numbers in the vertical line represent the measurement result for each time.

	Results (ms)		
Times	**60Hz**	**120Hz**	**150Hz**
1	70	60	80
2	80	70	60
3	70	70	60
4	70	70	80
5	70	70	60
6	70	60	70
7	80	60	70
8	60	80	60
9	80	60	70
10	80	60	80
11	100	80	60
12	80	70	60
13	70	70	60
14	90	60	50
15	90	70	60
16	70	70	70
17	60	70	70
18	70	60	70
19	70	70	70
20	80	70	50

Table 52: Optical Delay With the Screen Refreshing Rate of 60Hz

	Results (ms)		
Times	**60Hz**	**120Hz**	**150Hz**
1	80	50	60
2	60	50	60
3	60	50	60
4	80	60	60
5	60	60	50
6	70	60	60
7	70	50	50
8	60	50	50
9	70	60	70
10	80	60	60
11	60	60	60
12	60	50	60
13	60	60	50
14	50	50	60
15	60	60	60
16	70	60	60
17	70	70	60
18	70	70	60
19	70	60	60
20	50	60	60

Table 53: Optical Delay With the Screen Refreshing Rate of 100Hz

	Results (ms)		
Times	**60Hz**	**120Hz**	**150Hz**
1	70	50	60
2	60	70	50
3	70	50	50
4	70	50	50
5	60	50	50
6	60	60	50
7	60	40	60
8	50	50	50
9	60	50	50
10	70	50	60
11	70	60	60
12	50	50	60
13	50	60	50
14	70	50	60
15	60	50	50
16	50	60	50
17	60	60	50
18	50	50	50
19	80	60	50
20	60	60	60

Table 54: Optical Delay With the Screen Refreshing Rate of 120Hz

Appendix B

Permission Contract

NTNU – Trondheim
Norwegian University of
Science and Technology

Permission is granted to Haijiao Zhang and Leif Arne R0nningen in order to publish the results of the performed tests during May, 2013 at the Department ofTelematics of Norwegian University of Science and Technology (NTNU).

Participants:

Wenjing Sun

Fenglin Han

Shuwen Dun

Leif Arne Rønningen

Haijiao Zhang

Appendix C

Results of Stimulus-reaction Time Measurements

This appendix contains all the results for stimulus-reaction time measurements.

No. of Times	**Tester 1 (ms)**	**Tester 2 (ms)**	**Tester 3 (ms)**	**Tester 4 (ms)**	**Tester 5 (ms)**
1	150	230	180	230	180
2	150	300	190	260	200
3	140	220	290	250	180
4	150	200	190	170	170
5	170	200	180	210	220
6	180	220	190	220	160
7	80	220	220	240	190
8	170	230	200	230	180
9	190	170	250	220	150
10	140	150	220	270	140
11	150	200	210	120	220
12	170	190	250	240	150
13	220	240	220	170	190
14	220	210	200	230	130
15	160	180	160	160	200
16	150	250	220	310	160
17	140	70	180	150	180
18	240	180	230	240	220
19	140	200	200	270	190
20	240	190	210	300	240

Table 55: Testers With a Fixed Interval

No. of Times	**Tester 1 (ms)**	**Tester 2 (ms)**	**Tester 3 (ms)**	**Tester 4 (ms)**	**Tester 5 (ms)**
1	310	290	260	240	270
2	310	270	340	210	260
3	280	330	290	220	210
4	260	280	300	200	310
5	270	280	290	260	210
6	290	270	280	250	320
7	240	330	340	280	220
8	210	260	300	200	210
9	280	290	330	240	170
10	260	260	310	220	230
11	230	270	300	280	180
12	230	250	280	210	230
13	280	360	360	210	300
14	280	270	350	290	260
15	280	360	350	230	210
16	310	250	320	220	220
17	230	310	350	260	170
18	290	270	300	200	280
19	310	360	330	220	270
20	310	240	330	190	210

Table 56: Testers With a Variable Interval

No. of Times	Result (ms)
1	4
2	6
3	4
4	6
5	8
6	4
7	6
8	6
9	8
10	6
11	6
12	3
13	4
14	4
15	6
16	3
17	4
18	4
19	3
20	6

Table 57: Delay Caused by Wireless Transmission

Appendix D

Survey

1. Do you think the perception and reaction time is directly related to age? (Yes, No, Uncertain)

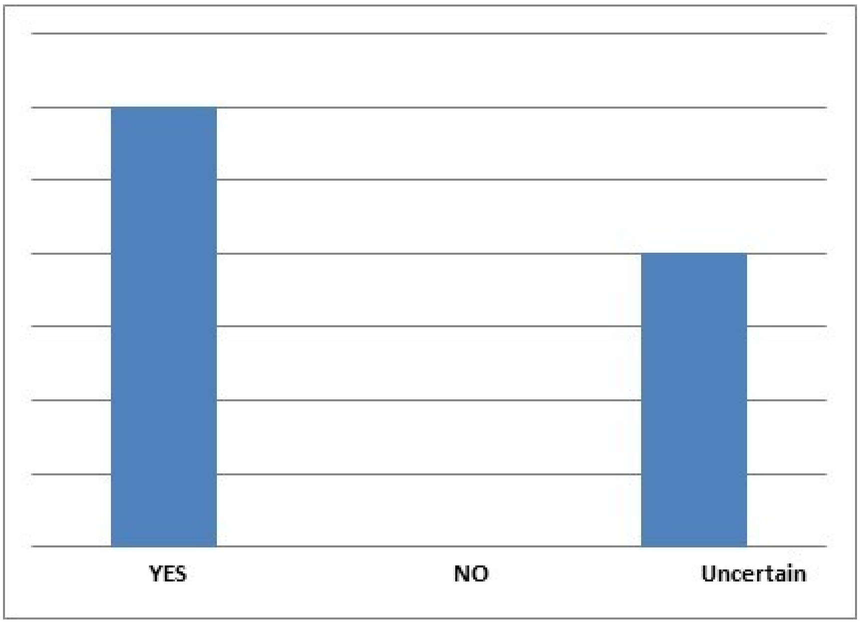

Figure 49: The Question About Age

2. Do you think the perception and reaction time is directly related to life habit, like sitting in front of laptop all the time? (Yes, No, Uncertain)

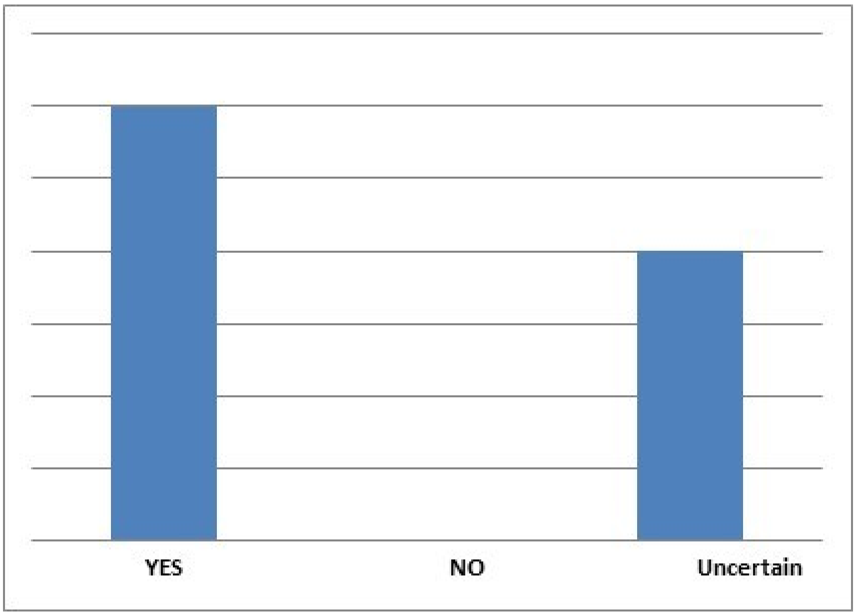

Figure 50: The Question About Life Habit

3. Do you think the perception and reaction time is directly related to tiredness?
(Yes, No, Uncertain)

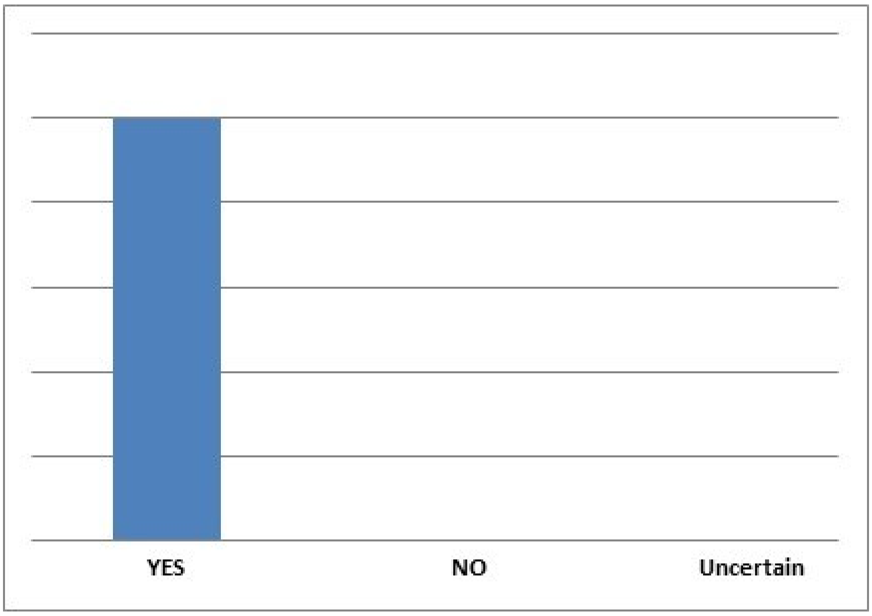

Figure 51: The Question About Tiredness

4. Do you think the perception and reaction time is directly related to education? (Yes, No, Uncertain)

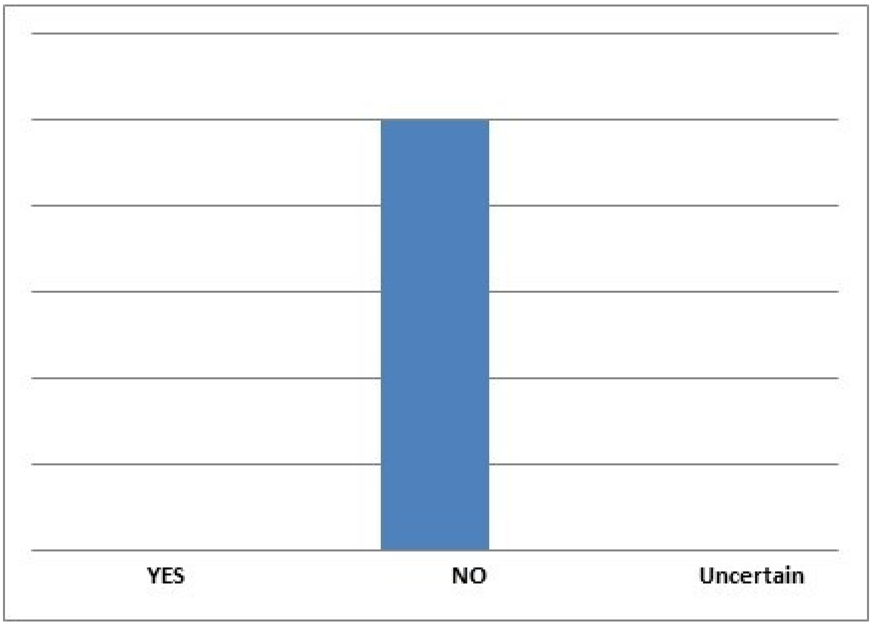

Figure 52: The Question About Education

5. Do you think the perception and reaction time is directly related to concentration? (Yes, No, Uncertain)

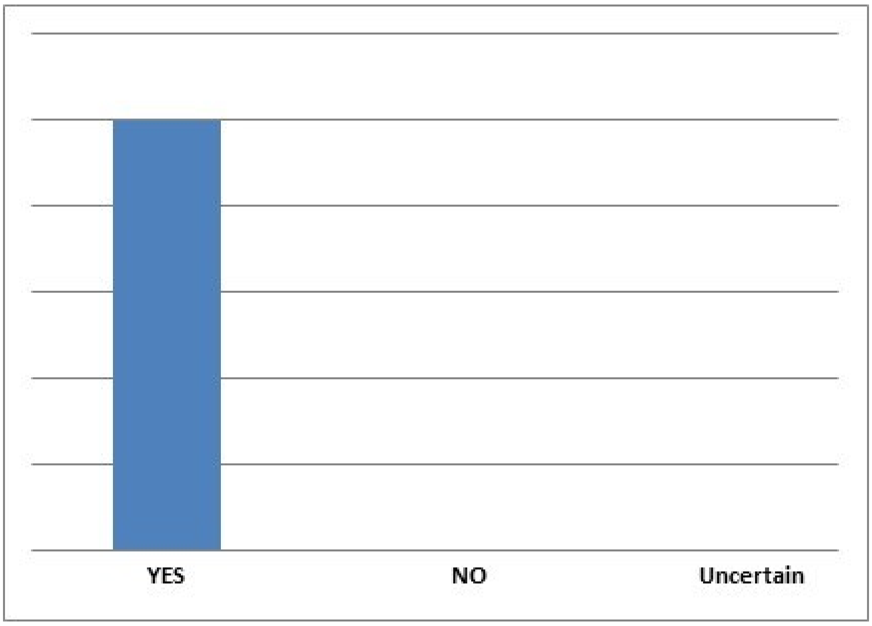

Figure 53: The Question About Concentration

6. Do you think the perception and reaction time is directly related to practice? (Yes, No, Uncertain)

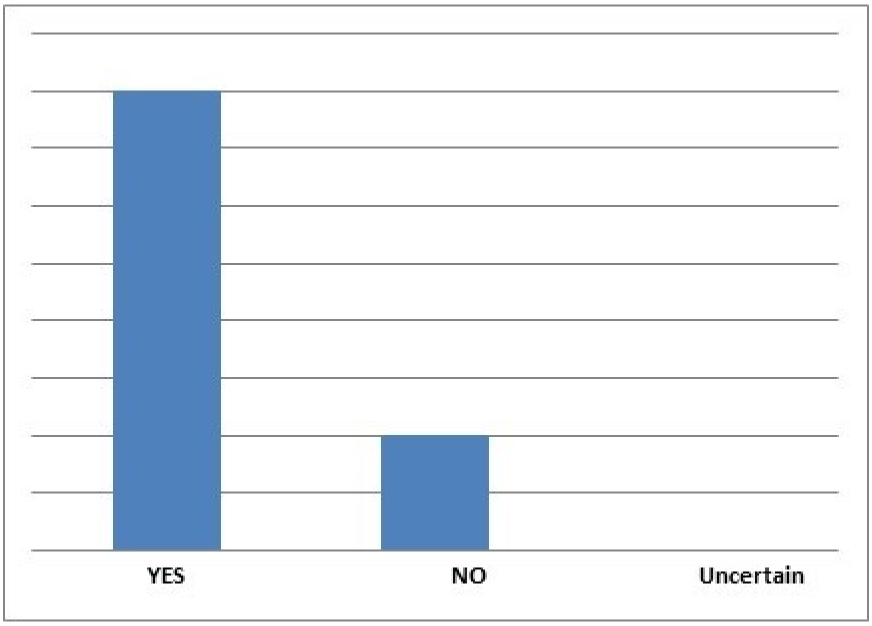

Figure 54: The Question About Practice

7. Do you think the perception and reaction time is directly related to gender? (Yes, No, Uncertain)

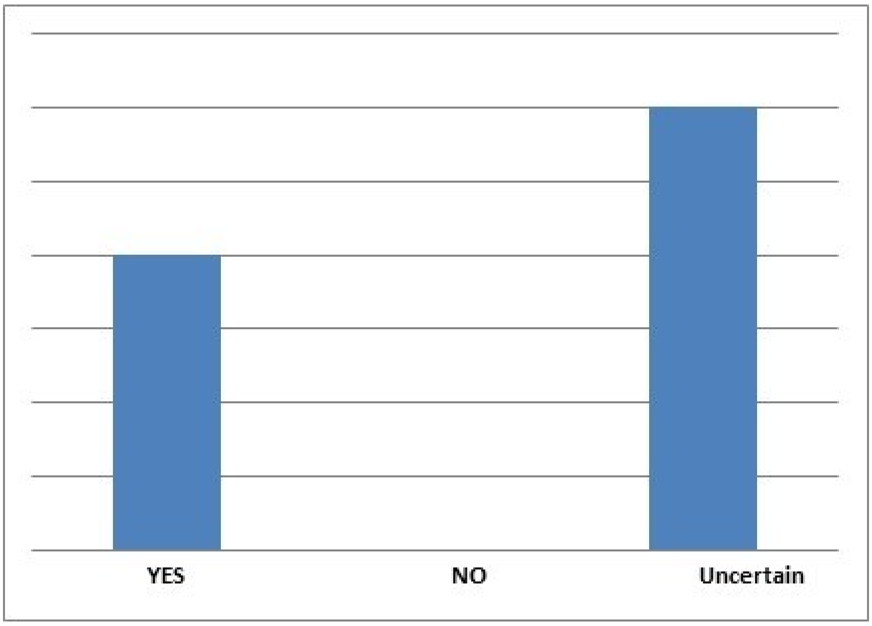

Figure 55: The Question About Gender

8. Do you think among all the factors that influence the perception and reaction time, concentration is the most important factor? (Yes, No, Uncertain)

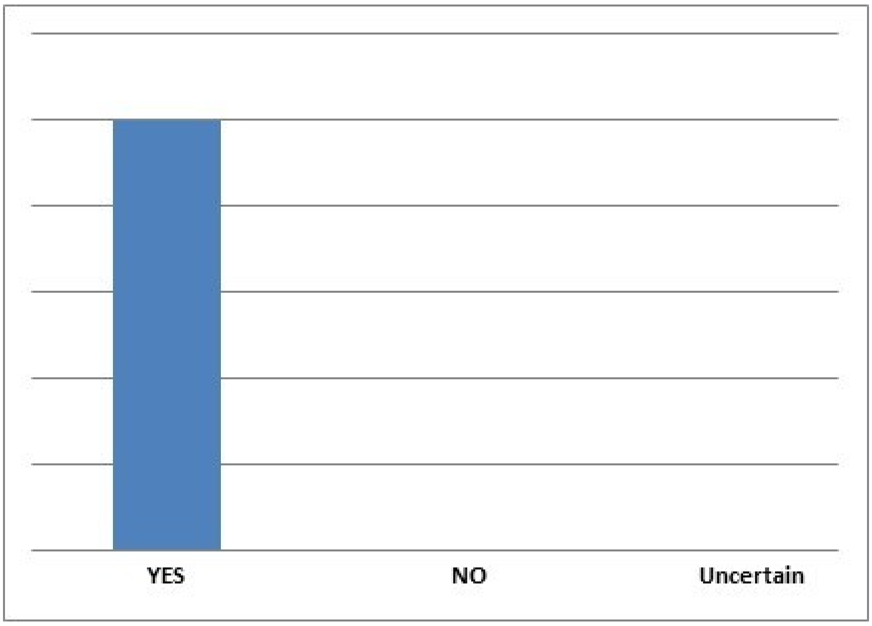

Figure 56: The Question About the Most Important Influencer

Appendix E

Results of 3D Measurements

All the measurement results about 3D measurements can be found in this appendix.

	Results (ms)			
No. of Times	**2D with 25Hz**	**2D with 50Hz**	**2D with 100Hz**	**3D with 25Hz**
1	160	40	120	148
2	220	90	80	156
3	160	60	108	156
4	200	80	100	52
5	150	150	32	152
6	240	70	80	160
7	140	120	64	70
8	220	40	56	64
9	110	90	76	200
10	110	150	148	180
11	180	64	64	135
12	140	64	60	168
13	220	32	36	170
14	230	92	52	60
15	180	80	84	120
16	280	112	96	60
17	120	80	80	160
18	220	64	60	152
19	180	112	72	64
20	160	96	28	156

Table 58: Constant Speed With a Frame

	Results (ms)			
No. of Times	**2D with 25Hz**	**2D with 50Hz**	**2D with 100Hz**	**3D with 25Hz**
1	-56	16	-90	-20
2	-124	8	24	10
3	-124	-120	60	-80
4	-80	80	48	20
5	80	-100	-40	50
6	-64	44	88	50
7	-12	60	12	20
8	-100	1	8	60
9	-16	20	60	70
10	-32	140	94	50
11	-44	-60	80	90
12	36	8	68	70
13	60	-24	88	60
14	32	160	80	40
15	-20	-20	52	60
16	60	-16	8	90
17	8	32	60	100
18	16	44	-60	60
19	32	36	-50	60
20	6	100	56	10

Table 59: Constant Speed Without a Frame

	Results (ms)			
No. of Times	**2D with 25Hz**	**2D with 50Hz**	**2D with 100Hz**	**3D with 25Hz**
1	48	10	220	90
2	124	110	10	160
3	55	100	270	100
4	23	60	130	200
5	28	60	20	80
6	154	180	170	10
7	35	35	5	20
8	10	120	260	50
9	53	90	25	190
10	36	96	28	50
11	20	170	130	10
12	60	80	60	10
13	60	120	120	5
14	30	90	260	40
15	30	180	50	140
16	140	80	10	10
17	40	120	230	60
18	70	10	25	40
19	40	9	45	170
20	180	220	60	80

Table 60: Variable Speed With a Frame

	Results (ms)			
No. of Times	**2D with 25Hz**	**2D with 50Hz**	**2D with 100Hz**	**3D with 25Hz**
1	120	70	-60	60
2	130	40	-30	0
3	130	10	-110	50
4	80	140	-30	120
5	90	70	100	-30
6	-30	50	80	50
7	50	160	100	30
8	-40	20	60	40
9	100	-80	30	-15
10	100	140	30	100
11	110	20	0	-30
12	60	30	-150	40
13	70	120	70	100
14	-50	50	80	70
15	100	40	70	10
16	140	50	30	0
17	2	40	100	-10
18	90	40	30	-30
19	70	40	30	-50
20	2	40	130	10

Table 61: Variable Speed Without a Frame

Appendix F

Interview

1. Can you tell a big difference between videos with different frame rates?
   No, it looks very similar. Probably for the 2D video with 25fps, it blurs a little. For the 2D video with 100fps, it does not blur.

2. What do you think of 3D?
   Well, it feels much more real compared with other videos. It makes me believe that I can almost touch the adviser's hand.

3. What advantages do you think the 3D introduce?
   It definitely introduces the space perception. It makes the system much more real.

4. What disadvantages do you think the 3D introduces?
   When I wear the shutter glasses, it makes me feel a little bit dizzy when I wear it for a long time. In addition, I have my own glasses, so putting the shutter glasses in front of my own glasses make me feel unconfortable.

5. Do you think 3D can influence the perception and reaction time?
   For this, I do not know

6. What do you think about the negative value of perception and reaction time?
   Well, in the case without a reference, I do not know when exactly I should react, so I just react to the point that I thought is the most far point. I do not want to predict, but it is unavoidable.