Gesture Based Interface for Asynchronous Video Communication for Deaf People in South Africa

By

Tshifhiwa Ramuhaheli

Supervised by

Prof. Edwin Blake

Thesis presented for the degree of Master of Science

In the Department of Computer Science

University of Cape Town

May 2010

Abstract

The preferred method of communication amongst Deaf people is that of sign language. There are problems with the video quality when using the real-time video communication available on mobile phones. The alternative is to use text-based communication on mobile phones, however findings from other research studies show that Deaf people prefer using sign language to communicate with each other rather than text. This dissertation looks at implementing a gesture-based interface for an asynchronous video communication for Deaf people. The gesture interface was implemented on a store and forward video architecture since this preserves the video quality even when there is low bandwidth.

In this dissertation three gesture-based video communication prototypes were designed and implemented using a user centred design approach. These prototypes were implemented on both the computer and mobile devices. The first prototype was computer based and the evaluation of this prototype showed that the gesture based interface improved the usability of sign language video communication. The second prototype is set up on the mobile device and it was tested on several mobile devices but the device limitation made it impossible to support all the features needed in the video communication. The different problems experienced on the dissimilar devices made the task of implementing the prototypes on the mobile platform challenging. The prototype was revised several times before it was tested on a different mobile phone.

The final prototype used both the mobile phone and the computer. The computer served to simulate a mobile device with greater processing power. This approach simulated a more powerful future mobile device capable of running the gesture-based interface. The computer was used for video processing but to the user it was as if the whole system was running on the mobile phone. The evaluation process was conducted with ten Deaf users in order to determine the efficiency and usability of the prototype. The results showed that the majority of the users were satisfied with the quality of the video communication. The evaluation also revealed usability problems but the benefits of communicating in sign language outweighed the usability difficulties. Furthermore the users were more interested in the video communication on the mobile devices than on the computer as this was a much more familiar technology and offered the convenience of mobility.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Firstly I would like to thank God for all his blessings and for affording me the opportunity to conduct this research. There were plenty of times when things were hard but eventually it all worked out for the best. I would like to express my deep appreciation to all those who supported me during the past couple of years. Without God and all these peoples' support none of this would have been possible. There are many I am grateful to but am unable to name here because of the sheer numbers so just to name a few I would like to take this opportunity to thank:

My family

My Supervisor Professor Edwin Blake

My friends, colleagues and everyone who helped me during this period

All the people involved with the Deaf Telephony Project

(Special thanks: Dr. Bill Tucker, Daniel Erasmus, Muyowa Mutenwa, Long Yi and Gary Ma)

All the people at Deaf Community of Cape Town *(Special thanks to: Stephen and Faith)*

Meryl Glaser at SLED

Michelle Lombard for all the interpretation

All the participants in the various user studies

Vhonani Matodzi for proof reading my dissertation

Sponsors

Telkom/Nokia-Siemens/Telescience and Thrip Centre of Excellence

I would also like to especially thank Telkom for their financial support throughout the whole period.

List of Acronyms

3G	Third generation cellular technology
API	Application Programming Interface
ASL	American Sign Language
BREW	Binary Runtime Environment of Wireless
DCCT	Deaf Community of Cape Town
FET	Further Education and Training
FPS	Frame per Second
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communication
HCI	Human-Computer Interaction
ICT4D	Information and Communication Technology for Development
IDE	Integrated Development Environment
IM	Instant Messenger
Java ME	Java Micro Edition
MMS	Multimedia Messaging Services
NGO	Non-government Organization
OS	Operating System
S60	A software platform for mobile phones that runs on Symbian operating system
SASL	South African Sign Language
SDK	Software Development Kit
SIP	Session Initiation Protocol
SLED	Sign Language Education and Development
SMS	Short Message Service
TCP/IP	Transmission Control Protocol/Internet Protocol
TV	Television
USB	Universal Serial Bus
WiFi	Wireless Fidelity (802.11 abg)

1. Introduction

Communication is a vital part of life, it enables people to share ideas, learn from each other and relay information to one another. Deaf people rely mainly on visual interaction as a means of communication whereas hearing people usually rely on their auditory sense when communicating. There are different types of communications; for example, sound based communication like talking used by hearing people and visual communication like sign language used by Deaf people. Hearing people can use their voice to communicate with each other over long distances via a telephone. For Deaf people this long distance communication requires video communication if they are using Sign Language. There are text based communications like email, Instant Messaging (IM) and Short Message Service (SMS) that can be used by both hearing and Deaf people.

The differences in the mode of communication often lead to challenging situations for Deaf people whilst interacting with hearing people. The majority of hearing people do not understand sign language and this makes it difficult for Deaf people to communicate with them. It requires considerable effort for Deaf people to learn spoken languages as speech mainly relies on sound which requires one to possess hearing ability. Some Deaf people are able to lip-read to understand spoken languages but it is sophisticated and resource intensive to acquire this skill. Thus, Deaf people prefer to use sign language as a means of communication but this is only effective if the parties involved both understand sign language.

Numerous problems can arise when hearing people and Deaf people are interacting with each other because they prefer to use different kinds of communication. This hindrance can only be eliminated if they find a common way of interaction like sign language, lip reading or using text. In public places like stadiums, train stations and taxi ranks Deaf people experience a lot of difficulties when announcements are made through the public address system. This works well for hearing people as they can easily hear the announcements being made but not for Deaf people.

These communication problems do not arise only when Deaf people are interacting with hearing people. Deaf people can only communicate with each other face to face if they are using sign language. On mobile phones they can only use text to communicate with each other and text based communication is not in sign language which is their preferred language (Tucker, 2003a). Video call on mobile phones is unsuitable for sign language video communication due to the poor video quality. Video call is also blurry if someone is signing fast which makes it difficult to see the gestures in the message.

In this dissertation we propose a way of providing an efficient method that supports sign language video communication for Deaf people. Before going any further we must first explain what sign language is.

1.1 Sign Language

Sign language is a combination of hand gestures, head movements, upper body movements, eye movements and also facial expressions and it is processed by the eyes (SLED, 2009b). The signing area is the front of the person from the waist to the head and people have to keep this region in each other's sight in order to communicate using sign language. Sign language requires both hands and signs can incorporate facial expressions and eye movements. Some signs only need one hand but both hands are needed to effectively use sign language as other signs require both hands. Sign language is not a representation of any spoken language but a language on its own. It has its unique grammar and vocabulary different from spoken languages.

There is finger spelling which is used to spell words from spoken languages that don't have a corresponding sign (SLED, 2009b). Finger spelling is also used to spell out personal information like first name, surname, address, etc. This is only used when there is no other option. People who use sign language have a name sign which is simply a gesture in sign language used to designate someone. These name signs are used to refer to people in a conversation instead of finger spelling the person's first name. Sign language differs from country to country and therefore signs meaning the same thing may differ depending on the variation of sign language that is used. There is no written form of South African Sign Language (SASL) but there is a SignWriting method for American Sign Language (Ahmed and Seong, 2006). In this dissertation the focus is on South African Sign Language.

1.2 Motivation

In 2001, the South African national census statistics showed that there are 453,104 people with a hearing disability in South Africa (Statistics SA, 2001). There is a difference between deaf and Deaf. Capitalized "Deaf" refers to people who are active in the signing Deaf Community and Deaf culture whereas lowercase "deaf" is typically people with a hearing disability (Cavender et al., 2006). In this dissertation, only Deaf people were considered therefore all the users involved in this research use sign language as a means of communication.

This research is part of the Deaf Telephony project that has been developing communication application for Deaf people together with a non-government organization (NGO) called the Deaf Community of Cape Town (DCCT) which is based at the Bastion in Newlands, Cape Town. DCCT have a video communication system implemented. This particular system uses a store and forward architecture where users have to sign in front of a camera and record a video. After the recording process is completed users can then send the video message to another user. The communication between the two users is not in real-time (Ma and Tucker 2007). All these operations are done by using a mouse and keyboard to select the desired options. This research builds on the video communication applications developed for Deaf people. The system that was in place is computer based and is discussed in detail in section 4.1.1 of this dissertation.

1.3 Research Objectives

The main objective of this research is to find out if a gesture based interface can improve the usability of an asynchronous or store and forward video communication for Deaf users. We wanted to investigate whether using an interface that can be controlled using hand gestures (without having to physically be in contact with a mouse or keyboard) would make it easier for Deaf people to communicate with each other. Whilst signing users have to sit at a reasonable distance from the camera and the computer requiring them to often move forward to use a mouse or keyboard. By implementing this interface the users are able to control the application from a comfortable signing distance.

We implemented a gesture interface where users just used pre-determined hand gestures in front of the camera to select the desired option. The final phase of this research involved investigating the feasibility of implementing this interface on mobile devices.

The overall aim of the research:

1. To understand how Deaf people communicate with each other and how technology can help in aiding the communication.
2. To design communication applications that takes into consideration the Deaf context and is suitable for sign language video communication.
3. To determine if using a gesture based interface will improve the usability of asynchronous video communication system.

1.4 Significance of Research

Deaf people in South Africa currently do not have efficient video communication on mobile phones and this is discussed in the next chapter in section 2.2. This research addresses this problem and a new way of interacting with a mobile phone is proposed. This allows users to utilise hand gestures to control the video communication application on mobile phones. It enables Deaf people to communicate with each other on mobile phones using video with acceptable video quality for sign language. The overall main objective of this research was to investigate if a gesture based interface will improve usability of an asynchronous (store and forward) video communication system for Deaf people. In addition, we investigate whether using an interface that can be controlled by hand gestures resulted in an improvement on the flow of the conversation.

1.4.1 Research Questions

The following questions were the focus of our research:

1. How can we use mobile phones to provide efficient and effective video communication for sign language?

Mobile phones are effective and convenient for communication among hearing users. Hearing users can use audio, video call and also text to communicate on these devices. Deaf people, on the other hand can only use text on mobile phones as these devices are currently not suitable for sign language video communication. The real-time video communication on mobile phones lacks sufficient video quality required to view all the sign language features clearly. In this research we investigate how these mobile phones can be used to provide effective video communication for sign language.

1.1 Can the main rear camera on the mobile phone be used in conjunction with an external display on a television screen?

Most mobile phones have two cameras with the camera having better video quality positioned at the back of the handset. This camera at the rear can record video with better quality in terms of resolution and frame rate, quality which is essential and suitable for sign language video communication. Unfortunately this introduces the problem of how to display the video considering the fact that when the camera faces the user, the phone screen faces away from the user. We have to investigate whether a television set can be used in conjunction with the rear camera to record video. Although this approach limits the mobility there is a gain in video quality. This provides an extra option for the users to communicate with each other efficiently when they are in an area with a television.

1.2 Can a gesture based interface be implemented effectively on a mobile phone?

Mobile phones have a lower processing power as compared to personal computers since they have smaller processors and memory. We investigate if mobile phones can process and recognize hand gestures at real-time while giving feedback within a reasonable response time when a user is signing.

2. Does a gesture based interface improve usability of a store and forward video communication for Deaf people?

Using hand gestures to control an application can be thought of as being similar to sign language communication as they both use the same mode of visual interactions. They both involve using hands to make gestures or signs. Deaf people prefer communicating in sign language so interacting with a mobile phone in a familiar way may have an effect on the usability. We investigate if using an interaction mode that incorporates gestures has an effect on usability.

1.5 Dissertation Outline

The rest of this dissertation is organized as follows. Chapter 2 gives an overview of the related literature and highlights the different aspects of this research. The chapter also gives a background on Deaf culture and communication methods used by the target users (Deaf people). The experimental design and the methodologies used are explained and discussed in Chapter 3. The process followed when gathering information and understanding the target users is covered in Chapter 4. The design and implementation of the prototypes is discussed in Chapter 5. The interaction method for controlling the interface using hand gestures is also introduced in this chapter. The process followed when evaluating the prototypes is discussed in Chapter 6. Furthermore presentation of results and analysis are also covered in Chapter 6. In the end the conclusions from the research are discussed in Chapter 7 together with possible future research related to this study.

2. Background Chapter

In this chapter, we discuss and review the existing literature with a focus on the Deaf culture and designing applications for the mobile phone platform. The different sub sections in this chapter cover the different aspects needed to implement a gesture based video communication system. The statistics about Deaf people in South Africa are also discussed in this chapter.

2.1 Deaf

There is a difference between "deaf" in lowercase and capitalized "Deaf". Capitalized Deaf refers to people who are active in the signing Deaf community and is associated with Deaf culture whereas lowercase deaf is typically people with a hearing loss (Cavender et al., 2006). Lowercase deaf even includes people who lost their hearing at old age and do not use sign language as a means of communication. People who are deaf are not necessarily considered to be part of Deaf Culture. Deaf Culture involves more than just sign language and is formed around shared experience, common interests, shared norms of behaviour and survival techniques **(**Deaf Linx, 2009). In South Africa Deaf people prefer to communicate with each other through South African Sign Language (SASL).

Deaf people are more economically disadvantaged than the hearing people in South Africa. It is estimated that around 65% of all deaf adults are unemployed and many of those who have jobs are underemployed (Glaser and Aarons, 2002; Glaser and Tucker, 2004) but according to DeafSA (2010) 70% of the Deaf community is unemployed. There is debate surrounding the exact figures from the different sources. Due to lack of adequate educational practices 30% of Deaf adults are functionally illiterate (Glaser and Aarons, 2002; Glaser and Tucker, 2004). There are inadequate educational facilities for the Deaf as the majority of Deaf children do not attend school or end up attending at a late stage (Asmal, 2001; Kiyaga & Moores, 2003). DeafSA (2010) states that there are only 12 schools for the Deaf that offer Grade 12 and these schools are concentrated in 3 provinces and only 2 FET Colleges employ Sign Language interpreters. 1 in 10 babies in South Africa is born with a hearing loss (DeafSA, 2010). In 2001, the South African national census statistics (Statistics SA, 2009) show that there are 453,104 with a hearing disability in South Africa. There are around 600

000 South Africans that use South African Sign Language as their primary language (DeafSA, 2010).

2.2 Deaf Communication

Deaf people are unable to use voice based communication methods on a telephone. They must employ either text or video communication on the phone. Deaf people can only communicate with each other by signing to each other face to face or use text based communication like fax, IM and SMS on mobile phones (Tucker et al., 2003a). There was a device called Teldem, a product invented by Telkom which uses text as a means of communication for Deaf people (Glaser and Tucker, 2001). The Teldem has a keyboard and a two line screen that can be used for typing messages. The device is connected to a telephone system similar to a normal telephone and communication takes place over standard Public Switched Telephone Network. The communication on Teldem is half-duplex therefore only one user can type at a time (Glaser and Tucker, 2001).

The use of SMS on mobile phones is unnatural and slow therefore it is not preferable. SMS cannot effectively substitute for real-time synchronous communication (Tucker et al., 2003b). The use of mobile phones with video cameras is becoming popular and these devices can be utilised to provide efficient means of communication for the Deaf community. Real-time video communication that is currently available on the mobile phone is not suitable for sign language communication due to the poor video quality which results in blurry motion and unclear facial expressions (Ma and Tucker, 2007). Asynchronous video is a favourable alternative for providing video communication suitable for sign language. Ma and Tucker (2008) determined that Deaf users prefer to communicate in sign language and are more concerned about the video quality than the delay.

2.3 Video Communication

When users are signing they have to sit at a reasonable distance from the camera and the computer in order to ensure that their body from the chest to the head can fit in the video frame, since body language and facial expressions are vital in sign language. Deaf people using a video chat application often have to move forward to use a mouse or keyboard in order to record and stop a video. When signing both hands are needed therefore the mobile phone has to be placed on a stand. Currently most mobile phones have cameras on the back and the screen on the front of the phone. This makes it easier for people to see what is being recorded but is unsuitable for the purposes of recording oneself which is what is essential in sign language conversation.

2.3.1 Video on Mobile Devices

Mobile devices have cameras that can be used for video communication but these cameras are not available on all the mobile devices. Cavender et al. (2006) conducted a study with a focus group consisting of four members of the American Deaf community to learn more about potential uses of mobile phone video technology. Participants recommended that the phone should have a way to prop itself up and mentioned that the camera and the screen should face the same direction to allow filming of oneself. Some phones have two cameras where one camera is on the same side as the screen and the other one faces away from the screen. An example of such a phone is the Nokia N96 (GSMArena.com, 2008c). The camera with the better video quality though, in most cases faces away from the screen. This poses a potential obstacle for Deaf people because the camera on the same side as the phone's screen has low video quality. To provide intelligible sign language communication, the video quality has to be sufficiently sharp in order for the facial expression to be clearly visible and the motion should not be blurry.

2.3.2 Video Compression

The inadequate video quality in real-time video results in the loss of some of the important sign language features like facial expressions and eye movements. Finger spelling in sign language needs good video quality as there is a need to see all the fingers and words are finger-spelled at a rapid rate. By using asynchronous video communication some of the problems experienced in real-time communication can be solved. Store and forward makes it easier to preserve video quality since there is enough time to apply better compression methods that will preserve sign language features (Ma and Tucker, 2007). The compression on the video also lowers data costs for video transfers by decreasing video file size. This makes video communication possible even when there is low bandwidth which is not suitable for real-time video communication. In store and forward we can also use a higher video frame rate which will reduce blurriness that can occur if someone is signing fast.

Video compression also plays a vital role in an asynchronous video communication application because there is a need to decrease the size of the video files that are transmitted over the network. Smaller file sizes will result in lower network costs and reduce the delay experienced by users while sending files. Ma and Tucker (2008) explained that Deaf users are more concerned with the video quality than the delay, since a tiny visual gesture may be a key to understating the entire sequence. Muir and Richardson (2002) propose a way to optimize video compression systems for a specific application aimed at personal communication for Deaf people. Instead of optimizing compression across the entire video scene which is adequate for general purpose applications, the video compression is optimized to take into account information carrying importance of different temporal and spatial components. Segmenting the video scene and employing object-based coding makes it possible to make better use of the bandwidth available by selectively optimizing the key features or important regions of sign language video sequence for the target user. This results in enhanced quality of important video regions that carry sign language features. Cavender et al. (2006) also uses a similar method where the face regions of the video are encoded at better quality than the other regions. They also used reduced frame rate encodings where fewer better quality frames are displayed every second.

MobileASL is a project which focuses mainly on video compression that enables wireless mobile phone communication for sign language (Cavender et al., 2006). Although this work is for compression of real-time video, a study was conducted with a certified sign language interpreter to test the different frame rates of sign language video. Preliminary tests showed that 10 frames per second (fps) and 15 fps were both acceptable for intelligible American Sign Language. Decreasing the frame rate further down to 5 fps made it nearly impossible to watch and understand finger spelling. However the difference between 30 fps and 15 fps was negligible.

2.4 Gesture Recognition

We are only interested in recognition of simple hand gestures which are not necessarily part of sign language. Designing the gesture interfaces can be a complex process as we need to use gestures that are easy to learn and gestures that are natural. For users to keep their arm extended repetitively and even for short periods of time can be very stressing and can affect the usability of the interface. There is a need to use simple gestures that will not tire the users of the system. Finger and arm pointing are probably one of the most basic gestures we all use (Cabral et al., 2005). Nielsen et al. (2003) lists the following principles in ergonomics to building a good gesture interface:

1. *Avoid outer position.*
2. *Relax muscles.*
3. *Relaxed neutral position is in the middle between outer positions.*
4. *Avoid repetition.*
5. *Avoid staying in static position.*
6. *Avoid internal and external force on joints that may stop body fluids.*

These principles help in developing gesture interfaces that are easy to use without putting too much stress on the user even if a gesture is repeated several times. If gestures are defined with care and taking into consideration the fatigue that can be caused by repetitive actions, the user experience can be enhanced without degradation over time.

There are perceptual user interfaces that use alternate sensing modalities to replace or to complement traditional mouse and keyboard input. Wilson and Oliver (2003) discuss a system called GWindows that uses a real-time stereo vision algorithm. The system used two inexpensive video cameras to sense the presence of a user, track the user's hands in order to control a cursor and also to perform commands with gestures. The cameras were used to detect depth information as well. The system was not developed specifically for Deaf users but it uses a video from the camera to pick up any input commands from the users. The use of cameras to control the cursor is an interesting mode of interaction worth considering when developing the gesture based video communication system. This makes it possible to have a system that can be controlled completely using hand gestures. The cameras used were inexpensive which makes it possible to implement such a system on a mobile phone. Ideally the gesture based video communication system has to be implemented on a mobile phone.

Another line of research involves controlling electronic products using hand gestures. For example Ike et al., (2007) describes a real-time gesture recognition system and its application to controlling consumer electronics. Freeman and Weissman (1995) presented a method of controlling a Television using hand gestures. The system used a video camera connected to a workstation to record the images. When a trigger gesture was detected, the television entered a control mode and a hand icon appeared on the screen which tracked the movements of the viewer's hand. This mode of interaction is worth considering on the mobile platform where the gesture based video communication has to be implemented. Deaf people can place the phone down and be able to control the application without touching the actual device. In Freeman and Weissman (1995) the hand gestures were detected using a workstation. Detecting the hand gestures on the mobile phone poses a greater challenge because these devices have smaller memory and processors as compared to the workstation.

2.4.1 Sign Language Recognition

There has been substantial research conducted in sign language recognition and most of these works make use of the Hidden Markov Models. In this project the aim is not to recognise sign language but certain predetermined hand gestures, which will be used to control the system. The initial idea was to have a certain programmed hand gesture that is used as a command for a certain function. The gestures did not necessarily have to be part of sign language. Starner et al. (1998) use a single camera to extract two dimensional features of continuous American Sign Language and use the features as input for the Hidden Markov Model. When the camera was mounted on a desk a word accuracy of 92% was achieved and when the camera was on a user's cap 98% word accuracy was obtained. All these results were achieved when recognising the sentences based on a 40-word lexicon.

Zhang et al. (2005) proposes applying adaptive boosting strategy to continuous Hidden Markov Model which is aimed at solving pattern recognition errors that arises when there exists some hard to classify samples in the sample space. These errors in sign language recognition arise as there are single hand signs or a similar sign pair which is usually regarded as a difficult problem. Experiments were conducted on the vocabulary of 102 frequently used Chinese Sign Language single handed sub word signs, where every two or more signs were similar. Experimental results showed that the boosted Hidden Markov Model had an accuracy improvement of 3%. Ramamoorthy et al. (2003) discuss the development of a Hidden Markov Model based gesture recognition system which uses both temporal and shape characteristics of the gesture. The system developed works in real-time and does not require special gloves to be worn. The recognition process works in a satisfactory manner even in the presence of background clutter so there is no need for a uniform background. The system also allows the change in hand shape which gives flexibility to the user to alter hand shape between hand gestures.

All these works mentioned above recognize hand gestures and use techniques that can also be applied to the video communication system. The system has to be able to distinguish between movements on the background and gestures that are aimed at controlling it. Deaf people use sign language on the video communication system and also hand gestures to control the system. The system has to be able to pick up gestures that are used to command it from gestures that are part of sign language. The two types of gestures should not be mistaken for each other. The gestures that are used for controlling the system are simple hand gestures and do not form part of sign language.

2.5 Mobile Device

The use of mobile phones has significantly increased over the past few years. ITU (2009) cited by Dörflinger et al. (2009) states that approximately 1 million people become mobile phone users everyday and 85% of these reside in the developing world. The number of people using mobile phones in the developing world is growing enormously. In 2002 44% of all mobile subscribers worldwide were from the developing countries and five years later in 2007 the number had increased to 64% mobile subscribers worldwide (ITU, 2009; Dörflinger et al., 2009). The mobile phone industry penetration in South Africa at the end of 2007 was about 83% percent (Telecom Week, 2008; Dörflinger et al., 2009). The biggest potential of future mobile development is in the market of about one billion potential users in emerging economies (Dörflinger et al., 2009).

2.5.1 Mobile Devices in ICT4D

The use of mobile phones can have as tremendous an impact in the developing world as the personal computer did in the developed world. With the increase in mobile phones in the developing world people can have access to information using their mobile devices. Dörflinger et al. (2009) states these users will experience the internet first on a mobile device rather than a computer. The availability of these devices in the developing world has lead to development of applications to alleviate poverty. Such an example of a successful m-banking application is M-PESA (Morawczynski and Miscione, 2008).

Mobile phones are used in projects like Cell-Life (Cell-Life, 2010) to provide information and for communicative services for people infected by HIV. In South Africa there are approximately 36 million active mobile phone users and 80% of the youth and adults have a mobile phone (Cell-Life, 2010), this makes the use of mobile phone technology to reach the larger masses an effective means. Communication applications developed for mobile phones have the potential to reach the 80% of adults and youth who have access to mobile phones. The availability of mobile phones makes the mobile phone a good platform for the gesture based interface developed in this project.

2.5.2 Mobile Device Limitations

Mobile phones have lower processing power compared to Personal Computers since they have smaller processors and memory. They usually have processors that are around 100 MHz or multiples of this whereas desktop personal computers have processors that in the GHz range. Smart phones have better processing power and some new phones like Nokia N96 have dual processors. The Nokia N96 has a Dual ARM 9 264 MHz processor, a storage memory of up to 16 Gigabytes, 128 Megabytes of RAM, 2.8 inch screen and a 5 mega pixels camera (GSMArena.com, 2008c). Mobile phones also have smaller display screens which support screen resolutions of (320 x 240) whereas Computers can support resolutions around (1280 x 1024). Top of the range phones like N96 has a 5 mega pixel which is good camera but other phones like Nokia 6120 classic (GSMArena.com, 2008a) and Nokia N70 (GSMArena.com, 2008b) have a 2 mega pixel camera which does not have good video quality. The poor video cameras found on these phones make it harder to recognise hand gestures.

Power is another limiting factor since mobile devices use a battery as a power source. Applications running on a mobile phone should be designed to use less power in order to pro long the battery life of the device. In order to develop real-time applications with reasonable response times in the mobile device platform, efficient algorithms and data structures are required. Efficient memory management and code optimizing can play a major role in development of usable applications as they speed up the processing time required by the application. The issue of limited resources is also addressed by Bowles (2007) in a study conducted with undergraduate students creating a mobile phone game using efficient data structures and algorithms. After observing the students developing the game it was clear that using mobile phone as a platform for computer game was a feasible option for a project. However it was also noted that the limitations on the mobile platform and the overhead of learning the development environment was not suitable for beginning student in a CS1 course but was feasible for a CS2 course.

The quality of mobile device applications should also be considered when assessing limitations on mobile devices. Therefore the limitation on resources should not result in development of applications that are not robust. Umphress et al. (2004) address this issue by developing software which conforms to industry standards. The software developed by students in this work was assessed using the Nokia OK Certification procedures. The Nokia OK Certification procedures are basically used to test if software is robust and conforms to the industry standards for wireless software development. This certainly helps in promoting the development of robust applications for mobile devices. The gesture based video communication system implemented on the mobile phone can be tested using these procedures in order to ensure that it is robust and conforms to industry standards.

2.5.3 Development Environment

Mobile phone manufactures enable for the deployment of applications on a wide range of mobile phones by providing an operating system which runs on the device. Developers have to create applications that will run on a specific operating system and then deploy it on any mobile phone that runs that operating system. These operating systems include Binary Runtime Environment of Wireless (BREW) and Symbian (Coulton et al., 2005). There are other mobile phone platforms like Windows Mobile (Windows Phone, 2010) and there is a new mobile phone developed by Google that runs on the Android platform (Android, 2010). There is a smaller version of Java called Java Micro Edition which makes it possible to develop applications with cross platform portability which can then run on any phone that supports Java regardless of the underlying operating system (J2ME, 2007).

There are several programming languages that can be used to develop applications that are able to run on the Symbian operating system. These languages include Symbian C++, Java ME, Flash Lite and Python for S60. These languages need to be analyzed carefully before choosing the language that is suitable for developing the gesture based video communication system. Criteria that can be used to compare these languages include camera support, communication, file system, multimedia support and life cycle of application.

Camera: Java ME and Symbian C++ applications can take a picture, record a video and also have the ability to process image data directly from the camera. Direct access to image data is very useful for motion detection applications. This is required in order to implement the hand gesture recognition system. Flash Lite does not support taking a picture or video recording. Access to image data directly does not come as a standard feature but developers can extend the runtime or call out the underlying native OS APIs (Symbian Developer Network, 2008).

Communication: Java ME and Symbian applications are capable of accepting Bluetooth connections, send data and receive data via Bluetooth. These applications can accept TCP/IP connection from another device, send data and receive data via TCP/IP. They are also able to control initiation of GPRS connection and WiFi connections. In a video communication application the video files have to be sent from one mobile phone to another. The application has to be able to send and receive these videos in order for the video communication to take place. Flash Lite does not support all these as a standard feature but developers can extend the runtime or call out the underlying native OS APIs (Symbian Developer Network, 2008).

File System: Both Java ME and Symbian C++ applications can read and write files on the device subject to underlying security permissions whereas Flash Lite does not support these features. Flash Lite and Java ME can read and write XML files subject to underlying security permission but Symbian C++ 's capability to read is only limited to XML files. Flash Lite has limited file sizes other than available storage space on the device whereas Java ME and Symbian C++ file sizes are restricted by the available storage space on the device (Symbian Developer Network, 2008). Java and Symbian are suitable for the video communication systems because they permit applications to create video files. The video file sizes are only dependent on the available space. The video files do not need restrictions on the length of the video since users can create space to accommodate bigger file sizes.

Multimedia: Java ME and Symbian C++ applications can play audio, record audio, play video, record video and also display an image. Flash Lite only supports playing audio, playing video and displaying an image as a standard feature but recording audio and video requires the developer to extend the runtime or call out the underlying native OS APIs

(Symbian Developer Network, 2008). The video communication system has to be able to play video and record video so that Deaf people can communicate with each other. This feature is essential in order for the system to function appropriately and serve its purpose. The recording of audio and audio playback is not necessary because the system is used by Deaf people. Sign language on video does not utilise any voice communication so the audio support is not needed.

Life Cycle: Java ME and Symbian C++ support background running where once an application has been started, it continues to run until the phone is shut down or rebooted. Flash Lite does not support background running. Symbian C++ applications can be started automatically when the phone is switched on without any user intervention whereas Java and Flash Lite do not support this feature. Java ME and Symbian applications can be started from a browser or by other applications while Flash Lite does not support this as a standard feature (Symbian Developer Network, 2008).

Java and Symbian C++ have all the features that are necessary for developing vision based systems that use the camera to get input aimed at controlling the application. To successfully develop a gesture based video communication system we need to be able to process images from the camera in order to determine what actions the user intends to perform. The system should also be capable of recording, playing, sending and receiving video messages. All these features are supported on Java and Symbian C++ which makes them good candidates for developing gesture based communication applications. Flash Lite supports some of these features but in most cases the developer needs to extend the runtime which will require more development time than in Java or Symbian C++ where all these features are already supported as a standard.

The following table summarizes the features supported by the different runtime environments which were discussed above.

Area	Feature	Explanation	Flash Lite	Java ME	Symbian C++
Camera	Take picture	The application can take a picture.	×	√	√
	Record video	The application can record video.	×	√	√
	Access image data directly from camera	The application has the ability to process image data directly from camera, e.g., for motion detection.	√*	√	√
Communication	Send and Receive via Bluetooth	The application can send data and receive data via Bluetooth.	√*	√	√
	Accept Bluetooth connection	The application can accept a Bluetooth connection from another device.	√*	√	√
	Send and Receive via TCP/IP	The application can send data and receive data via TCP/IP.	√*	√	√
	Accept TCP/IP connection	The application can accept a TCP/IP connection from another device.	√*	√	√
	Initiate GPRS and Wi-Fi connection	The application can control initiation of GPRS and Wi-Fi connection.	√*	√	√
File system	Read and Write files	The application can read files and write files on device subject to underlying security permissions.	×	√	√
	Unlimited file size	The file size is unlimited other than by available storage space on device?	×	√	√
Multimedia	Play audio	The application can play audio.	√	√	√
	Record audio	The application can record audio.	√*	√	√
	Play video	The application can play video.	√	√	√
	Record video	The application can record video.	√*	√	√
	Display image	The application can display an image.	√	√	√
Lifecycle	Always on	Once the application has been started, it continues to run until the phone is shut down/rebooted (also known as background running).	×	√	√
	Start on boot	The application can be started automatically when the phone is switched on.	×	×	√
	From browser From other application	The application can be started from browser or from other applications.	×*	√	√

Table 1: Comparisons of Runtime Environments.

√ ***- In standard platform release.***

×* ***- OEM / device / platform dependent.***

√* ***- The feature does not come as standard, the developer needs to extend the runtime or call out to the underlying native OS APIs, e.g., C and C++.***

× ***- Not in standard platform release.***

The original table was obtained from (Symbian Developer Network, 2008)

2.5.4 Image Processing Performance on Mobile Phones

Application speed is an important factor when developing real time applications so there is a need to compare the speed for applications developed in different languages. Kerr et al (2009) conducted a study to test three programming languages which were Symbian C++, Python for S60 and Java Micro Edition. The tests were conducted on a Nokia N95 which has Dual ARM 11 332MHz processor including a 3D Graphics accelerator and 64MB of RAM. The test involved processing images from the camera in order to detect a certain block of colour in the middle of the image, this involved a six part binary ‘OR’ comparison for the purpose of detecting the dot. Other test involved a full scan of the image where every pixel in the image is scanned and then writes each pixel back onto itself without conducting any binary comparisons. All images were taken at a resolution of 640×480 and converted to bitmaps before they were processed by the detection algorithm.

The results obtained showed that programming at the low level in C++ produced the fastest results but limited the deployability of the applications to just a series of devices running on the same version of Symbian. Symbian applications can only run on phones with the same environment. They are further limited to a certain series of smart phones (e.g. Series S60 on Nokia) and have to be rewritten to cater for other series of smart phones. The images were processed very rapidly and C++ provided the highest number of frames per second that were obtained from the camera. The number of frames per second was just over 18 fps. When performing a full scan where every pixel in the image is scanned and written back to itself the frame rate dropped to just over 14 fps. It took 0.05 seconds to track and scan half of the image to detect a red dot in the middle of the camera’s field of view (Kerr et al., 2009). This makes it possible to develop video communication applications for Deaf people as the video frame rate is not affected drastically by the image processing to a point where the sign language video is unusable.

Python was able to obtain images from the camera with a speed of over 10 fps but had lowest image processing speed. The images were obtained from the camera at reasonable rate but the time taken to process every pixel in the image was too high thereby the application ended up with a frame rate of 0.05 fps. The number of images that were scanned in 1 second was 0.05 which means it took more than 1 second to scan all the pixels in the image. It took python 9.9 seconds to detect a red dot in the middle of the camera’s field of view. Python showed moderate deployability and low speed when compared to the Symbian C++ and Java ME. The application created in Python was able to obtain frames from the camera by implementing a wrapper to the Nokia CV class (Kerr et al., 2009). The drop in frame rate when doing image processing makes it more difficult to use for sign language video communications applications. What is required is a system capable of detecting hand gestures at a usable response time, whilst maintaining a good frame rate for the sign language video.

Java ME was limited by its ability to get frames from the mobile phone's camera as there was no proper function to directly access the frames from the camera. A snapshot had to be taken at 640×480 pixels by using the camera shutter which takes a longer time than the methods used in C++ and Python. Java ME was able to obtain the images from the camera at a frame rate of 0.2 fps. When performing a full scan the frame rate dropped to 0.02 fps. The time taken to scan half an image in order to detect a red dot in the middle of the camera's field of view was 3.9 seconds which is 6 seconds faster than in Python (Kerr et al., 2009). To successfully develop a gesture based system the system should be able to process every video frame from the camera in order to detect any hand gestures aimed at controlling it. Taking a snapshot is not suitable for video communication as it takes too long to acquire the image, the system needs to have reasonable response times and have the ability to record the video messages. To develop a gesture based video communication system requires the ability to process video frames and record the video simultaneously. This renders the Java approach inadequate for developing the system.

There are more Java ME compatible phones which provide a wider coverage of handheld devices. Java has moderate speed and highest deployability as it can run a wider range of devices including the Symbian platform. Java application can be deployed on mobile phones that do not run Symbian because it only requires a device which has the Java run time environment. Symbian C++ application can only run on the mobile phones with the Symbian platform. Davis et al., (2005) demonstrate the possibilities of application tailoring. This is very useful and helps in porting software to a specific mobile device without manually rewriting the code in Java ME. Java ME applications run slower than applications developed for Symbian and this difference in speed can be explained since Java applications run on the Kilobyte Virtual Machine (KVM) whereas Symbian applications are compiled to machine code which runs faster (Coulton et al., 2005).

2.6 Chapter Summary

This chapter gives an overview of the literature that covers the different aspects of the gesture based video communication system. The main aspects covered include sign language, video communication, gesture recognition and mobile devices. The current situation faced by Deaf people in their daily lives is discussed with a focus on the communication problems. The different communication difficulties faced by Deaf users are presented here in this chapter. Furthermore video communication is also discussed as it solves some of the problems experienced by Deaf users when using text based communication.

The mobile device was deemed as a suitable communication platform as it is more accessible by the target group therefore the different aspects of the mobile platform had to be discussed in detail in this chapter. The notable outcomes of this chapter were that a video communication approach would solve most of the communication problems and deploying this communication on mobile devices result in more users having access to the technology.

3. Experimental Design

This chapter gives an overview of how the research was conducted. The users involved and all the other parties that played a role in the project are also highlighted in this chapter. The methodology used to design and evaluate the prototypes developed is discussed as well.

3.1 Research Questions

This research was conducted with the aim of answering a couple of research questions. The research questions have been discussed already in the introduction chapter under section 1.4. These were the research question that guided our research:

1. How can we use mobile phones to provide efficient and effective video communication for sign language?
   1.1 Can the main rear camera on the mobile phone be used in conjunction with an external display on a television screen?
   1.2 Can a gesture based interface be implemented effectively on a mobile phone?

2. Does a gesture based interface improve the usability of a store and forward video communication application for Deaf people?

The rest of this chapter discusses the process followed in this research in order to answer the research questions. All the stages followed in this research contributed in finding the solutions needed to answer our research questions.

3.2 Methodology

To develop usable interfaces, researchers in Human Computer Interaction (HCI) have developed various techniques that take into account the factors that influence how people will interact with these interfaces. A user centred design approach is one of these techniques that can be used to develop usable interfaces. In this research a user centred design approach was employed in order to gain insight on the Deaf context and to produce a usable design. This involved following an Interaction design model where the main focus was on the Deaf users and their needs. Jones and Marsden (2006) identified the three main stages of interaction design which are:

1. Identifying needs and establishing requirements.
2. Developing alternative designs.
3. Evaluating designs.

The three basic activities listed above are executed using an iterative development methodology. The iterative stages involve requirements gathering, designing and evaluation of the prototypes on the computer and on the mobile phone platform. Deaf users were involved in this project from the requirements gathering stage, in the design of prototypes and all the way to the evaluations of the designs. The requirements gathering stage is discussed in detail in Chapter 4 and the prototype designs are discussed in Chapter 5. The four main iterative stages were:

1. Requirements gathering
2. Computer Prototype
3. First Mobile Phone Prototype
4. Second Mobile Phone Prototype

The prototypes stages involved establishing requirements, developing alternative designs and evaluations of the design. Therefore the three stages of interaction design were repeated for all the prototypes developed.

3.2.1 Requirements Gathering

The researcher needs to communicate with the end users in order to successfully build applications when using a user centred design approach. The target users in this research were Deaf people who use sign language as means of communication. Unfortunately the researcher could not communicate using sign language in the beginning of the project so we enrolled in a sign language course at Sign Language Education and Development (SLED) (SLED, 2009a). This was done in order to gain more understanding on how sign language works and to improve the communication between the researcher and the end users. The course involved weekly lessons for a period of six months at SLED offices. The sign language course was very useful to our understanding of important sign language features.

The researcher had to make weekly visits to DCCT to meet with the users throughout the whole research duration. DCCT is at the Bastion for the Deaf in Newlands. These visits made it easier for the users to get better acquainted with the researcher and as a result they became comfortable and were able to open up more. The more information can be obtained from the users the more researchers can understand the reasons behind the user's actions and how they interact with technology. Learning sign language also made it easier for the users to communicate with the researcher using their preferred language instead of resorting to text based communication. This further helped in improving communication during the weekly visits in the absence of the interpreter who was only available during the main studies. Unfortunately the interpreter could not be used during the weekly visits because of the cost and availability of the interpreter. Although the researcher was learning sign language a professional interpreter was used during interviews, focus groups and evaluation session. Communication with the Deaf users using an interpreter was more effective as both parties were able to use their preferred language.

During the requirements gathering stage a focus group study was conducted with Deaf users to obtain feedback on the current video communication systems. Two systems were used in this study with one being a real-time video communication system and the other an asynchronous video communication system. The users had to evaluate the two systems which were computer-based and give feedback on their usability. The users had a discussion afterwards about how to design a new video communication system. This discussion was used to establish what features should go into the prototypes designed in this research. This study was used as a starting point for gathering requirements for the prototypes produced in this research in order to determine if we should build on what is already there or start over from the beginning using a completely different approach. The study aided in getting insight on how the users want the communication application to be structured.

Furthermore a survey was conducted in order to investigate which communication methods the target users are comfortable with. The survey was conducted with Deaf users who were participating in the prototype evaluation studies. It was conducted as an ongoing process where users were asked to fill a questionnaire with communication related questions. The survey covered communication technology issues to determine what kind of tools they used for communication. The users were asked questions to determine what methods of communication they were using on mobile phones and on computers. The results from the survey are discussed in Chapter 4 in section 4.2.

3.2.2 Computer Prototype

The second stage involved developing a computer prototype for the gesture based asynchronous video communication system. The prototype was designed using the requirements that were obtained from the requirements gathering stage. The prototype had to perform the basic features that were specified by the target users. Computer prototypes are easy to implement so this made it possible to test the concept at an early stage. The users were consulted during the design phase and had to determine what gestures are suitable for controlling the interface. The computer prototype was evaluated with Deaf users to test its suitability for sign language video communication. The gestures were also evaluated to determine if they were useful. The users had to evaluate the usability of the prototype and also appraise whether the video quality was enough for sign language. The feedback from the users about which features should be added was analysed and the users had to determine which features were more important.

3.2.3 First Mobile Phone Prototype

The first mobile prototype was implemented entirely on the mobile phone. The prototype was designed based on the previous computer prototype. The mobile device platform is less powerful when compared to the computer as it has lower processing power and less memory. At this stage the aim was to test if the gesture based video communication can be implemented successfully on the mobile phone. The users were also involved in the process of redesigning the computer prototype to work on the mobile phone. The users were consulted again in the design phase because some of the features and design aspects had to be revised before the prototype was implemented on the mobile platform.

The new design had to cater for the mobile device limitations which were not considered when designing the computer prototype. However this mobile prototype was not evaluated with the Deaf users because it did not meet the performance requirements required in order to produce a usable interface. For an interface to be usable and effective the users need to be able to get a response from the system in real-time. The problems experienced when conducting performance tests for this prototype are discussed in Chapter 5 under section 5.3.2. The mobile platform was unable to produce acceptable response times so the prototype had to be revised to meet the performance requirements.

3.2.4 Second Mobile Phone Prototype

The final prototype used both the computer and the mobile phone platform. The computer was used to do all the processing needed by the application therefore it was simulating a mobile phone with more processing power. The prototype was implemented the same way as the first mobile prototype with the exception that the processing was done on the computer in order to get a reasonable response time. To the user it appears as if the video communication prototype is running on the mobile phone. The prototype was evaluated with 10 Deaf users in order to determine if it achieved its goals. The main goal of the project is to provide a usable video communication application so the users had to evaluate the usability of the application and also evaluate the quality of the video. The users had to evaluate whether the prototype was usable and effective in facilitating sign language video communication. The prototype was evaluated using a questionnaire, observations and interviews. The questionnaire was used to rate the usability and efficiency of the different aspects of the prototype. The observations were used to analyse the way in which the users interacted with the prototype. The interviews were used to get feedback on how the users perceived the prototypes and also to give the users a chance to make comments and suggestions.

3.3 Prototype Designs

Three prototypes were developed; the first one on a computer, the second one on the mobile phone and the last one a mobile phone simulation that used both the computer and the mobile phone. Prototypes make it possible to get feedback on the designs and enable users to give feedback from an early stage of the design process**.** Jones and Marsden (2006) discussed the different types of prototypes which are low-fidelity prototype and high-fidelity prototype. Low-fidelity prototypes use materials that are different from the final product and are supposed to be cheap and easy to produce. High-fidelity prototypes use materials that may be in the final product and are usually more complex to produce as compared to low-fidelity as they are supposed to resemble how the final product will look like.

In this project both low-fidelity and high-fidelity prototypes were used but only high-fidelity prototypes were successful in obtaining insight on how the end users want the system to operate. The gesture based interface was used to control the video communication application. The Deaf users could not understand how such an interface would work. The users had not used any sort of gesture based interface before so they were not familiar with this kind of interaction making it unsuitable to use a paper prototype (low-fidelity). This is due to the fact that in order to use a paper prototype, the user has to imagine what the final interface would look like but without any previous exposure to this kind of interface they could not relate to the concept.

The Deaf users preferred a working prototype instead of a paper prototype because they did not really understand how things worked in theory. They preferred to interact with the system in order to fully understand what was going on. All the users were familiar with a touch screen and wanted something similar to a touch screen as they all knew how it worked. Maunder et al (2006) suggests that many techniques that originated from usability experience conducted in the developed world environments may not be relevant in the developing world. These techniques often have to be re-evaluated before being applied to the developing world context.

High-fidelity prototypes were used in this project with the first prototypes implemented on a computer. Computer-based prototypes are easier to develop as compared to mobile phone prototypes which need complex implementation as these devices have lower processing power. The gesture based interface was implemented as a prototype on a computer first and then tested with Deaf users before moving to mobile phone prototypes. Development of computer prototypes was quicker and this made it easier to get feedback more rapidly and to serve as a proof of concept.

3.4 Evaluation of Designs

Evaluation of designs is very important in testing whether the design accomplishes what it was designed for. This also gives insight into the applicability and usability of the design. Usability can be measured both quantitatively and qualitatively in order to evaluate if a design is usable. User evaluations can be used to test the usability and intuitiveness of the design. In order to measure the usability accurately the evaluation has to include the target users of the system. Using the actual users provides insight into how the users will interact with the design in real world situations and can reveal problems that could not be discovered during the design process. Evaluating designs with actual users can also reveal the strength of the design thereby confirming the usability of the system. However the evaluation can also reveal the weaknesses in the design that hampers its usability. System evaluation is also useful in testing whether the system actually performs its functions correctly.

As mentioned before in the background chapter there is low literacy and low computer literacy levels in the Deaf community so applications developed for these users have to be simple and easy to use. All the prototypes were evaluated with Deaf users from DCCT and this helped in refining the aspects of the design which were not suitable for the target users. The users were given the opportunity to use both the computer and mobile prototypes. Questionnaires were used to get feedback on what aspects of the design were not suitable. User feedback and changes made to the prototypes are discussed in detail in Chapter 5 and the evaluation of the designs is discussed in Chapter 6. There were observations during the evaluation sessions in order to uncover parts of the design which users were struggling with. The observations also helped to determine which aspects of the design needed to be revised. Problematic parts of the prototypes that are not mentioned when users are giving feedback can also be discovered during observations.

3.5 Participants

This research focuses on developing communication applications for Deaf users. There is a need to test these applications with the target group in order to successfully evaluate the efficiency and effectiveness of the developed products in a real world situation. Deaf people who use sign language as their means of communication were consulted in the different stages of this research in order to gain insights on the applicability of this research. Deaf users involved in this research were recruited through the Deaf Community of Cape Town which is an NGO based at the Bastion for the Deaf in Newlands, Cape Town, South Africa. The weekly visits to the Bastion also involved participation of the Deaf users in the project. This made it possible to have a continuous interaction with the target group.

3.6 Technology and Tools

In order to design and implement the prototypes different tools and development environments had to be evaluated. The development environment can affect the design and implementation of the prototype due to the functionality that is supported by that particular environment. In Chapter 2 section 2.5.3 the different development environments were discussed. Based on that discussion and comparisons Symbian was chosen as the development environment for the mobile gesture based asynchronous video communication application. It was chosen as the development environment because it provides faster image processing capabilities needed to detect hand gestures. Symbian OS runs on a limited range of devices which limits the number of devices the application can run on but it provides a superior environment for applications that require speed. However the platform independence provided by J2ME can therefore determine the success of the application. This is due to the wider coverage of handheld devices that can run Java applications. Java applications run on the Kilobyte Virtual Machine (KVM) which makes them run slow.

The IDE used for developing the application is Carbide C++ 2.0 with S60 3rd Edition SDK Feature Pack 2. The SDK includes the default phone emulator even though it does not support all the basic features that are supported by the mobile phones. Video capture on the emulator is not supported, however the IDE supports on-device debugging. The on-device debugging allows for quick debugging of applications on the actual target device instead of using the mobile phone emulator. The IDE also supports self signing of applications as all applications that are installed on the Symbian OS need to be signed. Self signing allows the application to get access to some of the Symbian API's but some of the API's like the MultimediaDD need a developer certificate that can be obtained from Nokia.

The computer gesture based video communication prototype was developed using C++. The application used to simulate data processing by a powerful mobile phone was also implemented in C++. The mobile phone simulation application was implemented as a result of the problems experienced when the gesture interface was implemented on the mobile device platform. The IDE used was Microsoft Visual Studio 2008 so the application was tested only in Microsoft Windows environment. The operating system used in all the testing scenarios was Microsoft Windows XP Professional.

In the first computer prototype that did not use a mobile phone, the library used for the image processing and for obtaining the images from the web camera was OpenCV (OpenCV, 2010). The OpenCV library was also utilised in the final mobile simulation prototype to get images from the mobile phone when it was connected to the computer using the TV out cable. The library allows the application to get the video from the web camera or the video from the mobile phone as individual frames. The individual frames can then be processed to detect the hand gestures.

3.7 Chapter Summary

In this chapter we discussed the process followed throughout this research. The four main stages of the research were also discussed. The overview of how the different stages of the project were performed is discussed from requirements gathering stage to the design stage and all the way through to the evaluation stage. The overall methodology used in the research is also discussed. One notable outcome of this chapter is that the user centred design methodology was used. Furthermore we also discuss the participants involved in the design process and also in the evaluation of the designs. This chapter summarizes the experimental design used for the project.

4. Understanding Users

In this chapter we go through the steps used in order to understand the user's needs. The video communication prototypes designed in this research are aimed at sign language communication so there is a need to include Deaf users in the design process. Therefore before any designs can be made we have to understand what the users expect from the video communication application. We must also understand how they interact with this kind of a system. In order to understand the users, experiments and user interviews were conducted with the target group.

4.1 Requirements Gathering

The researcher has to comprehend the context in which Deaf people use communication applications and how they interact with interfaces in general. User centred design attempts to understand the user and the tasks they need to perform by using multiple data gathering techniques like questionnaires, interviews, focus groups, workshops and naturalistic observations (Jones and Marsden, 2006). There were weekly visits to DCCT and the visits involved observing Deaf people using the computers in the computer room. The users were not given any tasks during these weekly visits, they used the computers just like they normally would on any other day. The observation helps in giving insight to the researcher on how the users interact with interfaces in general. A focus group study was conducted with the users at DCCT in order to determine what features are needed in the video communication system.

4.1.1 Focus Group Study

To acquire information on video communication, a focus group study was conducted with Deaf users. In this study we expected to gain feedback about the usability and how to improve the two systems that have been developed for the Deaf users at DCCT. The system is computer-based and is a store and forward video communication system. The system was designed by Ma and Tucker (2007) and the focus was on the video compression. Users have to first record a video then send it after they finish recording the message so communication is not in real-time. The system interface has two video windows where one window is used to display the video being recorded and the other displays the video messages that are received.

To control the interface users have to use a mouse to select the desired options. The following screenshot (Figure 4.1) shows the interface of the system used in the focus group study.

Figure 4.1: A screen shot of the asynchronous video communication application.

"On the left hand side the user can see a video of themselves and on the right hand side they can play the video messages they have received. The screen shot was taken from the system designed and implemented by Ma and Tucker (2007)."

The second system used is a synchronous or real-time video communication system called Kiara (Yi and Tucker, 2008). The real-time system is also computer based but users can communicate with each other in real time. The video messages are streamed to the other user instantaneously and are therefore not stored first and then sent after recording. Kiara is a SIP-based communication tool that supports synchronous communication of text, voice and video. Only the synchronous video communication was used in this study. The Kiara interface has a big video window which shows the second user and there is a smaller video window within the bigger window. The smaller video window is at the bottom right corner within the bigger window. The video is displayed as a picture within a picture. The following screenshot (Figure 4.2) shows the Kiara interface with the two video windows.

Figure 4.2: A screen shot of the Kiara video communication application.

"The bigger video window shows the video message from the other user and the smaller video window on the bottom right corner shows the video of oneself. The screen shot was taken from the system designed and implemented by Yi and Tucker (2008)."

Six Deaf users from DCCT participated in the study and four users were asked to use the current asynchronous video communication system for about five minutes. The other two Deaf users where asked to use synchronous video communication system for five minutes. Afterwards the two groups where then asked to switch and the users who were using the asynchronous system switched to synchronous system and vice versa. The interface of the systems used for this study is not gesture based so users had to use a mouse or keyboard to interact with the interface.

Users were able to ask questions and also provide comments when they were using the two systems. After using both systems users were asked if they thought converting the current systems to using hand gestures for controlling the interface instead of using a mouse and keyboard would improve the usability of the system. The types of gestures were supposed to be a sign that might mean something in sign language or it could have merely been any hand gesture. A certain predetermined gestures were going to correspond to a certain function in the video communication system. They also had to propose any hand gestures they thought would be appropriate for the different functions in the system. Users were also requested to give suggestions on what could be improved on the current system or suggest any functionality that could be added to the system.

4.1.2 User Feedback

The users could not reach a consensus on whether or not using a gesture interface was going to help make the system more usable. Some users thought that a gesture based interface would make the system usable but others did not consider it a good idea especially on the synchronous (real-time) system since there is not much to do after the conversation has started. The option available on the real-time system is ending a conversation once you start a conversation. On the asynchronous system they thought the gestures would make the interface easier to use considering they have to keep on recording, sending video messages and also playing incoming video messages repeatedly.

Type of Gestures

The use of gestures to control the system raised issues such as what kind of gestures ought to be used between simple hand gestures like just raising one’s hand as compared to a proper sign that actually meant something in sign language. Users indicated they would prefer to use a sign that has a meaning but the system should also be able to distinguish between a sign in a conversation and a sign which is a command to the system. For example to stop recording a video they would prefer to use a sign or gesture that means stop. The system should distinguish between a user using the gesture for *stop* in a conversation and when the user wants to stop recording the video. If the gestures used to control the system formed part of the sign language therefore they can be part of a conversation and the system should have a certain designated area within the video window where these gestures are detected. All the other areas on the screen are used for conversation. Gestures used to control the system are only considered if they are within the designated area and if outside this area they are regarded as part of the conservation.

A suggestion to invent new signs that are not part of the sign language for the different functions was made but the participants could not agree on the signs. This posed a bigger problem as some users did not think it was a good idea. They did not think introducing new signs was a better alternative because it involves learning the new signs that are used to control the system. This would necessitate an extra step of training the users to learn the new signs before they can use the interface and it also pose a problem of users forgetting the signs while they are using the interface. For example, if a user forgets the sign for stopping or sending the video message they cannot stop the video recording once it is started.

Touch Screen

The use of a touch screen was among some of the suggestions made and all users thought it would be much easier to have a touch screen as users just have to select the option they wish on the screen without having to use a mouse. The touch screen was considered to be simple to use and all the users knew exactly how it worked. This did not actually solve all the problems seeing as whilst using the video communication system, users have to stop signing to each other and move forward to touch the actual screen. They have to move forward to touch the screen due to the fact that they usually have to sit at a reasonable distance from the camera which is mounted on top of the screen.

When using a touch screen users still have to move forward like before when they were still using a mouse, however the introduction of gestures eliminates the need for that as they can just continue on signing. With a gesture interface all that is required is for the users to just switch between hand gestures for the conversation and gestures used to control the system without the need to move from where they are sitting. Using a touch screen was also not a good alternative when the interface had to be deployed on the mobile phone because not all phones have a touch screen. The design of the interface needs to be simple enough to be implemented on both the computer and the mobile phone. There was no touch screen monitor available so implementing a touch screen interface was not a practical solution. A touch screen interface will not be accessible to the Deaf users due to lack of equipment.

System Interface Improvements

Users recommended some improvements that should be implemented on the current interface. The labels on the system should be changed to words that are easy to understand. The current system used words like *“Capture”* for starting to record a video and *“Transmit”* to send the video. Words like *“Record”* and *“Send”* were suggested as alternatives. The system made use of buttons with labels only and there were no icons, so users felt the use of icons will make the interface more user-friendly and it would be much easier to determine what the different functions are used for. The system should also give feedback and notifications clearly after completing an operation. The current system displayed messages at the bottom of the interface in a small black font as a way for notifying users of new incoming messages and confirmations of sent messages. The interface should represent these notifications and confirmations in a better way that can be easily noticed by the users.

Video Improvements

The asynchronous system uses two windows for displaying the video. One window is for viewing the other user’s video message and the other window is for viewing oneself. The users suggested having one big window which shows the other user and a smaller video window in a corner within the bigger window. This would be similar to the synchronous system which uses this approach which appears as a small picture incorporated within a bigger picture. Users wanted a video within a video approach instead of having two separate video windows. They also suggested combining both the asynchronous and synchronous video communication systems into one system. The two systems should be able to stream video in real-time and also use store and forward when real-time is not possible. The system could also support text communications so that users can also communicate with hearing people thereby allowing users to switch between the three different communication methods.

4.2 Communication Survey

In this research the main focus is investigating the use of video communication for sign language. Since we were implementing video communication applications for Deaf people we had to understand how Deaf users communicated with each other. A survey was conducted with 12 Deaf users who were participating in the various user studies. The survey was conducted as an ongoing process throughout the different user studies. In the different user studies conducted there were a mixture of new people and people who had participated in previous user studies. The new Deaf participants were asked to answer the questions in the survey only once so if a person had participated in the survey before they did not have to answer any survey questions in future user studies. The survey was conducted during the design and evaluation of the mobile phone prototypes.

The number of people who participated in this survey was small hence the results obtained from this study are not statistically significant. Even though the results are not statistically significant they provide insight on the Deaf users who participated in the design and evaluation of the prototypes. Since the results are not statistically significant, they cannot be used to make statements about the Deaf population. However they are relevant and useful in the context of this project.

4.2.1 Mobile Phone Ownership

The video communication application has to be accessible to the target users. Therefore it should be implemented on a device that is available to the users. All the 12 users who participated in the survey owned a mobile phone. The mobile phones brands were as follows: 7 Nokia phones, 3 Samsung phones, 1 HTC phone and 1 LG phone. There were 6 people who owned computers but only 5 of these 6 people's computers were in good working condition. This was due to the fact that one of them had a computer which was not functional at the time. There were 2 people who had a computer that had internet access. The other 4 had computers but did not have an internet connection, therefore they cannot use their computers for any video communication that requires an internet connection. A video communication application implemented on the mobile device platform would be accessible to more Deaf people because they all owned mobile phones.

The mobile phone seems to be more favoured for video communication but owning a mobile phone is not enough. The phone has to have a camera and be able to record and play video. There were 4 people who had a phone that did not have a camera and the other 8 had phones with one or two cameras. The 8 people with phones containing a camera were split into two groups of users who had phones with two video cameras and 6 with phones with just one video camera. On the phones with two cameras the video camera with better video quality was at the back of the phone. The phones with a video camera have a potential to be used for video communication. The phones with two video cameras are capable of making video calls which use the camera on the front of the phone. Video calls can be made from these phones without installing any additional applications provided it is supported by the mobile phone network in use.

4.2.2 Mobile Phone Usage

The users had to answer questions about how they use the mobile phones for communication purposes. There were questions about text based communication and video communication. The most popular method of communication was SMS because all the 12 users used SMS on their phones. There was one user who also used MMS on their phone. In an MMS users can attach a picture or a small video clip. None of the users had ever used video call before to communicate with each other. This might be due to the fact that only two out of the twelve people had phones that had the capability of making video calls. There were four people who used other text based communication like instant messengers, facebook[1] and Gmail[2] (Google mail). The instant messengers used were 2go[3], Mxit[4] and Fring[5]. All these instant messengers' support text based communication but Fring also support video calls but none of the users used the video call feature.

None of the users had ever communicated through sign language on their mobile phones. Most of the phones were low-end phones so implementing the gesture based interface on these devices is a challenging task. The users also mentioned that they wanted video communication on their phones but it had to work on the low end phones since they are unable to afford the more expensive phones. Expensive phones have more memory and more processing power which makes it easier to provide acceptable video quality. The final video prototype has to work on mobile phones with limited resources and provide video that is suitable for sign language communication.

---

[1] Facebook is a social network - http://www.facebook.com
[2] Gmail is an email service by Google - http://www.gmail.com
[3] 2go is an instant messenger - http://www.2go.co.za
[4] Mxit is an instant messenger - http://www.mxitlifestyle.com
[5] Fring is an instant messenger - http://www.fring.com

4.3 Chapter Summary

In this chapter we discussed the steps followed in order to understand what the users expected from the video communication applications. A focus group study was conducted with Deaf users using two systems. A real-time video communication system and an asynchronous video communication were evaluated by the users. These two systems were used as a starting point in order to determine what features should go into the gesture based interface. The feedback obtained included discussions about what gestures should be used and what improvements should be made. A survey was also conducted with 12 Deaf users in order to gain more understanding about Deaf communication on mobile phones.

5. Design and Implementation

In this chapter we discuss the process followed when going through the iterative cycles. The cycles involved using various prototypes on different platforms. A computer prototype was developed first and evaluated with Deaf users. The second prototype was developed for the mobile platform and the third prototype used both the computer and the mobile phone. Section 5.2 discusses the design of the first computer-based prototype, section 5.3 discusses the mobile prototype design and section 5.4 discusses the design of the prototype that uses both the computer and mobile phone.

5.1 Design

The goal of this project is to implement a gesture based interface that uses a mobile phone. The design of this interface needs to be usable and easily understood by the target group. The interface ought to be simple enough so that users can interact with the system without requiring excessive training or manuals. There is substantial work done on how to design usable interfaces and some of these works include Norman's (1988) four principles for good practice that promote ease of use:

1. *Ensure a high degree of visibility.*
2. *Provide feedback.*
3. *Present a good conceptual model.*
4. *Offer good mappings.*

There are other principles and rules of design that were considered when developing the gesture based interface but other notable ones are Ben Shneiderman's eight golden rules of interface design (Shneiderman, 1998).

1. *Strive for consistency.*
2. *Enable frequent users to use shortcuts.*
3. *Offer informative feedback.*
4. *Design dialogs to yield closure.*
5. *Strive to prevent errors and help users to recover from errors quickly.*
6. *Allow 'undo'.*
7. *Make users feel they are in control of a responsive system.*
8. *Reduce short-term memory load.*

The rules mentioned above were considered when designing the interface for the video communication application. These rules are supposed to improve the usability of the interface so that the users can achieve their goals without experiencing any difficulty when interacting with the interface.

5.1.1 System Requirements

When signing, users have to sit at a reasonable distance from the camera. Deaf people using an asynchronous video communication application often have to move forward to use a mouse or a keyboard in order to record and stop a video. By implementing this gesture based interface the users are able to control the application by using hand gestures without having to move forward to use a mouse or keyboard.

Sign Language Video

Sign language communication on a mobile phone needs a device that has a video camera that can be used to record the video. When Deaf users are signing they need to use both hands so the mobile phone has to be placed on a stand. Most current phones have cameras on the back and the screen on the front of the phone. This makes it easier for people to see what is being recorded but is unsuitable for the purposes of recording oneself which is what is needed in sign language conversation, Cavender et al (2006) also raised this issue**.** Some phones have two cameras where one camera is on the same side as the screen and the other one faces away from the screen but the camera with better video quality is in most cases always on the reverse side of the device, facing away from the screen. This poses a potential problem for Deaf people because a sign language conversation needs crisp video quality.

The solution for the camera and the screen being on opposite sides of the phone is to connect a mobile phone to a Television (TV) for display purposes so that the camera at the back of the phone with better video quality can be used to record the video. Some new phones like the Nokia N96 (GSMArena.com 2008c) and Samsung i8510 Innov8 (GSMArena.com 2008d) have a TV out port which can be used to connect the phone to a TV set. In the case where the camera with better video quality is on same side as the screen there is no need to connect to a TV, the phone’s screen can be used instead. The following Figure 5.1 shows a Nokia N96 connected to a TV.

Figure 5.1: A mobile phone connected to a television.

"The mobile phone (The Nokia N96) is connected to a TV using the TV out cable. The TV is displaying the mobile phone's menu. This makes it possible to use the camera at the back of the phone and at the same time be able to see what is on the mobile phone screen"

In Figure 5.1 the mobile phone's camera is facing the user and the TV is used for display purposes. This setup makes it possible to record oneself using the mobile phone's back camera which has better video quality. In this setup the mobile phone's keypad is facing away from the user so the gesture interface can be used to control the video communication application.

Store and forward Video Architecture

A sign language conversation needs good video quality and current real-time video available on mobile phones is unsuitable. A store and forward video architecture was chosen for the gesture based interface because mobile phones do not have enough bandwidth for a real-time video communication. Available real-time video on mobile phones is not fitting for sign language video due to the poor video quality and as result some of the important sign language features like facial expressions are lost. Ma and Tucker (2007) states that store and forward makes it easier to preserve video quality since there is enough time to apply better compression methods that will preserve sign language features. The compression also lowers data costs for video transfers by decreasing video file size. This makes video communication possible even when there is low bandwidth. Low bandwidth is not suitable for real-time video communication. In store and forward, higher video frame rates can be achieved which reduces blurriness that can occur if a user is signing fast.

Video Messages

In store and forward video communication applications, users should be able to record a video message and also be able to send the video messages. The user should have an option to cancel the recording process if they wish to do so. Play back of video messages that are received should be possible too. The video quality should be clear enough for users to see all the hand gestures and facial expressions.

Interface Layout and Feedback

To design a usable system the interface should use concepts that users are familiar with. This makes it easier for the users to relate to the interface. The interface should also be clearly labeled and avoid any ambiguities. This can be achieved by using words that are simple to comprehend for labels and combining them with icons. For users to keep track of their actions and kept up to date the system should give feedback and notifications clearly when operations are completed. The interface should represent these notifications and confirmations in a manner that can easily be noticed by the users.

Error handling should be implemented. Whenever an error occurs the application should have error prevention and error recovery measures. The application should be able to recover from any erroneous input and should give users an option to recover from any errors that may occur. The error messages need to be clear and informative so that users will know exactly how to fix the problems they experience.

Gesture Recognition

The gesture based interface uses simple hand gestures which are not necessarily part of sign language. The gestures used must be uncomplicated and easy to remember in order to enhance the interaction of the users with the system. Nielsen et al. (2003) lists the following principles in ergonomics to building a good gesture interface:

1. *Avoid outer position.*
2. *Relax muscles.*
3. *Relaxed neutral position is in the middle between outer positions.*
4. *Avoid repetition.*
5. *Avoid staying in static position.*
6. *Avoid internal and external force on joints that may stop body fluids.*

All these principles have to be taken into consideration when designing a gesture interface. The interface for this project was deployed on a mobile phone so the image processing has to be kept to a minimum. This can be achieved by using certain predetermined sections of the screen to detect hand gestures that will be used as commands in the application. Gestures used to control the system are only considered if they are within the designated area whereas if they are not they are regarded as part of the conservation. The image processing needed for the recognition of gestures is reduced to only the predetermined section in the video instead of processing the whole video.

5.1.2 Interface Design

The interface was designed using a user centred design approach in order to obtain feedback and input from the users. A user centred design approach allowed users to have a say on the design aspects of the prototypes. To find out the requirements needed for a usable interface user studies were conducted with a focus group of Deaf people and sign language users from the Deaf Community of Cape Town. The process involved going to DCCT weekly in order to gain more knowledge of Deaf culture. The requirements gathering stage was very useful in acquiring information on how the design for the gesture interface should look like. These studies made it easier to determine the most significant features that are required for the interface as discussed already in section 5.1.1.

The interface was designed using three prototypes. These prototypes were designed using an iterative approach with the first prototype based on the computer, the second prototype based on the mobile phone and the third prototype using both the computer and the mobile phone. The computer prototype was implemented in order to test the concept and also to get feedback from the Deaf community promptly. The design and evaluation of the three prototypes are explained in detail in the next sections of this chapter.

5.2 Computer Prototype Design

Computer based prototypes are easy to develop and test quickly as compared to mobile phone prototypes. This makes it possible to test the concept and evaluate the design in its early stages. In Chapter 4 users suggested a touch-screen interface as a possible design alternative as they were all familiar with how the touch-screen works. The touch-screen design was later considered cumbersome as it did not solve the problem of having to move forward to select an option. Another problem with the touch-screen was that on most mobile phones it is on the opposite side of the preferred video camera.

The gesture based interface was designed to be similar to a touch screen in the sense that users just have to move their hand to a certain marked area displayed on the screen. Instead of touching the screen at the marked area they just have to move their hand in front of the camera. Norman (1988) suggests that the ease of use can be promoted by presenting a good conceptual model. The gesture interface achieves that by emulating the interaction method similar to that of touch screens. The background of the screen displays a video of the user and the marked areas so that when a user moves their hand they can see its corresponding position on the screen at real time. Once their hand is on the desired marked area they have to hold it up for a second until it is detected. This is implemented in order to cater for the situation where a user accidentally put their hand on the marked area while signing. If a hand is moved too swiftly the gesture is ignored as it is assumed that the user is signing and they accidentally moved their hand to the marked area. Figure 5.2 shows the interface and all the marked areas which are enclosed in a blue rectangle at the top, left and right side of the screen.

Figure 5.2: View of the Interface showing all the marked areas at the (Top, Left and Right).

In Figure 5.2, all the marked areas are displayed in blue rectangles. All the options that are available have an icon and text which describes the function that will be activated if a hand is moved to that position. In Figure 5.2 there is only one option available which is *record* at the top of the screen. While the user is recording a video message the only marked areas that can be activated are for sending a video message and for cancelling the recording process. Figure 5.3 shows these two options that are available if the user selects *record*.

Figure 5.3: View of the Interface while recording.

In Figure 5.3, the two options that are available are *cancel* which cancels the recording if it is selected and *send* which sends the video to the other user. All the marked areas are on the sides of the screen so as to give the user a larger area at the centre of the screen for signing purposes. Users have to sign at the centre of the displayed video for conversation purposes and use hand gestures on the edges in order to select the available options. When a user selects an option a confirmation message is displayed. This is in accordance with both Norman's (1988) and Shneiderman's (1998) principles which state that the users have to be provided with informative feedback. In Figure 5.3 above the message at the bottom of the screen is informing the user that the application is recording a message. In Figure 5.4 shows a user selecting an option to start recording a video.

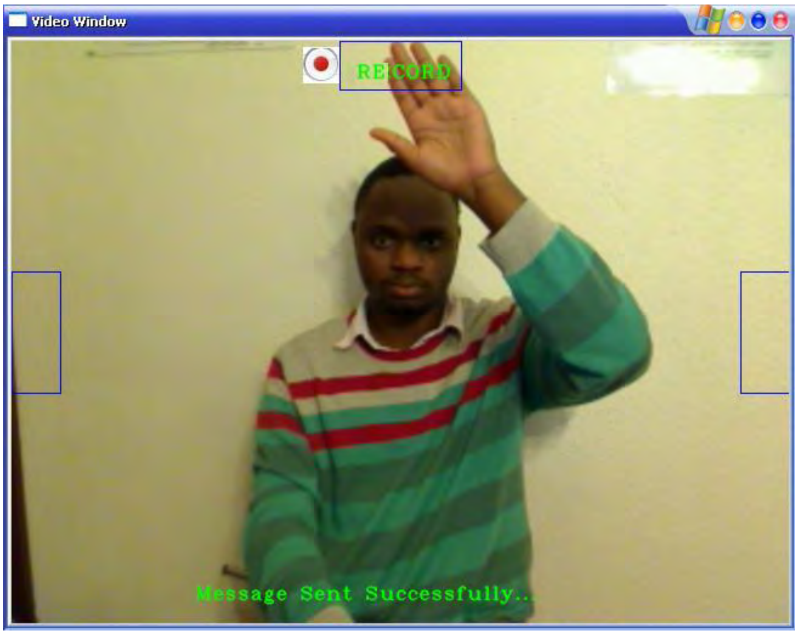

Figure 5.4: View of Interface while a user is selecting *record* **at the top.**

In order to enable the user to watch video messages and still be able to control the system simultaneously, the centre of the screen is used for displaying video messages that are received. The marked areas on the sides always display the footage of the user watching the video. This makes it easier for the user to cancel or reply without having to use a mouse or keyboard to stop the video playing. Figure 5.5 shows a user watching a video message they have received.

Figure 5.5: View of the Interface while a user is watching a video message.

In Figure 5.5 a user can either stop the video from playing or record a reply. If the user does not select any option, the video plays until it is finished and the small video window will disappear so that the application can return to idle state where the user can see a video of themselves at real-time, as in Figure 5.2.

5.2.1 Computer Prototype Evaluation

A user study was conducted using a prototype implemented on a computer in order to get feedback from the users before the system was implemented on the mobile phone. The study was conducted in order to evaluate the usability of the gesture interface. Users used the application by simply moving their hands in front of the camera to select the menu options in the application.

The difficulty experienced to perform all the basic operations of the application were measured using a questionnaire. The users had to rate their experience using a Likert scale ranging from 1 to 7 with 1 being very difficult and 7 being very easy. The users were also observed while using the interface to spot if they had any problems.

Participants

The study was conducted with Deaf users from the Deaf Community of Cape Town. Seven Deaf users were considered in testing the usability of the system and they were split into two groups. There were only six computers available at the Bastion for the Deaf so the users had to be split into two groups. We used only four computers per group therefore there was a maximum of four users per group. The first group consisted of four new users. The second group had three new users and one user from the previous session. One user was asked to participate twice so that the third user in the second group can have a partner. The system needs people to be in pairs so that they can communicate with each other.

Method and Procedure

Communication with the users was through a professional interpreter. The users were given five minutes first to learn how to interact with the interface and after they were comfortable with using the system they were given 20 minutes to communicate with each other in pairs. Users were given the tasks that had to be completed while they were chatting to each other. These tasks included:

1. *Sending a video message,*
2. *Reading a video message,*
3. *Playing a video message they received,*
4. *Cancelling a recording of a video message,*
5. *Stopping a video message while it's playing*.

The users were asked to rate the difficulty or ease experienced when performing each task they were set. They were asked to comment on the usability of the system. The questionnaire used had all the tasks mentioned above and users had to rate them. There was also a discussion with all the participants about how to improve the usability of the prototype and future systems.

Results

The users from the two groups were asked to rate the difficulty or ease experienced while performing each task on a Likert scale with 1 being very difficult and 7 being very easy. Most of the users completed the tasks given with ease and were able to use the system on their own without any assistance.

Table 2 below shows the results obtained from the evaluation with the individual user's feedback in each column.

Tasks	Users						
	A	***B***	***C***	***D***	***E***	***F***	***G***
Recording a video message	3	7	7	4	7	7	7
Sending a video message	7	7	7	4	7	7	7
Playing a video message they received	3	7	6	3	7	7	7
Cancelling a recording of a video message	7	7	6	3	7	7	7
Stopping a playing video message	-	7	6	4	7	7	7

Table 2: Results from questionnaire.

Two of the users (*user **A** and user **D***) struggled to use the gesture based interface and based on the feedback given on the table above they both had the most difficulty in using the system.

Discussion of Results

Most participants were able to use the interface without any difficulty and without any assistance from the researcher. In Table 2 only two participants, user A and user D rated some of the tasks to be a little difficult. The two participants gave the tasks a value of 3 on a scale of 1 to 7. User D gave the rest of the tasks a rating value of 4 which is half way from very easy and very difficult. The user seemed to have had problems with interface and based on the observations it was noticed that the user often tried to touch the screen to select an option instead of just moving the hand in front of the camera. It took a while for this user to familiarize themselves with the gesture based interface. The other five were satisfied with the interface and managed to learn how to use it quite fast. All the users had comments on how some aspects of the interface and the system as a whole could be improved.

The users all welcomed the concept and were very interested in how it would work on a mobile phone because while they all had mobile phones, not all of them had computers. The implementation of this interface on a mobile phone seems to be more practical than on a computer since most Deaf people have access to mobile phones.

Usability

The users felt that the system was very straightforward to learn except for one user who struggled a little in the beginning but the same user was really keen on the fact that they did not have to move back and forth to use the mouse. The menu options in the system were kept to a minimum and the users were quite pleased with this. They felt that too much text confuses them, so they advised on using more icons than text.

The messages that are displayed when a video message was sent or deleted successfully were kept short and simple but the users wanted icons in addition to the text. The use of icons seemed to be preferred over written text this may be due to the low literacy levels among the Deaf. One user also pointed out that there was no confirmation dialog if a user chooses to cancel a recording of a video message. This meant that if they accidentally move their hand over to cancel they would lose all the information in the video, so they wanted a confirmation before the message was deleted.

Video Messages

The users were satisfied with the video quality and the flashing icons and notifications that were displayed when a new video message was received. They suggested adding a pop up message so that the notifications will still be visible in case the application was minimized. They also liked the way the video was displayed on the computer but were concerned about how the video will be displayed on the mobile phone since they have smaller displays. Connecting the mobile phone to TV set would solve the display problem but the users wanted an option to use the mobile phone without any connections to a TV.

The prototype did not have any options to rewind or fast forward the video, the addition of this feature in future prototypes was suggested. Other suggestions included having offline storage in case messages are sent to an offline user. They wanted the user to be notified through an SMS about the new message. The prototype discarded all messages after they were played but the users wanted all messages received to be saved. Storing all these messages will pose a potential storage problem on mobile phones since they have small amounts of memory.

Multiple Users

The prototype allowed only two users to communicate in pairs so once connected to each other they were not able to connect to other users unless if they disconnect first. This meant you could only chat with one user at a time so the users suggested that there should be an option to connect to multiple users. They should be able to initiate communication with multiple users and be able to switch between them just like in most text chatting (instant messenger) applications. They also wanted an option that enables them to send broadcast messages which can be sent to all the users they are connected to. This will make group chat possible for Deaf users.

5.3 Mobile Phone Design

The second prototype was designed using the feedback obtained from the users during evaluation of the computer based prototype. The prototype was designed and implemented entirely for the mobile device platform. The users wanted the interface to be implemented entirely on the mobile device since most of them do not have a computer but they have mobile phones. Implementing the video communication application on the mobile device makes it possible to reach a higher number of the target group users because they have access to the mobile phone technology. The following diagram shows the overview of how the mobile video communication application works.

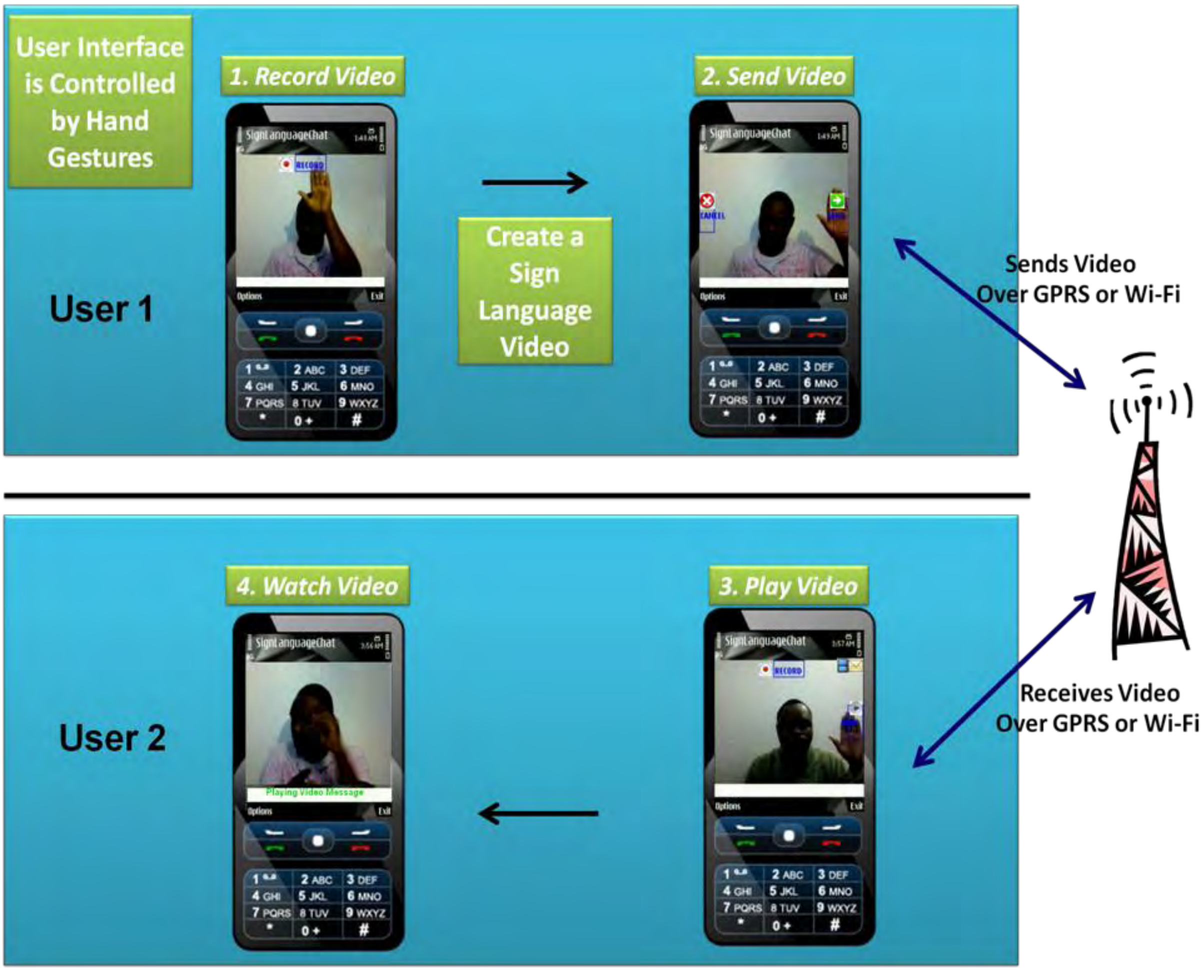

Figure 5.6: Mobile prototype overview.

The mobile prototype interface looks similar to the computer prototype interface but some of the design aspects had to be revised.

5.3.1 Changes Made

Computers have higher processing power and more memory as compared to mobile devices. Therefore some of the design aspects could not be implemented on a mobile device since the resources needed were not available on this platform. On the computer prototype users had an option to stop a video while it is playing by either selecting *stop* or by selecting *record* in order to reply to the message. Selecting *stop* or *record* while the video is playing resulted in the video playback being stopped. The following diagram shows the screenshot of the computer prototype playing a video on the left and the mobile phone prototype playing a video on the right.

Computer prototype playing video. **Mobile phone prototype playing video.**

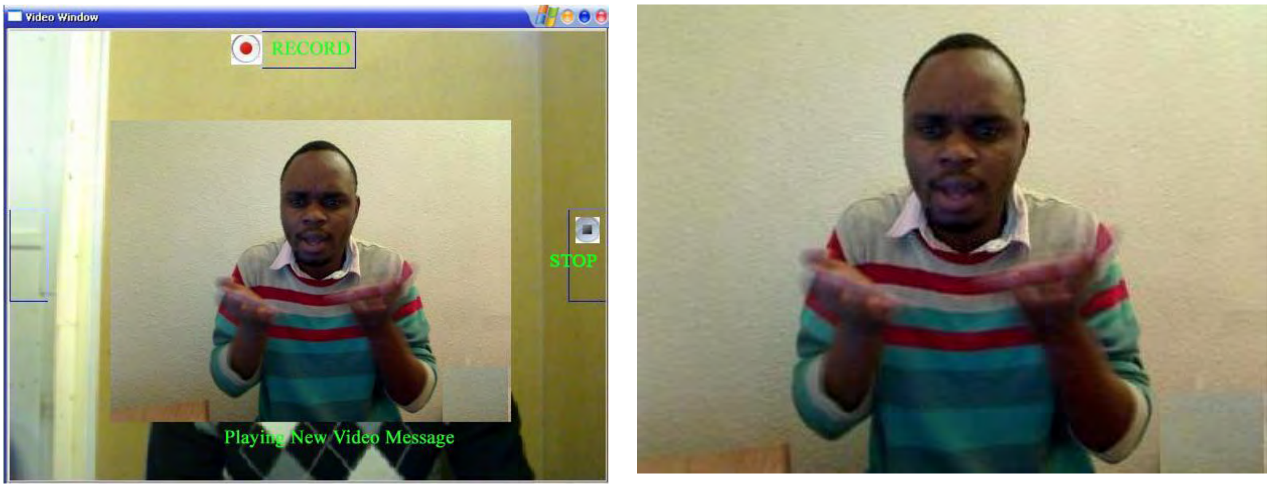

Figure 5.7: Difference between playing video in computer prototype and mobile prototype.

"The video on the computer prototype plays as a video within a video whereas in the mobile prototype only one video is displayed. On the computer the video message is played on the smaller window and users can see themselves on the bigger video window."

On the mobile device platform this was not possible because the device did not have the function needed to obtain individual frames from for the playing video. The video playing can only be rendered directly to the video window. This makes it impossible to insert the playing video frame into the video frame obtained from the camera. As a result the picture in picture effect as in the computer prototype could not be implemented. Only one video can be displayed at a time on the mobile phone. It is not possible to display both the user's video and the incoming video messages. The mobile prototype had to suspend previewing the user's video to play video messages and then resume previewing afterwards. When playing a video message the users are not able to see themselves so this made it harder to *stop* the video because the user could not select the stop menu option. The video was displayed in full screen on the mobile phone and the user has to watch the whole video before returning back to seeing themselves.

The mobile phone prototype had a confirmation screen for cancelling the recording of a video message. In the computer prototype if a user selected *cancel* while recording the video message, the recording was cancelled immediately. This posed a problem when users accidentally selected cancel while they were recording a video message. This means if a user selected cancel by mistake the video message was lost and they had to start all over again. The mobile phone application solved this problem by adding a confirmation screen which asks the user if they are sure they want to cancel recording the video message. The user can then select *yes* to cancel or select *no* if they want to continue recording the video message. Shneiderman's (1998) golden rules of interface design states that dialogs should be designed to yield closure. The rules also state that errors should be prevented and users should be able to recover from errors promptly. The confirmation screen provides a way for the users to recover should they accidentally select the wrong option. The following screenshots show a user selecting *cancel* on the left and the confirmation screen is shown on the right.

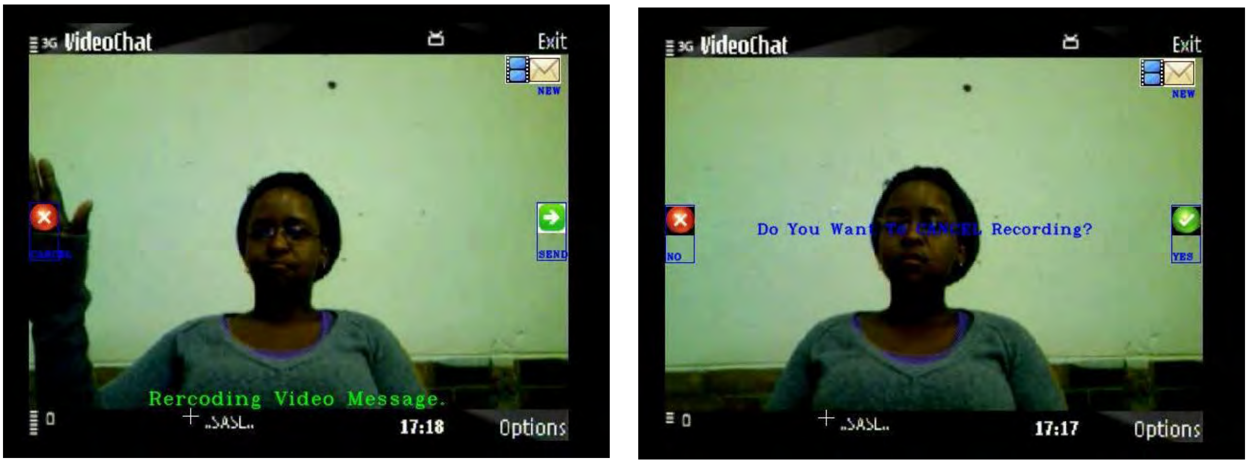

Figure 5.8: Cancelling confirmation screens on mobile device.

"The screen sequence when a user selects cancel while recording a video message. The screenshot on the left shows a user selecting cancel and the screenshot on the right shows the cancel confirmation screen. The question asking the user if they want to cancel recording is at the centre of the screenshot on the right."

5.3.2 Mobile Phone Design Problems

Some design and performance problems were encountered when the gesture interface was implemented on the mobile device platform. The gesture interface was implemented and tested on several Symbian S60 mobile devices. The devices used include the Nokia 5800 XpressMusic (GSMArena.com, 2008a), Nokia 6120 Classic (GSMArena.com, 2008b), Nokia N96 (GSMArena.com, 2008c) and also the Samsung i8510 Innov8 (GSMArena.com, 2008d).

Problems on Samsung Innov8

The video communication application with the gesture interface requires the device in use to have the capability to record and encode video. When the mobile application was installed on the Samsung Innov8 device the video frames from the camera where rendered successfully on the screen and the video was visible on the preview window. All the icons and menu items were also rendered successfully and displayed on the preview window. However when the video recording process was initiated the application did not record the video. The device initiated the recording process and the video preview was halted. The process of allocating the resources needed to record a video message was never completed so the application remains in a pause state until the application is manually closed. This problem made it difficult to use this mobile device for video communication as no video message could be recorded.

The mobile device did not support mirror images of the camera frames therefore the images that were obtained from the video camera were not mirrored. If the user sitting in front of the camera moves their hand to the right on the screen the hand is displayed moving to the left. For the gesture interface to work properly the images need to be mirrored, this means they should be displayed on the screen as though the user was standing in front of a mirror. For example if the menu option for sending a video message displayed on the screen is on the user's right hand side the user should be able to only move their hand to the right. However if the images are not mirrored this action is not displayed as moving to the right hand side it is displayed as moving to the left hand side. This effect is caused because normally cameras preserve the direction and when the image is displayed it will be moving to the right hand side from the screen perspective but from the user's perspective it is the left hand side.

This is similar to when two people are facing each other and they both stretch out their right hands, from one person's perspective the other person appears to be stretching their hand in the opposite direction which is to the left. If the same person stretches their right hand in front of the mirror the reflection on the mirror will have their hand stretched out in the same direction. The mirror effect is what is required by the gesture interface so that users can select the application's menu options. The mirror image feature returned a warning message saying the feature is not supported on the device whenever it was called from the application. This made the Samsung Innov8 unsuitable for the gesture based interface video communication application as it did not support mirror images and was unable to record video messages.

<u>Problems on Nokia 6120 Classic</u>

The Nokia 6120 Classic was also used to test the application. This mobile device's camera has the lowest resolution when compared to the other phones used for testing. This mobile phone had a 2 megapixel camera. The mobile device was able to render the camera frames to the screen without any problems but when the icons and menu text were loaded directly there was flickering on the screen. The application had to be modified to use double buffering. The icons and menu texts are rendered to the back buffer and any notifications that has to be displayed was also rendered to the back buffer. The back buffer is then displayed on the screen only after everything has been delivered to it. This helped in removing the flickering when video from the mobile device's camera was previewed.

Furthermore this mobile device did not support mirror images of the camera frames so the mirror images had to be created within the application. The amount of time needed to create these mirror images and also load icons and draw all the text was too long. The was a noticeable delay of more than 1 second consequently by the time an image was displayed on the screen the user had already changed the position of their hands. For example if a user put up their hand they could only see their hand being raised on the screen a second later. This made it hard to use the gesture interface as the users were not able to see their hand movements in real-time. There were also other delays whenever recording was initiated or video playback was initiated. The delay was due to the amount of time required by the device in order to allocate the resources needed for the process being started. The lag experienced when mirror images are being created made the mobile device unsuitable for the gesture interface video communication application.

Problems on Nokia N96

The application implemented and tested on the Samsung Innov8 was also tested on the Nokia N96. The modifications made for using the double buffering were not necessary on the Nokia N96. The Nokia N96 supported mirror images therefore the images from the video camera were rendered directly into the screen without any modifications. The mobile device was also capable of rendering menu icons and text directly to the screen without any flickering on the display. Video recording was also initiated without any problems and different encoding levels were also available so the application could record video with high quality or low quality depending on the video encoding scheme selected.

There was a noticeable 2 seconds delay when user selected record. The delay was a result of the time needed by the mobile device to allocate resources needed for video recording. There was also a delay of over 3 seconds when the user selected the option to play a video message. Other than the delays experienced before recording and playing, both processes were completed successfully.

The mobile device was able to record, play video messages and also detect any hand gestures aimed at controlling the application. However all these could be performed without any impediment when the mobile device was not connected to a TV. If the mobile device was connected to the TV, video playback was not possible but the rest of the options such as video recording and gesture detection were performed without any problems. If a user selected play in the application when the video out cable is connected to a TV the application crashes and it is terminated. However if the cable is removed the video is played with ease. This made it difficult to use the gesture based video communication application without using a TV for displaying the video because the camera and the phone screen are on opposite sides of the phone. This makes it impossible to record oneself and also see the screen concurrently.

Even though the Nokia N96 was able to perform all the essential functions when not connected to a TV it failed to perform the very same functions when used in the setup that included the TV. The setup which includes a TV for display purpose is what is needed for the gesture based video communication application so this made the Nokia N96 unsuitable for this application.

Problems on Nokia 5800 XpressMusic

The Nokia 5800 XpressMusic mobile device was used in testing the feasibility of using a gesture interface on a mobile device. The Nokia 5800 unlike the Nokia N96 did not support mirror images therefore the application did not get the mirror images directly from the mobile device's camera. The mirror images had to be created within the application. There was flickering on the phone's display when icons and menu texts where rendered directly to the screen. The double buffering technique used in the Nokia 6120 Classic had to be implemented in order to prevent this problem. All the icons, menu text and all notifications were rendered to the back buffer and then displayed on the screen.

Video recording and video playback was performed without any problems on the mobile device. The mobile device was able to play video even when TV out cable was connected to a TV. The problems experienced when using the Nokia N96 did not occur during video playback. The mobile device could be used to record and play video messages while using the camera at the back of the phone and using a TV for display purposes.

The only problem that was encountered was the lag caused by creating the mirror images. There was a delay of more than 1 second from the time the image was obtained from the camera to the time it was displayed on the screen. This meant that when a user made a movement it was displayed on the screen a second later so the response times of the application were no longer in real-time. The 1 second lag was experienced if the gesture detection algorithm was not running. When the gesture detection algorithm was enabled and performed on all the images from camera the lag increased to over 2 seconds.

When the images were not mirrored and used as they are, the user's movements were displayed as they were happening. The only problem with this is that users will have to move their hand to the opposite direction in order to select an option. For example if a user wants to select an option on the right, they have to move their hands to the left. This complicates the process of selecting an option on the screen. This made the Nokia 5800 XpressMusic unsuitable for the gesture based interface as the camera did not support mirror images and creating mirror images did not result in real-time responses from the application.

Table 3 below summarise the problems experienced on the different mobile devices.

Mobile Phone	**Problems Experienced**
Samsung Innov8	Failed to record videos messages. The mobile phone did not support mirror images of the camera's video frames.
Nokia 6120 Classic	The mobile phone did not support mirror images of the camera's video frames.
Nokia 5800 XpressMusic	The mobile phone did not support mirror images of the camera's video frames.
Nokia N96	Failed to play video when connected to a TV.

Table 3: A summary of mobile device problems.

"The table shows the summary of the all the problems experienced when the prototype was evaluated on the different mobile phones. All the mobile phones used had problems."

The gesture based video communication application was tested on 4 different mobile devices namely the Samsung Innov8, Nokia 6120 Classic, Nokia N96 and the Nokia 5800 XpressMusic. All the mobile devices tested were not suitable for the gesture interface due to the complications experienced when using these mobile devices. These problems prevented certain critical functions of the application to be performed effectively.

5.4 Mobile Phone Simulation Design

The final prototype used both the computer and the mobile phone platform. The prototype simulated the scenario where the mobile device had enough processing power needed by the different aspects of the gesture based video communication application. To the user the application appeared as if it was running entirely on the mobile phone whereas in reality the processing was done on the computer. The mobile phone had a Symbian application running that was responsible for rendering the video frames to the phone screen.

All the image processing, recording, playing and sending of video messages was performed on the computer. The hand gesture detection was executed on the computer since the mobile phone platform was not suitable for this task as explained in the section 5.3.2. All the tasks that required more processing power than is available on the mobile device were performed on the computer. The computer made it feasible to simulate how the interface would operate when running on a phone with more processing power. The following diagram shows the overview of how the overall video communication works.

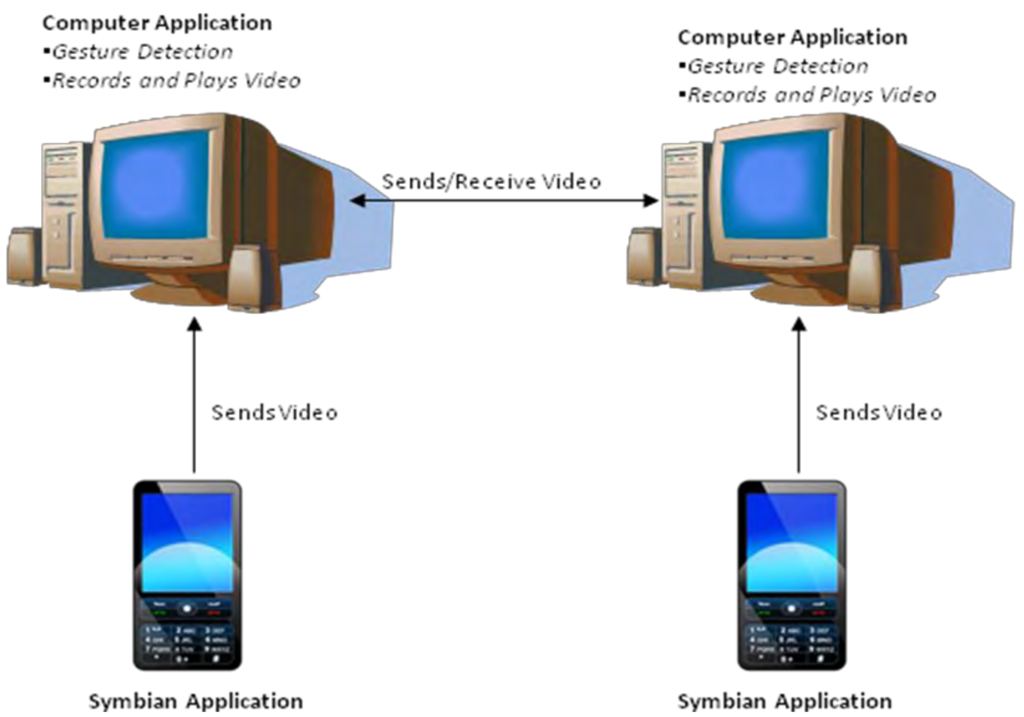

Figure 5.9: Overview of how the computer and mobile prototype work.

5.4.1 The Mobile Application

The mobile phone application was responsible for getting the video and sending the video to the computer application. The mobile phone application was implemented using Symbian C++. The Symbian application is connected to a computer using the TV out cable. The TV out cable is in charge of sending the phone's screen as a video to the computer application. On the mobile application the video was obtained directly from the phone's camera and rendered to the screen as mirror images if the function is supported by the phone. The phone that supports mirror images which is the Nokia N96 was used to run this application.

This application does not perform any image processing functions as a result the video is rendered as it is after it has been obtained from the camera. Therefore if the phone does not support mirror images for the camera video frames then the images attained from the camera are rendered without any changes. This application requires a mobile phone that supports mirror images as this is a requirement for the gesture based interface video communication application.

5.4.2 The Computer Application

This application simulates a mobile device with more processing power. The computer application acquires the video from the mobile phone application through a TV tuner card. The mobile phone is connected to the TV tuner card through a TV out cable. The video is obtained as composite input on the TV tuner card. The TV tuner card then sends the individual video frames to this application. This application does all the image processing required to detect the hand gestures aimed at controlling the application.

The notifications, icons and menu items are rendered to the screen by this application. A computer monitor can be used to display the video from this application as it runs on the computer. A TV connected to the computer can also be utilised to display the video from this application. This application is also responsible for recording, playing and transfer of all video files. The computer application also makes use of the store and forward architecture hence users have to record the video first and then send it. The video files are sent to a similar application with the same setup running on another computer.

The following screenshots (Figure 5.10) show the video that is displayed by the mobile prototype on the left and on the right the video that is displayed by the computer application after all image processing has been completed.

Mobile phone prototype view

Computer prototype view

Figure 5.10: Difference between mobile phone prototype and computer prototype view.

In Figure 5.10 the difference between the view of the mobile prototype and the computer prototype is shown. There are no icons and menu items on the left and on the right the icons and menu items are loaded to the screen by the computer prototype.

5.4.3 Task Flow

There are several steps that a user has to perform in order to successfully complete a certain task. In order to complete these tasks the user needs to carry out these steps in order, over several screens. To record and send a video message assuming they are already connected to the second user they want to communicate with, the user needs to follow the following steps:

1. Start the application.
2. Select record on the menu option.
3. Record the video message.
4. Send the video message.
5. Get confirmation for successful completion.

Figure 5.11 shows all the screens the user has to go through to perform all the steps listed above.

Figure 5.11: Recording and sending a video message.

Figure 5.12 shows all the screens the user has to go through in order to play a video message they have received.

Figure 5.12: Playing a video message.

The user has to follow the 3 steps in Figure 5.12 to play a video message. To cancel recording a video message the user had to go through all the 6 stages shown in Figure 5.13.

Figure 5.13 shows all the screens the user has to go through in order to cancel recording a video message.

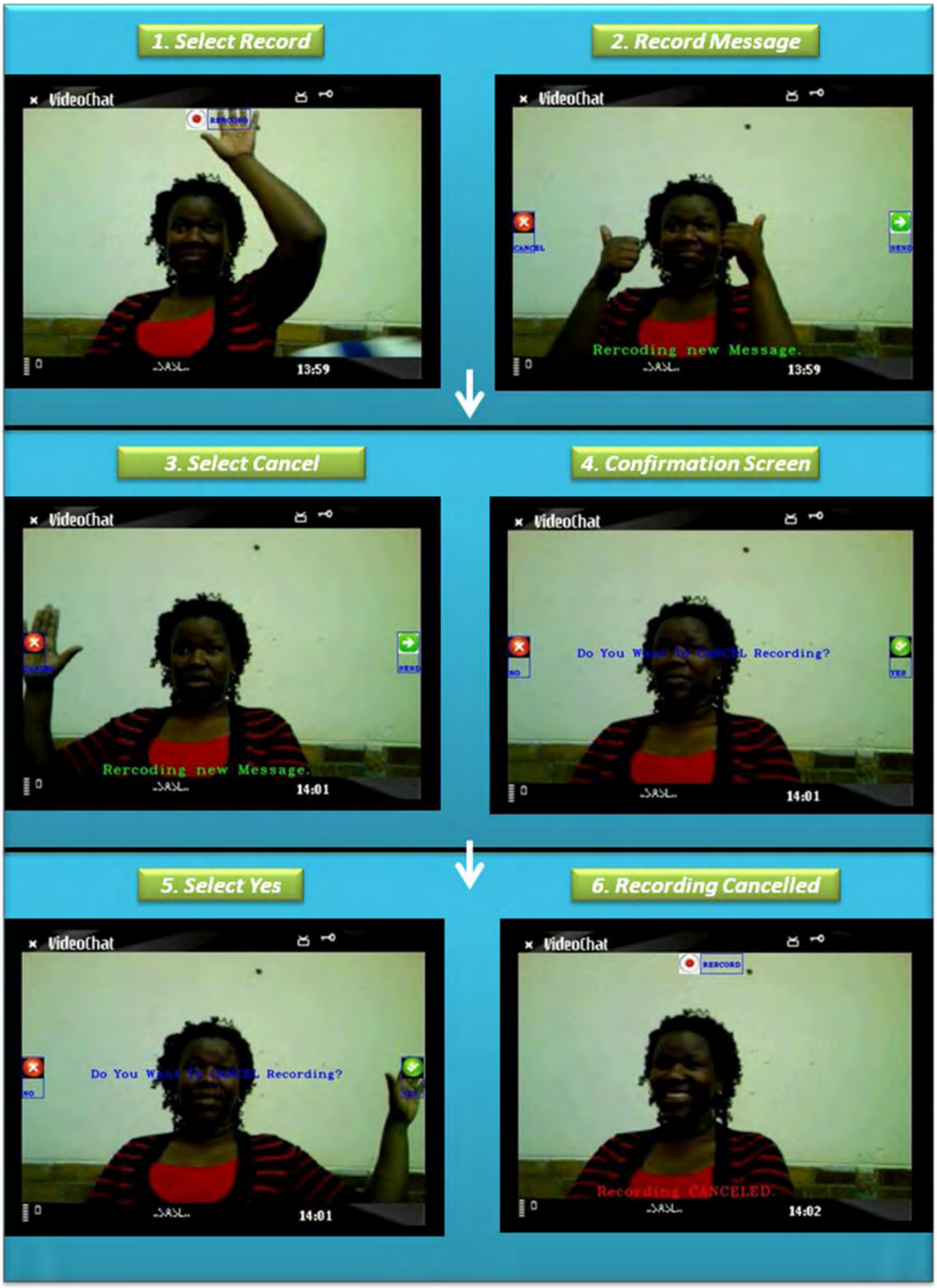

Figure 5.13: Cancelling recording a video message.

5.5 Implementation of Computer Application

The application is composed of 5 different modules which are the Display Manager module, Image Processing module, Recognition module, File Management module and the Network module. The following diagram shows the overview of the system and how the different modules interact with each other.

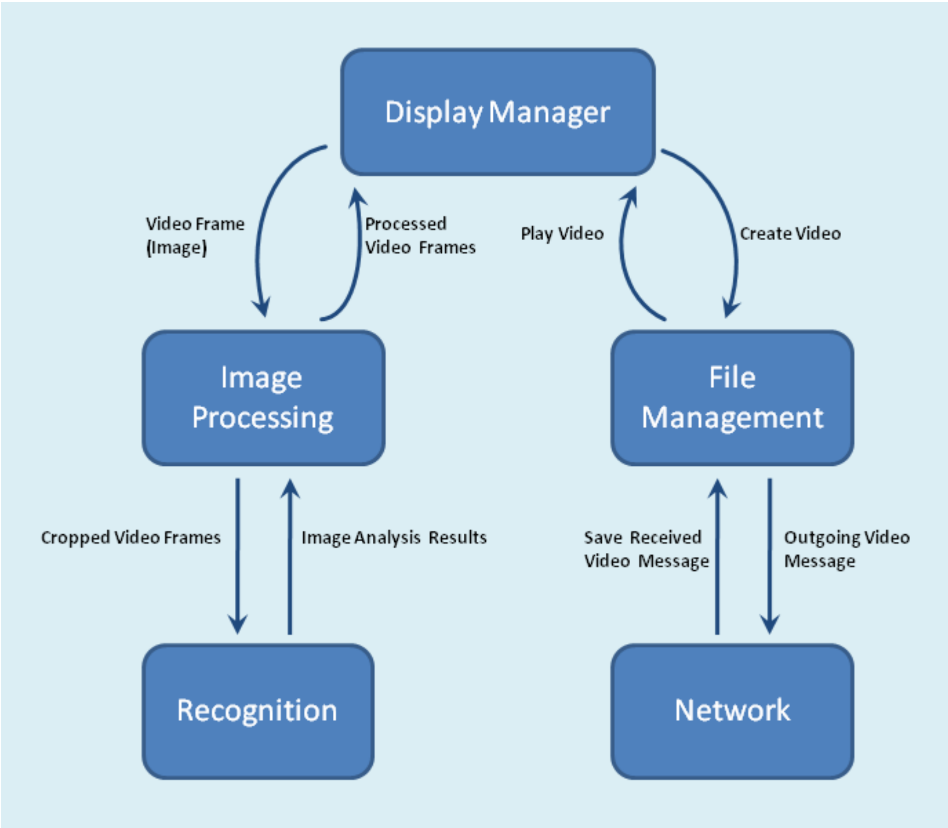

Figure 5.14: Overview of the computer application modules.

5.5.1 Display Manager Module

This is the main module and controls the flow of events in the application. It is used to manage all the other modules in the application and is responsible for displaying all the icons and videos on the phone screen. This module uses the image processor for processing the images obtained from the phone's camera in order to detect hand movements. It is also responsible for loading icons and menu items directly into the video window. So this module controls all objects that have to be added to the screen including application notification and messages that are displayed on the screen.

5.5.2 Image Processing Module

This module is used for processing the images obtained from the phone's camera. The images obtained from the camera are then cropped to get the sections of interest from the image as we only look for hand gestures at marked predetermined sections of the image. The cropped sections are then analyzed and compared to previous images to determine if there are any changes in the video being displayed. This module makes use of the Recognition module in order to detect the actual hand movement on the image's sections of interest.

When the video from the camera is being displayed not all the video frames are analyzed for performance reasons therefore only two frames per second are used. If a video has a frame rate of 30 frames per second (fps) then the images will be analyzed after every 15 frames but all the frames are displayed on the screen whether analyzed or not. A video with a frame rate of 15 fps will have the frames analyzed after every 7 frames. Analyzing only a couple of frames per second helps to keep a higher frame rate because analyzing every frame requires greater processing power than what is available on a mobile device. Although this module was implemented on the computer it had to simulate a mobile device. The approach used had to be realistic and practical enough to be implemented on a mobile device with more processing power.

5.5.3 Recognition Module

This module is used for detection or recognition of the hand on the images displayed on the screen. There are three sections of interest on the images that are obtained from the phone's camera. The three sections are at the top, left and right sections of the image. These sections are clearly marked and are assigned functions like recording a new message, playing a video message, etc. Whenever a hand is detected at the specific section the function assigned to that section is then executed. The application takes reference images and stores them in memory as soon as the camera has been started and the light settings have been adjusted. These images are then compared to newer images obtained from the camera in order to detect for any changes in the marked area. The reference images are continuously updated in order to cater for changes in the light settings or changes in the environment.

A simple algorithm is used to do the comparison between these images. The images are converted to grey scale images first then a contour match is performed to determine if the images are similar to a certain threshold. The function used to carry out the contour match was provided by the OpenCV library (OpenCV, 2010). If the difference between the results was too great it is assumed that a hand is causing this change. If the results of the contour match function are very similar then it is regarded that there is no change so no hand has been detected. The use of this function required a uniform background because when there were changes in the background, the results of the contour match function were affected.

5.5.4 File Management Module

This module is tasked with storing the video files that are received or the video files that the user records. Whenever a user wishes to record a new video this application creates a new file and also ensures that enough space exists to store the new video message. Video playback for all incoming new video messages is also managed by this module. This module makes use of the Network Module which uses the underlying low level protocol for the actual transfers of outgoing and incoming video messages.

5.5.5 Network Module

This is the underlying low level module that is used for listening to incoming connections and also for connecting to other peers. The actual transfers of files and application messages are handled by this module. It is responsible for sending the video files to another application. Furthermore the network module handles the new video messages that are being received and then notifies the File Management module which then plays the new videos. Once the clients are connected to each other the messages are sent directly from one client to another. This module can use WI-FI or any network connection available on the computer to transfer the video messages.

5.6 Chapter Summary

This chapter has presented the 3 different prototypes used in the iterative cycles. The prototype designs were discussed in detail and the process used to design the prototypes was discussed as well. Furthermore the problems experienced in each design cycle were also highlighted and the solutions for the problems were discussed as well. In this chapter we discussed the steps followed in order to come up with the design for the final prototype. One notable outcome of this chapter is the final prototype that simulates video communication on a powerful mobile phone. The final prototype uses both the computer and the mobile phone in order to achieve its goal which is to allow Deaf users to communicate through sign language. Ideally the gesture based interface is supposed to work entirely on a mobile phone but the problems on the mobile platform influenced the final design to use the computer in order to combat those anomalies. The evaluation of this prototype is discussed in the next chapter.

6. Evaluation

In this chapter the evaluation of the prototypes is explained and the results are discussed. In Chapter 5 the process followed when designing the prototypes was described. The prototypes were designed with usability in mind and although the process followed included the users in the design process this does not always result in a usable design. Evaluations make it possible to test the usability of the prototype and also to determine if the prototype achieves its goal.

6.1 System Evaluation

The system was tested in order to determine whether it could complete all the tasks correctly. System tests also included the handling of errors. The system should be able to recover from any erroneous input and should give users an option to recover from any errors that may occur. The error messages need to be clear and informative so that users know how to deal with the problem. A set of use cases covering all the menu options available on the interface was used to test if the system can perform all the functions correctly. The system was tested to see if it was able to execute the following tasks:

1. *Recording a video message.*
2. *Sending a video message.*
3. *Playing a video message they received.*
4. *Cancelling a recording of a video message.*

These are the main tasks that users need when communicating with one another. The system was able to complete all these tasks successfully. The users can communicate with each other using sign language because they can send video messages to each other. The system was successful in achieving its goal of facilitating video communication for sign language.

6.2 Heuristic Evaluation

Heuristic evaluation according to Nielsen (2005a) is defined as *"a usability engineering method for finding the usability problems in a user interface design so that they can be attended to as part of an iterative design process"*. This evaluation method involves using a small set of expert evaluators judging the interface with accordance to the usability principles called the heuristics (Nielsen and Molich, 1990; Nielsen 1994). Each expert conducts an independent evaluation and compiles a list of usability problems. The usability problems are then mapped to the heuristic being violated. The ten heuristics according to Nielsen (1994) and Nielsen (2005b) are**:**

1. *Visibility of system status.*
2. *Match between system and the real world.*
3. *User control and freedom.*
4. *Consistency and standards.*
5. *Error prevention.*
6. *Recognition rather than recall.*
7. *Flexibility and efficiency of use.*
8. *Aesthetic and minimalist design.*
9. *Help users recognize, diagnose, and recover from errors.*
10. *Help and documentation.*

No usability experts were available for evaluation in this project therefore four computer science masters students who are learning South African Sign Language were asked to carry out the heuristics evaluation. The four students were involved in the Deaf Telephony project and where designing different applications for the Deaf. None of the four evaluators had done heuristic evaluation before. The final prototype that uses the mobile phone to capture video and the computer for hand detection was used. The prototype was modified to enable it to function as a standalone application therefore the video messages were not sent to another application but redirected back to the same application. Consequently when a video message was recorded and sent it was redirected to the same application as an incoming message. This made it possible for the evaluators to play the video message they had just recorded.

The evaluators were asked to use all the menu options on the screen and to also give feedback as they were busy evaluating the application. This was done so that the evaluators could comment on the aspect of the design or on the events as they occurred. This helps in removing the risk of evaluators forgetting some of the events that occurred if they had to report their findings at the end of the session. The other option was to let the evaluators write down events as they occurred but this is slower as it requires the evaluator to pause what they are doing in order to note their observations down. Letting the evaluators talk about the event as they occur was deemed as a suitable alternative as it speeds up the evaluation process and the evaluators have more time to explore the system. The observer is responsible for recording the evaluator's feedback.

6.2.1 Feedback from the evaluators

The evaluators did not map any of the problems they found to the heuristics. The evaluators had not conducted heuristic evaluation before so they simply pointed out potential problems that may arise from the design and did not map the problems to one of the heuristics. One evaluator mentioned that the users do not have the freedom to pause the application and leave. The application would continue to detect any movement if the user stands up and accidentally covers one of the menu options. Therefore the users had to sit in the middle of the video being previewed in order to avoid accidentally activating any menu option. Once the user started communicating they cannot pause and resume, they can only end the conversation. To continue they have to restart the application from the beginning. The evaluator also suggested having two video windows where one window would be for recording video and the other for controlling the system. Incorporating this feature on the application would be unsuitable for a mobile device, however on a computer it might be useful. The ideal outcome is for the entire application to run on a mobile phone so this feature is not suitable for this application.

Another evaluator had an issue with the size of the icons and the size of the text used in the application. The evaluator felt that the icons should be more visible. The visibility of the icons makes it easier for the users to know the function of the menu option without even reading the text. Furthermore the icon used for *send* and the icon used for *yes* were too similar because of the size that was used during the evaluation. The icon for *send* was used for sending video messages and the icon for *yes* was used as a confirmation when users chose to cancel a video message. The user has to choose either *yes* or *no* when asked whether they are sure they want to cancel recording. Enlarging these two icons might result in the users being able to differentiate between the two icons easily. The menu text and notification were deemed to be too small for users to see because the users usually sit at a distance from the camera and the screen. The users sit at a reasonable distance in order to have their hands, head and torso in the video as these are needed in sign language conversation.

The third evaluator encountered problems with the manner in which the video messages were played. The users did not have an option to scroll through the video messages and choose which video message they wanted to play. However there was a replay option which allowed users to replay the last message they viewed. The evaluator suggested that the users should be able to replay any video of their choice not only be restricted to the last video. Furthermore the evaluator pointed out that the video was too slow and would not be suitable for a user who signs fast. The last evaluator only had a problem with the visibility of the facial expressions in the video. The light in the video was not bright enough to enable the facial expressions to be discerned clearly should the user happen to be of a dark complexion. Sign language relies on facial expressions so the evaluator suggested introducing more light in the video.

Heuristic evaluation can be used to identify errors early in the design process. The evaluation did not highlight as many problems in the design as anticipated. However the fact that the evaluators did not discover a lot of problems does not necessarily mean that the design is usable. Desurvire et al. (1991) noted that heuristic evaluation works well when usability experts are conducting the evaluation and that non expert user results were sometimes not reliable. In this evaluation students who had never conducted heuristic evaluation before were used and that might explain why a lot of usability problems were not discovered.

The changes suggested by the evaluators that are suitable for a mobile device were incorporated in the prototype after the heuristic evaluation was completed. The size of the icons, menu texts and notification texts were enlarged in the prototype. Some of the suggestions for example using two video windows and allowing the user to choose which video message should be replayed were not incorporated into the prototype before the user evaluation. After the changes were completed the user evaluation was conducted with the updated prototype.

6.3 User Evaluation

This evaluation method allows users to have a say and judge the usability of the design. Three methods were used during these evaluations which are observation, questionnaires and interviews. Users are observed while they are interacting with the system in order to gain insight or deeper understanding of how they perceive the system. Users can also rate the usability of the system using questionnaires in this evaluation method. Interviews during the user evaluation give the users a chance to make comments about the system.

The user evaluation was conducted using the final prototype discussed in section 5.4. The prototype that uses both the computer and mobile phone was used in the evaluation because the stand alone mobile prototypes were unable to perform all the necessary functions needed for the gesture based interface video communication application. The computer and mobile phone prototype is suitable for this process as it removes the specific device dependent variables like the computational time needed to render the icons and menu text to the screen. It also removes the lag experienced as a result of creating the mirror images at the same time the hand detection algorithm is running. By using the prototype that performs the hand detection on the computer; the users can judge the usability of the application without experiencing the performance loss caused by the mobiles device’s limitation. This makes it possible to evaluate the usability of the design without experiencing any device specific performance losses.

6.3.1 Aim

The main objective of this research is to uncover whether gesture based interface will improve usability of an asynchronous video communication for Deaf users. In this evaluation we wish to investigate if using an interface that can be controlled using hand gestures makes it easier for Deaf users to communicate with each other without having to physically touch the computer or mobile phone’s keypad. Usability testing on the application’s interface was conducted with Deaf users.

6.3.2 The Experimental Setup

The evaluation took place at the Bastion for the Deaf which is situated in Newlands, Cape Town, South Africa. The Bastion for the Deaf is where the Deaf Community of Cape Town offices are based. The users evaluated the system in the computer lab at the Bastion. Two computers, two Nokia N96 mobile phones and two USB TV Tuner cards were used. The mobile phone was running the Symbian application which sends the video to the computer application. The computer application does all the image processing needed for hand detection. The mobile phone was connected to the computer through the USB TV tuner cards.

The USB TV Tuner cards were used to capture the video from the mobile phone and the captured video was utilised by the computer application for image processing. The computer application also loads the icons and menu text before displaying the video on the computer monitor.

A screen capturing software was employed to record all the users' activity and interaction with the prototype. The video from the screen capturing software was used for observation purposes after the experiment thus making it possible to observe the users' actions even after the evaluation process had been completed. The video can be replayed and watched several times to observe the aspects of the prototype design that were not easily understood by the users. A video camera was also used to record the users' feedback during the interviews conducted after they finished interacting with the prototype.

6.3.3 Method and Procedure

Ten Deaf users from the Deaf Community of Cape Town were used to evaluate the usability of the interface. There were five male participants and five female participants. Each participant received R40[6] remuneration for their participation. Two users were considered per session therefore five sessions were needed to evaluate the interface with the ten users involved. There was no grouping of the users so the two users for each session were determined by the availability of the participants. The participants were not all available at the same time so the two users in each session were determined by whoever was available and willing to participate at that particular time. Some of the participants worked at the Bastion for the Deaf and other participants were there to attend a literacy class so the participants arrived at different times of the day. Each evaluation session with the two participants was approximately 40 to 50 minutes depending on the amount of feedback the participants gave during the interviews.

Users were given a chance to learn to operate the new gesture based interface system until they were comfortable with using the interface on their own. Users were then asked to use the new gesture based video communication system for approximately 10 to 15 minutes to communicate with each other. The exact amount of time users communicated depended on how much they wanted to say to each other. Users were not given any specific topics that

[6] The currency is the South African Rand

they had to use while they were communicating to each other. They had the freedom to have any conversation they desired. They were asked to use all the basic options supported by the application while they were communication with each other. The following Tasks had to be performed by the users during the evaluation:

- *Recording a video message.*
- *Sending a video message.*
- *Playing a video message they received.*
- *Cancelling a recording of a video message.*

Before concluding the session the users were asked if they had cancelled a message they were recording. This task would not be completed by simply communicating with each other. The first three tasks would have been completed if users communicated with each other successfully. The users needed to be reminded to perform the last task if they had not done so already. At the end of the session a questionnaire[7] was completed to rate some aspects of the design. The participants were also interviewed in order to get their thoughts on the interface and the system as a whole. This afforded the users a chance to pose any questions as well as give general comments. Communication with the users was through a professional interpreter. Although the researcher completed a stage one sign language course, the researcher was not fluent enough to have a sign language conversation without the interpreter present. A professional interpreter was used in order to make communication with the Deaf participants easier as both the researcher and the participants were afforded a chance to use their preferred language.

6.3.4 Results[8] and Analysis

The results from the user evaluation were both quantitative and qualitative. The users were asked to rate the different aspects of the prototype on a questionnaire and also give additional comments about the whole system. The results from the questionnaire are presented below and followed by discussions of the results. The quotes in the discussion were obtained from the interpreter so there might be a slight variation from what the Deaf user actually said. A professional interpreter was used in order to minimize the variation as much as possible.

[7] The questionnaire used is in Appendix B2.
[8] The data collected during the experiment is in Appendix D.

Statistics

Ten users were involved in the evaluation of the prototypes therefore the sample size was not large enough to perform thorough statistical analysis. For that reason the results of the evaluations cannot be used to make statements about the Deaf population. However the results from these evaluations provided insight on the usability of the prototype developed. The quantitative data obtained was based on the Likert scale so the results might be misleading if the data is analysed without looking at the qualitative data. There was more emphasis on how the users interacted with the interface rather than looking at the statistical data alone.

Usability

The users were asked to rate how they felt when performing certain tasks using a Likert scale. Before communicating with each other they had a practice run to familiarise themselves with the gesture interface. The histogram below in Figure 6.1 shows the results of how the ten Deaf participants felt when they were learning to use the gesture interface. The scale was a 5 point Likert scale with 1 being *very difficult*, 2 being *difficult*, 3 being *neutral*, 4 *easy* and 5 *very easy*. None of the ten users felt that learning the system was *very difficult* or *difficult*. The responses ranged from *neutral* to *very easy* and three users felt that the experience was *neutral*, two users felt that it was *easy* and five users felt that it was *very easy*. The mean was 4.2 which suggest that on average the experience can be classified as being easy. The following histogram shows the ten users' response:

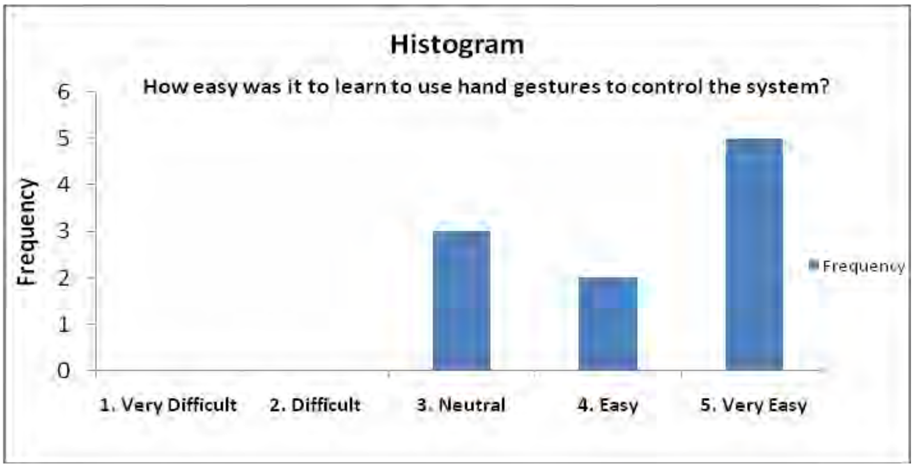

Figure 6.1 Learning to use the gesture interface.

"*The users' responses when asked how easy it was for them to learn to use the system.*"

When users were learning to use the system the female users had more difficulty learning the gestures than the male users. This might have been influenced by the fact that most of the female users had not participated in the project before. Only two out of the five female participants had been involved in the project prior. All the five male users had been involved in either the design phase or had participated in the evaluation of the first prototypes discussed in Chapter 5. The sample size was too small to differentiate between the two groups with any statistical significance. More users are required before any claims can be made about how the different sex groups interact with the interface.

The combination of using the hand gestures to control the application and using the hands gestures for sign language in a conversation was deemed as an interesting factor to take note of. Therefore the users rated how they felt about using both the hands gestures to control the system and also for signing. The following histogram (Figure 6.2) shows the users response:

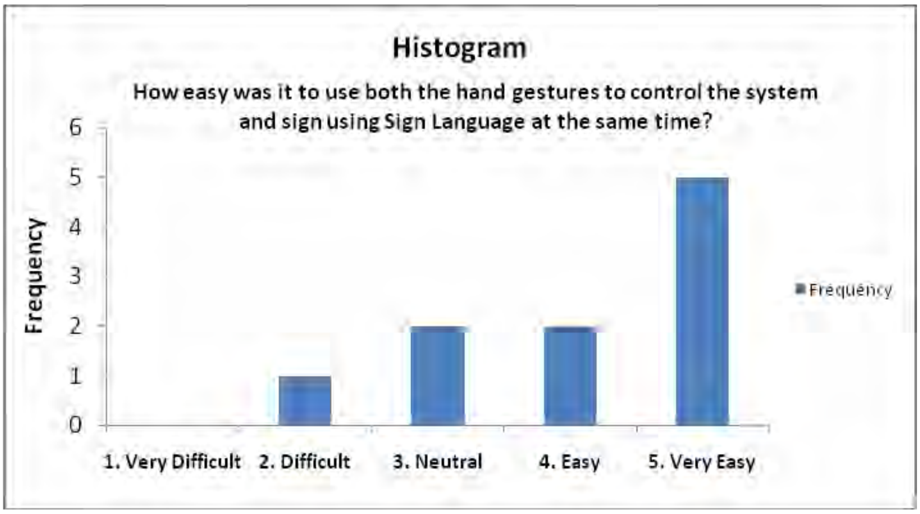

Figure 6.2: Controlling the gesture interface and signing.

Based on the results in Figure 6.2, only one user felt that it was *difficult* to switch between hand gestures for controlling the system and signing but no one felt that it was *very difficult*. However two users felt that it was *neutral*, two more users felt it was *easy* and five users felt that it *very easy*. The mean was 4.1 therefore the overall experience can be classified as being easy. However the mean for learning to use the gesture interface is 4.2 and this is slightly higher than the mean for using both the hand gestures and signing (which is 4.1). So overall learning to use the system was rated as being easier than using the hand gestures for controlling the system and signing. The difference is small enough to be insignificant.

For a store and forward video communication the users have to able to record and send video messages to each other. The way in which these messages are recorded and sent has to be clear and easy enough so that the users can achieve their goal which is to communicate with one another. The video messages that are sent need to be played on the receiving end in order to view the contents of the video messages. Therefore there is a need to measure how the users felt about recording, sending and playing the video messages. The following histogram shows the results for recording, sending and playing side by side:

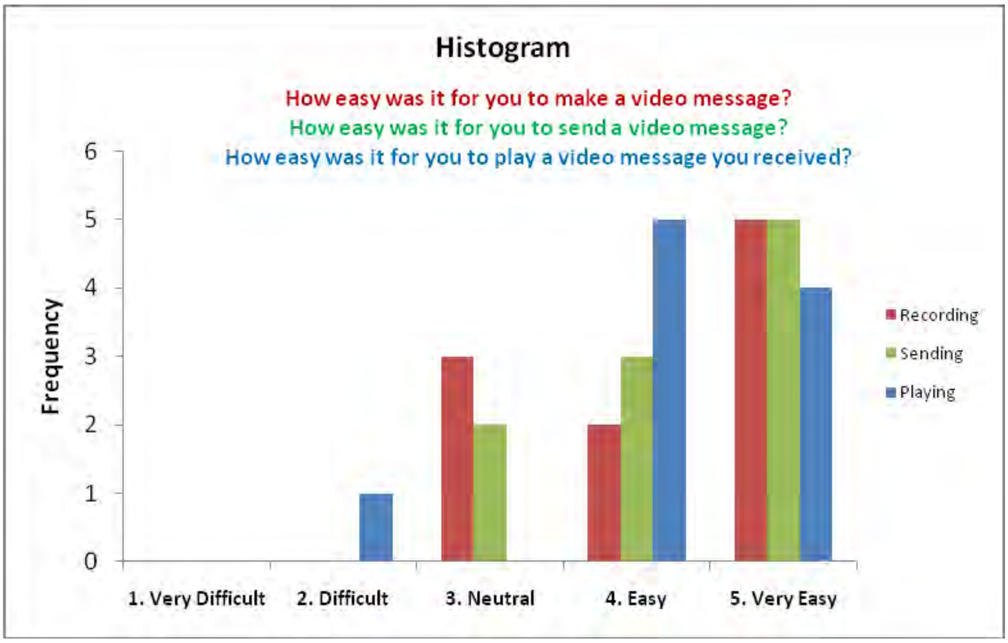

Figure 6.3: Recording, sending and playing a video message.

None of the users rated the experience to be *difficult* or *very difficult* for recording and sending but one user felt that playing a video message was *difficult*. Three users felt that recording was *neutral*, two users felt that it was *easy* and five users felt that it was *very easy*. However for sending two users felt that it was *neutral*, three felt that it was *easy* and five users felt it was *very easy*. Both recording and sending had half of the users rating the experience to be *very easy*. There was not too much variation in the ratings for both sending and recording of video message. All the users rated the experience to be on the positive side of the Likert scale starting from *neutral*. The mean for sending was 4.2 and the mean for recording was 4.3 so both can be classified as being *easy*. However on average recording can be classified as being easier than sending based on the means.

For playing a video message only one user in Figure 6.3 rated the experience to be *difficult* and the rest of the 9 users felt that it was either *easy* or *very easy*. Out of the remaining 9 users 5 rated the experience to be *easy* and 4 rated it to be *very easy*. The mean was 4.2 which means on average the experience can be rated as being easy. None of the 10 users felt that playing a video message was *neutral* or *very difficult*.

The users found that recording, sending and playing a message was easy. One user, when asked to rate how they felt mentioned that *“They were all the same because in the beginning I struggled but it was much easier with practice… but now it’s okay so it’s fine. It was easy now”* When observing the videos of the users interacting with the prototype some of the users did not put their hands at the menu option for playing but chose to put their hands over the message icon that indicated that there was a new message. When a new message is received the play icon is visible on the right hand side of the screen but at the top right corner of the screen there also appears an icon with flashing text indicating that there is a new message. When users saw the icon for announcing the presence of a new message they chose to put their hands over this icon instead of the play menu option. The following screenshots (Figure 6.4) shows the users selecting the icon indicating that there is a new message instead of selecting the menu option for playing.

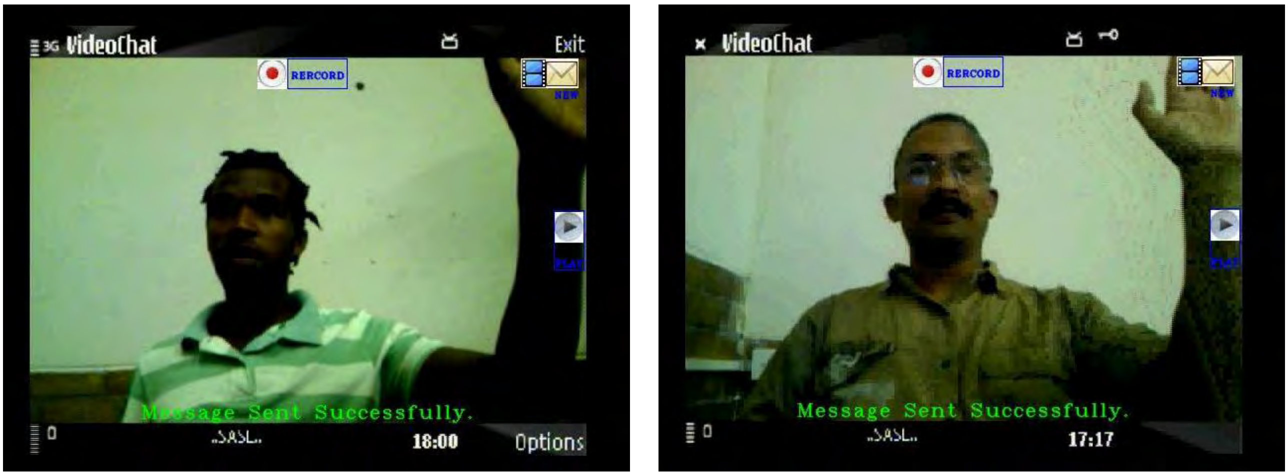

Figure 6.4: Users selecting the new message icon.

“The users placing their hands over the icon that shows that there is a new message instead of selecting the play menu option.”

In Figure 6.4 the user is selecting the icon at the top right hand corner in order to play a message and this does not play the message as there is no gesture detection performed at the right hand corner. The gestures for playing are only considered if they are within the rectangle on the right hand side where the play menu option is located. This led users to the false assumption that the system was not responding to their requests. However in other occasions as in the screenshots in Figure 6.4 a message was played even though their hand was not selecting the right option. This was due to the fact that the elbow or the other part of the hand was over the play menu option. The play menu option was activated by the other parts of the hand instead but to the users they thought the new message icon was the one playing the message. The prototype perhaps should have used both the play menu option and the icon indicating that there is a new video message to play the video message. If hand detection was performed at the new message icon and the menu option then the users will be able to play the message by selecting either of the icons. Therefore users would have been able to play the video message regardless of the position of their elbow. This was not anticipated during the design stages when the icon for indicating that there is a new message was placed at the top right hand corner.

All the other options in the menu of the video communication application were straight forward because if a user selected a particular option it was immediately executed. However when cancelling a video message that was being recorded the users had to go through a confirmation dialog. The confirmation was instituted in order to prevent a user accidentally cancelling a recording of a video message. The users were asked to rate how they felt when they cancelled a recording of a video message. The following histogram in Figure 6.5 shows the users' response:

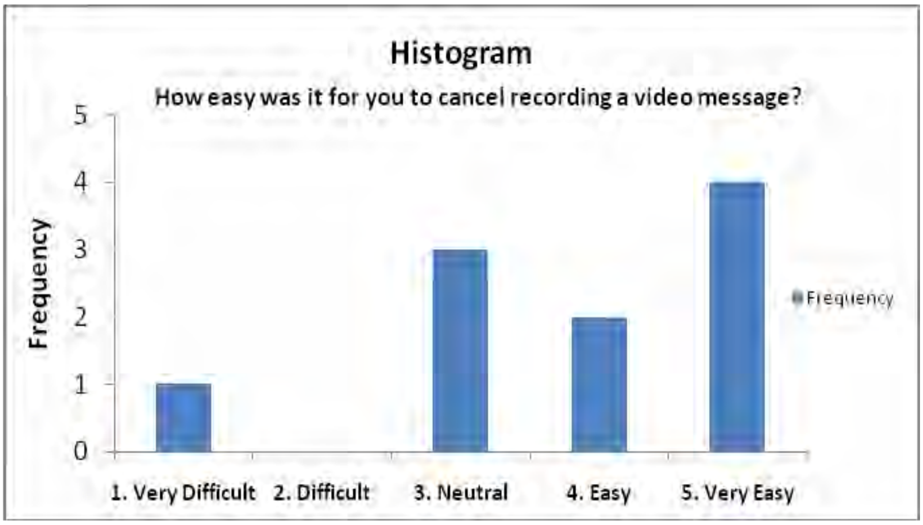

Figure 6.5: Cancelling when recording a video message.

Only one user felt that it was *very difficult* to cancel recording a video message. The rest of the users' responses were on the positive side of the Likert scale. Out of the 9 users in the positive side of the Likert scale three felt that cancelling was *neutral*, two felt it was *easy* and the last four felt that it was *very easy*. Although most of the responses were in the positive side of the Likert scale all the users paused when they saw the confirmation screen. At the confirmation screen the user had to select *yes* in order to cancel recording the video message. If the user selected *no* the application went back to recording the video message because the user would have chosen not to cancel recording. Figure 6.6 shows the confirmation screen shown to the users with *no* on the left hand side of the screen and *yes* on the right hand side of the screen:

Figure 6.6: Cancel Confirmation screen.

"The message at the centre of the screen is asking the user if they want to cancel recording a video message. The user can either confirm by selecting yes or continue to record the video message by selecting no"

When the users saw the conformation screen, the pause might have been a result of the users reading the message. After the pause, majority of the users asked if they should choose *yes* in order to cancel. They were not certain if choosing *yes* would cancel the recording and so opted to ask before selecting either of the menu options.

Video quality

The video quality of the messages that were recorded and sent was evaluated by the users in order to determine if the video was suitable for sign language communication. The users had to evaluate the clarity of some of the sign language features and also rate how easy it was to understand the video messages. The following histogram in Figure 6.7 shows the users' responses to how well they understood the content on the video messages.

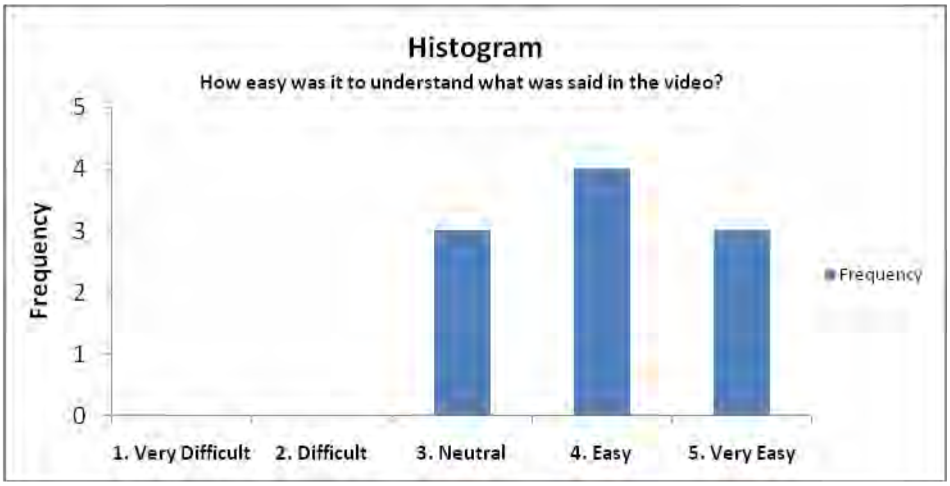

Figure 6.7: Understanding what was said in the video message.

None of the users experienced difficulties understanding the content of the video messages they received, as no one felt that it was *difficult* or *very difficult*. There were 3 users who felt that it was *neutral*, 4 users felt that it was *easy* and the remaining 3 felt that it was *very easy* to understand what was being signed in the video. The mean for the 10 users was 4 therefore on average it can be said that it was fairly easy to understand what was signed in the video. Although the users understood what was said in the video, the signs became unclear when users were signing too fast. This resulted in users having to sign slowly prompting one user to point out that a signer signing slowly is perceived somehow negatively. This is what the user said:

"*I could understand the signing but the signs were slow and staggered. I think that sign language is very upbeat and there is lots of action happening so when you getting the slow signs it's basically perceived that the sender is slow or the person is not happy so we need to differentiate between whether the person is slow or whether the person is feeling that way because that's the way we perceive the people to be*."

The users had to rate the clarity of both the hand gestures and the facial expressions. The clarity of the hand gestures and facial expressions was rated on a Likert 5 point scale but with different labels from the ones used in the previous histograms. The Likert scale used here had labels were 1 is *not clear at all*, 2 is *not clear*, 3 is *neutral*, 4 is *clear* and 5 is *very clear*. The following histogram in Figure 6.8 shows the users responses for both clarity of hand gestures and facial expressions plotted side by side:

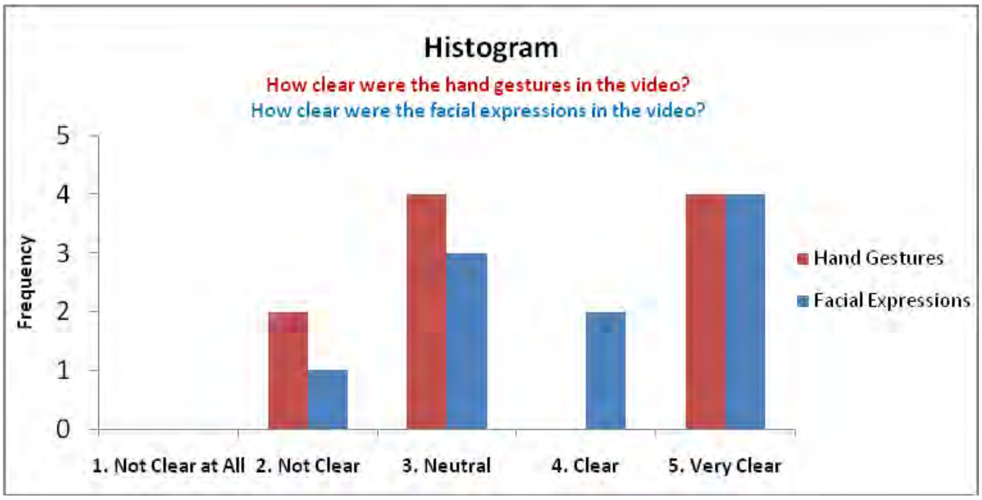

Figure 6.8: Clarity of the hand gestures and facial expressions in the video.

In Figure 6.8 none of the users felt that any of the hand gestures and facial expressions was *not clear at all*. For the hand gestures in the video messages 2 users felt that they were *not clear* and 4 users felt that they were *neutral* and 4 users felt that they were *very clear*. No user felt that the hand gestures were *clear* so the users were split among *not clear*, *neutral* and *very clear*. Only 1 user felt that the facial expressions were *not clear*, 3 felt that they were *neutral* and 2 users felt that they were *clear* and the remaining 4 felt that they were *very clear*. The mean was 3.6 and 3.9 for hand gestures and facial expressions respectively. Based on the means the hand gestures can be said to be in between *neutral* and *clear* whereas the facial expressions were almost *clear*. On average facial expressions were clearer than hand gestures on the video.

Facial expressions are a very essential part of sign language and they can determine the tone of the conversation. This means that sign language video has to be of a quality good enough for the users to be able to clearly discern the facial expressions. One user had this to say about facial expressions:

"I think it will be a problem if you can't see the face if the picture is dark because we rely a lot on facial expression and you need to see the face. If you can't see the face then there is problem because I don't know if **(the other user's name)** *has an attitude or what's going on with her face so it's very important for us to see the face."*

During the interviews the users also pointed out that the video was streaky when someone was signing fast. There were lines in the video displayed on the screen which were caused by the USB TV tuner cards used to get video input from the mobile phone. This is what the users had to say about the streakiness in the picture:

"*The hand shape wasn't clear because of the streakiness on the hand itself so in terms of that it wasn't clear*."

"*It was a bit staggered I can say the picture, the quality is good of the picture but as soon as there is a movement then there is a bit of a problem there."*

"The movement wasn't as great as before so there was not enough facial expression so it was medium. It wasn't that wonderful."

"I think that the picture quality is not that great either because it's not clear it is very staggered"

The following screen shot shows the lines that were formed when a user was signing fast.

Figure 6.9: Interlaced video frame

"An interlaced video frame showing lines over the hand"

When the mobile phone was connected to the TV set directly without any connection to computer using the USB TV tuner card the lines were not present in the video. In the evaluation the video from the mobile phone was sent to the TV tuner card as interlaced frames. The video was passed to the computer application as it is without any de-interlacing done on the video. This resulted in lines on the video when users where signing fast.

The users have to be willing to accept the video quality in order for a video communication application to be successful. If users are not satisfied with the quality of the video messages they might not be willing to use the application. The users were asked if they would be happy to use sign language video of this quality on a mobile phone. A 5 point Likert scale was used to rate how happy they were with the video quality. The labels on the scale this time were 1 *not happy at all*, 2 *not happy*, 3 *neutral*, 4 *happy* and 5 *very happy*. The following histogram in Figure 6.10 shows the users response:

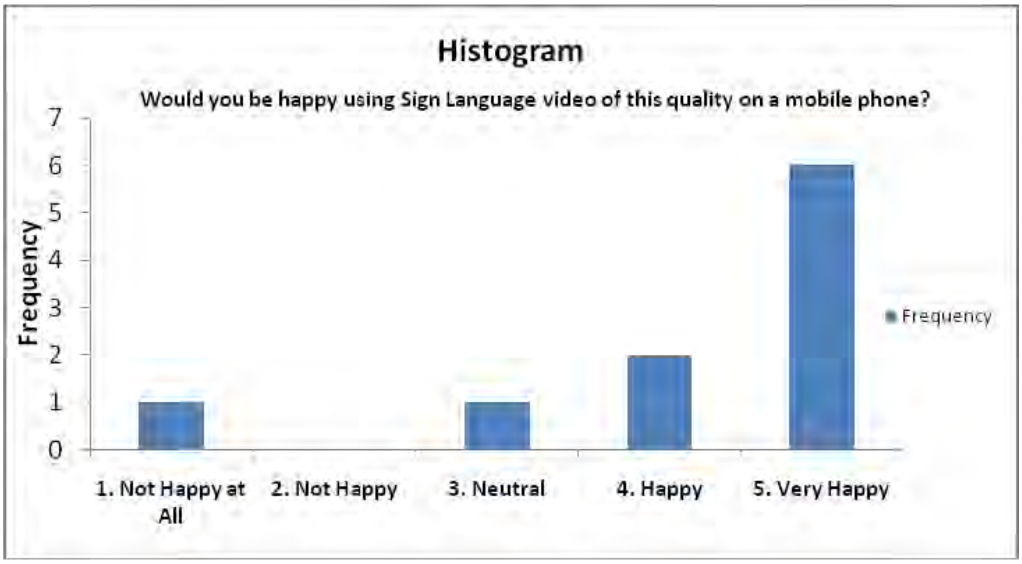

Figure 6.10: The quality of the sign language video.

Most of the users were very happy to use video with this quality on a mobile phone. Out of the 10 users 6 were *very happy*, 2 were *happy*, 1 was *neutral* and 1 was *not happy at all*. The mean was 4.2 so on average the users were happy with using video of this quality on a mobile phone. The problem the users had with the video was that there was not enough lighting in the video. This is some of the comments the users made about the lighting:

"It's clear but it depends on the area if there is a whole lot of lighting then it might have been better but it's very dark here in this area but if there was good lighting then it would have been okay it just depends on the situation because it's a bit dark here."

"… colour was very dark the screen was very dark and the picture quality was very dark."

"The lighting isn't so wonderful so I need to control where I stand to be in a good lighting so I don't know if it's dark maybe or something happens or if you in an area where it's very dark is it not possible to have a light attached to the camera so it will be easier for the receiver to see the person so maybe if there is a light on the camera or something you know it might just work very well. Just like a torch on the phone maybe that could be developed as well because I don't think it has to be a strong light maybe a normal light it's just something that I'm concerned about."

The users were happy to use the video with this quality even though they were not happy about the light conditions in the room. The computer room had a light which was turned on but in the video the faces appeared a bit dark. Participants had to sit at a certain angle to allow for some light to be cast on their faces. The light however was still not bright enough to clearly discern the facial expressions on the video so they still had problems with the video. Even though it was dark in the room it was clear enough to see the signs. The users suggested adding a light source on the phone's camera or using the flash light when the video is recorded. The problem could also be solved by using a contrast/brightness correction on the video but this will require more processing power. This can only be implemented if the application is running on a powerful mobile device.

Hand gestures

The gesture interface has to use gestures that are easy to use and are not stressing to the users. Thereby allowing users to feel more comfortable and be able to use the system effectively and efficiently. The gestures should not be tiring to the user because it can affect the user experience if the user suddenly becomes tired while using the gesture based interface. The users had to rate whether the gestures aimed at controlling the application were difficult to make and also if the gestures were tiring when used repeatedly. A 5 point Likert scale was used in this question as well but the labels used were 1 being *definitely no,* 2 being *no,* 3 being *maybe*, 4 being *yes* and 5 being *definitely yes*.

The following histogram shows the users' response to whether any of the gestures were difficult or tiring:

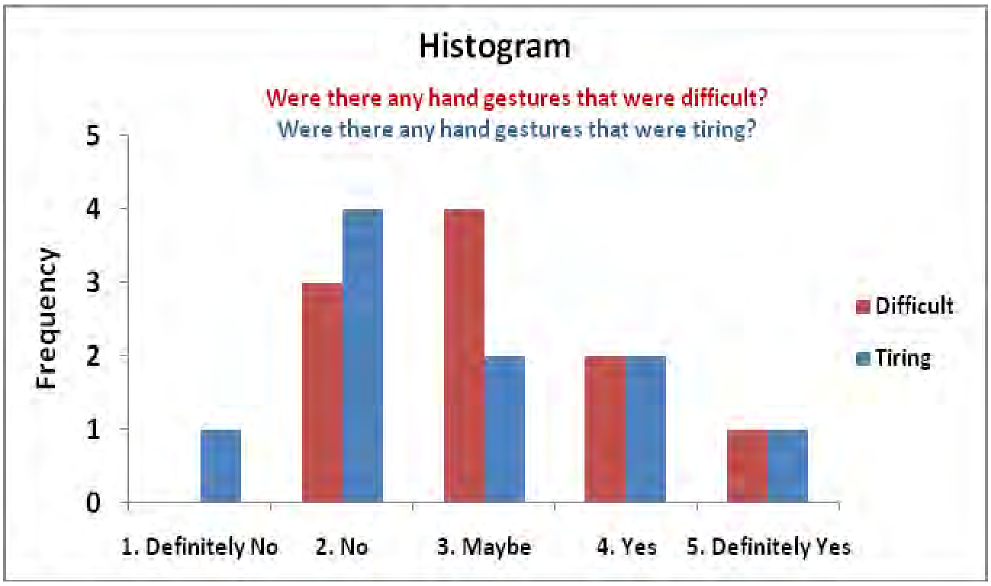

Figure 6.11: Difficult and tiring hand gestures.

The results for difficult gestures had 3 users who said *no* the gestures were not difficult. Four users were undecided they said *maybe* the gestures were difficult. Two users said *yes* the gestures were difficult with 1 user saying *definitely yes*. The mean was 3.1 so overall maybe the gestures were difficult as the user's responses are halfway on the scale. Half of the users did not think that the gestures were tiring with 4 users saying *no* and 1 user saying *definitely no*. The other half had 2 users undecided saying *maybe* the gestures were tiring, 2 users saying *yes* the gestures were tiring and 1 user saying *definitely yes*.

During the evaluation some of the users had a problem when selecting the menu options. The hand detection algorithm could not sense an open hand quickly enough. This only happened to fair-skinned users as their hands were not that different from the background in the room. However it was discovered during the training session when they were learning to use the system that if they start with an open hand and then close their hands to make a fist the detection time was quicker. One user who had a problem selecting the menu option with an open hand said the following:

"*I thought that it was very nice but I had a bit of a problem with the background. If it was blue maybe or red you would be able to see better because then you are not able to see the eyes or hair that well with such a background because it's white on white, white against white so maybe if you had a different colour background it might have worked better and with the hand gestures it became difficult."*

Another solution was for users to move their hands over the menu option until the wrist was over the menu option. As soon they moved their wrist over the menu option the hand detection algorithm was able to pick it up. These problems resulted in some of the users feeling that the gestures were difficult and tiring. This is what some of the users who had problems had to say about the gestures:

"You know it became an issue trying to use the gestures do we have to use a fist or can't we even upgrade it to having to use your finger because I don't like it, it should have been more like a touch screen, you just touch it and it goes off. It shouldn't be you know us having to struggle using our fist to try and get the icons to work."

"For me I found it very hard to get the icons to work I think it would be best if you could just use your finger ..... Maybe we should just use a finger and just pin point to the icon and it would go off immediately."

"It wasn't comfortable for me the gestures aren't comfortable for me trying to get it to work wasn't comfortable. It wasn't comfortable."

This made the users who had to use a fist to select the menu items feel like the gestures were not comfortable. They wanted the gesture interface to work quickly like in a touch screen where you touch an item and get an instant response. The hand detection algorithm in the prototype only detected a hand if it was placed over the menu option for at least a second. Not all the frames in the video were analyzed since some frames were skipped in order to minimize the amount of processing needed. The hand had to be detected over a couple frames that are analyzed. This was designed to disregard the events where a user accidentally places their hands over the option when they are signing. If a hand was held up for a while, it was considered as a gesture aimed at controlling the system else it was regarded as part of the signing conversation.

The users however wanted the detection to be fast and to recognize fingers just like in a touch screen. In one of the evaluation sessions one user even mentioned that *"I would have preferred to rather use my index finger"* and the other user in the session said *"Just like touch screen I think that you should maybe try to copy that kind of technology because it would be much easier if we are just able to use our index finger."* Making the prototype work on a mobile phone will be more challenging if the detection algorithm has to cater for fingers.

This results where unexpected as all the users used sign language so the assumption was that they are already familiar with and used to using their hands to make hand gestures. There was a further difficulty experienced when the users were using gestures to control the system. The users mentioned that the hand gestures were distracting as they were part of the video message. The problem was more apparent when the hand detection took longer at the end of the video message. The users were selecting the menu option with an open hand or making a fist and these actions ended up being part of the video message. The received video messages had these actions causing the other user to be uncertain if the user sending the message was signing or controlling the system. This is what the users had to say:

"The gesture itself was distracting that's what I can say the fact that because you think they are saying something but they are not. They are trying to record or they are trying to do something so that was a bit distracting."

*"Because of the sign **"s"** people would think that they are saying stop or **"s"** or something so it could be a sign so maybe if you can just eliminate that completely from the video it might just help."*

"it's just that when someone is using the hand gestures you think they are waving or signing to you so once it's done so if you send they think you saying goodbye but you not saying good bye maybe you want to continue with the conversation. That gesture is not a good gesture because it either says goodbye so the waving of the hands indicate goodbye and that's not a good thing."

The users indicated they preferred for the gestures aimed at controlling the system to be completely removed from the video message. They suggested that the last couple of seconds of the video be removed completely or to have an area of the video at the centre as it is and the area where the icons and menu options are situated to be then cropped from the video. If the surrounding area of the screen was cropped out from the video message the gestures aimed at controlling the system would not be part of the video. The following screenshots in Figure 6.12 show the gestures that were aimed at controlling the system but were in the beginning and also in the end of the video messages.

Figure 6.12: Hand gesture for controlling the application that were in the video message.

"On the left video window a hand is shown at the top of the video and this was not supposed to be part of the video message because the user was selecting the record menu option which was at the top. On the right video window a hand was selecting send menu option and this action appeared in the video message."

The gesture on the left where a user has their hand raised up is not supposed to be part of the video message but was rather aimed at initiating the recording processes. The gesture on the right was aimed at sending the video message. When a user raised their hand to select record at the top of the screen the video recording process was started before the user can lower their hand and start signing. The video message recorded included all these actions where the user had their hands up. On the right after the user finished signing they had to put their hands over the send menu option so before the option was activated the hand gesture for sending was recorded in the video message. The users did not want these gestures in the video message so the application should have a way of cutting out these gestures completely from the video message.

The users had an opportunity to rate whether there were any gestures they liked. The following histogram in Figure 6.13 shows the results of the users' responses:

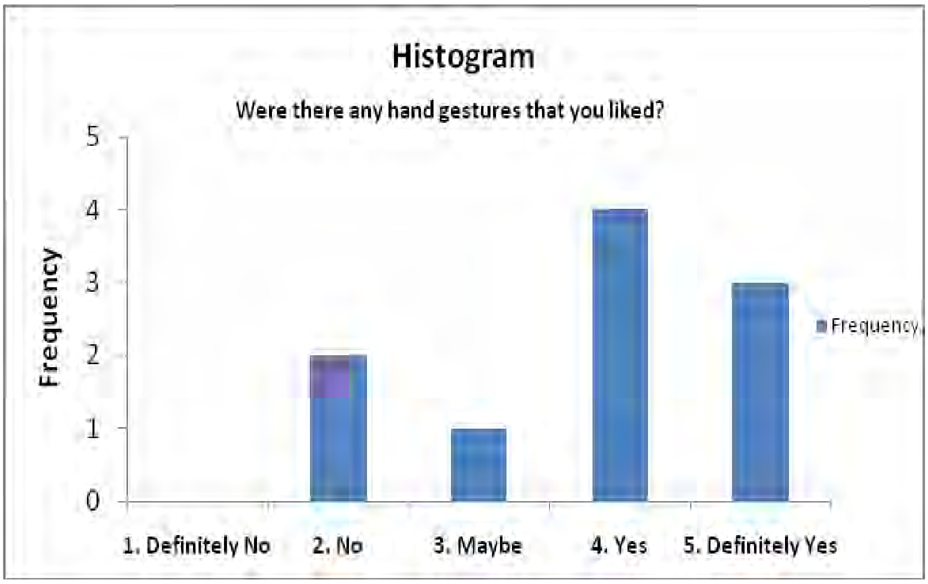

Figure 6.13: Likable hand gestures.

In Figure 6.13 seven users found gestures they liked that were aimed at controlling the prototype. The remaining 3 users had 1 user undecided and 2 users saying they did not like the gestures. Out of the 7 users who found gestures they liked, 4 users said that they liked all the gestures aimed at controlling the system.

Using the system

A system can be designed using the principles of design that help in creating usable designs but the target group of the system should be willing to use it. Coming up with good designs alone is not enough, the system needs to help users achieve their goals and they have to be stimulated to use it. The users should have the will to use the system on their own and not use it only in evaluations. The ten 10 Deaf users were asked to rate on the Likert scale how willing they were to use the system if it was available.

The video communication application uses the network for transferring the video messages from one client to another. There is a network charge associated with the transfer of video messages and the cost depends on the type of network the users are using. The user’s will to use the system is not enough if the users are not prepared to pay the data charges associated with the transfer of the video messages. If the users are not willing to pay for the data charges then they won’t use the system as there is a cost associated with video communication.

The following histogram in Figure 6.14 shows the users response on their willingness to use the system if it was made available:

Figure 6.14: How willing are you to use the system?

More than half of the users were willing to use the system if it was made available. There were 7 users in total who were willing to use the system. Two users were not sure as they said maybe they will use the system and only one user was not willing to use the system. The video communication application has data charges associated with it so the users were asked how much they were willing to pay for the data charges. The cost of the data charges was explained in terms of how much they were willing to pay for a minute of video sent. The amount the users chose is not for the whole video message but for every minute in the video. The user’s response was recorded using a 5 point Likert scale with different prices as labels.

The following histogram in Figure 6.15 shows how much the users were willing to pay to use the system:

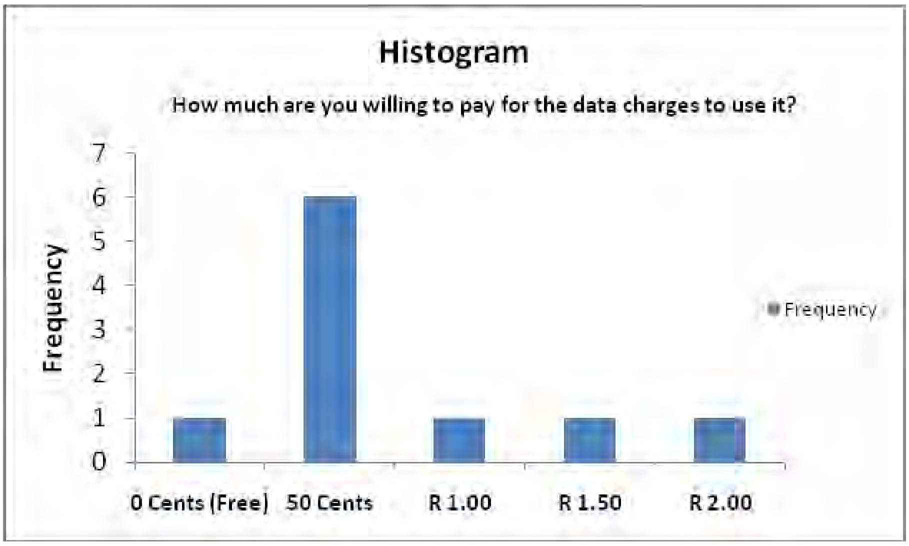

Figure 6.15: Data charges for using the video communication application.

Only one user wanted to use the video communication for free and the other 9 users were willing to pay 50 cents or more per minute of video to use the system. Six users were willing to pay 50 cents and the remaining 3 users were each willing pay R1.00, R1.50 and R2.00 per minute of video. The user who wanted the video communication application to be free said that it should that way because Deaf people want things for free. Interestingly enough the user who was not willing to use the system chose that they were willing to pay 50 cents per minute of video even though they were not willing to use the system. The reason the user was willing to pay 50 cents per minute of video was that it is just like SMS so the cost should be similar to that of an SMS.

6.3.5 Discussion

The users liked the new way of interacting with the video communication prototype. The interaction method made it easier to communicate using video. There were concerns with the quality of the video when users were signing fast but they were able to see the signs in the message. The video quality was not what the users would have preferred but they thought that it was good enough to communicate with each other as long as the users were not signing fast. Even though there were complaints regarding the quality of the video especially with the lines that were visible when users were signing fast, the users were just pleased that they could use sign language to communicate with each other. They accepted the video quality because they were comparing it to SMS, at least the video communication allows them to use sign language which is their preferred mode of communication. If they are using SMS they have to deal with the problem of understanding the different English words which they are not familiar with. Some of the users made comments such as the following:

"It's much easier to sign it's much easier to communicate via sign language. If I SMS using the words it's a struggle so if I'm able to use this type of technology it will be great."

"if a word comes up and you say you don't understand what that word is you ask them what does this word means or you have to go research what this word means and they say you know very well and I don't know what you are talking about. There is a whole lot of miscommunication and wasting of money so if it's face to face and that will be great. It's much better in sign language."

The users were happy to be using sign language to communicate with each other so they looked past some of the problems they experienced when using the gesture interface. They were still willing to go through some of the usability problems to communicate with each other. Some of the usability problems were caused by the background, in the end some of the users could not get a quick response from the system when they were selecting a menu option with an open hand. The users also liked the fact that they did not have to move forwards in order to select a menu option. They had plenty of positive feedback about the system in general but most of it regarded the fact that they did not have to use SMS. Other comments made in the evaluation session included:

"It's a wonderful system and its something new I'm sure that will be great for Deaf people especially because Deaf people don't understand SMS you know so it's a better way of communication sometimes there miscommunication with SMS even so this is a better communication."

Although the evaluation was done using both the mobile phone and the computer the users felt that it would be much better if the system can run entirely on the mobile device. The problems experienced while implementing the prototype as already discussed in Chapter 5 were also mentioned to the users, hence why both the computer and the mobile device were used in the evaluation. However the users were still interested on how the video communication application would work on the mobile device.

The other concern was which devices would be able to run the gesture based communication application. One user raised the issue whether the application would be compatible with all the phones. When it was explained that the prototype was tested mostly on high-end phones and the prototype still had problems even on these more expensive phones there were concerns on when Deaf people would finally be able to afford those phones in order to access this technology. The user said:

"*The thing is you know I don't think that as Deaf people would be able to afford something like that if you are talking about high end phones.*"

Even though they thought that this kind of communication was very useful to Deaf people they felt that it is not within their reach. The user also mentioned that *"I will be too old to even enjoy this technology then".* The other issue was that in future they will always be better mobile phones and they will not be able to afford them so by the time they can afford the phones the technology would be outdated as there will always be new and better things coming out. The general feeling was that if the gesture based interface can be implemented on the cheaper phones it would be helpful to many Deaf people as they will be able to afford it. The concept provides an alternative communication method apart from SMS that Deaf users can use to communicate with each other. More work has to be done on the system before it can reach the greater masses in the Deaf community.

6.4 Chapter Summary

In this chapter we evaluated the gesture based interface video communication prototype that was implemented on both the mobile phone and the computer. Two evaluation methods were used to evaluate the usability of the prototype. Heuristic evaluation was conducted first in order to uncover any usability problems before the user evaluation was conducted. The heuristic evaluation did not discover as many problems as anticipated due to the fact that no experts were used in the evaluation. However some problems were uncovered and corrected before the user evaluation was conducted.

The user evaluation was conducted with the target group of users in order to gain insight and more understanding on how they interact with the prototype. The evaluation was conducted with 10 Deaf users although it would have been preferable for the sample to be larger. The availability of participants willing to participate in the evaluation determined the total number of participants involved. The evaluation highlighted some aspects of the design that needed to be revised in order to increase the usability of the interface. The evaluation also revealed that the users were more satisfied when they communicated in sign language and were willing to look past the usability problems they experienced. The benefit of using sign language outweighed all the usability problems.

7. Conclusion

This dissertation was aimed at finding better communication methods for Deaf people. The communication method explored was video communication and a gesture interface was developed in order to investigate if it can increase the usability of the video communication. Gesture based video communication prototypes were developed both on the mobile platform and on the computer. These prototypes were developed using a user centred design methodology where users are involved from the requirements gathering stage to the design process and all the way to the evaluation process. Working with Deaf users in the development of the prototypes was a challenging task due to the communication problems but it made it easier to understand what the users needed from the prototypes.

Involving the users in the design and evaluation of the prototypes was more challenging than expected as the evaluation revealed problems that were not encountered during the design process. Although the users were involved in the design process, during the evaluation some of the users did not like some of the design aspects of the prototypes. Different users had different reactions to the interface and some users preferred a different approach from the one used in the prototypes. For example while other users were happy with the current interaction method some of the users preferred to use an index finger to interact with the gesture interface as already discussed in section 6.3.4. The users wanted rich interactive interfaces with quick response times and they wanted all these functionality on the low end mobile phones. They preferred to have all these functionality on the low end mobile phones as they all owned one. Users did not take into considerations the device limitations but wanted all the functionality that they thought would make the interfaces more usable. This made it more challenging to decide what should be taken into consideration and what should be left out of the designs in order to ensure the designs can be successfully implemented.

In order to answer the research questions of this project, the interface was evaluated using two evaluation methods (heuristic evaluation and user evaluation). However only one evaluation method provided more feedback about the usability of the gesture based video communication prototype. The heuristic evaluation did not yield expected results because experts were not used in this evaluation. The user evaluation with the Deaf users was able to highlight aspects of the design that needed to be revised. The results of the user evaluation showed that the Deaf users were much more satisfied to be using video communication to communicate with each other. During the user interviews the users had problems with current text based communication methods that are available and showed preference to the video communication. Although the gesture interface had problems which were discussed in section 6.3, the users were still happy to be using sign language to communicate with each other. The users were willing to use the system even though the video quality was not perfect. The video quality was good enough for the users to communicate with each other. They were willing to look past the problems as long as they communicated with through sign language.

The rest of this chapter is arranged as follows: the research questions and reflections are discussed in section 7.1, general conclusions are discussed in section 7.2, research limitations are discussed in section 7.3 and future works are discussed in section 7.4.

7.1 Research Questions

The approach used in the design and development of the prototypes in this project was aimed at answering the following research questions.

1. **How can we use mobile phones to provide efficient and effective video communication for sign language?**

Sign language is a visual communication language so people need to see each other in order to communicate. Video communication makes it possible to communicate with one another through sign language. The device limitation associated with mobile devices makes it more challenging to provide real-time video communication. Video call on mobile phones does not have the video quality that is suitable for sign language communication. Therefore a store and forward video communication is the alternative as it allows for better video compression. This preserves the important sign language features in the video and makes it possible to provide an efficient and effective video communication application. Although there were video problems during the user evaluation as discussed in Chapter 6 the users were able to understand what was said in the video. Therefore by using store and forward video communication Deaf users can communicate effectively using sign language.

1.1 Can the main rear camera on the mobile phone be used in conjunction with an external display on a television screen?

The mobile prototype developed in this project showed that it is possible to use the rear camera and a television for display purposes. However even though this was possible not all the mobile phones were successful in performing all the tasks that are needed by the video communication application. Unfortunately when implementing the system on mobile devices only, there was not a phone that was able to run the gesture based interface effectively and also use the television for display purposes. As discussed in section 5.3.2 the mobile devices that were able to use the rear camera for capturing video and displaying the video on a television set had a longer response time than what was acceptable for the gesture interface. Although the rear camera could be used in conjunction with the television the response times were not suitable for the gesture based video communication application. When simulating a mobile device with more processing power, the rear camera on the mobile phone was successfully used in conjunction with a TV.

1.2 Can a gesture based interface be implemented effectively on a mobile phone?

The gesture interface was implemented effectively on the mobile device platform but it did not function well on all of the devices. All the phones tested could detect the hands gestures which were aimed at controlling the system. The gesture interface was implemented and tested on different phones and only one phone met all the requirements needed. The Nokia N96 was able to detect hand gestures and respond in reasonable time however there were problems when playing the video on the television. Although the gesture interface was implemented effectively on this mobile device other aspects of the video communication prototype did not work. Testing on the Nokia 6120 Classic and Nokia 5800 mobile devices revealed that the mobile device’s camera did not support mirror images as discussed in section 5.3.2. The mirror images are an essential part of the gesture interface so the interface was not implemented successfully on the other devices tested except for the Nokia N96. The Nokia N96 had better features so as a result the gesture interface was implemented successfully on this high end device. The gesture interface was implemented successfully on the high end device but did not work on the other devices. This makes it inaccessible to the greater masses of the Deaf people as it works on the high end phones which are more expensive.

2. Does a gesture based interface improve the usability of a store and forward video communication application for Deaf people?

The gesture based interface allowed the users to use video communication prototypes without physically touching the mobile phone or the computer to record, send and play video messages. The results indicated that the users were happy and willing to use the gesture based video communication application. Even though some users experienced detection problems the overall results obtained from the questionnaires was positive and most of the tasks were rated as being easy to complete. There were difficulties with some of the gestures and some aspects of the interface needed to be revised in order to be more usable but most of the users were happy with the way they interacted with the system. The majority of the users felt that the system was easy to use and liked the concept of using hand gestures to control the system.

7.2 General Conclusion

This dissertation presented an alternative method for controlling a video communication application other than using the conventional keyboard on a computer or keypad on the mobile phone. The gesture interface was developed for a video communication application for Deaf people thereby allowing Deaf people to communicate with others using their preferred language. The results from the evaluation showed that the Deaf users were interested in using the prototypes and were satisfied with using the gesture interface. Although most of the users wanted the whole system implemented on the mobile phone there were no solutions found to overcome the mobile device limitations. As a result the video communication prototype was implemented and tested using both the computer and the mobile phone.

7.3 Research Limitations

The work presented here was a result of involving the target users in the development process of the prototypes. Therefore the outcomes of the research were largely influenced by the users' perception of how video communication application should work. Therefore replicating the results obtained in the studies conducted here is unlikely if a different group of Deaf users with a different background are used. The users used in this project were recruited from the Deaf Community of Cape Town; some of them were attending a literacy class there and had been involved in other projects so their knowledge of video communication technology would be different from other Deaf people. The number of users involved in this project was low so the results might have been different if a bigger group was available. The reason for using a small number of users was due to the availability of Deaf users because getting more Deaf people involved in the project was challenging.

7.4 Future Work

There are several ways in which the work done in this project can be extended. This involves additional features that can be added to the gesture based interface as well as new ways in which the video communication can be implemented.

7.4.1 Refining the hand gestures

The gesture interface currently uses hand gestures to select menu options. The interface can be revised to use fingers instead of hands. The current detection algorithm would not be suitable for detecting fingers so alternative detection algorithm would have to be used instead. The current detection requires a uniform background in order to detect any hand gestures, this can be updated to work in any environment. Making the gesture interface work without a uniform background makes it much more applicable and practical as it can be used in any situation. However adding all this functionality will result in more computing resources being utilized.

7.4.2 Deploying the application on mobile devices

The current setup used for the gesture based video communication application is not suitable for everyday use and it is too complicated to be used by any Deaf person. The solution is to deploy the whole system on the mobile device as most Deaf people have a mobile phone. Therefore if the system could be deployed entirely on the mobile platform it would reach a greater number of Deaf users. The application should work on low end cheaper mobile devices as most Deaf people cannot afford the high end phones. However connecting the mobile device to a TV would still limit the use of the application to places where a TV is available.

7.4.3 Using the front camera on the mobile device

The use of a TV is a limiting factor as users can not use the communication application on the mobile devices while on the move. Using the front camera on the mobile device would enable users to communicate easily as they can see the mobile phone's screen and also record the video. The video quality on the front camera would have to be improved in order to preserve the important sign language features. New methods of achieving this would have to be explored. If the front camera can be used effectively the gesture interface can be eliminated completely from the application as the users can simply use the phone's keypad to select any menu options. Using the phone's keypad makes it possible for the users to use the video communication application even when they are in public where there is no uniform background. Although they have to lean forward to press buttons on the phone, they will be able to use the application in situations where the gesture interface is not suitable.

7.4.4 Using better compression methods for Real-time video

Real-time video communication would be the best solution for sign language communication. Poor video quality in current video call on mobile phone results in the loss of some of the important sign language features like facial expressions. By using better compression methods the quality of real-time video can be improved. The video compression can be optimized to preserve important sign language regions in the video similar to work done in MobileASL project (Cavender et al., 2006). If the compression used preserves this features the video transmitted will be suitable for sign language communication and users can communicate in real-time.

Appendices

Appendix A: General Questionnaire

1. Do you own a mobile phone?

Yes	
No	

If yes above:

1.1 What make and model is your phone?

____

1.2 Does the phone have a video camera?

Yes	
No	

If yes above, specify the number of cameras: ________

1.3 How do you communicate on a mobile phone?

SMS	
Video Call	
Instant Messenger (e.g. mxit)	
Other (Specify:______________________)	

2. Do you own a Personal Computer?

Yes	
No	

If yes above does it have internet?

Yes	
No	

Appendix B: Evaluation Questionnaires

B1: Computer Prototype Questionnaire

For the following questions, please circle the number which best represents how difficult or how easy it was to perform that particular task with 1 being very difficult ad 7 being very easy.

1. How difficult or easy was it for you to record a video message?

1	2	3	4	5	6	7
Very Difficult						**Very Easy**

2. How difficult or easy was it for you to send a video message?

1	2	3	4	5	6	7
Very Difficult						**Very Easy**

3. How difficult or easy was it for you to play a video message you received?

1	2	3	4	5	6	7
Very Difficult						**Very Easy**

4. How difficult or easy was it for you to cancel a Recording of a video message?

1	2	3	4	5	6	7
Very Difficult						**Very Easy**

5. How difficult or easy was it for you to stop a video message while it was playing?

1	2	3	4	5	6	7
Very Difficult						**Very Easy**

6. Do you have any suggestions on how we can improve the usability of the interface?

____
____
____
____
____

B2: Mobile Prototype - Gesture Interface Questionnaire

1. **How easy was it to learn to use hand gestures to control the system?**

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

2. **How easy was it to use both the hand gestures to control the system and sign using Sign Language at the same time?**

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

3. **How easy was it to understand what was said in the video?**

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

4. **Would you be happy using Sign Language video of this quality on a mobile phone?**

1	2	3	4	5
Not Happy at All	**Not Happy**	**Neutral**	**Happy**	**Very Happy**

5. **How clear were the hand gestures in the video?**

1	2	3	4	5
Not Clear at all	**Not Clear**	**Neutral**	**Clear**	**Very Clear**

6. **How clear were the facial expressions in the video?**

1	2	3	4	5
Not Clear at all	**Not Clear**	**Neutral**	**Clear**	**Very Clear**

7. **Please specify what you did not see clearly on the video?**

____

____

____

____

8. How easy was it for you to make a video message?

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

9. How easy was it for you to send a video message?

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

10. How easy was it for you to play a video message you received?

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

11. How easy was it for you to cancel recording a video message?

1	2	3	4	5
Very Difficult	**Difficult**	**Neutral**	**Easy**	**Very Easy**

12. Were there any hand gestures that were difficult?

1	2	3	4	5
Definitely No	**No**	**Maybe**	**Yes**	**Definitely Yes**

If there were which ones were difficult?

____

____

____

13. Were there any hand gestures that were tiring?

1	2	3	4	5
Definitely No	**No**	**Maybe**	**Yes**	**Definitely Yes**

If there were which ones were tiring?

____

____

____

14. Were there any hand gestures that you liked?

1	2	3	4	5
Definitely No	**No**	**Maybe**	**Yes**	**Definitely Yes**

If there were which ones did you like?

____

____

____

15. If the system was available are you willing to use it?

1	2	3	4	5
Definitely No	**No**	**Maybe**	**Yes**	**Definitely Yes**

16. How much are you willing to pay for the data charges to use it?

(The cost is measured per minute of the video message sent i.e. how much are you willing to pay for 1 minute of video)

1	2	3	4	5
0.0 Cents (Free)	**50 Cents**	**R1.00**	**R1.50**	**R2.00**

17. Please specify what you liked about the System?

____

____

____

____

18. Please specify any problems you experienced and what you did not like about the System?

____

____

____

____

19. General Comments.

____

____

____

____

Appendix C: Consent Form

I, ______________________________, fully understand the Gesture Based Interface for Asynchronous Video Communication project and agree to participate. I understand that all information that I provide will be kept confidential, and that my identity will not be revealed in any publication resulting from the research unless I choose to give permission. Furthermore, all recorded interview media and transcripts will be destroyed after the project is completed. I am also free to withdraw from the project at any time.

I understand that a South African Sign Language interpreter will provide sign language translation. That person is bound by a code of ethics that does not allow him/her to repeat any information that is given during the session. This means that my identity will remain confidential.

For further information, please do not hesitate to contact:

Tshifhiwa Ramuhaheli

Department of Computer Science

University of the Cape Town

Email: tramuhaheli@gmail.com or tramuhah@cs.uct.ac.za

Name: ____________________________________

Signature: __________________________________

Date:______________________________________

Appendix D: Experimental Data

D1: Mobile Prototype Questionnaire Data

These are the users' response to questions in Appendix B2: Computer and Mobile Phone Gesture Interface Questionnaire

Questions	Participants									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Q1	3	5	5	5	5	5	3	4	4	3
Q2	3	4	5	5	5	5	4	5	3	2
Q3	4	4	4	5	3	5	3	5	3	4
Q4	4	5	5	5	5	5	4	5	3	1
Q5	3	5	3	5	2	5	3	5	2	3
Q6	4	3	5	5	4	5	3	5	2	3
Q7	***This question required qualitative data***									
Q8	3	5	5	5	5	5	4	4	3	3
Q9	4	5	5	5	4	5	4	5	3	3
Q10	4	5	4	5	4	5	4	5	4	2
Q11	4	5	5	5	3	5	4	3	3	1
Q12	4	2	2	4	3	5	2	3	3	3
Q13	2	2	1	4	4	2	2	3	5	3
Q14	3	4	5	4	4	5	5	4	2	2
Q15	3	4	5	5	4	5	4	5	2	3
Q16	1	5	2	4	2	2	2	2	2	3

Table 4 : Data collected during the final prototype evaluation.

D2: Mobile Prototype Interview Data

The following responses were obtained using the questionnaire in Appendix B2.

-- Interview Session 1 -------------------------------------

Q5. How clear were the hand gestures in the video?

User1: sometimes it's fine sometimes it's not

Q7. Please specify what you did not see clearly on the video?

User1: The hand movement sometimes wasn't clear

User2: The hand shape wasn't clear because of the streakiness on the hand itself so in terms of that it wasn't clear. Other than that facial expression was fine, everything was fine. I just think that the background next time you should have a better background.

User1: I'll vouch for that too (referring to User2's comment on hand streakiness)

Q12. Were there any hand gestures that were difficult?

User1: I had a copper band and it didn't work so when I took it off. When you said I should try the wrist it was better. Sending was difficult.

User2: No I had no problems I was fine with the gestures.

Q14. Were there any hand gestures that you liked?

User1: Cancelling was quick so I like that one.

Q16. How much are you willing to pay for the data charges to use it?

User1: I say free. Deaf people they want things for free.

Q17. Please specify what you liked about the System?

User1: I thought that it was very nice but I had a bit of a problem with the background. If it was blue maybe or red you would be able to see better because then you are not able to see the eyes or hair that well with such a background because it's white on white, white against white so maybe if you had a different color background it might have worked better and with the hand gestures it became difficult. When signing the streakiness of the hands and sometimes the hands not being visible was an issue.

User2: It's a wonderful system and its something new I'm sure that will be great for Deaf people especially because Deaf people don't understand SMS you know so it's a better way of communication sometimes there miscommunication with SMS even so this is a better communication.

Q18. Please specify any problems you experienced and what you did not like about the System?

User1: No nothing else that I would like to add.

User2: I don't see any problems with the system other than what we have spoken about.

Q19. General Comments.

User2: I'll just like to suggest maybe try and hmm for example User1ny was having problems with the hand gestures and trying to get the computer to record maybe if you have done some training on it like you did earlier was to say maybe use your wrist instead of the hand because of the background. Such things like that would be able to help. But other than that everything was okay.

User1: I would reflect the same as User2 has said.

-- End of Interview Session 1 -------------------------------

-- Interview Session 2 -------------------------------------

Q7. Please specify what you did not see clearly on the video?

User3: I could understand the signing but the signs were slow and staggered. I think that sign language is very upbeat and there is lots of action happening so when you getting the slow signs it's basically perceived that the sender is slow or the person is not happy so we need to differentiate between whether the person is slow or whether the person is feeling that way because that's the way we perceive the people to be.

User4: I think the facial expression was not clear. That was all in terms of the facial expression.

User3: I would just like to add to that as well. I think it will be a problem if you can't see the face if the picture is dark because we rely a lot on facial expression and you need to see the face. If you can't see the face then there is problem because I don't know if **User4** has an attitude or what's going on with her face so it's very important for us to see the face.

Q12. Were there any hand gestures that were difficult?

User3: They were all the same because in the beginning I struggled but it was much easier with practice so no. Why are you asking the question? **(After the question was answered)** but now it's okay so it's fine. It was easy now.

User4: No

Q14. Were there any hand gestures that you liked?

User3: All of them were okay, I like all of them. So can I write all? And where do I write? **(Question was referring to where to write on the questionnaire)**

Q16. How much are you willing to pay for the data charges to use it?

User3: Maybe picture is expensive. Cheap, something cheap.

Q17. Please specify what you liked about the System?

User3: It's much easier to sign it's much easier to communicate via sign language. If I SMS using the words it's a struggle so if I'm able to use this type of technology it will be great.

User4: I would just like to add to that it's much easier for Deaf people to understand via sign language than via SMS because it waste time because if a word comes up and you say you don't understand what that word is you ask them what does this word means or you have to go research what this word means and they say you know very well and I don't know what you are talking about. There is a whole lot of miscommunication and wasting of money so if it's face to face and that will be great. It's much better in sign language.

Q18. Please specify any problems you experienced and what you did not like about the System?

User3: When User4 sent me her message it was very staggered and very slow. I don't know about my message but hers was very staggered and slow. **(This comment was made before we started answering the questionnaires)**

User4: Do we have to change because I saw that I was also very staggered when I saw my message so when it was sent I saw User3 message but it was very staggered. **(This comment was made before we started answering the questionnaires)**

User3: The reply or the message that I got was not of good quality. What I expected was not what I wanted. It was not clear. It should be like how I'm standing here signing that how the quality should be or the speed as well because when I received Ntomibi's message the signs where very very slow compared to the way she was signing towards the camera so it should be the same way as I'm signing right now or the way the person is signing should reflect within the message that's being sent.

-- End of Interview Session 2 -------------------------------

--- Interview Session 3 -------------------------------------

Q7. Please specify what you did not see clearly on the video?

User5: The gesture itself was distracting that's what I can say the fact that because you think they are saying something but they are not. They are trying to record or they are trying to do something so that was a bit distracting.

User6: You know it became an issue trying to use the gestures do we have to use a fist or can't we even upgrade it to having to use your finger because I don't like it, it should have been more like a touch screen, you just touch it and it goes off. It shouldn't be you know us having to struggle using our fist to try and get the icons to work.

User5: because of the sign "s" people would think that they are saying stop or "s" or something so it could be a sign so maybe if you can just eliminate that completely from the video it might just help.

Q9. How easy was it for you to send a video message?

User5: I still think that we should just be able to use our index finger to try and send something off instead of using our fist.

User6: Just like touch screen I think that you should maybe try to copy that kind of technology because it would be much easier if we are just able to use our index finger.

Q11. How easy was it for you to cancel recording a video message?

User5: It's the same things that we are reflecting we don't want to use the fist.

Q12. Were there any hand gestures that were difficult?

User6: For me I found it very hard to get the icons to work I think it would be best if you could just use your finger because when we using the hand try to you know maybe we should just use a finger and just pin point to the icon and it would go off immediately.

Q13. Were there any hand gestures that were tiring?

User6: It was just boring maybe you could just you know the touch everything maybe I would have been happy instead of using my hand to try to get this thing to work.

User5: No we all used to using our hands so it's fine.

Q14. Were there any hand gestures that you liked?

User5: I would have preferred to rather use my index finger to use the gestures you know the icons.

Q15. If the system was available are you willing to use it?

User5: Yes I would if it was available but would we be able to afford it though?

Q17. Please specify what you liked about the System?

User5: My only worry is if the phones would be compatible if you are going to use this in future would the phones be compatible and that is my only worry.

(After explaining to them it's very unlikely for it to work on all the phones)

User5: The thing is you know I don't think that as Deaf people would be able to afford something like that if you are talking about high end phones.

(After explaining about phone prices going down for current phones)

User5: I will be too old to even enjoy this technology then. What about government in terms of uplifting something like this but it might just delay things?

User6: Could I just ask a question on a phone like this is that an old phone or very recent phone? **(Referring to the Nokia N96)**

(After explaining the phone is 3 years old)

User6: You see so it's much cheaper now is it not or is it still expensive?

(After explaining it's around R5000)

User6: So if we want such how are we going to afford something like this? After 3 years already which means we have to wait another what? 3 years in order to afford something like this because Deaf people can't afford to have such high end phones we have to wait for so long until the amount of such a phone will be substantial for us to purchase.

(After explaining that in future this phones will be much cheaper)

User6: Because remember you just said that that is 3 years old and it will cost around R5000 right now. In future there will be new upcoming things and there might be new upcoming technology and then we will still struggle to have that type of technology.

(After explaining that N96 was a high end phone with lots of features and more memory)

User5: Seeing that it's got lots of memory and what not so is it susceptible to viruses?

Q18. Please specify any problems you experienced and what you did not like about the System?

User6: About the stands the phone stand, remember previously we just did the phones and we just did video calling on the phones. Now we doing the technology that's on the computer screen remember we had problems we had problem with the clarity of the phones and what not. On the computer was much better the quality of the picture was much better than on the phone.

User5: If there is an emergency situation and you want to develop something like this. It's taking already so long to record something and send it off and its already taking so long for the recipient to receive the message. What if there is an emergency?

User5: okay that's fine it might just take long. What I'm trying to say is that I feel that this technology is much better than SMS due to the fact that many. If I have to send an SMS to User6 maybe User6 won't understand because of the English and because they might not understand what message I'm trying to relate to User6 so the thing is that I'm pressing on the issue on the video calling and if there is an emergency situation we really need to up the game in terms of getting it to the person efficiently as possible and for the quality to be good and for it to go quickly.

User5: So do I have to have a 3G phone in order to video call or even to use this technology?

(After explaining that for a video call they need 3G and for store and forward it works even on GPRS)

User5: So it's still a no. I'm dying here.

-- End of Interview Session 3 --------------------------------

--- Interview Session 4 ---------------------------------------

Q7. Please specify what you did not see clearly on the video?

User7: It was a bit staggered I can say the picture, the quality is good of the picture but as soon as there is a movement then there is a bit of a problem there.

User8: It's clear but it depends on the area if there is a whole lot of lighting then it might have been better but it's very dark here in this area but if there was good lighting then it would have been okay it just depends on the situation because it's a bit dark here.

User8: I was quite surprised that you had a web cam and the phone so I didn't understand what is it about now. Why do have both? **(Referring to the webcam that was on top of the computer: the web cam was not connected to the computer)**

(After explaining the experimental set up)

User8: So you can use the phone and connect it to the screen why can't you use the phone screen instead of the computer?

(After explaining that the camera and phone screen are on opposite sides)

User8: Can't you use the camera in front? I have experience video calling on a Nokia there is a small picture and a bigger picture. I'm sure that sometimes you can close it and open it so why can't you use that the camera over there in front?

Q10. How easy was it for you to play a video message you received?

User8: Hahh it's so easy you repeating the same thing. I can understand it.

Q12. Were there any hand gestures that were difficult?

User8: To receive a message was difficult. To send is easy but when receiving a message and trying to play the message was quite difficult but I can't remember though. Cancelling that's easy and receiving is easy, sending is easy.

(After being asked which part of the gestures was hard)

User8: Cancelling and receiving a message was a bit a difficult to open a message.

Q13. Were there any hand gestures that were tiring?

User8: It wasn't comfortable for me the gestures aren't comfortable for me trying to get it to work wasn't comfortable. It wasn't comfortable. Could I make a suggestion maybe if you can touch it maybe it could work. I feel that maybe a keyboard might have even worked better than what we are trying to do right now. If you are able to do it like a touch screen or touch the icon then it would have been much better.

User8: Touch it on the screen. Touch it its nice its better because if they saying to me you know what are you doing there you waving at me or you. I can see it

User7: It's okay it's just that when someone is using the hand gestures you think they are waving or signing to you so once it's done so if you send they think you saying goodbye but you not saying good bye maybe you want to continue with the conversation. That gesture is not a good gesture because it either says goodbye so the waving of the hands indicate goodbye and that's not a good thing.

(After explaining about cutting out the part of the user selecting the system from the vide)

User7: My idea is if I'm going to sign a message and I send that message of then the facial expression is still there because I'm waiting for it to send. What's best is maybe to have something lower down the record icon is fine okay the send is fine but maybe take the send on the right hand side and put it at the bottom of the screen. It will make things so much idea.

User8: That's a good idea I can understand what he is trying to say it's because maybe there is something black and you can't see you know cancel down here at the bottom you know record here at the bottom where you can't see anything so put the icons at the bottom where it's not part of the screen so maybe you can zoom in and have all the icons at the bottom and it's not part of the picture. That's just an idea.

User7: I'll just like to add to that as well maybe have the icons at the bottom where it's not part of the message or part of screen.

User8: It will be easier instead of having your hand all over the place.

Q15. If the system was available are you willing to use it?

User8: No No No. I'm happy with the system but the way it's laid out it's not very great they way it's been processed is not very great maybe if you put the icons at the bottom it might just work out.

User7: Yes maybe maybe.

Q16. How much are you willing to pay for the data charges to use it?

User8: I don't know 50 cent it depends on Vodacom and MTN.

User7: It's just like SMS how much does SMS cost?

User8: 38 cents

User7: So normal

User8: It's normal for me so 50 cent I don't know.

Q17. Please specify what you liked about the System?

User8: I don't know it's normal.

User7: It's good but it needs more development and it needs to be more clear.

User8: I like the way it's structured I'm not sure if I'm right but you can use it on the computer can you use it on the TV as well?

User7: DTV

User8: DTV. TV not DTV on the TV if you at home can you plug it in on the TV?

(After explaining ideally we want it to work on TV without the computer)

User8: Okay if that's an option that will be better.

(After asking her if she means using the TV and phone without a computer)

User8: No not on the computer only on the TV when it's connected to a TV will it show up on a TV? Maybe I don't want to do this at home and I connect it to the laptop or the TV will it work?

(After explaining that ideally we want it to work on any TV that can take input from the phone)

User8: Okay we will check the TV out maybe we should do the TV. When are you going to do the TV?

User7: Go to …

User8: Now but I want to know on when they going to be testing it out on the TV.

User7: But they have all the work …

User8: But how do I know how to connect it that's the thing I want training on it I'd like to know how.

(After explaining about the problems on the phone and that there won't be any more tests)

Q18. Please specify any problems you experienced and what you did not like about the System?

User8: I don't know when you come and sit here I can understand maybe that's the issue I don't have a problem with the system. It was quite interesting to find out what was going on. I don't think I have any problems about it It's interesting.

User7: I think that you need to teach Deaf people on how to use the system to make them more motivated and interested in order to learn something like this because it's quite interesting and we would like to use this system so could you please train Deaf people on how to use it. That's my only issue.

Q19. General Comments.

User8: I want to continue with it and ask you on more issues but it was so short I don't think we had enough time to explore. Is this the final experiment that you are doing or are you going to do more experiments on this thing or what?

(After explaining this is probably the last one I'll be conducting)

User7: I just want to inform you that all in all it's great and in terms of confidentiality and consent form is great. Last week you had the cameras we basically did it on the phones and then afterwards a Deaf person has a got a video camera and they got that technology available so he has already known about how to video call. I haven't been prepared on how to communicate on video calling so how do they know that video calling is available? They say have known it from their friends and what not it was quite interesting.

(After explaining video call has been available for a number of years now)

User7: Yes but many Deaf people don't know about it. How is it that this person knows and hearing people know? Deaf people don't know about mxit they are like oh oh do you use mxit now? Thank you very much for the information you know. Twitter, 2go all of these things are coming up and we don't know about it. I understand what you are trying to say that we are behind in terms of it it's very bad.

(After telling them I'll let the other people in the project know so that they can update them on these things)

--- End of Interview Session 4 -------------------------------

-- Interview Session 5 ---------------------------------------

Q7. Please specify what you did not see clearly on the video?

User9: The movement wasn't as great as before so there was not enough facial expression so it was medium. It wasn't that wonderful.

User10: For me it was okay but the color was very dark the screen was very dark and the picture quality was very dark. Maybe if we can have a sort of a dull or alteration of the lighting because with the other one it was great it was clear but the color on the screen is not so wonderful so maybe we can change the lighting on the screen.

User9: I think that the picture quality is not that great either because it's not clear it is very staggered.

Q12. Were there any hand gestures that were difficult?

User9: No it was fine.

Q14. Were there any hand gestures that you liked?

User9: I can't choose one.

Q15. If the system was available are you willing to use it?

User10: Yes.

User9: On the computer or on the phone?

(After explaining that if it's available on both the phone and the computer)

User9: Yes

Q16. How much are you willing to pay for the data charges to use it?

User9: For one message?

(After explaining that per minute of a video message)

Q17. Please specify what you liked about the System?

User9: It's flexible there is nothing okay you are not able to touch anything but I think it's better because you don't have touch anything and make things dirty so in terms of that its better. So if you are sick you can communicate from far and there won't be any germs on the equipment.

User10: Like he said maybe you are sick you or don't know you have one arm or something you don't struggle to still touch anything so it's easier that way with the icons being that way. You can work with facial expression because it's live and you can use your facial expressions as well so it's simple and easy and comfortable to use.

User9: and if a person is not strong enough to press any buttons that helps a lot with the icons being you just wave your arms in the air or something.

User10: I like the fact that I like instant messaging you know I like instant messaging that can work very well. I just want to be clear on one issue if I want to contact User9 and I dial his number will it come up on the main screen or what on the software?

(After explaining that it works like mxit and the other user needs to be logged into the system before they can send the message)

User10: Do you think that for example would you be able to chat to many people at once such as mxit and when you know that they are logged in or off can I contact more than one person at a time if they are all logged on into the same site.

(After explaining that the aim is to make it work similarly to mxit but using video instead of text)

User10: Just one more question I'm a bit concerned do you think that if for instance seeing that the icons are so close to the person if I touch the icon and it just sends off but I didn't mean to send it off or I did something by mistake will it mess up the phone in any way.

(After explaining that it won't mess the phone and that the hand has to be over the icon for over a second before the option is activated so if they accidentally put their hand when signing it won't activate the option)

User10: I like that.

Q18. Please specify any problems you experienced and what you did not like about the System?

User9: The picture quality is not so good and the picture is staggered so if you want to sign fast the picture is staggered.

User10: Well the picture quality is a problem I would say that as well. The lighting isn't so wonderful so I need to control where I stand to be in a good lighting so I don't know if it's dark maybe or something happens or if you in an area where it's very dark is it not possible to have a light attached to the camera so it will be easier for the receiver to see the person so maybe if there is a light on the camera or something you know it might just work very well. Just like a torch on the phone maybe that could be developed as well because I don't think it has to be a strong light maybe a normal light it's just something that I'm concerned about.

(The following statement was made under the next question but it was a follow up on the statement made above)

User9: I would like to make a comment it's not a comment but we need to communicate with light you know it's just a comment that we need to communicate with light upon us because Deaf people can't communicate in the dark so there is need to be good lighting so that's a good idea but it mustn't be bright you know that you can't see what you are doing.

Q19. General Comments.

User9: I would like to make a comment it's not a comment but we need to communicate with light you know it's just a comment that we need to communicate with light upon us because Deaf people can't communicate in the dark so there is need to be good lighting so that's a good idea but it mustn't be bright you know that you can't see what you are doing.

--- End of Interview Session 5 --------------------------------

References

Ahmed, A. S. and Seong, D. S. 2006. SignWriting on mobile phones for the deaf. In Proceedings of the 3rd international Conference on Mobile Technology, Applications &Amp; Systems (Bangkok, Thailand, October 25 - 27, 2006). Mobility '06, vol. 270. ACM, New York, NY, 28.

Android. 2010. What is Android? http://developer.android.com/guide/basics/what-is-android.html. Last Accessed 24 February 2010.

Asmal K (2001). Plenary address. Proc. International Conference On Deaf Education: Building Bridges To Literacy, Johannesburg, South Africa.

Bowles J. B. (2007). Cell phone games for a CS2 data structures course. In *Proceedings of the 45th Annual Southeast Regional Conference* (Winston-Salem, North Carolina, March 23 - 24, 2007). ACM-SE 45. ACM Press, New York, NY, 300-303.

Cabral, M. C., Morimoto, C. H., and Zuffo, M. K. 2005. On the usability of gesture interfaces in virtual reality environments. In Proceedings of the 2005 Latin American Conference on Human-Computer interaction (Cuernavaca, Mexico, October 23 - 26, 2005). CLIHC '05, vol. 124. ACM, New York, NY, 100-108.

Cavender, R. E. Ladner, and E. A. Riskin, (2006). MobileASL:: intelligibility of sign language video as constrained by mobile phone technology. In *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA, October 23 - 25, 2006). Assets '06. ACM, New York, NY, 71-78.

Cell-Life (2010). Cellphones 4 HIV. http://www.cell-life.org/cellphones-4-hiv. Last Accessed 7 March 2010.

Coulton P., Omer Rashid , Edwards R., Thompson R. (2005), Creating entertainment applications for cellular phones, Computers in Entertainment (CIE)

Deaf Linx (2009). Deaf Culture. http://www.deaflinx.com/. Last Accessed 10 September 2009.

DeafSA (2010). Deaf Education: The current situation… http://www.deafsa.co.za/index-2.html. Last Accessed 7 March 2010.

Davis V., Gray J., and Jones J. (2005), Generative approaches for application tailoring of mobile devices. In Proceedings of the 43rd Annual Southeast Regional Conference - Volume 2 (Kennesaw, Georgia, March 18 - 20, 2005). ACM-SE 43. ACM Press, New York, NY, 237-241.

Desurvire, H., Lawrence, D., and Atwood, M. (1991). Empiricism versus judgment: comparing user interface evaluation methods on a new telephone-based interface. SIGCHI Bull. 23, 4 (Oct. 1991), 58-59.

Dörflinger, J., Friedland, C., Merz, C., and de Louw, R. 2009. Requirements of a mobile procurement framework for rural South Africa. In *Proceedings of the 6th international Conference on Mobile Technology, Application & Systems* (Nice, France, September 02 - 04, 2009). Mobility '09. ACM, New York, NY, 1-4.

Freeman W. T., Weissman C. D. (1995) "Television control by hand gestures" IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, Zurich

Glaser, M. (2000). A Field Trial and Evaluation of Telkom's Teldem Terminal in a Deaf Community in the Western Cape, Proceedings of the South African Telecommunication Networks and Applications Conference (SATNAC 2000), Stellenbosch, South Africa.

Glaser, M., and Aarons D. (2002). A Deaf adult literacy collective. *Stellenbosch Papers in Linguistics*, (34), pp. 1-18.

Glaser M. and Tucker W. D. (2001). Web-based Telephony Bridges for the Deaf. *Proc. South African Telecommunications Networks & Applications Conference (SATNAC 2001)*, Wild Coast Sun, South Africa, 672-677.

Glaser M. and Tucker W. D. (2004). Telecommunications bridging between Deaf and hearing users in South Africa. *Proc. Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI 2004),* Granada, Spain. (CD-ROM publication).

GSMArena.com 2008a. Nokia 5800 XpressMusic - Full Phone Specification. http://www.gsmarena.com/nokia_5800_xpressmusic-2537.php. Last Accessed 28 August 2009.

GSMArena.com 2008b. Nokia 6120 classic - Full Phone Specification. http://www.gsmarena.com/nokia_6120_classic-1958.php. Last Accessed 28 August 2009.

GSMArena.com 2008c. Nokia N96 - Full Phone Specification. http://www.gsmarena.com/nokia_n96-2253.php. Last Accessed 28 August 2009.

GSMArena.com 2008d. Samsung i8510 INNOV8 - Full Phone Specification. http://www.gsmarena.com/samsung_i8510_innov8-2471.php. Last Accessed 28 August 2009.

Ike T., Kishikawa N., Stenger B. (2007) A Real-Time Hand Gesture Interface Implemented on a Multi-Core Processor. Proc. Machine Vision Applications (MVA), pages 9-12, Tokyo, Japan, May 2007.

ITU. (2009) Measuring the Information Society - The ICT Development Index. Technical report, ITU, 2009.

J2ME (2007). Java ME Technology. http://java.sun.com/javame/technology/index.jsp. Last Accessed 5 July 2007.

Jones, M., & Marsden, G. (2006). Mobile Interaction Design. Chichester, West Sussex: John Wiley & Sons.

Kerr S, Thinyane H, Foster G, (2009). Mobile Phone Performance Analysis for Camera Based Visual Interactions. SAICSIT 2009.

Kiyaga NB and Moores DF (2003). Deafness in sub-Saharan Africa. American Annals of the Deaf.

Ma Z.Y. and Tucker W.D. (2007). Asynchronous video telephony for the Deaf. Proc. South African Telecommunications Networks and Applications Conference, (SATNAC 2007), Mauritius, 134-139.

Ma Z.Y. and Tucker W.D. (2008). Adapting x264 for asynchronous video telephony for the Deaf. Proc. South African Telecommunications Networks and Applications Conference, (SATNAC 2008), Wild Coast Sun.

Maunder, AJ, WD Tucker and Gary Marsden (2006) Evaluating the relevance of the 'Real Access' criteria as a framework for rural HCI research. In Proceedings 5th Conference on Human Computer Interaction in Southern Africa, Cape Town, South Africa.

Morawczynski O. and Miscione G., (2008). Examining Trust in Mobile Transactions: The Case M-PESA in Kenya. Pretoria: Human Choice and Computers 2008, pp. 287-311.

Muir L. J. and Richardson I. E. (2002). Video telephony for the deaf: analysis and development of an optimised video compression product. In *Proceedings of the Tenth ACM international Conference on Multimedia* (Juan-les-Pins, France, December 01 - 06, 2002). MULTIMEDIA '02. ACM, New York, NY, 650-652.

Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods. John Wiley & Sons, New York, NY.

Nielsen, J. (2005a). How to Conduct a Heuristic Evaluation. http://www.useit.com/papers/heuristic/heuristic_evaluation.html. Last accessed 25 February 2010.

Nielsen, J. (2005b). Ten Usability Heuristics. http://www.useit.com/papers/heuristic/heuristic_list.html. Last accessed 25 February 2010.

Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces, Proc. ACM CHI'90 Conf. (Seattle, WA, 1-5 April), 249-256.

Nielsen M., St¨orring M., Moeslund T., and Granum E (2003). A procedure for developing intuitive and ergonomic gesture interfaces for human computer interaction. In Proc. of the 5th Internation Gesture Workshop, GW 2003, Genova, Italy, April 2003.

Norman D.A. (1988). The Psychology of Everyday Things. Basic Books.

OpenCV (2010). OpenCV (Open Computer Vision). http://opencv.willowgarage.com/wiki/. Last Accessed 5 April 2010.

Ramamoorthy A, Vaswani N., Chaudhury S. and Banerjee S. (2003), Recognition of dynamic hand gesture, *Pattern Recogn.* **36**, pp. 2069–2081.

Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human and Computer Interaction. Addison-Wesley.

SLED (2009a). Sign Language Education and Development. http://www.sled.org.za Last Accessed 24 February 2010.

SLED (2009b). South African Sign language Stage 1 Learner Manual. Sign Language Education and Development.

Starner T. , Pentland A., Weaver J. (1998). Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.20 n.12, p.1371-1375.

Statistics South Africa (2001). http://www.statssa.gov.za/. Last Accessed 9 March 2009.

Symbian Developer Network (2008). Runtime Environment Feature Table. http://developer.symbian.com/main/documentation/runtime_environments/ . Last Accessed 26 August 2009.

Telecom Week. (2008) African Mobile Factbook 2008. Blycroft Publishing, 2008.

Tucker W. D., Glaser M. and Lewis J. (2003a). SoftBridge in Action: The First Deaf Telephony Pilot. *Proc. South African Telecommunications Networks & Applications Conference (SATNAC 2003)*, George, South Africa, II-293-294.

Tucker W., Blake E. and Glaser M. (2003b). Building Bridges for Deaf Telephony in South Africa: A Community -centred Approach. *IFIP Working Group 9.4 Newsletter*, 13(2), November 2003.

Umphress, D. A., Cross, J. H., Jain, J., Meda, N., and Barowski, L. A. 2004. Bringing J2ME industry practice into the classroom. In *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education* (Norfolk, Virginia, USA, March 03 - 07, 2004). SIGCSE '04. ACM Press, New York, NY, 301-305.

Wilson A. and Oliver N. (2003). GWindows: robust stereo vision for gesture-based control of windows. In *Proceedings of the 5th international Conference on Multimodal interfaces* (Vancouver, British Columbia, Canada, November 05 - 07, 2003). ICMI '03. ACM, New York, NY, 211-218.

Windows Phone. 2010. Windows Phone. http://developer.windowsphone.com/Default.aspx. Last Accessed 24 February 2010.

Yi L. and Tucker W.D. (2008). Kiara: an open source SIP system to support Deaf telephony. Proc. South African Telecommunications Networks and Applications Conference, (SATNAC 2008), Wild Coast Sun.

Zhang L., Chen X., Wang C., Chen Y., and Gao W. (2005). Recognition of sign language subwords based on boosted hidden Markov models. In *Proceedings of the 7th international Conference on Multimodal interfaces* (Torento, Italy, October 04 - 06). ICMI '05. ACM, New York, NY, 282-287.

Plagiarism Declaration

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own. Each contribution to, and quotation in, this document from the work(s) of other people has been attributed, and has been cited and referenced.

I acknowledge that copying someone else's work or part of it, is wrong, and declare that this is my own work.