Non-Visual User Interfaces

by

Clare Gwyneth Drake

A report submitted for the degree of

Bachelor of Engineering in Computer Systems Engineering

1 Palm Place

Wilson WA 6107

31 October 2003

Dr Syed Islam

Head of Department

Department of Electrical and Computer Engineering

Curtin University of Technology

GPO Box U1987

Perth WA 6845

Dear Dr Islam,

I hereby offer this thesis in partial satisfaction of the requirements of the degree of Bachelor of Engineering (Honours) in Computer Systems Engineering.

Yours sincerely,

Clare Drake

Abstract

This thesis investigates the necessity for, feasibility and construction of non-visual user interfaces to improve computer access for the visually impaired. The senses of hearing and touch are both appropriate for non-visual human-computer interaction, and this report presents and analyses different ways to use these senses to access a modern graphical user interface. A set of guidelines for the optimum way in which this can be done is followed by a specification of a complete non-visual user interface that meets these requirements.

Acknowledgements

Firstly to my supervisor, Iain Murray, thank you for all your guidance and encouragement.

This work would have been much the poorer without the help of Scott, Kerry, Ryan, Bianca, Gemma and Peter. Thank you so much for helping me understand the necessity of and requirements for non-visual user interfaces.

Thank you also to my mother Georgina for encouragement and superb editing, Martin for patience, Paul for support, and all three for helping me work through some of the knottier problems.

Nomenclature

Audicon	A non-speech sound used in a non-visual user interface. Can be either an auditory icon or an earcon.
Auditory icon	An everyday non-speech sound used in a non-visual user interface.
Earcon	A musical tone used in a non-visual user interface.
Filtear	A filter applied to an audicon to subtly change the sound, in order to provide additional information.
Force-feedback	A device providing output by means of forces.
GUI	Graphical User Interface
Haptic	Relating to the sense of touch. Includes both kinesthetic and tactile factors.
Kinesthetic	Relating to perception of the position of the body. This sense originates from receptors in the joints, muscles and tendons.
Manipulandum	The object providing the force in a force-feedback device. Usually held in the hand. Common examples are a thimble, stylus (pen-like object) and mouse.
NVUI	Non-Visual User Interface
Tactile	The sense originating from receptors in the skin. Usually used to mean the sense of pressure.
Vibrotactile	Stimulation of the tactile sense by a small vibrating object, usually a small metal rod.

Contents

List of Figures

1. INTRODUCTION

The rise in popularity and omnipresence of the personal computer is testament to its utility for a large majority of the population of developed countries. It has become essential equipment in most workplaces and many homes. Computer access can be particularly valuable for people who are visually impaired because it offers access to communication and information that is otherwise only available in print. However, people who are visually impaired can be further disadvantaged by modern reliance on personal computers. The graphical user interfaces of modern computers make computer access very difficult for these people.

Currently, many blind people use a screen reader and speech synthesiser as their only means of computer access. This provides access to information presented as text, but much of the information presented by a graphical user interface (or GUI) is non-textual. This makes computer use extremely difficult for those relying on synthesised speech as the only output.

Past research has frequently focussed on improving various components of non-visual human-computer interaction. Examples include optimisation of displays of virtual geometric shapes and particular applications such as graphing tools. Other research has investigated possible building blocks of non-visual user interfaces such as auditory icons or navigation tools. In contrast, relatively little research has been concerned with putting it all together into a complete interface allowing non-visual access to graphical user platforms such as Windows. Additionally, no comprehensive surveys of previous research in this field are available.

This thesis investigates alternative methods of human-computer interface that can improve computer access and make optimum use of the human senses other than sight. It begins with the provision of some more detail about computer access methods currently used by the visually impaired and why computer access is so important to this user group. In order for the interface to promote optimal use of the senses, it is necessary to have a basic understanding of their capabilities. Therefore, chapter three gives an introduction to the capabilities of the human sensory system.

Chapters four to six fill a major gap in the literature by providing a summary of previous research on auditory, haptic (relating to the sense of touch) and multimodal human-computer interaction. This includes a comparison and critical analysis of a broad sample of recent and past research on these subjects. This analysis leads to a set of considered guidelines for non-visual interface design in chapter seven, which essentially form the requirements for such an interface. This is followed by a specification of a complete non-visual user interface in chapter eight.

This thesis bridges a gap in non-visual human-computer interaction research by providing an analytical survey of many of the other research documents available. Its central purpose is to investigate a much-needed solution to the problem of providing the visually impaired with access to graphical user interfaces. The investigation begins in chapter two with a further description of this problem and its importance.

2. COMPUTER ACCESS BY THE VISUALLY IMPAIRED

2.1 The importance of computer accessibility

Access to computers can be particularly valuable for visually impaired people. It is a useful substitute for the sighted person's ubiquitous pen and paper that is faster, easier and much more easily copied than embossed braille. It is also a means of written communication with sighted people who cannot read braille. As previously mentioned, computers are often essential equipment in a workplace environment; so visually impaired people must be able to access computers as a part of meaningful employment.

Another very important facility provided by computers is access to the internet. Through email, the internet reduces isolation of people who may find mobility difficult. The independence provided by online instant access to information, some of which most people gain from print media, is also extremely valuable. Other facilities such as online shopping have the potential to make the lives of visually impaired people a lot easier.

In the early days of the personal computer, user interfaces were exclusively text-based. This text could be read from the screen by a screen review package and converted to synthetic speech. This provided visually impaired people complete access to the interface, as its content was entirely text and mostly serial.

The advent of graphical user interfaces (GUIs) made personal computing available to a larger user group. Most people find this method of interaction much easier and

more powerful. Unfortunately, however, the prevalence of GUIs has made computer access difficult for the visually impaired. Most personal computers now use an operating system that includes a graphical user interface, such as the various versions of Windows and the Macintosh operating system. It is rarely feasible to change the operating system of a computer in a classroom or workplace to accomadate a blind user. Additionally, most modern applications are based on these operating systems and therefore have graphical user interfaces. As discussed below, the current methods of access to GUIs are very difficult to use and omit much important information.

There are some circumstances under which a visual interface is less useful than an interface in another modality, even for a sighted user. When a lot of visual information is presented simultaneously, or when the eyes are otherwise occupied (for example when driving or adjusting lighting) it can be difficult to assimilate further visual information. Buxton (cited in Brewster 1991) suggested that "we are all visually disabled at various times": an overloaded or otherwise occupied visual channel impairs our ability to comprehend visual information. In addition, large visual displays are impractical in the rapidly growing field of mobile computing. Smaller, more practical screens reduce the usefulness of graphical interfaces. Walker et al (2001) proposed that this was the case for a calendar in a palmtop computer. Their alternative interface, using 3D sound, produced higher accuracy of recall and lower perceived workload than its small visual equivalent. It therefore appears that non-visual user interfaces could be extremely useful for sighted people as well as those with a visual impairment.

2.2   Commercial accessibility products

Several screen readers exist that convert text from the screen into braille or synthesised speech output. These provide basic interpretation from the GUI into a non-visual medium. Examples include JAWS from Freedom Scientific and GW Micro's Window-Eyes. The synthesised speech is usually provided by a software synthesiser such as Eloquence, although a hardware synthesiser can also be used. braille output can be produced using a refreshable electronic braille display.

Although screen readers allow blind people to use computers, much of the scope of their use is unavailable or extremely difficult. The advent of graphical user interfaces, while revolutionising the way sighted people use computers, made matters much more difficult for people who rely on speech synthesis of text to navigate the operating system. Graphical user interfaces present vast amounts of non-textual information. In addition, the text that exists is often positioned according to its context – text used to display the name of the active application is much different to the text contained within the document. Also, input to the interface is optimised for the use of the mouse, a device unusable for those who cannot see its cursor.

Roth et al (2000) point out that a major problem with existing access methods (such as braille displays or text-to-speech engines) is their linear nature. Graphical user interfaces, including that of the internet, are presented in a non-linear, two-dimensional manner. Another problem with screen readers is that they do not function well with non-standard application interfaces. A good example of this is the

numerous websites that use a graphical navigation method and have no plain text equivalent.

Additional hardware and specialised applications have been another method of attempting to provide a partial solution to the problem of making a visual interface accessible to other senses. Examples include keyboards that enable braille input using a standard QWERTY keyboard and stand-alone braille note-takers that can be connected to a PC. Keyboards with large (4 times normal size) keys can be useful for the partially sighted and those with movement or multiple disabilities. A range of talking devices and software is also available. Examples include keyboards and stand-alone typewriters; talking word processors, newspapers, web browsers and email clients. Some versions of the equipment mentioned above are available for quite reasonable prices (Silver & Gill, 2003).

The two major access methods for computer text are speech synthesizers and braille displays. Speech synthesizers do not require the user to learn braille, can present information very quickly, and can be implemented without purchasing any additional hardware. As a result, speech synthesis appears to be the prevalent method of text access in North America (Mynatt & Weber, 1994) and Australia. However, in Europe, braille displays are more popular due to "the lack of affordable speech synthesizers for many languages as well as the notational imprecision of speech" (Mynatt & Weber, 1994). Programmers also sometimes prefer braille as it forms a much clearer representation of code.

Unfortunately, both text access methods have other disadvantages. Users have reported confusion when synthesised speech is used for interface messages as well as content (Mynatt, 1997), even when very different 'voices' are used. A braille display used as an alternative to speech occupies the dominant hand of the user, making it difficult to provide any additional haptic feedback. It is clear from these examples that the capabilities of the senses must be considered when developing non-visual computer interfaces.

3. THE SENSES AND PERCEPTION

3.1 Introduction

The human senses are very different. For most people, sight is the primary sense that is used to gain information about our environment. Blind and partially sighted people have to rely more on their other senses for information. Each sense has its own advantages and disadvantages which should be acknowledged and considered when designing a non-visual user interface.

This chapter gives a brief summary of some characteristics of the senses of hearing and touch. Although much more is known about these senses, a full explanation of this knowledge would be outside the scope of this thesis and irrelevant to its subject. Therefore information relevant to the development of non-visual user interfaces has been summarised.

3.2 Speed of perception

One clear difference between the senses is the speed with which information can be assimilated. In computing, an information transfer speed is frequently called bandwidth, a term which can be extended to apply to the senses. Vision is the strongest sense in this regard; an estimate has put its bandwidth limitations at around 10^6 bits per second (Kekjer, cited in Way & Barner, 1997a). Information bandwidth limitations for the ear and skin were given at 10^4 and 10^2 bits per second respectively. However, this last value was obtained for vibrotactile stimulation (as of a vibrating pin) which may not give an indication of the tactile sense's full potential. In addition, other research has put this value at no more than 10 bits/s (Way &

Barner, 1997a). On the other hand, Mandic et al argued that the maximum information flow of the tactile sense is in fact around 10^6 bits per second at the receptor level and around 50 bits per second at the cognitive level, and can in fact be greater than that of the ear (Mandic et al, 2000).

Perhaps a more definitive comparison of the capabilities of the senses could be achieved through comparison of respective reading rates. Although not the most efficient way of utilising each sense, reading is an example of a task that requires full attention. Users commonly achieve braille reading rates of around 100 words per minute (Way & Barner, 1997). This can be compared to an average visual reading rate of around 250 words per minute and preferred rates of around 200 words per minute of synthetic speech (experienced users may prefer higher rates). These values could be converted to around 60, 150 and 120 bits per second respectively.

This obviously cannot give an indication of the full capacity of the senses, but does give a comparitive measure of comprehension speeds. A surprising result of this analysis is that it does not seem that differences between input rates are as large as previously expected. Therefore, major differences in computer access speeds must be largely due to user interface design.

3.3   Factors concerning perception

3.3.1   Cognitive

It is well-known that upon being presented with a new scene, people will firstly gain a quick impression of the whole before paying more attention to specific parts. This

breadth-then-depth search approach applies equally to the visual and tactile senses, and also to the auditory sense in a slightly lesser degree. Roth et al (2000) term this two-stage approach macro-analysis or "where" stage and then a micro-analysis or "what" stage to gain more detail. The serial approach to information created by the use of screen readers inhibits this two-stage search process.

The haptic and auditory senses of blind people have been shown to be no more acute than those of the general population. This disproves a common misconception that the other senses to some extent compensate for blindness. It is true, however, that blind people make better use of their other senses. A large amount of anecdotal evidence suggests that blind people often learn to extract more information from their haptic and auditory senses. For example, sighted people may find it very difficult to distinguish the patterns of dots that form braille code, but many blind people can read braille very easily. Blind people may also have learned to rely more on their memory than sighted people and thus may have better recall abilities.

3.3.2   Haptic

The word "haptic" is generally used to mean a combination of the sensing capabilities of the skin and the information about position and movement of the body. Oakley et al (2000) provided a good summary of terms relating to haptic perception, an excerpt of which follows as Table 3.1. In this report, "haptic" will be used to indicate a combination of tactile and kinesthetic capabilities. This section gives some information about the way in which haptic information is perceived, including some basic parameters of the haptic sense, the stages of perception

involved, its resolution and recognition capabilities, and the way the sense is affected by the number of contact points.

Table 3.1: Definition of terms relating to haptic perception

Proprioceptive	Relating to sensory information about the state of the body (including cutaneous and kinesthetic sensations).
Haptic	Relating to the sense of touch
Cutaneous	Pertaining to the skin itself or the skin as a sense organ
Tactile	Pertaining to touch (cutaneous)
Kinesthetic	Meaning the feeling of motion. Refers to sensations originating in muscles, tendons and joints.

The haptic sense is hampered by a comparatively low rate of information perception. In addition, the scope of haptic human-computer interaction is (mostly) limited to the pressure sensors in the fingertips and the kinaesthetic receptors in the hands and forearms. Generally, tactile perception has been found to vary little from person to person (Way & Barner, 1997). However, persons with diabetes, a common cause of blindness, sometimes experience reduced tactile sensitivity. Table 3.2 gives a summary of parameters relevant to tactile perception, an excerpt from a table found in Way & Barner (1997).

Table 3.2: Summary of parameters relevant to tactile perception

Factor	Parameters
Ratio of tactual to visual bandwidths	1:10000

Minimum discernible seperation of two points (static)	about 2.5mm
Minimum discernible displacement of a point on a smooth surface	0.002 mm
Height of braille dot	0.2-0.5mm
Minimum discernible seperation of grooves in grating (dynamic)	1.0mm
Resolution of expanded microcapsule paper	1-5 capsules/mm
Displacement of expanded microcapsule paper	0.2-1.0 mm
Resolution of human fingertip	about 1 dot/mm
Best size for tactile image	3-5in to a side

A haptic recognition task can be divided into four stages: detection, discrimination, identification and comprehension (Way & Barner, 1997a). This is a specific instance of the breadth-then-depth search as described above, but broken down into more detail. Detection and discrimination form the breadth search, in identifying the presence of an object and its independence from other objects. The depth search is made up of identifying the object and then understanding what it means.

Due to the low bandwidth and resolution of tactile perception, it is important to avoid information overload on this sense. If many objects are presented for tactile exploration, a person may easily forget the earlier objects due to the time taken to explore them all. The low bandwidth causes objects to be explored almost serially. This can cause difficulties in comparing dimensions or relative positions of objects (Magnussen et al, 2002). It also puts a larger demand on memory as an

understanding of a complex object must be constructed from serial understanding of its parts. The low resolution puts limits on the amount of detail that can be comprehended in a tactile manner; for instance, a detailed tactile image can be difficult to understand and much information is lost (Way & Barner, 1997a).

Several researchers have concluded that although many tactile patterns are possible and recognisable, only around eight to ten can be easily distinguished at one time (Challis, 2000). This has strong implications for the design of possible haptic icons in an interface.

When a three-dimensional object is explored haptically, usually all the fingers or even both hands are used to increase the bandwidth experienced. Unfortunately, most haptic human-computer interaction devices have restricted the user to a single point of contact, such as a fingertip or stylus. It has been shown that this restriction results in higher exploration times and error rates when exploring objects haptically (Kirkpatrick & Douglas, 2000).

3.3.3 Auditory

A major advantage of the human auditory sense is that it is inherently parallel and omnidirectional. A change in auditory stimulus is noticed regardless of direction and attention. Current use of audio in non-visual user interfaces uses synthetic speech almost exclusively. It is not possible to listen to more than one stream of synthetic speech at once. However, while listening to speech it is possible to notice a non-speech alert and decide whether to respond, without losing the sense of the speech.

This demonstrates a parallelism of the auditory sense that is not currently exploited in auditory user interfaces.

Sound has three basic components: pitch, intensity and timbre (Brewster, 1991). Pitch is related to frequency, intensity is the amplitude or volume of the sound, and timbre is determined by the shape of the waveform. Timbre is the quality which allows people to distinguish between different musical instruments. In a complex sound (such as that from a musical instrument) consisting of many frequencies of sine wave, the pitch is defined by the lowest frequency wave. In a complex sound, the frequencies are all multiples of the lowest, called the fundamental frequency. However, if the fundamental frequency is removed, the perception of pitch does not change. The fundamental frequency is somehow inferred by the listener (Brewster, 1991).

The relationship between pitch and frequency is non-linear. Pitch can also be affected by intensity. The pitch resolution of the human auditory system also varies with frequency. Brewster (1991) gave the approximation that a human can hear frequency changes of less than 1% at low frequencies, although this resolution increases with frequency.

The intensity range of human hearing varies from 6.5dB (the quietest sound that can be heard at 1kHz) to 120dB (louder sounds can cause damage to the ear). A 10dB increase in intensity of a sound causes its volume to double. Decibels, or dB, are the

logarithm of a ratio of the air pressure to a reference air pressure. 0dB is defined as the quietest sound a human can possibly hear, and occurs at around 3kHz.

Frequency also changes the perceived loudness of sounds. At frequencies lower than 100Hz, intensities of 20dB and more are required for the sound to be heard. Frequency perception also varies with age. A good range for sounds to be used as interface elements might be between 200Hz and 5kHz.

It is also possible to determine the direction of a source of a sound, which forms a fourth basic parameter to the sound. Perceived direction is generally on the horizontal plane and is perceived due to a combination of the difference in intensity and delay between the two ears. However, this is also dependent on frequency and can be ineffective at around 1.5kHz to 2kHz (Stevens & Newman, cited in Brewster 1991). In addition, it can be difficult for the listener to tell if a sound source is directly behind or in front of them.

The pinnae, the external part of the ears, can enable a listener to determine the vertical height of a sound source. This can be measured in the form of a Head-Related Transfer Function (HTRF) which is then used to simulate vertical positioning (Roth et al, 2000). Software libraries which include generalised HRTFs are available from Intel and Microsoft and these have been used with some success. Unfortunately, these functions are dependent on the individual user's head shape and thus it could be difficult to use this effect in a commercial software package.

It has been found that pitch can be successfully used to represent vertical position. Mansur, Blattner & Joy (cited in Brewster 1991) found that people naturally perceive a higher pitch to come from a spatially higher source. Some researchers have used this relationship between pitch and perceived height to add a third dimension to artificial sound (for example, Kamel, Roth & Sinha, 2001; Yu & Brewster 2002).

Different sounds will tend to be grouped by the listener into different 'sources'. The parameters by which listeners distinguish sounds are listed by Brewster (1991) as follows: fundamental frequency, location, rhythm and musical key. Perhaps the most powerful distinguishing parameter is the spatial location of the sound. The well-known "cocktail party effect" is based on this premise; a person can very effectively concentrate on one conversation in a noisy room with many conversations.

The human auditory sense is very temporally sensitive. Responses to auditory stimuli are often faster than those of any other sense. This is aided by the omnidirectional nature of sound, in that a person will notice a sound even if their attention is elsewhere. Small changes in frequency are easily noticed, provided the change is abrupt. Changes in rhythm are also very noticeable.

In summary, there are many facets of sound, all of which can give different messages to the listener. In addition, the auditory system allows for specific attention focus, fast reaction time, and a parallelism which allows a listener to respond concurrently to different stimulae. Only a very small part of these abilities are used by listening to

synthetic speech. Therefore, the next chapter details ways in which an interface can make better use of the human auditory sense.

4.   AUDITORY HUMAN-COMPUTER INTERACTION

4.1   The use of the human auditory system

Synthesised speech from a screen reader is currently the most common method of non-visual computer access. As described in chapter two, this approach takes a long time to deliver interface information. In addition, much information is not sonified.

Synthesised speech is, however, not the only method of auditory output of which a computer is capable. Numerous researchers have investigated the use of non-speech sounds to assist or replace speech in delivering interface information. This would take better advantage of both the capabilities of computer output and the human auditory sense.

While synthesised speech usually takes too long to deliver an interface message, it is unambiguous and the meaning does not have to be learned. This is a strong argument for presenting synthesised speech messages as an optional addition to the interface - the speech output can aid new users in their learning of an auditory environment.

Another aspect of the auditory system is that it is very sensitive to changes in status. In particular, a listener can become accustomed to a noise and ignore it; however, it is a noticeable event when it stops (Raman, 1997; Brewster, 1991). This encourages the use of continuous sounds for monitoring continuous processes, such as an internet download.

Brewster, Wright & Edwards (1995a) suggested that sound could be used to represent otherwise hidden information, even in GUIs intended for sighted users. Information might be hidden because it is not available on the display (due to screen size or other hardware constraints), it requires actions to access, it is ignored because the visual channel is overloaded, or it is not noticed because the user's attention is elsewhere on the screen. The authors also suggest that sound feedback could reduce user errors, particularly when the user's attention is elsewhere. This frequently occurs, for example, when scrolling through a document, or if a button is clicked and the visual focus quickly changes to another area of the screen.

The remainder of this chapter will describe some of the research that has been conducted on auditory human-computer interfaces. Firstly, individual applications which have been developed with auditory interfaces will be considered. This will be followed by a summary of some methods and techniques that can be used in auditory interfaces, and then some developments designed to enhance a GUI with auditory output will be summarised. Navigation in a purely auditory interface is often found very difficult, so some tools which have been developed to aid navigation are presented. A complete description of the large body of research on this subject is outside the scope of this paper. What is given here is therefore a summary of the most relevent investigations and developments.

4.2   Auditory interfaces for specific applications

Roth et al. (2000) developed several tools to enable access to the internet using 3D audio. Their software is used in conjunction with a touch screen or graphic tablet to

give spatial meaning to input. The system can then adjust the location of its feedback according to the user's finger position on the tablet. Different sound characteristics of the feedback (for example pitch and timbre) are used to indicate different user interface elements or HTML tags. The authors have used a software library based on Head Related Transfer Functions to simulate the vertical dimension of the 3D sound.

Roth et al. (2000) also produced several sonic games to provide a learning tool for the computer system. The games take the form of identifying locations and sequences, join-the-dots, and Memory/Concentration where a user has to pair auditory representations of geometric shapes. Another of their developments was an audio drawing tool. This made continuous sounds which changed when the cursor was on a line or inside a closed element.

Brewster (1991) provided an overview of Soundtrack, a word processing application with an auditory interface. The initial screen of Soundtrack was divided into a 4x2 grid. Operated using a mouse, the interface produced a different tone when the cursor entered each area. A warning sound indicated the edge of the screen, and if the user 'got lost' at any time, a click would read the name of the object under the cursor. A double-click was used to execute a command or open a menu. Although Soundtrack was one application and thus did not allow access to an entire user interface, it demonstrated that it is possible to understand and navigate in a purely auditory interface. It also demonstrated the feasibility of the mouse as an input device to non-visual user interfaces.

Another application was described by Edwards, Stevens & Pitt (1995). It consisted of two tools that can help blind people with mathematics. Mathtalk enables meaningful access to algebraic equations, using structured speech to present the notation in a way that can be explored and repeated, and also allows manipulation of the equations. Soundgraph is a simple auditory graphing program that allows reading and creation of simple line graphs. Partially sighted users have also found Soundgraph useful, as the screen display is supplemented by the audio output.

Drewes, Mynatt & Gandy (2000) created an audio-only game with a detective theme. They carefully considered the difficulties of an audio environment, particularly the amount of detail. The game was structured like rooms in a house, with footsteps, turning sounds and a collision sound for navigation. Ambient noise provided a means of identifying the room, and dialogue provided information about the non-player characters in each room. User testing revealed that subjects enjoyed the game, but found that it was difficult to identify the rooms. It seems that this is due to sounds with unclear meanings, rather than a lack of information. Every subject reported using the collision sound to find walls, which emphasises the need for navigational reference points.

4.3   Methods of audio presentation

A large amount of research on auditory interface development has concentrated on different ways to sonically represent user interface elements such as buttons, menus and scrollbars. Auditory icons and earcons are the two most common approaches to the solution of this problem.

Auditory icons were used as early as 1986 (Gaver, cited in Brewster 1991). Previous work on everyday sound recognition showed accuracies of as much as 95%. It is suggested that everyday sounds can, through various properties, convey information about the object that made the sound: "Another important property of everyday sounds is that they can convey multidimensional data. When a door slams a listener may hear: the size and material of the door; the force that was used; and the size of room on which it was slammed" (Brewster, 1991). Brewster suggests that similar properties could be employed to give detailed information about user interface elements.

Earcons are abstract musical tones that are used to sonically represent messages. They can vary by timbre (musical instrument), register, pitch, rhythm, duration, intensity, and location. Guidelines written after much research on earcons by Brewster and his colleagues (Brewster 1991; Brewster, Wright, & Edwards 1995a) are as follows:

- Timbres should be from distinctly different instruments to aid recognition. Timbres should have multiple harmonics (reasonably complex sound).

- Pitch and register cannot be relied upon to distinguish earcons. Large differences (for example two octaves) can be sufficient for absolute recognition, and smaller differences can be used for relative recognition (forming a tune).

- Very different rhythms make earcons very easy to tell apart.

- Earcons should be short so that they can keep up with system events and user interaction.

- Intensity should be kept to a small range so that users can control the volume without losing any information.

- Spatial stereo location of earcons can be particularly helpful if they are played in parallel.

Brewster, Wright & Edwards (1995b) noted that earcons take around 1.3 – 2.6 seconds to play, which is too slow to keep up with user interaction. They proposed parallel earcons as a solution to this – if two earcons are sufficiently different, they can be played at the same time. Previous research (Gerth, cited in Brewster, Wright & Edwards 1995b) had shown that playing two sounds at once did not significantly reduce recognition rates, but playing three or more sounds did. This suggests that two earcons played in parallel could be effectively recognised.

Brewster , Wright & Edwards (1995b) tested prototype parallel earcons to confirm this hypothesis. Statistical analysis of their results indicated that the parallel earcons were recognised just as well as their serial equivalent. There was a significant increase in recognition rates when the test was repeated, due to the increased experience of the subjects. However, three participants were discarded from the testing because they did not reach required recognition levels in the training phase. Two of these reached the required level after more training, but one did not. The authors suggest that this may be due to tone-deafness; however, this is an important consideration when designing an auditory interface to be used without visual backup. Of the results from the accepted participants, there was no significant difference between musically trained people and non-musically trained.

A strong argument for earcons is their structured composition. In this way, they show similarities between elements or events. This is particularly valuable when learning the interface, as (for example) a newly encountered error noise will sound similar to other errors and thus will be interpreted as an error, even if the user does not yet know what type. Brewster (1991) proposes that the strongly structured and rhythmic composition of earcons will also aid the users' memory when learning the earcons.

There is continuing debate amongst researchers as to whether earcons or auditory icons are preferable for use in auditory interfaces. Auditory icons are easier to learn, but can become annoying (Sikora et al, cited in Bussemakers & de Hann 2000); in addition, as Brewster (1991, p. 26) says, "some interface actions and objects have no obvious representation in sound." Earcons are abstract sounds and thus can take longer to learn, but have the advantage of a firm structure which can group similar elements or events conceptually and in the interface. In addition, user interface concepts are largely abstract and thus may be suited to the more abstract tones.

Brewster (2002) suggests that earcons and auditory icons could be used simultaneously. For instance, if auditory icons were used to represent interface elements that are unstructured and have an easily associated real-world sound, while earcons were used to represent the more structured components of a user interface (for example a scrollbar) the advantages of both approaches could be exploited. Brewster did not mention that this flexible approach could also be of benefit when adapting the interface to beginner or expert users.

A number of researchers have suggested that an interface sound can be altered to give information about the properties of that object (Gaver, and Ludwig, Pincever & Cohen, cited in Mynatt 1997). These modifications are termed filtears. Mynatt (1997) used low-pass filters and high-pass filters to give the effects of muffling and thinning respectively. Muffling is a useful technique because it can give the impression of an object being unavailable without removing it and changing the layout; just as objects can be 'greyed out' in a graphical user interface. Thinning was used in this case to indicate that an object is selected, which is also a very important piece of information. Mynatt also modified the sounds of objects such as menu items and buttons to indicate their position in a list.

Ramloll (2001) tested comprehension of tabular data using synthetic speech output with and without pitch feedback. The pitch output was relative to the value in the current table cell, and was played in stereo depending on the current x position in the table. Use of the NASA task load index indicated that the addition of pitch dependent on the value of the table cell significantly decreased mental demand, frustration and overall workload. Accuracy was also significantly increased.

Petrie et al (1997) constructed a prototype auditory interface (which they called DAHNI) to meet the multimedia access requirements they researched and defined (see chapter seven). 'Buttons' for navigation and other functions were grouped in the shape of an 'H' on sideways. These could be accessed using a keyboard, joystick, or touch tablet with tactile overlay. The interface was tested with blind users in a

sample hypermedia environment. They found it easy to use and particularly easy to learn.

4.4   Audio feedback as a supplement to a visual interface

The SonicFinder was an auditory enhancement to the graphical Finder tool on an Apple Macintosh (Brewster, 1991). It used auditory icons such as tapping noises when something was selected. The tone of the sound (paper, wood, metal) gave the type of object, and the deepness indicated its approximate size. Copying was given the sound of pouring liquid into a receptacle - the copying was finished when the receptacle sounded full. Other sounds were added to delete, dragging, scrolling and zoom actions.

Gaver, Smith & O'Shea (cited in Brewster 1991) created and tested auditory enhancements to the interface to the control of a bottling plant. They found that auditory queues were very useful in monitoring continuous processes, and that the auditory feedback enabled users to monitor the many machines in the plant simultaneously. However, sometimes users did not notice that a machine had stopped working because its noise had simply stopped. Gaver et al also found that users moved quickly to stop failure sounds that were demanding without firstly considering their cause, and this could cause the user to delay attention to the cause of this and more serious problems.

An earcon-enhanced scrollbar was constructed by Brewster, Wright & Edwards (1995a) using the guidelines presented by Brewster (1991). This scrollbar was tested

using two tasks – a search task, to find a significant feature in a document, and a navigation task to find a particular point in the document. A visual-only scrollbar was also tested as a control. Total task time and number of errors were recorded, as was subjective workload using the NASA Task Load Index. This requires participants to estimate the mental demand, physical demand, time pressure, effort expended, performance level achieved and frustration experienced. Brewster et al added a seventh measure: annoyance, since this is a common concern with auditory interfaces. They also asked participants to rate which scrollbar they preferred overall.

The subjective results showed that there was a significant decrease in mental demand using the auditory-enhanced scrollbar as opposed to the purely visual case. Mental demand was also rated the highest of the seven workload measures. There was no significant difference in the other task load measures, which indicated that the auditory scrollbar was not found significantly more annoying than the purely visual one. Overall the auditory scrollbar was rated as significantly better.

The earcon-enhanced scrollbar experiment was valuable in its own right, but in this context it formed an example of how to apply the guidelines presented in Brewster's paper and how to test auditory-enhanced interfaces. Although the author concentrated on auditory feedback as an enhancement to a visual interface, the guidelines would also apply to audio-haptic or purely auditory interfaces. The feedback required would obviously be more detailed, but the principles of earcon design presented here can be applied to any audio component of an interface.

In summary, Gaver showed that different real-world sounds (auditory icons) could be used to represent complex attributes of abstract objects. However, it is important that sounds are not more demanding than is merited; additionally, demanding error sounds should automatically cease after a short period so that users' attention can be focussed on solving the problem. Auditory implementations of abstract user interface objects can also use earcons to beneficial effect.

4.5   Auditory navigation tools

Kamel, Roth & Sinha (2001) investigated the use of 3D audio to provide meaningful spatial representations in the use of non-speech audio. They used stereo sound to represent geometrical shapes including triangles, squares and circles. The stereo effect was deemed to be sufficient to represent the horizontal plane, whereas the vertical plane was supplemented by increasing or decreasing the frequency of the sound. Beeps were added to represent line junctions and corners.

Firstly, an input tablet was used. Each shape was given a unique physical position and could be activated by touching that area with a stylus (pen-like pointing instrument). An audio grid was then added to the interface. This was constructed by using different musical timbres to represent horizontal and vertical axes. In the 3 by 3 grid, two consecutive notes were used to represent crossing a grid line, depending on the line. For example, crossing the line from (1,1) to (1,2) gave a clarinet tone of C then G, and crossing from (1,2) to (1,3) gave a clarinet tone of G then high C. Thirdly, the presentation of the geometric shapes were available over the whole

tablet, with their sounds getting softer as the user moves away, and at the loudest when the stylus is in the centre of the shape.

Test participants were asked to explore the virtual environment (which contained the three shapes and one of the three test cases – no feedback, grid or shapes) and then locate one of the shapes. Overall, the grid condition was found to be best in terms of exploration and relocation tasks, user confidence, and accuracy. The authors therefore concluded that it is beneficial to partition work space into subsections to aid navigation.

An interesting idea proposed by Gaver & Smith (cited in Brewster 1991) is that of "soundholders". These function like a sonic bookmark - they can be placed anywhere and constantly produce a simple noise, which gets louder as the interaction point moves closer to the soundholder. Gaver & Smith suggested that environmental noises such as a bird call or burbling stream could be used in this way. Whatever sound is used, soundholders could provide a novel navigational aid in an auditory environment.

It can be seen through observation of the use of the current speech-only interfaces that navigation is extremely difficult. The preceding researchers have attempted to improve this problem by implementing different navigational tools. However, in a user interface it is important that the non-visual implementation corresponds to the graphical version to enable collaboration between blind and sighted co-workers. This would not be possible in a grid-orinted interface like that proposed by Kamel, Roth

and Sinha (2001) or as used in Soundtrack (Edwards, cited in Brewster 1991). One technique that has not been well investigated is the possibility of a sound which changes continuously in pitch and stereo position as the mouse cursor moves across the screen.

5. HAPTIC HUMAN-COMPUTER INTERACTION

5.1 Mice

5.1.1 Devices

Much effort has been put into the possibility of a haptically enhanced replacement for the common mouse. This has ranged from a vibrating mouse which reflects sounds in multimedia applications and is primarily intended to enhance the experience of a sighted user, to devices that attempt to represent any screen display in a tactile manner. Haptic mice can be divided into two categories: mice with tactile displays mounted on them, and force-feedback mice that simulate objects and textures using forces.

A good example of the second is ScreenRover from Betacom Corporation Inc. This is a motorised mouse which uses force-feedback to help the user explore the 'screen': "Screen Rover generates a variety of pull-push effects to distinguish icons, controls and text; guides the user to the active area of windows and their elements; lets the user position and size jpeg's and bitmaps; distinguishes frames on websites and multiple windows; [and] reads any text on the user's screen." (Silver & Gill, 2003)

Another force-feedback mouse, the Virtual Reality Mouse, was released in 1995 (Cullen 1997; Sjöström 1999). It consists of a mouse attached to a force generator by a bar at the front (where the cord usually is). This allows output of up to 1kg of force. The Virtual Reality Mouse was designed to work with Windows 95 and 3.x, haptically rendering Windows display elements. It is compatible with several common screen readers and magnification programs.

The Feelit mouse from Immersion Corporation is a serious attempt at haptically enhancing Windows. It directly represents all display elements via a force-feedback mouse on a mouse-pad sized workspace. The Feelit mouse has a small workspace and low force capabilities; however, it is intended for the general market and is thus made affordable. The concept was taken over and commercialised by Logitech for their Wingman force-feedback mouse, which is very similar.

Figure 5.1: Logitech Wingman force-feedback mouse

O'Modhrain and Gillespie (1995) constructed a prototype force-feedback mouse for user interface display to the blind called the Moose. Windows elements are represented using variable resistance to create haptic icons ('hapticons'): individual force characteristics that identify the element. In addition, actions such as 'drag-and-drop' can be represented haptically – in this case the puck seems heavier when an item is being dragged. The mouse is perpendicularly connected by flexible spring steel to a workspace tablet. The effective workspace is 3cm square.

A very early example of placing a tactile display on top of a mouse was developed by Terry and Hsaio (1988). A small pin, mounted on the top of a mouse, vibrated when the cursor was over a predefined screen target. Experimentation showed that the tactile feedback slightly reduced targeting time when used as a supplement to visual targeting tasks.

A few years later, the Virtac system was designed by Jubinski (1992). The goal was to provide tactile access to the entire screen of a IBM PC by attaching a rectangular array of braille-like pins to the top of a mouse. In this way, the user could explore the whole screen a bit at a time. Virtac was designed to eavesdrop on the visual memory of the computer without adding any extra load to the CPU. However, no mention of optical character recognition was made, so it is assumed that it would not be possible to read text using this system, unless the text was large enough for its pictorial representation to be recognised.

Another example is the VTPlayer from VirTouch (2003). This looks like a normal mouse except that it has two 16-pin displays, similar to those on a refreshable braille display, mounted where the user's first two fingers sit on top of the mouse. These form a tactile representation of the screen environment under the cursor. Input functions like a normal mouse are still available. The predecessor of the VTPlayer was called the VirTouch Mouse and followed the same principle, except that 32 pins were used under each of 3 fingertips, giving a better resolution. A wider range of specifically designed software (including tutors and games) is available with the more compact VTPlayer.

Figure 5.2: VTPlayer

5.1.2   Research

Wies et al (2000) presented an attempt to use a force-feedback mouse to improve introductory understanding of electric field theory. Testers reported that their understanding was significantly improved by the program.

Sjöström used the FeelIt mouse prototype, extending its basic graphic-to-tactile conversion to make the most of the haptic modality (Sjöström, 1999). He proposed several virtual haptic search tools to improve the difficulty of finding isolated objects, such as icons on a desktop. Sjöström implemented a search tool in the form of a virtual cross, which provided feedback when the cursor was horizontally or vertically lined up with an object. Users found the tool helpful, but found the small workspace of the Feelit mouse very difficult. Other proposed search tools included a magnet-style tool, which pulls the cursor towards a nearby object, and a 'ball' tool that allows users to feel objects at a distance but with less detail.

The Feelit and Wingman mice were designed to enhance a visual interface. Advertising material suggests that users will find games more enjoyable and find

work easier with haptic feedback. This second suggestion was confirmed by Oakley et al (2000) who conducted usability testing on haptically-enhanced scrollbars and buttons using the Phantom. They found that effort and frustration were significantly reduced when using a haptic recess effect.

Yu & Brewster (2002) used standard interface methods on the Phantom and Logitech Wingman mouse to render line graphs. They added optical scanning and image filtering software to enable blind users to explore conversions of printed graphs. The line was represented as a groove in the back wall of the Phantom's workspace and as a 'sticky' line for the Wingman mouse. The authors noted that although the rendering was quite different, the devices required the same data input - the start, end and intermediate points of the lines.

5.2 Three-dimensional haptic devices

5.2.1 Devices and overview

Three-dimensional haptic devices have proved very attractive to researchers due to the flexibility implied by the use of three dimensions. In particular, they can present all possibilities of virtual haptic environments for investigation. The Phantom and Impulse engine are the most commonly-used devices in this category. All three-dimensional haptic devices are very expensive, however, and thus are not feasable for widespread use for user interface access.

The Phantom, from SensAble Devices, is a robot that provides one-point force-feedback in three dimensions. Three motors provide applied forces based on the

manipulandum position to give the effect of a solid object in 3D space. The Phantom workspace volume is 13x18x25 cm with a position resolution of less than 1 mm. Optical sensors monitor the user's movements; this information can then be used to calculate feedback. The feedback is given by a manipulandum attached to the end of the robot arm via a freely rotatable gimble. The manipulandii provided are a stylus and a thimble (Sjöström, 1999).

Figure 5.3: The Phantom

The Immersion Impulse Engine consists of a pen-like probe which can move in three dimensions. The device provides up to about 8N of force-feedback dependant on the position of the probe (Colwell et al, 1998).

Figure 5.4: The Immersion Impulse Engine
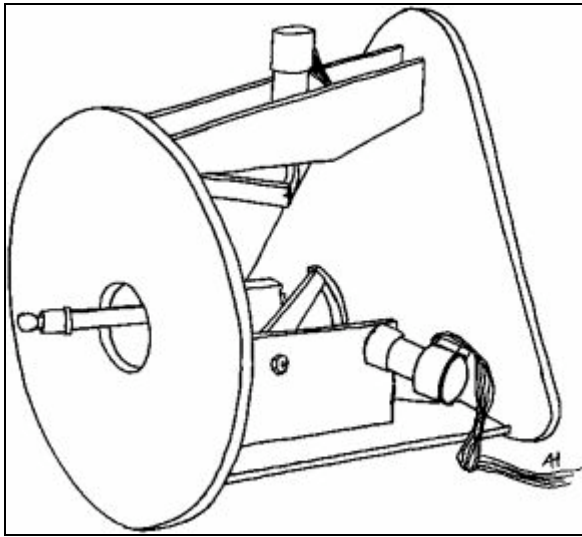
5.2.2   Research

The Phantom has been used as the hardware for a large amount of research into virtual haptics and their application. No software was provided with the first version of the Phantom, allowing a large scope for haptic programming research. However, two commercial software development kits for the Phantom have since been developed. This has allowed greater and easier experimentation using this device.

One of the earliest significant researchers who used the Phantom was Minsky (1995), who wrote programs to simulate complex virtual textures such as sandpaper. Numerous researchers have since investigated other virtual textures. Extensive research, programming and usability testing with the Phantom has been conducted by Sjöström and colleagues at the Certec rehabilitation research centre at Lund University of Technology, Sweden. Fritz & Barner (1999) investigated textures, force fields and mathematical plots, whereas Oakley et al (2000) researched 3D haptic augmentations to a graphical user interface.

Fritz & Barner (1999) aimed to make mathematical graphs accessible to blind people. They considered one, two and three dimensional data. A detailed description of the mathematical process required to translate an idea into a force profile suitable for implementation on the Phantom was given. This included force vector derivation, texture, and resolution. Modelling of objects is also considered; the authors here use a line of attractive force to render a line. This attractive force is similar to that which would be felt if the line was an iron wire and the manipulandum a magnet. Three dimensional vector fields were represented directly as force fields. These are useful to aid in the understanding of subjects such as electrical fields, fluid dynamics and gradients.

In the basic Phantom toolkit, 3D surfaces can be modelled but they all have the same 'slippery' surface texture. This does not provide any information about data values on a mathematical surface graph and it can be easy to slip off the surface. Fritz & Barner (1999) continued by suggesting grid lines or tick marks to provide a frame of reference in a haptic graph. Data values could be estimated using these marks or presented in a different way, for example speech. Surface friction can be used to improve users' adherence to the surface or to represent other information such as emulate physical resistance in a fluid. Fritz & Barner also consider various methods for creating virtual texture. These textures can be used to give a sense of colour depth, a smooth texture indicating a light colour and a rough texture indicating darker areas. When implementing lines and surfaces to represent mathematical equations, Fritz & Barner used virtual walls to make it easier for a user to find the relevant part of space.

Oakley et al (2000) used the Phantom to investigate the use of touch as a way to reduce visual overload on conventional user interfaces. They represented a screen by using the 'back wall' of the Phantom's operating space. This was used in conjunction with an adjacent visual monitor. The experiment was concerned with enhancing standard user interface buttons using haptics to decrease error rates and increase usability. Four effects were used to augment the buttons: texture (a ripple pattern), friction, recessing (or engraving) and gravity.

Oakley et al (2000) conducted usability testing on their button enhancements, measuring task time, error rates, and subjective workload via a modified version of the NASA Task Load Index. They found that the gravity effect significantly reduced user errors when compared to the control (no haptic effect), while the texture effect actually increased errors. Correspondingly, the gravity effect reduced subjective workload measures, while the texture effect increased workload.

A second experiment was conducted by Oakley et al (2000) to measure the usability of the gravity condition in a more realistic task. A gravity and recess enhanced scrollbar was tested in an experiment designed to simulate normal operation, where a scrollbar is frequently needed but the visual focus is elsewhere on the screen (in the text area). The authors found that the haptic enhancement decreased perceived effort and frustration, but had no significant effect on fatigue or mental demand.

Sjöström and colleagues at Certec research centre, Lund University, Sweden, have spent many years on rehabilitation engineering research. Much of this research has used the Phantom as part of numerous tests and evaluations to clarify the extent to which a virtual environment is understandable and useful to blind people. These have involved programs for games, mathematics, drawing and navigating. The programs are also seen as investigatory steps on the way to a future complete haptic interface called Touch Windows.

One of their first programs was called the Memory House. This was an array of buttons embossed on the 'back wall' of the workspace which made a sound when pressed. The object of the game was to eliminate buttons by pressing buttons with matching sounds, similar to the game 'Concentration'. Tests conducted on the Memory House showed that almost all blind users were able to complete the game. Two out of the nine were unable to finish, although they did find some pairs. These users may have benefited from more practice with the device and/or a tutorial on its use. It is especially interesting to note that several blind users finished the game using a similar number of button presses to that of a sighted group of test subjects using a mouse. This, along with the fact that seven users finished the game, would seem to show that it is possible to understand and remember a reasonably complex virtual haptic environment.

Table 5.1: Applications developed and tested by Sjöström and colleagues

(from Sjöström, 2002b)

Application	Description

Submarines	A haptic version of the Battleships game
Paint with your fingers	Simple drawing program
Early mathematics program	Simple mathematical curve representation, using a ridge or groove
The Memory House	Spatial and memory game
Haptics in Collaborative Virtual Environments	Collaborative haptic manipulation
Radial Haptic Menus	A menu organised like a clock face
Virtual Haptic Search Tools	A 'cross' that gives feedback when the cursor is in line with a screen object
Mathematics – Herbivores and Carnivores	Mathematical simulation of population relationships
Textures	Simulation of textures such as sandpaper, corduroy and wood
Line Drawings	Monochrome line drawings
Floor Plans	Raised floor plans of building interiors: rooms and corridors
Geometrical Objects	Recognition of cubes, cylinders, spheres.
VRML Objects	Recognition of complex realistic objects
Traffic Environment	Training and game with buildings, roads, cars.
Sound Memory Game	Haptic and audio sound matching game
Mathematical Surface	Graphing program where equations can be entered as text

These programs were tested by blind and sighted users. Attention was paid to the user's actions, ideas and feelings rather than objective measures (eg task completion time) in order to improve the perceived usability of the applications and tools. The results from these tests are presented in the form of guidelines for haptic interaction design, and are therefore presented in chapter seven.

Penn et al (2000) experimented on roughness perception using the Phantom. They concluded that there was a negative relationship between groove width and perceived roughness; that is, the surface seemed rougher as groove width decreased. This is contrary to experiments with real textures which have been proven to have a positive relationship. This result is a demonstration of how different narrow one-point interaction can be from conventional means of haptic exploration. In addition, the authors found that the thimble created greater 'sensitivity' than the stylus – perhaps due to the users expecting one-point interaction when holding a pen-like object, but being accustomed to an area of sensations when using a fingertip.

Yu, Ramloll & Brewster (2000) investigated two simple line graphs using the Phantom. The lines were modelled as embossed cylinders, using 'sticky' or 'slippery' effects to distinguish between them. Blind and sighted participants tested the graphs, with mixed results. Participants confused the lines, found it difficult to keep the pointer on the line (especially at corners and ends) and many could not understand the intersection points between the lines. However, most users were able to distinguish between the textures used and gain some impression of the shape of the graph.

Gridlines were not very effective and added further confusion. It was found to be very difficult to provide estimates of coordinates, for example, maximum or intersection points. The authors recommended that additional information be available; for example, a speech output of the current cursor position available on request. The results made it clear that the embossing technique is not suitable for force feedback devices. The authors suggested the use of a groove instead. Additional features such as selective display of lines could also be helpful. A non-speech audio 'overview' feature was proposed.

The Phantom can be used to construct very complex virtual environments. It has been used for scientific visualisation of complex models (Hollerbach, 2000). Ruspini et al (1997) used polygonal rendering techniques with the Phantom to construct complex haptic models. For instance, a virtual AT-AT walker consisted of 11088 polygons, and a very realistic-looking teapot was constructed from 3416 polygons.

Colwell et al (1998) used the Impulse Engine to investigate users' perception of virtual textures and 3D objects. A similar experiment to that of Penn et al was conducted to determine relationships between perceived roughness and simulated groove width of a sinusoidal texture. Colwell et al found, however, that some people perceived a negative relationship and some a positive. This would seem to be a caution to researchers that not all people respond the same way to virtual environments.

Colwell et al also conducted experiments on size and angles of virtual objects. They found that haptic size estimation was approximately correct from the outside of the object, but sizes were significantly overestimated when the inside of the object was presented. (The authors nicknamed this the 'Tardis' effect.)

5.3 Tablets

5.3.1 Overview

Tactile graphics tablets have received much attention as the most direct way to represent the information on a visual monitor. They are frequently comprised of a large rectangular set of pins which can be raised and lowered like those in a refreshable braille display. However, other approaches have used vibrating pins or plastic bumps.

An important benefit of tactile tablets is that they often support multipoint interaction like that in the real world. That is, objects can be felt with all fingers of both hands, rather than with the point of one stylus or by the movement of a mouse. This allows maximum use of the small bandwidth capabilities of the sense of touch.

Although this would seem to be a very promising means of tactile display, commercialisation of this type of device has been severely hindered by the large cost of the displays. Much research has been conducted in the creation of different types of display, however, only the Optacon has been commercially available. This is partly due to the small size of its display area.

5.3.2   Devices

The Optacon is a vibrotactile display available since the 1970s that was designed to allow access to printed text. It consists of an optical scanner and an array of pins which vibrate according to light and dark areas under the scanner. It is possible to read normal text using the Optacon, but reading speeds are very slow and the vibration can temporarily reduce sensation in the user's finger after much use (Way & Barner, 1997a). However, many blind people have found the portable Optacon to be invaluable for activities such as reading labels on supermarket items.

System 3 (Vanderheiden et al, 1992) included a graphics tablet which was used with a special puck which had a set of 100 vibrating pins on top. Its position on screen was directly related to its position on the tablet, and it had zoom capabilities. It provided direct access to the whole screen. Despite the potential of this hardware device, the idea was unfortunately not developed further.

Shaped memory alloys have also been used in the construction of tactile graphics displays. Two major disadvantages of this approach are that the metal is activated by heat and reacts in a non-linear manner. Howe, Konatarinis & Peine (1995) created a tactile display with a pneumatic cooling system and feedforward derivative sensing and position sensing to improve accuracy. This allowed the display to be used accurately at 6Hz; their goal (based on research on tactile perception) was 10Hz. They proposed that this could be acheived through liquid cooling and improved control models.

Wagner, Lederman & Howe (2002) constructed a 6x6 array of pins which could be mechanically raised to between 0 and 2mm. Resolution was around 3mm. This implementation required the array to be mounted on a reasonably large column of motors (around 15cm tall), and this box would increase if resolution or array size was increased, making the device very bulky. However, the authors achieved their aims of relatively low cost and large vertical displacement.

Shinohara et al (1998) developed a prototype tactile tablet that could present three-dimensional objects. Its 64x64 array of pins could each be raised in 0.1mm steps to a maximum height of 10mm. The third dimension can be used to represent shade in a monochrome image, for example. The authors found that blind testers could recognise particular parts of the brain on an image of the cerebral cortex which was represented on the prototype. Unfortunately, the high complexity of this device causes a very bulky package and high cost.

Fricke & Baehring (1993) realised that tactile tablets had great potential but that the costs of electromechanical tablets were prohibitive for the average user. Therefore, their research concentrated on a tablet which consisted of a board with small circular channels. The ends of these channels were closed by an elastic membrane that could be raised by the activation of an electrorheological fluid in the channels. They believed that this method could be used to create a comparitively low cost tactile array (an estimate gives a cost per pin of 5% of that of a piezoelectrically driven pin, with a total cost of less than twice the prices of an 80-character braille display) with no failure prone mechanical pins. In addition, a better resolution could be achieved

without the mechanical elements. A very good technical description of the display and controller is given.

The authors proposed that a blind person could use the tablet to mark areas of text and graphics on a display for character recognition. This approach has been largely outdated by the provision of software interfaces such as Microsoft Active Accessibility, but could still be useful for the numerous times when text is provided embedded in an image. The tablet could also be used for presenting mathematical formulae, directly or using braille-like systems such as dots+.

A later addition to the previous device was a high-sensitivity electrical impulse sensor (Fricke, 1997). This detected the electrical impulses in the user's fingers and thus their position. The tablet could then be used for input as well as output; hopefully allowing direct manipulation of display elements.

Taylor et al (1997) also researched an electrorheological fluid based tactile array. Their array was 5x5 with a vertical force of around 150 grams. Each cell was activated by pressure using a stylus or similar, implying that the array could also detect the user's movements. However, each cell was large: 11mm square. The principles they used could conceivably be applied to much smaller cells and larger arrays, and it is anticipated that this type of display could be comparatively robust and inexpensive due to the small number of mechanical parts.

Yobas et al (2001) have developed a miniature pneumatic valve that can deliver sufficient air pressure under a polyurethane sheet to emulate a braille dot. This would be a novel way to implement a refreshable braille display or tactile tablet, but unfortunately the device is in its early stages and the authors do not give any indication of possible resolution or cost.

Watanabe and Kobayashi (2002) developed a tactile tablet that was designed to allow blind people to draw and erase raised lines. A matrix of 768 pins 3mm apart could be raised or lowered electronically, according to the movement of a stylus attached to a 2-axis positioning arm. The tablet acted simultaneously as an input and output device, as the user could draw using the stylus in one hand and feel the result using the other. This could enable direct manipulation in the user interface such as drag-and-drop tasks. User trials indicated that the tablet had good potential, but the pins were too far apart to accurately represent a line. In addition, a left-handed user found the stylus difficult to use due to the structure of the positioning arm.

5.4   Static displays

Graphs and diagrams are often presented using static means, for example using braille and raised lines on microcapsule (swell) paper (Yu, Ramloll & Brewster, 2000). Problems with this include the small proportion of braille literacy, the low resolution and accuracy of these displays. In addition, dynamic data cannot be displayed, and it is impossible to change the data: a new sheet must be constructed.

Kociolek et al (1999) developed a method of printing braille on memory shape plastic films. These could be embossed and also erased using heat of around 40 degrees celsius. This would provide a way for blind people to draw and erase lines. Although this provides a static display, it may be worth investigating a similar material for a dynamic display.

Way & Barner (1997a) developed a method for automatic production of comprehensible tactile images. They used various image processing techniques to reduce detail in an image such as a photograph to make it suitable for tactile perception. This image could then be printed on swell (microcapsule) paper. They also provided a good specification of microcapsule paper and some of its applications.

Challis & Edwards (2001) developed an application to improve the accessibility of music notation. Conventional non-visual music notation (eg. braille music and Talking Scores) is almost serial. This means that it is not easy to ignore irrelevant information, for example other musical lines. In addition, they form a different mental model to that of graphical music notation, which makes communication between sighted and blind musicians difficult. This project, dubbed Weasel by its authors, used vacuum-formed PVC overlays which were designed to correspond with their visual counterparts. Different heights were used to convey different information. The overlay was placed on a touchpad which could provide audio descriptions of elements and a control section at the bottom to change settings.

User assessment was by a group of six competent musicians, five sighted and blindfolded and one blind. The group was trained using five introductory overlays, then asked to describe the information given in a new overlay and use the control section to change some settings. The tests were assessed in a subjective manner to gain information about users' understanding and ease of use of the Weasel system.

The authors concluded that the following principles should be used when designing a static tactile display:

- Consistency of mapping between visual and non-visual displays

- Use tactile representation for static data only

- Height should be used to discriminate between information types

- An excess of empty space should be avoided

- A tactile representation identical to its visual equivalent can cause difficulties

- The design should encourage a specific exploration strategy

- Double-clicking is inappropriate without haptic feedback, i.e. in static displays

- A display should be sized appropriately. A suitable maximum size is A4 in landscape orientation

- 'Tactile objects should be simple'

The overlays were changed to meet those criteria.

However, these principles apply to the group who took part in the testing. If the tests had been conducted with a group of people who are more accustomed to using tactile input (all blind people, not necessarily musicians), different results may have been achieved. For example, the authors found that people 'got lost' in large areas of empty

space. It is expected that this would also apply to blind people but perhaps to a lesser extent.

5.5   Other forms of tactile display

The Tactile Vision Substitution System, developed in the 1970s, was an early commercial tactile device (Way & Barner, 1997). This was similar in principle to the Optacon, except on a larger scale: the vibrating image was displayed on the user's back.

The TASO (Tactile Acoustic Screen Orientation), another tactile exploration device, was available from Frank Audiodata (Poll & Waterham, 1995). This was designed to explore a character-based text screen and consisted of a vertical and a horizontal sliding key to indicate the cursor's position, along with non-speech audio augmentation of the position information. It then used synthesised speech to present the text at the current position.

A fairly new material, Ionic Conducting Polymer gel Film, was used by Konyo, Tadokoro & Takamori (2000) to produce a tactile display capable of subtle changes in surface texture. The film changes shape in electric fields, and high-frequency vibration of shape can give subtle surface differences. In this manner it is possible to recognisably represent the feel of different types of cloth, such as carpet, denim, and linen. This method of tactile display could have great potential, but unfortunately it is presently extremely expensive.

Experiments have been conducted on electrical stimulation of fingertips and other areas of the body such as the abdomen. One example is given by Kaczmarek et al (1997), who proved that simple geometrical patterns can be identified from mild electrical stimulation of the fingertip. However, pattern recognition was more accurate with embossed braille-like mechanical stimulation. In addition, more than half of the possible test subjects were disincluded because of high skin resistance or a small comfortable current range. This would suggest that this approach is not useful to all people.

A serious problem with electrostimulation is the widely varying resistance of skin, related to its moisture content. A system which circumvents this difficulty is the Tongue Display Unit (Kaczmarek, n.d.). The human tongue is very sensitive, with many nerve endings, and saliva is always present to provide electrical conductivity. The TDU consists of 144 gold electrodes which deliver a small electric current to the tongue. This is attached to a camera which can be worn on the forehead. Test users found that, with practice, the TDU could enable them to navigate virtual mazes and real environments.

5.6   Summary

Various hardware has been developed to enable haptic human-computer interaction. These devices can loosely be classified into mice, three-dimensional devices, tablets, static displays, and others. Unfortunately, the high cost and developmental nature of most of these devices make them unavailable to the majority of visually impaired computer users. Only the Wingman force-feedback mouse and the static display

methods are affordable. However, the large body of research that has been conducted on all forms of hardware has dramatically improved understanding of haptic human-computer access, and is therefore extremely useful for future development.

6. MULTIMODAL HUMAN COMPUTER INTERACTION

6.1   Combining haptic and auditory access

It can easily be seen from the previous two chapters that both auditory and haptic interfaces have their strengths and shortcomings. Given the amount of research that has been undertaken to explore different aspects of auditory and haptic human-computer interaction, surprisingly little work has been done to combine the strengths of each approach. In addition, most of this has been concerned with providing access to graphs; some examples follow.

Yu, Cheung & Brewster (2002) constructed an application which dynamically produces audio-haptic line graphs from a web browser. The haptic effects are accessed using a Logitech Wingman mouse. Tests showed that the mixed modality graphs were much easier to use and understand. Overall interpretation accuracy was 79% for the audio-haptic condition, only 10% for the purely haptic condition, and 36% for the audio-only condition. This emphasises the importance of using all available modalities for non-visual computer access.

Brewster (2002b) added non-speech audio graphs to a table of data values. Tests with blind users showed that perceived workload and task time were significantly reduced in the audio-and-speech condition as compared to a speech-only condition. Accuracy was also significantly improved.

For a haptic graph, Brewster (2002b) found that different friction effects were effective at distinguishing multiple lines on a graph. However, they found that high

friction could cause users to mistake a sharp bend for the end of a line. In a pilot study which did not use friction effects, some participants did not notice line intersections or even changed lines without noticing.

Ramloll et al. (2000) gave a detailed description of their audio-haptic approach to graph access, including the rationale, hardware and software considerations involved. They used the Phantom and 3D audio to present the graph. Ramloll et al (2000) proposed an interesting extension to their system in which a user is assisted to explore the graph. When the cursor is on a line, the cursor will follow the line of its own accord. Applying force in the same direction will speed it up, force in the opposite direction will slow down the cursor or change its direction. Force in other directions will take the cursor off the line. The authors propose that this could facilitate users' exploration, and additionally reduce errors when users mistake one line for another at intersections. Although this could be a very useful tool for quickly exploring graphs, it seems important that it be possible to turn off the automated tracking if a user prefers. For instance, if a user wants an exact data value read from a graph, the cursor will need to stay still while this is requested and retrieved.

It has been shown by several researchers (such as Ramloll et al. 2000; Sjöström 2002b; Yu & Brewster 2003) that a purely haptic or audio line graph can give the user a basic understanding of the shape of a graph, but sighted test subjects' sketches of the graphs show a lack of accuracy in exact and relative positioning of the lines. It can therefore be seen that it is important to provide speech output to enable more accurate details to be extracted from a graph. Constant speech feedback would be

annoying and unnecessary, but it appears to be essential that the exact current position of the stylus or cursor be provided on demand.

Yu & Brewster (2002) found that using multiple modalities to provide access to graphs reduced the importance of a high-quality force feedback device. They compared the Phantom and the Logitech Wingman mouse for use with a multimodal graph and found little difference in users' performance between the two devices. Using each device, audio and haptic combined feedback was easier to comprehend and produced better accuracy than haptic or audio feedback alone. Therefore, the use of multimodal presentation can enable understanding at affordable cost.

Yu & Brewster (2002) also found that sighted test subjects exploring multimodal graphs used non-speech audio to initially overview the data and speech to provide exact data values. The haptic feedback was also used to gain an impression of data trends, but its main use was in navigation. Blind users tended to use speech output both for navigation and to access the data, as this is often their normal method of computer access. These observations further emphasise the importance of allowing subjects practice with a new interface before conducting user testing, to ensure that the interface is used to its best advantage.

Treviranus (2000) used geometric virtual shapes, environmental noise and speech audio to present a map used in teaching 4th grade geography. Feedback from blind and partially sighted primary school students indicated that they all thought that the

combination of the three feedback types was most important in their comprehension of the map and that no one type was more important than the others.

6.2   Multimodal enhancements to graphical interfaces

Some researchers have argued that multimodal interfaces are justified and useful for all computer users. Although the current graphical user interfaces are useful and functional for mainstream computing, they attempt to present so much simultaneous information to the visual channel that it can become overloaded and users begin to miss messages. In addition, they require a large screen for effective use, and this is not practical in many mobile applications.

Brewster, Wright & Edwards (1995a) worked to enhance visual interfaces with non-speech audio. They presented a good argument for multimodal interaction even when single-mode is adequate. The real world is multimodal, so to make the best use of our senses it is logical to use more than one of them. It can also use the individual advantages of each sense. For example, unlike a visual alert, it does not require localised attention to notice an auditory alert.

Crease, Brewster & Gray (2002) created a toolkit of multimodal user interface widgets. These were designed to be consistent across the multiple modalities and no modality was assumed to be of more importance than any other. The widgets can adapt to their environment, determining which modalities are appropriate and using all appropriate modalities to their best effect. The authors developed this software to be adaptable to many uses: they propose its use for mobile applications such as on

palmtops with limited screen space and mobile telephones where the screen cannot be seen while on a call as well as for the visually impaired.

Although the above research concentrated on adding multimodal interaction to graphical interfaces, it is also useful for non-visual user interfaces. If multimodal interfaces become better understood and more mainstream through such research, it would be easier or perhaps even unnecessary to create dedicated non-visual user interfaces.

7. GUIDELINES

7.1   Introduction

It is extremely important to learn from the research findings of others before attempting to create a new computer access method. Almost every researcher has their own ideas about how a non-visual interface should be designed. However, many of their guidelines are in fact in agreement or complementary.

In this chapter, the recommendations of many researchers for non-visual interface design are collected and analysed. Additional guidelines developed during research for this thesis are also included. This is followed by a summary of the most important points.

7.2   Flexibility of presentation

One almost universal characteristic of successful software is flexibility. It is vitally important that a user be able to adapt a user interface to his or her needs. In addition, interfaces and applications are usually constructed so that basic tasks can be accomplished after very little learning. More advanced features do not intrude on a beginner user, but are available for when the user acquires more confidence and experience.

Blind and partially sighted people are by no means a homogenous user group. They are as diverse in their previous experience of computers as the rest of the population. In addition, different levels and types of visual impairment cause people to have different requirements from an interface. People with additional disabilities, such as

deafness or reduced tactile perception, add further requirements. It may well be impossible to build a human-computer interface that meets everyone's needs, but by including as much flexibility as possible in the interface it is more likely to be useful to more people.

Ludi (2002), himself a visually impaired computer user, strongly advocates flexibility of interface. He emphasises the importance of providing "the ability to customise an interface so that it can accomodate a variety of needs and thus address many groups of users" (Ludi 2002, p. 10).

The following table gives a summary of the main ways in which an interface could offer flexibility. These possibilities offer many choices for the end user of the system. It is therefore important to facilitate the user's learning of the possible functionality and to supply appropriate defaults so that a user can start without learning all the options.

Table 7.1: Targets for flexibility in a non-visual user interface

Method of content presentation	braille Synthetic speech Magnified display Non-speech audio
Type of presentation	Grade of braille Speed of synthetic speech Use non-speech audio for:

	_ All events
	_ New events
	_ System events
	_ Critical events only
	_ None
Amount of information presented	Letter
	Word
	Sentence
	Paragraph
	Document
Amount of detail	Application name
	Name of document/web page etc
	Headings
	Menus
	Links
	Full text

7.3   Use of multiple modalities

It is evident from the preceding chapters that no single sensory modality has enough bandwidth to replace the sense of sight. However, by combining the other modalities which are possible in human-computer interaction, the best use of available modalities can be achieved.

One of Sjöström's (2002b) five central guidelines for haptic interface design recommends the use of all available modalities. This includes the use of auditory labels and braille or speech captions to provide users with a context to aid in their identification of haptic objects. He also suggests that environmental sound can be used to provide an overview when using single-point haptic interaction.

7.4   Manipulation of the interface

There should be clear differences between the actions of input and navigation. Users need to feel confident that they can explore the interface without changing it or attracting other unwelcome consequences. However, a commonly known human-computer interaction guideline is to avoid modes in interfaces. Therefore, a seperate 'exploration' mode and 'input' mode are inappropriate and alternatives must be found. One can compare the way this issue is addressed in common GUIs – exploration is generally accomplished by looking at the screen and sometimes by single clicking the mouse, whereas input usually takes the form of double-clicking, some single-clicking, and all keyboard input. Clearly, in a non-visual interface, exploration cannot be undertaken without some form of input to the computer, which complicates the issue of providing clearly different navigation and input.

Mynatt lists direct manipulation as one of the five important goals of screen reader interface design. It "is achieved when the user is able to directly interact with objects of interest to the task at hand, and output in the interface is expressed via these objects" (Mynatt, 1997). In a graphical interface, direct manipulation and interaction are supported through the use of the keyboard and mouse. The mouse and its

associated actions (such as clicking and drag-and-drop) are unavailable in a non-visual interface, so a replacement means of interaction must be found (Mynatt & Weber, 1994). Challis (2000, p. 222) agrees that "whenever possible, a user's intuition to interact directly with tactile objects should be accomodated". It is therefore evident that a successful non-visual user interface should support direct manipulation.

7.5 Congruence with the graphical equivalent

A large proportion of personal computers currently in use run a Microsoft Windows operating system. As a result, a successful and portable means for the blind to access a computer must accurately interface with the Windows environment and the software which uses this platform. Important examples of these include Web browsers, email clients, word processors and document displayers such as Adobe Acrobat. It is necessary to focus on software that is widespread for purposes of affordability. The most useful software presently available seems to (mostly) conform to this requirement; good examples are Jaws and Magic (a popular screen magnification program). Ideally, to increase the scope and utility of a non-visual computer access program it should be designed so that it can be adapted to run on multiple operating systems.

It is vitally important that all functionality and information given by a graphical interface be also available in the non-visual version. Brewster (2002a) agrees that no information should be hidden completely. This includes portions of interfaces that

are not designed to accept user input and thus cannot be navigated to in some screen readers (Mynatt, 1997).

When a screen reader does allow access to a user interface element, it usually gives some information about its function and title, but omits much information given implicitly in a visual interface (Mynatt, 1997). This information includes the size of the element (for example a menu) and how it is grouped with other elements (such as information about containers).

Mynatt (1997) emphasises the importance of collaboration between blind users and their sighted colleagues. Such collaboration requires that the users all have the same mental model of the interface in order for communication about the interface to be possible. This also facilitates learning the non-visual interface, as its users can receive help from people using the graphical interface and also use any prior knowledge they have about the graphical interface. As Challis (2000) agrees, "A consistency of mapping should be maintained such that descriptions of actions remain valid in both the visual and the non-visual representations".

The interface should therefore be presented in a parallel way for sighted and non-sighted users. In addition, the respective presentations of the visual and non-visual interfaces "must be synchronised to support joint operation" (Mynatt & Weber, 1994). In other words, it should be possible to extract all information from the interface using only the screen and only the non-visual presentation at the same time.

Challis and Edwards (2001) agreed that it is necessary to have consistency of mapping between a static tactile display and its visual equivalent. However, they noted that a tactile display that is identical to its visual equipment can cause problems due to the intrinsic limitations of the different sensory modality (see also section 7.8.2 on the use of haptic interaction).

Mynatt and Weber (1994) noted that many user interface elements (for example, icons, buttons and scrollbars) are identified by their appearance. Additional attributes of these elements (such as greyed-out, meaning unavailable) are also often represented graphically. Alternative indicators for these concepts must be developed for use in a non-visual environment.

One further important requirement of a non-visual interface is that it be 'transparent' to the user. That is, it must convert the information and functionality of a graphical user interface to another modality without the user's intervention in this process. Ludi (2002) suggests that accessibility features should be integrated into all interfaces. This would significantly reduce the cost of the accessibility software and make it available on all computers.

7.6   Facilitate navigation and overview

As discussed in chapter three, the sense of sight provides an excellent means of overviewing information which cannot be replicated using the other senses. It is therefore the task of the interface designer to provide as much assistance in exploring, navigating and overviewing the output information. "Good design practice

should, where possible, encourage a specific strategy for exploration in a particular display" (Challis 2000, p. 220). The strategy used should be that which is most intuitive. Further research will be needed to determine which strategy to encourage, as it is beyond the scope of this paper and has not been well explored in previous research.

Another of the major guidelines proposed by Sjöström (2002b) for haptic virtual environments was to facilitate navigation and overview. He advocated well defined, easy-to-find, consistent and efficient reference points to aid users' navigation. He also proposed and investigated some haptic search tools which received positive responses from test subjects (see also chapter five).

Sjöström (2002a) conducted several experiments to confirm some of the guidelines he proposed in his doctoral thesis. He had suggested that walls and corners of a virtual environment are important as navigation tools. Sjöström also wanted to confirm the usefulness of gridlines as a navigational feature. Test results highly emphasised the usefulness of sidewalls as a navigational aid. Out of several test cases, the condition with no sidewalls was found to be far worst both subjectively and objectively. Most users commented that the gridlines disturbed their navigation and that they could not initially tell if the object they were touching was a gridline or a button. Overall, a task with gridlines took longer to complete than similar tasks without the gridlines. Sjöström therefore concluded that gridlines should not be used unless exact positioning is required (for instance, when exploring a graph), and then they should be optional. He also concluded that walls are critical as reference points.

Petrie et al (1997) researched hypermedia interface requirements for visually impaired people. The design solutions proposed included methods for overviewing information (lists of headings or links) and methods for effective navigation, including:

- Flexibility in method of moving through content

- Flexibility in methods of following links

- Provision of a history list and 'forward' and 'back' capabilities

- Provision of methods of returning to the top of a page and to the home page

- Provision of a logical path through content (even when this spans many pages)

Mynatt & Weber (1994) agreed that graphical user interfaces are able to present a great deal of information simultaneously and can easily be overviewed. This bandwidth cannot be replaced in other media and thus additional functionality should be provided to aid exploration.

The Mercator interface (Mynatt, 1997) provided a "preview" facility which gives a quick overview of the current portion of the interface, using short auditory cues. A sighted computer user can gain an overview of the current interface simply by glancing at the screen. However, a blind person's mental model of the current state of the computer must be constructed from their memory of previous events and actions. An auditory overview can therefore reduce mental demand on the user.

Sjöström (2002b) suggested that an overview of a web page (or similar, for example a document) could be provided by adding haptic cues to different types of text. A form of skim reading could be provided by using a coarse texture on headings and a fine texture on body text, allowing the user to quickly identify headings and titles for conversion to speech.

7.7 Facilitate learning of the interface

Since any entirely new software development will be unfamiliar to all users, it is important to provide some kind of assistance for learning the interface. Specific training environments have had some success in gaming, but even then people are more likely to complete a 'tutorial' if it is interactive in the same way as the gaming environment. Commercial applications more often allow the user immediate access to all functionality, but ensure that the application can be used on a basic level for users to get started on their task immediately. More advanced functions can be introduced later as a user becomes more confident and explores the environment.

Hints can also be provided in the form of a 'tutor' - an example being the context-sensitive Microsoft Office Assistant. Sjöström (2002b) recommends the use of a "virtual guide" to help users explore virtual haptic objects. Another example is the assistance for exploration of a line graph developed by Ramloll et al (2000). However, as any experienced user of Word knows, it is very important to be able to turn off the assistance.

Experience in teaching blind people to use Windows has shown that it is very valuable to directly teach users the applications they will be using, rather than using training programs (Rasmussen, 2002).

(Mynatt & Weber, 1994) suggest that "The introduction of nonvisual user interfaces for GUIs can be successful only if an upgrade path from existing technology and metaphors is provided through the user interface." This is an important consideration which holds true for most aspects of interface and applications design – people will not use a system that is completely different from what they are used to. It also reflects experienced computer users' desire to start working immediately, without having to spend time specifically learning a new interface.

Some researchers have emphasised the importance of real-world metaphors to facilitate learning of new interfaces. A visual interface uses iconic representations of real objects to give clues to users as to the function of the icon, for example a notepad or rubbish bin. Petrie et al (1997) researched hypermedia interface requirements. They proposed that such real-world metaphors can quickly assist users' understanding of interaction methods. Mynatt (1997) found that blind users had "expressed an overwhelming positive response to the use of everyday sounds in screen reader interfaces" (Mynatt 1997, p. 45). It is very important, however, that use of metaphors be consistent across the entire interface to avoid causing confusion.

An extremely important functionality of the interface is that of explaining its interaction methods using unambiguous synthetic speech or braille on request. This

can provide 'captions' for interface elements and events to aid users' learning. Sjöström (2002b) used this method to help users understand a virtual haptic environment. As Mynatt (1997) discovered, such a function is frequently used by test subjects in preference to asking another person about the interface elements. This displays users' wishes for independence, a consideration that is very important when designing an interface for blind people.

Sjöström (2002b) added several more suggestions that would facilitate users' learning process:

- Provide contextual information from different starting points.

- Present the haptic model or environment in its natural context.

- Provide information about the purpose of the program.

- Provide information about possibilities and pitfalls in the environment.

- Be consistent; limit the number of rules to remember.

- Give clear and timely feedback on the user's actions.

- Facilitate imitation of other users and situations if possible.

7.8   Appropriate use of modalities

7.8.1   Introduction

A graphical user interface provides much more information than is conveyed in its text. Important information of this sort includes grouping and contextual information. For instance, a dark bar at the top of a rectangular window probably contains information about the currently active application, and the items in the list immediately below this can probably be clicked upon to reveal menus of functions.

Non-visual user interfaces must find alternative means to convey this information. Petrie et al (1997) suggest that it is important to have "Clear differences between interface information and content". When determining suitable alternative means to convey this interface information, it is important to recognise the strengths and weaknesses of haptic and auditory modes of interaction. This will enable them to be used to best advantage.

7.8.2 Haptic

There is some debate as to whether haptic virtual objects should be detailed or simple. Challis (2000) conducted an experiment which showed that increasing tactile complexity dramatically increased error rates. He therefore suggested that "combined auditory-tactile tasks should be tactually simple". However, Magnussen et al (2002) found that complex 3-dimensional haptic representations of real-world objects were easier to identify than simple or abstract shapes, because they offered more clues as to their identity. Sjöström (2002b) suggested that complex objects can be easier to identify if they are more familiar than simple objects or because they may contain more clues to their identity. It is anticipated that the amount of detail that is useful will depend on the size of the object and the accuracy of the hardware used. For an icon that is to be frequently used in an interface, detail and complexity should be minimised. The more effective icons in a GUI are simple and the lower bandwidth of haptic interaction further encourages the use of simple virtual haptic objects in user interface design.

In a multimodal non-visual user interface, therefore, it would make sense to develop a system of simple symbols to represent frequently encountered interface elements. For congruency with GUIs (see section 7.5), the symbols could replace common icons. It is important to keep these symbols simple and easily distinguishable, and there should not be too many different symbols. Eight to ten tactile patterns have been found to be easily distinguishable, but other techniques such as texture and geometric shapes could be used to extend the number of haptic interface 'keys'. Whenever possible, haptic icons should be intuitive and encourage appropriate interaction. For example, a pushbutton should feel pushable.

An important consideration in the haptic aspect of a non-visual interface is whether the haptic interface should directly reflect the spatial presentation of its equivalent GUI. Many researchers have found that users get lost in a virtual haptic environment if there is empty space between objects (Challis and Edwards 2001; Sjöström 2000; Mynatt 1997). Challis (2000, p. 218) agrees that "[a] literal visual-to-tactile mapping is likely to produce problems and is therefore unlikely to be the most efficient design strategy". However, if the spatial layout is changed to better suit non-visual interaction, this presents a barrier to congruence of the interface.

The design of the physical means of contact with dynamic haptic output devices is also important. Sjöström argues that "The Wingman Force Feedback Mouse forces the user into a power grasp for what is essentially a precision task since the mouse must be grasped with the whole hand." (Sjöström, 2002b). Despite this, millions of computer users find a mouse appropriate and easy to use for precision pointing tasks.

It would be a little more difficult to sense precise haptic output with the whole hand on the mouse, but if the mouse were designed to be held with the fingertips (perhaps if a wrist rest were provided) this difficulty would become insignificant.

It has been shown that the thimble and stylus of the Phantom produce slightly different texture perception (Penn et al, 2000) and thus manipulandum (physical pointer) design can affect haptic output. However, the choice of contact point makes makes no difference in recognition of 3d geometric shapes (Magnussen et al, 2002). Some users feel that their performance is better using the thimble (Sjöström, 1999). It is recommended that haptic output devices be designed for ergonomic considerations first. It should be ensured that any perceptual effects of the particular manipulandum design are either unimportant or accomodated for in the software.

It is important that hardware intended for use in a haptic human-computer interface be comfortable to use for extended periods. Some researchers have experimented with touch screens as an input method, but it is evident that the vertical screen could get very tiring to use. A better workspace is horizontal like a mousepad, or even adjustable in angle like many keyboards. The size of the workspace should also be appropriate – a user should not be expected to overextend their arms to use it. Suggested workspace sizes range from A4 landscape (Challis & Edwards, 2001) to 7cm by 6cm (Sjöström, 1999).

Current user interfaces are presented in a two-dimensional manner. Unfortunately, a purely 2D haptic output device would be insufficient due to the need to represent the

information conveyed by colour on a screen. Sjöström (2002b) found that the three-dimensional capabilities provided by the expensive Phantom are a little extra to what is needed. He suggested a "2.5D" device, with a large workspace in 2 dimensions and about 10mm in the third dimension. This third dimension could be used to convey basic colour information. A sufficiently accurate device could convey many different levels of information using height.

7.8.3 Auditory

As discussed previously, the auditory sense has a capacity for immediately noticing changes in stimuli, irrespective of their physical location and (mostly) irrespective of the number or type of other stimuli. It therefore provides an excellent substitute for the equivalent facilities of vision. It also provides the capacity for monitoring a large number of sounds at once. Like vision, auditory attention can be concentrated on one source, filtering out all others. However, changes in other sound sources are easily noticed and can be ignored or focussed on as necessary. For example, unlike a visual alert, a user does not require localised attention to notice an auditory alert.

An important aspect of auditory stimulation is its annoyance. Repetitive or continuous sounds, particularly artificial ones, can quickly become annoying. It is therefore important that a designer of an auditory interface use continuous or repetitive sounds only when necessary, allow users to turn them off, and try to use sounds which do not become annoying.

Examples of noises which are not usually found annoying include fan and hard drive access sounds from a computer, or the noise of someone typing. By contrast, the noise made by someone sniffing irregularly or playing a mobile phone game with keypad sounds turned on can quickly become infuriating, particularly to someone who is concentrating on writing or some other mentally demanding task. It is recommended that proposed user interface sounds be user-tested over a reasonably long period of time to determine appropriate sounds.

Additional contributors to the annoyance of sounds include intensity and demanding qualities. As mentioned in the chapter on auditory interfaces, high volume should not be used to get a user's attention, as it quickly becomes annoying. Demanding sounds such as that of smashing bottles or ICQ's 'uh-oh' should be used with caution, only for important or urgent events, and should not be long in duration.

7.9   Testing

As with any computer system, testing of a non-visual user interface is critically important. In particular, thorough testing at the specification and design phases is essential because this is a relatively new field of development. Ideally, testing of a non-visual interface should include test subjects from many groups: blind/partially sighted/sighted, experienced/beginner, and even musically trained/non-musical.

It is important to test objective measures of usability to provide a firm comparison with current approaches to non-visual computer access. Almost every researcher in this field has used measures such as task completion time and accuracy for this

purpose. However, it is also imperative that users' perceptions of an interface indicate that it is easy to use. Brewster and colleagues (Brewster, Wright & Edwards 1995a; Walker et al 2001; Oakley et al 2000; Challis 2000 etc.) used a subjective testing system based on the NASA Task Load Index. This required test subjects to rate the following on a subjective scale of 0 to 20: mental demand, physical demand, time pressure, effort expended, performance level achieved and frustration experienced. Brewster et al added a seventh measure of annoyance, as this is often a concern with auditory interfaces.

When testing an interface, it can be expected that users' performance will improve with practice. It is therefore recommended that test subjects be allowed some practice before final usability testing is conducted. Another suggestion has been that users' performance be tested before and after practice, to allow measurement of 'learnability' of the interface. Jansson & Ivås (2000) tested users' identification of small (5 – 15mm) virtual 3d geometric shapes produced by the Phantom to confirm that recognition of simple haptic objects improves with practice. There were 4 shapes and each tester was presented with 216 objects over 3 days. The proportion of objects correctly identified approximately doubled after the three days, but seemed to reach an asymptote at that point. No participant identified all objects correctly even after that amount of practice.

If tests are to be performed comparing blind and sighted subjects' interaction with an interface, the test procedure must take into account the fact that blind users are often less confident about their performance than sighted users. Many researchers have

noticed this tendency, and it is confirmed in subjective ratings of an interface. For instance, Ramloll et al (2001) tested auditory enhancements to tabular data presentation. It can be seen from their results that the subjects' self-ratings using the task load index in the performance category were significantly lower for blind subjects. This was despite the fact that the task was purely auditory. Any difference between blind and sighted subjects could reasonably be expected to favour the blind subjects, as they are often more accustomed to receiving synthetic speech feedback from a computer interface.

7.10   Summary of Guidelines

Table 7.2: Summary of Guidelines

Flexibility	Allow flexibility in method of presentation, type of presentation, amount of information presented and amount of detail presented. Allow immediate use of simpler functionality without confusing the user with details. Provide easy access to more complex functionality for advanced users.
Multimodality	Make the best possible use of all available modalities.
Manipulation	Keep manipulation separate from exploration. Allow direct manipulation in an intuitive manner.
Congruence	Develop for the most common operating system. Develop in such a way that the interface can be transported to other WIMP-based (Windows, Icons, Menus, Pointer) operating systems. Do not completely hide any information or functionality.

	Have a consistency of mapping between the NVUI and the GUI. Facilitate collaboration with users of the GUI.
Navigation and Overview	Encourage a specific, intuitive navigation strategy. Consider the use of haptic search tools. Use consistent, easy to find 'landmarks'. Facilitate overview.
Learning	Use training assistants, tutorials or virtual guides to help users learn using the interface directly. Provide an upgrade path from existing common technology. Use intuitive, real-world metaphors. Provide explanation of (auditory or haptic) symbols on request, using synthetic speech. Provide contextual information whenever appropriate.
Appropriate use of modalities	Provide clear differences between interface information and content. Use a small number of easily distinguished haptic symbols. Avoid large empty spaces and disconnected objects in haptic displays. Make icons and other interface devices as intuitive as possible. Work around hardware limitations. Use the capacity of hearing for noticing changes. Do not overload the senses. Ensure sounds are not annoying and can be turned off.
Testing	Use appropriate, diverse test audiences. Test subjective factors as well as objective.

	Allow for differences between test subjects when assessing subjective factors. Test before and after practice. Use statistical analysis to determine significance of results.

8. SPECIFICATIONS

8.1 Introduction

In this chapter, a complete, adaptable and functional user interface following the guidelines in the previous chapter will be specified. This interface will sit transparently on the GUI of the operating system, supplementing its visual output with speech, non-speech audio and haptic output. A new haptic force-feedback device to provide this output and enable exploration is described.

This system should be developed in a very modular way, so that no rewriting is required for a change in hardware or operating system. New modules would be added to the system in either case, but it should not be necessary to rewrite existing software. For a diagram of this modular architecture, see Appendix 1.

As previously discussed, there is continuing debate about whether earcons or auditory icons are preferable for use in non-visual user interfaces. It has also been suggested that they could both be used in the same interface, taking advantage of the strengths of each approach. Such a combination is recommended in this paper. In the following, the word 'audicon' will be used to indicate any non-speech sound used in an interface. This can be chosen to be an earcon or auditory icon at the discretion of the developers.

It should be possible to operate the interface with or without specialised haptic hardware. If only the keyboard is used for navigation, this system will be very similar to current screen readers, only slightly enhanced by the use of the audicons.

A normal mouse could also be used, but it would be very easy to get lost. An auditory navigation tool could be developed to aid in this situation. If a force-feedback haptic device such as the Phantom or Logitech Wingman mouse is available, it should be possible to add a small software interface to adapt the system for use with this hardware. Ideally, however, the system would be used with the new haptic force-feedback device proposed in Section 8.6.

8.2 System events and responses

8.2.1 Event table

Table 8.1: System events and responses

Event	Response		
	Synthesised speech	**Non-speech audio**	**Haptic**
Windows starts	"Start"	-	Cursor magnets to Start button
Cursor moves over desktop icon	-	Audicon for application, folder or file as appropriate. If desktop icon is a shortcut, the audicon is relevent to the target, not the shortcut.	Embossed 1mm
Select desktop icon	Name of icon	Audicon for application, folder or	-

		file type as appropriate	
Application starts	Text on title bar	Audicon for application start	Cursor magnets toward cursor of application, or to upper-right corner of its workspace if there is no cursor
Cursor over menu	-	Menu	Embossed 0.5mm
Click or hold on menu	Name of menu	Menu opened	Sink 0.5mm, then rise 0.5mm, then cursor magnets to first menu item on submenu
Query menu	Name of menu	Menu	Embossed 0.5mm
Cursor over menu item	Name of menu item	Menu item	Embossed 0.5mm
Cursor over menu item that calls submenu	Name of menu item	Menu	Embossed 0.5mm
Query menu item	Name of menu item	Menu item	Embossed 0.5mm
Query menu	Name of menu item	Menu	Embossed 0.5mm

item that calls submenu			
Click menu item	Name of menu item	Button clicked	Sink 0.5mm, then rise 0.5mm
Click menu item that calls submenu	Name of menu item	Button clicked, menu opened	Sink 0.5mm, then rise 0.5mm, then cursor magnets to first menu item on submenu
Cursor over toolbar button	-	Toolbar button	Embossed 0.5mm
Click toolbar button	Tooltip	Button clicked	Sink 0.5mm, then rise 0.5mm
Query toolbar button	Tooltip	Toolbar button	-
Pop-up window – error	Title of window, then error message, then labels of buttons	Error, pop-up	Magnet to default answer
Pop-up window – question	Title of window, then question, then labels of buttons	Question, pop-up	Magnet to default answer
Select anywhere that	-	Text field	Magnet to text cursor

expects text input			
Cursor at edge of display	-	Out-of-bounds	Cursor cannot move further
Window in focus contains a progress indicator bar	-	Progress indicator	-
Query progress indicator	"progress indicator, x% complete"	Progress indicator	Recessed 0.5mm
Cursor over image	-	Image	Low friction
Query image	Tooltip and/or ALT text	Image	Low friction
Click on image	Tooltip and/or ALT text	Image	Low friction
Cursor over text	-	Text	Sandpaper
Cursor over heading	-	Text	Coarse sandpaper
Query text	Reads text	Text	Sandpaper
Click on text	-	Text	-
Cursor over	-	Link	Embossed rectangle

link		85	around link text
Query link	Reads link text, then target address	Link	Embossed rectangle around link text
Click on link	Title of new page	Link followed	-
Cursor over radio button	-	Radio button	Small round embossed area
Query radio button	Title of group, then title of control, if available	Radio button	Small round embossed area
Click radio button	-	Radio button selected	Small round recessed area
Cursor over check box	-	Check box	Small square embossed area
Query check box	Title of group, then title of control, if available	Check box	Small square embossed area
Click check box	-	Check box selected	Small square recessed area

8.2.2   Further information on system responses

The Query function is called by the third button of the haptic interaction device, or by pressing Ins on the numpad. The 'Ctrl' key will stop any audio currently running and any current magnetic haptic effect.

The word "magnets" is used here to indicate an effect similar to that if the cursor and object concerned were magnetically attracted to each other. This effect lasts for x seconds, where x would be the time taken for the cursor to travel the entire diagonal of the screen under its own power, if no resistance were met.

Any clickable area feels embossed. Recessed buttons improve targeting, but embossed buttons give the impression of being clickable. This metaphor can be continued across the entire user interface as well as elements of web pages. There should be a very narrow depression between adjoining buttons (as on a toolbar), but it should be distinctly noticable. A suggested depth for this is about 0.25mm. It may be necessary to investigate forces that are appropriate for this depression so that a user does not slide off the other side of a small button when trying to overcome them.

Distinctly different voices should be used for interface speech and content speech. A female voice is suggested as a default for interface speech, with a male voice for content. The user should be able to seperately choose the voices for each type.

8.3   Overview function

The + key on the numpad can be pressed at any time to give an overview of current options. Ctrl will always stop this process.

Table 8.2: Overview function

Context	Speech	Audicon	Haptic
Menu	All items in menu	Menu	-
Desktop	Names of all	Audicon for each	Magnet to nearest

	desktop items	item as they are read	desktop icon
Toolbar	Tooltips from all toolbar buttons	Toolbar button	-
Title bar	Text on title bar	-	-
Taskbar	All text and tooltips from taskbar	-	-
System tray	All tooltips from system tray	-	-
Web page			

8.4 Further specification of non-speech audio responses

Table 8.3: List of required audicons and suggestions

System ready after startup	"Ta-da!"
Application	Short drum riff
Folder	Drawer opening on metal filing cabinet
File	Flipping through papers
Application start	Creaky door opening
Application close	Creaky door closing (including thud)
Menu	Pitch depends on which menu, depth depends on length
Menu item	Pitch changes slightly according to position in menu
Toolbar button	Sound of saw on wood

Button clicked	Sound of button being pushed
Menu opened	Swoosh, range depends on size
Question	"Huh?"
Error	"Uh-oh!"
Pop-up	A rising whistling sound
Text field	Old-fashioned typewriter
Out-of-bounds	Sound of ball bouncing
Progress indicator	Sound of receptacle being filled with liquid. The pitch of the sound is dependent on the amount of progress.
Image	Swish as of a brush on paper
Text	Quiet typing, as of a modern keyboard
Link	Two quick footsteps going away from user
Link followed	Two quick footsteps going towards user
Radio button	Sound of a physical radio button being pushed twice
Check box	Sound of pencil on paper writing a tick mark

Care must be taken to ensure that sound clips used for audicons do not become overly annoying after repeated use.

Audicons should be as short as possible to convey the meaning. 0.5 seconds is a good length. They should also be interruptable with new actions as soon as possible. Audicons should be played serially. They may be interrupted by a new user action but not by a new system event; this should wait until the current audicon has finished. If there are several audicons waiting to be played, the priority list given

below will define their order. Some may need to be cancelled if they are no longer relevant.

Priorities for auditory output:

1. Audicon caused by user action

2. Error audicon

3. Speech giving interface information

4. Speech giving content information

Causality is an important property of any user interface. This means that input by the user should trigger an immediate response or acknowledgement by the interface. In this case, the audicons detailed above should start as soon as the triggering action is completed. Other audio output, such as speech interface information, can occur later. For example, an audicon for an application start is played immediately after the request is complete; the speech on title bar should be read when application has been loaded successfully.

Filtears are applied to the audicons if appropriate. An audicon for a "greyed out" or unavailable option is passed through a low-pass filter to make the sound seem muffled. The audicon for a selected object or option is high-pass filtered. The sounds for menus, menu items and buttons depend on their position in lists. The sound for a menu or list of options is also varied according to how many options it contains.

8.5   Haptic device

The suggested optimal haptic interaction hardware for user interface exploration is a 2.5D force-feedback device with a thimble as a manipulandum. It is termed 2.5D because it has a horizontal workspace a little larger than a common mousepad, but the available height output ranges within only 1cm. An A4 landscape sized workspace area has been suggested for this kind of device; this recommendation has been adopted here but narrowed slightly to correspond with normal screen proportions.

Single-point force-feedback devices seem to be the only ones that have a chance of being affordable. A thimble provides greater sensitivity than a mouse or puck, as a user is less likely to damp the force-feedback effects by holding the object with their whole hand. A hip-joint type connection below the thimble allows the user's finger and the tablet of the device to be held at a comfortable angle. Buttons provided on the device enable input and thus direct manipulation. A picture of the suggested device is given below, along with a table of core parameters.

Figure 8.1: A new haptic interaction device.
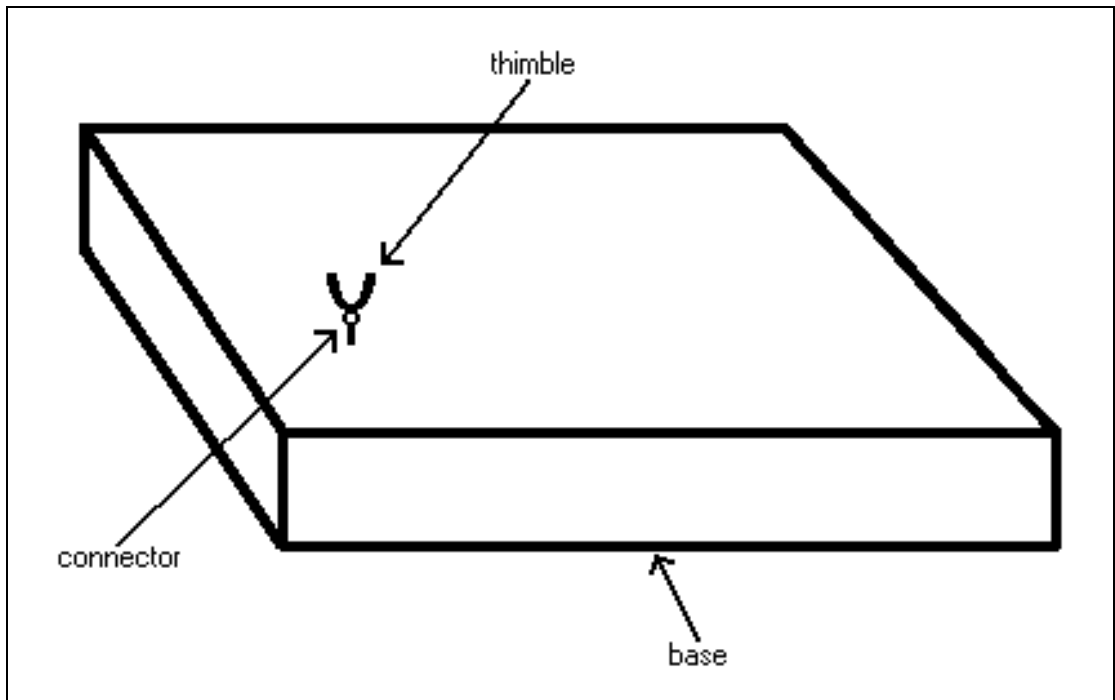
Table 8.4: Parameters of the new interaction device

Workspace width	270mm
Workspace depth	200mm
Workspace height	10mm
Resolution	1mm
Input buttons	3
Maximum force capability	5N

8.6 Rationale

By far the most commonly used non-visual computer access method today is the use of screen readers such as Jaws. This indicates that congruence and portability to commonly used operating systems is more important than a non-spatial presentation,

such as the hierarchical presentation used for the Mercator system. An interface that combines the transparency and portability of Jaws with the additional outputs of non-speech audio and haptic force feedback would therefore seem to be the best option.

An upgrade path from existing technology is provided by the fact that this system can be used like a common screen reader. The haptic effects and auditory icons are in addition to the normal speech feedback. They can therefore be useful when a user has learnt their function, while allowing people who are accustomed to screen readers to begin using the system immediately.

For people who are not used to computer access using current screen readers, this system will be easier to learn and to explore. In addition, its high congruence with the graphical user interface makes assistance by sighted users much easier.

Tutorials, virtual guides and virtual search tools are excellent ideas to help users learn the new interface and facilitate navigation. However, their specification is beyond the scope of this project. It is suggested that these may be good targets for further research.

Navigation is not entirely intuitive due to the constraint of consistency with the graphical user interface. However, this is improved by the inclusion of the 'overview' feature. In addition, users with a good mental model of the graphical user interface will find the transition easy.

Manipulation is kept separate from exploration by providing a third button on the haptic interaction device for the purpose of querying objects. Therefore, users can be confident that movement of the device, querying or overviewing will not be interpreted as manipulation by the interface. Direct manipulation is supported in the same way as with a mouse. All actions are confirmed by an auditory or haptic effect. Additionally, force feedback effects can be used to make the puck seem heavier when an object is being dragged.

While an ideal, realistic haptic display would enable multiple points of contact, the provision of this ability is very complex and expensive. Single-point haptic interaction devices are sufficient for user interface access. Multiple contact points are valuable, but used mostly for judging relative dimensions, proportions, and distances. In a user interface, detection and recognition are far more important tasks. These have been shown to be very possible using a single-point haptic device. Other necessary information such as size can be given using non-speech audio.

Flexibility in this interface is a little similar to that provided by Jaws. It is anticipated that advanced users will use the query and overview functions less as they become accustomed to the interface, location of elements and audicons. As with Jaws, speech rate needs to be flexible as it improves with practice.

9. CONCLUSION

9.1 The need for a non-visual user interface

Almost all applications in common use run only on operating systems based around graphical user interfaces. These are accessible only with great difficulty by visually impaired people. Screen readers provide access to the text on the screen, but miss the large amount of non-textual information. Therefore, a complete non-visual user interface (NVUI) is needed. This should run transparently on top of the graphical user interface and provide non-visual access to its content.

Although sight is the primary human sense, the senses of hearing and touch can also be used to gain a great deal of knowledge about our environment. These can provide alternative means of human-computer interaction, an area which has not been well explored. Much research has been done on haptic-only or auditory-only interaction, but little on a combination of the two. In addition, this is comprised mostly of work on a specific interaction method or tool; no comprehensive studies are available. In particular, the concept of a complete multi-modal non-visual user interface has not previously been investigated.

9.2 Construction of a non-visual user interface

As a result of their investigations into particular tools, methods, applications and hardware, researchers have proved the feasibility of the separate components that could be used to make up a non-visual user interface. They have also accumulated much knowledge about how such an interface (and its components) should be designed. Chapters 4, 5 and 6 give details of many examples of this research.

94

Chapter 7 details the most useful and relevant guidelines that have been gleaned from the literature. In summary, the guidelines concern:

- Flexibility

- Multimodality

- Manipulation

- Congruence

- Navigation and Overview

- Learning

- Appropriate use of modalities

- Testing

One of the chief requirements of a user interface is that it be flexible, adapting to users' varying experience levels and requirements. Beginner users must be able to accomplish tasks immediately without worrying about the advanced features of the system. Advanced users, however, will want to optimise the system for ease of use. Flexibility in a non-visual user interface should also allow a user to choose the amount of detail and amount of information that is presented at once.

To convey information without the sense of sight, the best possible use must be made of the other available senses. It is imperative, therefore, to use both hearing and touch to best advantage. It is also important to allow users to interact intuitively with the interface. Direct manipulation can be supported by a haptic input/output device.

As evidenced by current screen readers, an extremely important requirement of a non-visual user interface is that it must provide transparent access to commonly used graphical operating systems. Although an exact non-visual interpretation of the GUI is unnecessary and may be difficult to use, close similarity between the mental models used when navigating the GUI and NVUI is important. This helps a user to learn the interface, aids collaboration with users of the equivalent GUI, and ensures that all functionality is available.

The high perception speed of the visual sense provides humans with an excellent means of overviewing the environment to determine possibilities for action. In a NVUI, this must be replaced by the lesser capabilities and speeds of the auditory and haptic senses. This can be made much easier if the interface is carefully designed to assist navigation and overview. Set navigation strategies, virtual tools and reference points can all be useful in meeting this requirement.

Any successful user interface must be easy to use and to learn. Learning is facilitated by similarities to other, more familiar technology; also by making the interface and its components as intuitive as possible. Information about the context of a haptic or auditory symbol can be important when determining its meaning, and an explanation using synthetic speech should always be available. Tutorials or assistants can also improve the learning process.

The development of graphical user interfaces is a well-studied and understood area. Most people are much less familiar with the specialised constraints and opportunities

inherent in the use of the haptic and auditory senses for human-computer interaction. Both senses have a much lower bandwidth than sight and therefore it is important not to present too much information at once. Hearing can easily and quickly notice changes in stimulus, irrespective of direction; this ability should be used to best advantage. The haptic senses are limited by available hardware, but are reasonably good at determining the position and basic characteristics of an object. It is important to ensure that it is easy to discriminate between interface information (such as 'this is a button') and content of an application. Finally, NVUIs should be tested appropriately with regards to their target audience, and testing should include both objective and subjective measures.

9.3 Specifications of a NVUI

This report includes the specification of a complete non-visual user interface, meeting all the requirements given in Chapter 7 in a way that capitalises on the abilities of the auditory and haptic senses and optimises these often conflicting goals. Audicons are used in conjunction with haptic force-feedback effects to convey most of the non-textual information in a GUI. This is complemented by synthetic speech to provide textual information. In order to provide the haptic effects, a new force-feedback input/output device is proposed and specified. An overview function assists navigation, and the interface is designed to be as intuitive as possible.

9.4 Summary

This thesis presents the results of research and design on the subject of non-visual user interfaces. This report initially gives details on current non-visual computer access methods and their deficiencies, which is followed by a summary of relevant

information about the capabilities of the human auditory and haptic senses. Next, a broad summary of research is presented that has been undertaken on auditory, haptic and multimodal human-computer interaction respectively. A detailed analysis of the guidelines and requirements to which a non-visual user interface should conform follows, and an interface has been specified to meet these requirements. The proposed interface could significantly reduce many of the difficulties of computer access for the visually impaired, fulfilling a real human need.

9.5 Further research

A large scope of further research could follow on from this work. It is suggested that detailed development and usability testing on the auditory icons and haptic device be undertaken seperately, with integration to follow. A third area for development is the development of the screen review portion of the system – the part which gleans information from the existing user interface. Further investigation could also focus more closely on the needs of partially sighted users, and those who prefer braille to synthesised speech.

References

Brewster, S. 1991, 'Providing a model for the use of sound in user interfaces', internal report, University of York, U.K. Retrieved: May 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Brewster, S. 2002a, 'Chapter 12: Non-speech auditory output', in Jacko, J. & Sears, A. (eds), The human computer interaction handbook, Lawrence Erlbaum Associates, USA, pp. 220-239. Retrieved: August 1, 2003, from http://www.multivis.org/.

Brewster, S. 2002b, 'Visualization tools for blind people using multiple modalities', *Disability and Rehabilitation Technology*, vol. 24, no. 11-12, pp. 613-621. Retrieved: August 21, 2003, from http://www.multivis.org/.

Brewster, S., Wright, P. & Edwards, A. 1995a, 'The application of a method for integrating non-speech audio into human-computer interfaces', internal report, University of York. Retrieved: August 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Brewster, S., Wright, P. & Edwards, A. 1995b, 'Parallel earcons: reducing the length of audio messages', *International Journal of Human-Computer Studies*, vol. 43, no. 2, pp. 153-175. Retrieved: August 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Bussemakers, M., de Haan, A. 2000, 'When it sounds like a duck and it looks like a dog - auditory icons vs. earcons in multimedia environments,' *Online Proceedings of the International Conference on Auditory Display*, April 2000. Retrieved: June 21, 2003, from

http://www.icad.org/websiteV2.0/Conferences/ICAD2000/ICAD2000.html.

Challis, B. 2000, *Design principles for tactile communication within the human-computer interface*, Ph. D. thesis, University of York. Retrieved: May 8, 2003, from http://citeseer.nj.nec.com.

Challis, B., & Edwards, A. 2001, 'Design principles for tactile interaction', Retrieved: May 8, 2003, from http://citeseer.nj.nec.com/460495.html

Colwell, C., Petrie, H., Kornbrot, D., Hardwick, A. & Furner, S. 1998, 'Use of a haptic device by blind and sighted people: perception of virtual textures and objects'. Retrieved: August 16, 2003, from Citeseer database, citeseer.nj.nec.com.

Crease, M., Brewster, S. & Gray, P. 2000, 'Caring, sharing widgets: a toolkit of sensitive widgets', *Proceedings of BCS HCI2000*, pp. 257-270. Retrieved: August 3, 2003, from http://www.multivis.org/.

Cullen, J. 1997, 'Virtual reality mouse', discussion list. Retrieved: August 16, 2003, from http://www.nfbcal.org/nfb-rd/1437.html.

Drewes, T., Mynatt, E. & Gandy, M 2000, 'Sleuth: an audio experience', Retrieved: May 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Edwards, A. , Stevens, R. & Pitt, I. 1995, 'Représentation non visuelle des mathématiques', in A. B. Safran and A. Assimacopoulos (eds), *Le Déficit Visuel*, Éditions Masson, Paris, pp. 169-178. Retrieved: May 1, 2003, from http://www.cs.york.ac.uk/maths/maths_publications.html.

Fricke, J. 1997, 'Different approaches to large tactile screens suitable for graphics', *IEE Colloquium on Developments in Tactile Displays*, vol. 12, pp. 6/1 -6/3. Retrieved: August 9, 2003, from IEEE Xplore database.

Fricke, J., & Baehring, H. 1993, 'Design of a tactile graphic i/o tablet and its integration into a personal computer system for blind users'. Retrieved: May 1, 2003, from http://www.rit.edu/~easi/easisem/tact-tab.html

Fritz, J., Barner, K. 1999, 'Design of a haptic visualization system for people with visual impairments', *IEEE Transactions on Rehabilitation Engineering*, vol. 7, no. 3, pp 372-384. Retrieved: May 8, 2003, from IEEE Xplore database.

Hollerbach, J. 2000, 'Some current issues in haptics research'. Retrieved: August 21, 2003, from Citeseer database, citeseer.nj.nec.com.

Howe, R., Kontarinis, D., & Peine, W. 1995, 'Shape memory alloy actuator controller design for tactile displays', Proceedings of the 34th IEEE Conference on Decision and Control, pp. 3540-3544. Retrieved: August 9, 2003, from IEEE Xplore database.

Jansson, G., & Ivas, A. 2000, 'Can the efficiency of a haptic display be increased by short-time practice in exploration?', in S. Brewster & R. Murray-Smith (eds), *Haptic Human-Computer Interaction: First International Workshop, Proceedings,* Springer, Berlin, pp. 88-97.

Jubinski, P. 1992, 'VIRTAC, a virtual tactile computer display', *Proceedings of the Johns Hopkins National Search for Computing Applications to Assist Persons with Disabilities, 1992*, pp. 208 -211. Retrieved: August 9, 2003, from IEEE Xplore database.

Kaczmarek, K., Tyler, M. & Bach-y-Rita, P. 1997, 'Pattern identification on a fingertip-scanned electrotactile display', *Proceedings of the 19th International Conference - IEEE/EMBS 1997*, vol. 4, pp. 1694 -1696. Retrieved: August 3, 2003, from IEEE Xplore database.

Kaczmarek, K. (n.d.), 'Tongue display unit', University of Wisconsin. Retrieved: August 1, 2003, from http://kaz.med.wisc.edu/tdu.htm.

Kamel, H., Roth, P., & Sinha, R. 2001, 'Graphics and user's exploration via simple sonics (GUESS) : providing interrelational representation of objects in a non-visual environment'. Retrieved: May 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Kirkpatrick, A. & Douglas, S. 2000, 'A shape recognition benchmark for evaluating usability of a haptic environment', in S. Brewster & R. Murray-Smith (eds), *Haptic Human-Computer Interaction: First International Workshop, Proceedings,* Springer, Berlin, pp. 151-156.

Kociolek, M., Wiecek, B., De Mey, G., Steenkeste, F. & Muntenau, C. 1999, 'A braille printer on reusable thermoplastic sheets', *Proceedings of the First Joint BMES/EMBS Conference: Serving Humanity, Advancing Technology*, vol. 2, p. 869-870. Retrieved: August 3, 2003, from IEEE Xplore database.

Konyo, M., Tadokoro, S., & Takamori, T. 2000, 'Artificial tactile feel display using soft gel actuators', Proceedings of the 2000 IEEE International Conference on Robotics and Automation, vol. 4, pp. 3416 −3421. Retrieved: August 9, 2003, from IEEE Xplore database.

Ludi, S. 2002, 'Are we addressing the right issues? Meeting the interface needs of computer users with low vision', *Proceedings of the 1st Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT),* pp. 9-12. Retrieved: August 3, 2003, from http://rehab-www.eng.cam.ac.uk/cwuaat/02/

Magnusson, C., Rassmus-Gröhn, K., Sjöström, C. & Danielsson, H. 2002, 'Navigation and recognition in complex haptic virtual environments - reports from an extensive study with blind users'. Retrieved: June 30, 2003, from http://www.certec.lth.se.

Mandic, D., Harvey, R. & Kolonic, D. 2000, 'On the choice of tactile code', *IEEE International Conference on Multimedia and Expo, 2000 (ICME 2000),* vol. 1, pp. 195 –198. Retrieved: September 24, 2003, from IEEE Xplore database.

Minsky, M. 1995, *Computational haptics: the sandpaper system for synthesizing texture for a force-feedback display,* PHD Thesis, Massachussets Institute of Technology. Retrieved: May 1, 2003, from http://theses.mit.edu/

Mynatt, E. & Weber, G. 1994, 'Nonvisual presentation of graphical user interfaces: contrasting two approaches', Retrieved: May 2, 2003, from Citeseer database, citeseer.nj.nec.com.

Mynatt, E. 1997, 'Transforming graphical interfaces into auditory interfaces for blind users', Retrieved: May 8, 2003, from Citeseer database, citeseer.nj.nec.com.

Oakley, I., McGee, M., Brewster, S. & Gray, P. 2000, 'Putting the feel in "Look and Feel"', *Proceedings of ACM CHI 2000*, pp.415-422. Retrieved: May 2, 2003, from Citeseer database, citeseer.nj.nec.com.

O'Modhrain, M., & Gillespie, B. 1995, 'The Moose: a haptic user interface for blind persons', Proceedings of the Third WWW6 Conference. Retrieved: May 1, 2003, from http://decweb.ethz.ch/WWW6/Access/ACC239.html.

Penn, P., Petrie, H., Colwell, C., Kornbrot, D., Furner, S., & Hardwick, A. 2000, 'The haptic perception of texture in virtual environments: an investigation with two devices', in S. Brewster & R. Murray-Smith (eds), *Haptic Human-Computer Interaction: First International Workshop, Proceedings,* Springer, Berlin, pp. 25-30.

Petrie, H., Morley, S., McNally, P., O'Neil, A. & Majoe, D. 1997, 'Initial design and evaluation of an interface to hypermedia systems for blind users', Retrieved: May 2, 2003, from Citeseer database, citeseer.nj.nec.com.

Poll, L., & Waterham, R. 1995, 'Graphical user interfaces and visually disabled users', *IEEE Transactions on Rehabilitation Engineering*, vol. 3, no. 1, pp 65- 69. Retrieved: August 9, 2003, from IEEE Xplore database.

Raman, T. 1997, *Auditory user interfaces: toward the speaking computer,*Kluwer, Boston.

Ramloll, R., Yu, W., Brewster, S., Riedel, B., Burton, A. & Dimigen, G. 2000, 'Constructing sonified haptic line graphs for the blind student: first steps', *Proceedings of ACM Assets 2000,* pp. 17-25. Retrieved: August 3, 2003, from http://www.multivis.org/.

Ramloll, R., Yu, W., Riedel, B. & Brewster, S.A. 2001, 'Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users', *Proceedings of BCS IHM-HCI 2001*, pp 515-530. Retrieved: August 3, 2003, from IEEE Xplore database.

Rasmussen, H. 2002, 'Training blind people to work with windows in a sighted environment', *Proceedings of Technology and Persons with Disabilities Conference 2002*. Retrieved: June 30, 2003, from

http://www.csun.edu/cod/conf/2002/proceedings/163.htm.

Roth, P., Petrucci, L., Assimacopoulos, A. & Pun, T. 2000, 'Audio-haptic internet browser and associated tools for blind and visually impaired computer users', Retrieved: May 2, 2003, from http://citeseer.nj.nec.com/roth00audiohaptic.html.

Ruspini, D., Kolarov, K., & Khatib, O. 1997, 'The haptic display of complex graphical environments', *Computer Graphics*, August 1997., vol. 31, no. 3A, pp. 345-352. Retrieved: August 21, 2003, from Citeseer database, citeseer.nj.nec.com.

Shinohara, M., Shimizu, Y., & Mochizuki, A. 1998, 'Three-dimensional tactile display for the blind', *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 4, pp. 249 –256. Retrieved: August 16, 2003, from IEEE Xplore database.

Silver, J., & Gill, J. 2003, *Tiresias – international information on visual disability*. Retrieved: May 1, 2003, from www.tiresias.org.

Sjöström, C. 1999, The IT potential of haptics - touch access for people with disabilities, Licentiate thesis, Certec, Lund University of Technology, Sweden. Retrieved: May 2, 2003, from Citeseer database, citeseer.nj.nec.com.

Sjöström, C. 2002a, *Follow-up experiments on haptic interaction design guidelines*, internal report, Certec Research Centre, Lund University of Technology, Sweden. Retrieved: June 30, 2003, from http://www.certec.lth.se/doc/followup/.

Sjöström, C. 2002b, *Non-visual haptic interaction design: guidelines and applications*, Ph.D thesis, Certec, Lund University of Technology, Sweden. Retrieved: August 1, 2003, from http://www.certec.lth.se.

Taylor, P., Hosseini-Sianaki, A., Varley, C., & Pollet, D. 1997, 'Advances in an electrorheological fluid based tactile array', *IEE Colloquium on Developments in Tactile Displays*, vol. 12, pp. 5/1 -5/5. Retrieved: August 9, 2003, from IEEE Xplore database.

Terry, J., & Hsiao, H. 1988, 'Tactile feedback in a computer mouse', Proceedings of the 1988 Fourteenth Annual Northeast Bioengineering Conference, pp. 146 −149. Retrieved: August 9, 2003, from IEEE Xplore database.

Treviranus, J. 2000, 'Adding haptics and sound to spatial curriculum', 2000 IEEE International Conference on Systems, Man, and Cybernetics, vol. 1, pp. 588 –592. Retrieved: August 16, 2003, from IEEE Xplore database.

Vanderheiden, G., Boyd, W., Ford, K., Kunz, D. & Mendenhall, J. 1992, 'Systems 3: multisensory access to the graphic based computers by persons who are blind', *Proceedings of the Johns Hopkins National Search for Computing Applications to Assist Persons with Disabilities, 1992*, pp. 197 -199. Retrieved: August 9, 2003, from IEEE Xplore database.

VirTouch, 2003. Retrieved: May 1, 2003, from www.virtouch.com.

Wagner, C., Lederman, S., & Howe, R. 2002, 'A tactile shape display using RC servomotors', Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 354 –355. Retrieved: August 9, 2003, from IEEE Xplore database.

Walker, A., Brewster, S., McGookin, D. and Ng, A. 2001, 'Diary in the sky: A spatial audio display for a mobile calendar', *Proceedings of BCS IHM-HCI 2001 (Lille, France), Springer,* pp. 531-540. Retrieved: August 16, 2003, from http://www.multivis.org.

Watanabe, T., & Kobayashi, M. 2002, 'A prototype of the freely rewritable tactile drawing system for persons who are blind', *Journal of Vision Impairment and*

Blindness, vol. 96, no. 6, pp.1-8. Retrieved: May 1, 2003, from http://www.afb.org/jvib/jvib960611.asp.

Way, T. & Barner, K. 1997, 'Automatic visual to tactile translation, part i: human factors, access methods and image manipulation', IEEE Transactions on Rehabilitation Engineering, vol. 5, no. 1, pp 81-94. Retrieved: May 1, 2003, from Citeseer database, citeseer.nj.nec.com.

Wies. E., Gardner, J., O'Modhrain, M., Hasser, C., & Bulatov, V. 2000, 'Web-based touch display for accessible science education', in S. Brewster & R. Murray-Smith (eds), *Haptic Human-Computer Interaction: First International Workshop, Proceedings,* Springer, Berlin, pp. 52-60.

Yobas, L., Huff, M., Lisy, F. & Durand, D. 2001, 'A novel bulk-micromachined electrostatic microvalve with a curved-compliant structure applicable for a pneumatic tactile display', *IEEE Journal of Microelectromechanical systems*, vol. 10, no. 2, pp. 187-196. Retrieved: August 4, 2003, from IEEE Xplore database.

Yu, W. & Brewster, S. 2003, 'Evaluation of multimodal graphs for blind people', *Journal of Universal Access in the Information Society 2003*, Retrieved: August 21, 2003, from http://www.multivis.org.

Yu, W. & Brewster, S., 2002 'Comparing two haptic interfaces for multimodal graph rendering', *Proceedings of the 10th IEEE Symposium on Haptic Interfaces for*

Virtual Environment and Teleoperator Systems, pp. 3-9. Retrieved: August 21, 2003, from http://www.multivis.org.

Yu, W., Cheung, K. & Brewster, S. 2002, 'Automatic online haptic graph construction', *Proceedings of Eurohaptics 2002*, pp. 128-133. Retrieved: August 3, 2003, from http://www.multivis.org.

Yu, W., Ramloll, R. and Brewster, S. 2000, 'Haptic graphs for blind computer users', in S. Brewster & R. Murray-Smith (eds), *Haptic Human-Computer Interaction: First International Workshop, Proceedings,* Springer, Berlin, pp. 41-51.