Stewart, Craig D. (2011) *Mobile navigation: A multimodal approach.* MSc(R) thesis.

UNIVERSITY
of
GLASGOW

Mobile Navigation: A Multimodal Approach

Craig David Stewart

MSc by Research

September, 2011

Abstract

The functionality and processing power of mobile devices has increased dramatically over the last few years. Location based services and rich interactions are feasible with the majority of smart phones available today. However, whilst the capabilities of current devices afford rich interaction tailored to the user in mobile situations, they are still linked with desktop style interactions.

Spatially situated virtual objects are used to represent multiple forms of information. Ranging from navigation beacons to places of interest and gaming objects. This thesis gives an review of the current literature of the use of virtual objects and examines the role of vibrotactile feedback for egocentric heading detection for virtual objects. Experiment results are also reported showing users can utilise vibrotactile feedback for heading acquisition. Possible future steps include combining directions and distance information for mobile navigation systems.

Acknowledgements

I am indebted to my long suffering colleagues in the wider GIST (Glasgow Interactive Systems Group) and my Swansea colleagues, Matt Jones and Simon Robinson. It is clear to me that none of the contents of this thesis would be possible without the support and conversations that I have had the enjoyment of having over the last few years. John Williamson, Andy Crossan, Eve Hoggan and Dave McGookin each brought their own experiences of post graduate research and I have benefited from their experiences.

Im grateful to my supervisors Stephen Brewster and Rodderick Murray-Smith, who have not had an easy task.

Declaration

I declare that this thesis was composed by myself, and that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified.

(Craig David Stewart)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The current generation of mobile devices shows a significant leap in both processing power and the number of sensors available to interaction designers. Motion sensors such as 3D accelerometers, magnetometers and gyroscopes, proximity sensors such as capacitive sensing and ambient light sensing, and applied force sensors such as Force Sensing Resistors (FSRs) are now found on mobile devices. Only a few years ago mobile devices augmented with such sensing would be the cutting edge of device development. Mobile phones have traditionally only supported communication such as phone calls and text messaging. More up-to-date smart phones have a relatively high amount of processing ability and this has allowed more features to be built into mobile devices. One of the first non communication-based features to be added was the ability to playback music files. The advent of cheap GPS modules allowed the GPS location to be built into smart phones thus offering turn by turn navigation without the need for a single use device. Positional data combined with motion sensing has allowed mobile Augmented Reality (AR) systems to be realised commercially. Additionally the advent of financially cheap mobile internet services has allowed the merging of digital information with the user's experience of their environment. Commercial systems, ie Layar [57] have already employ this approach.

Initial research exploring methods of combining digital information with our perception of our environment has its roots in augmented reality and situated information spaces [22]. Initial research prototypes required physical tokens [3] to denote items with associated virtual information. Digital information was presented to the user by using visual overlays with a strong dependency on Head Up Displays. These systems all share a common difficultly of producing and placing the physical representations of the tokens. This problem has largely been solved with the usage of GPS coordinates; it is now possible to mark locations without the use of a physical marker.

Being able to virtually tag locations with information is not a new approach. Mobile augmented reality systems have been constructed using visual overlays to show the user information about the features, buildings and natural objects in their environment. Augmented reality systems all share a common root. Fitzmaurice [22] was one of the first to introduce the concept of situated information spaces where objects around people can be augmented with representative digital data. Physical objects were augmented by a visual display, placed in between the user's line of sight and the physical object. The visual device acted like a porthole into a world where the virtual aspects of the object fused with the physical. While most of the scenarios suggested by Fitzmaurice were indoors interacting with office furniture (i.e. fax machines in the office, or shelving units in the library) there is still scope for designing interactions based on augmenting larger static objects; for example buildings, landmarks and mobile objects, public transport and individual persons. Systems have been built to explore virtual objects placed in our environment using visual augmentation [20].

Questions arise over the ability of users to be as aware of dangers in their environment when they use a visual-based augmented reality system. The use of non-visual feedback such as vibrotactile and audio has been used to overcome such problems of placing too much visual demand on users [38].

1.2 Research Scope

This work is concerned with using low attentional demand, non-visual interactions to allow users to explore their physical environment. Our experiences are increasingly being stored digitally, either to augment our own memories of prior experiences or to mediate our current experience. Building systems that allow ad hoc interactions with such spatially situated information would seem to be of benefit.

1.3 Research Aims

This work aims to contribute to the larger literature by:

- Providing an overview of the current literature concerning designing and implementing a novel interaction metaphor for user interactions of virtual objects in a situated information space.
- Exploring and evaluating non visual feedback for use in heading acquisition of virtual objects.
- Suggesting interaction metaphors that could be used to extend this work.

1.4 Overview

The remainder of this thesis has three chapters:

- Chapter 2 contains the literature review of mixed reality systems and an overview of non-visual feedback techniques
- Chapter 3 describes the logic behind the heading acquisition system design, evaluation and possible extensions of the work described
- Chapter 4 explores the potential expansion of the work detailed in Chapter 3.

Chapter 2

Literature review

Systems built on the basis of typical human computer interactions are largely focused on the user issuing commands to the computer and the computer dutifully carrying out those commands. This method for interaction is useful when there are few competing demands on the user's attention and minimal uncertainty of user intent and system state. This interaction approach also tends to support closed ended interactions, where only simple command and response interactions are required, for example creating a line in a drawing application. Interactions in mobile scenarios introduces uncertainty of state and intent for both the system and the user [16]. Without taking such knowledge of the nature of the noise in the user's input, mobile systems cannot interpret user intent correctly. Thus the difference between what the user believes the results of their input and the system's belief of the intention of the input becomes larger.

Another key aspect of designing mobile interactions is the provision of noninvasive presentation of information. It is certainly possible with current technology and interaction techniques that when a user walks by an area of pre-defined interest that the user will be alerted to information that is linked to that physical space. Digital information presented as virtual objects can allow for a blurring of the perceived separation between information stored digitally and the physical world around us.

This literature review will offer an overview of mixed reality systems, identify some examples of visual, audio and tactile based augmented reality systems, discuss the definition of virtual objects and provide an overview of audio and tactile feedback techniques for mobile systems.

2.1 Mixed Reality

The concept of conveying virtual information as objects to the user is not new and there is an extensive literature dealing with the interaction with virtual objects in mixed reality systems. Approaches investigated have ranged from immersive virtual reality scenarios [14] to environments that use multimodal feedback to allow users to hear sounds spatially around them and feel external objects using haptic devices in the virtual environments [9] [49] [47] [5] [21]. While many of these approaches work well in their given scenarios they are not easily transported into mobile settings. It is prudent to have a review of the literature focused on the methods used to convey virtual information to users and systems that have been built to explore the uses of such systems. Mixed reality systems allow physical objects to be extended by virtual attributes. Systems using this approach could help users involved in complex construction or maintenance tasks [21] or train surgeons by overlaying images of the patients scans [6] and showing them the techniques needed to do the surgery.

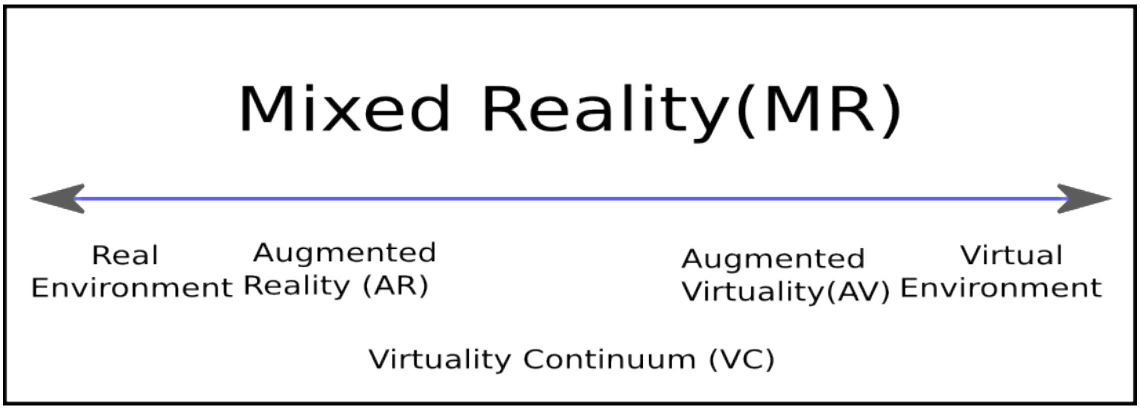


Figure 2.1: Mixed reality continuum, taken from [36]

.

A continuum of mixed reality systems can be found by Milgram [36] (Figure 2.1). His distinction between augmented reality and virtual reality is that while for virtual reality the surrounding environment is artificial, for augmented reality the environment is real. Mixed reality systems aim to fuse computer-generated stimuli and the natural stimuli produced by the environment in which the user is situated. The majority of mixed reality systems are hugely dependant on visual stimuli. However [4] used visual and audio to create more immersive and believable mixed reality systems. Mynatt *et al.* [38] used only audio cues in Audio Aura , while tactile feedback has been used for navigation aides, i.e. [18].

2.1.1 Visual Based Augmented Reality Visual Displays

As described by Bimber *et al.* [4] augmented reality visual displays are image-forming displays that generate images between the physical object and the observer. These display devices take numerous forms including projectors, head-mounted displays and handheld devices. They can be classified into three major groups: Video-see-through, Optical see-through and Head-mounted projection display. Video-see-through systems use a video camera to capture a video stream of the real environment and merge the virtual objects in real time with the video stream. Optical-see-through use a variety of different optic techniques to merge the two scenes visually. In contrast head mounted projection displays, instead of directing images directly the eyes of the user, take the opposite approach and project onto the physical environment [1] [28].

Head Mounted Displays

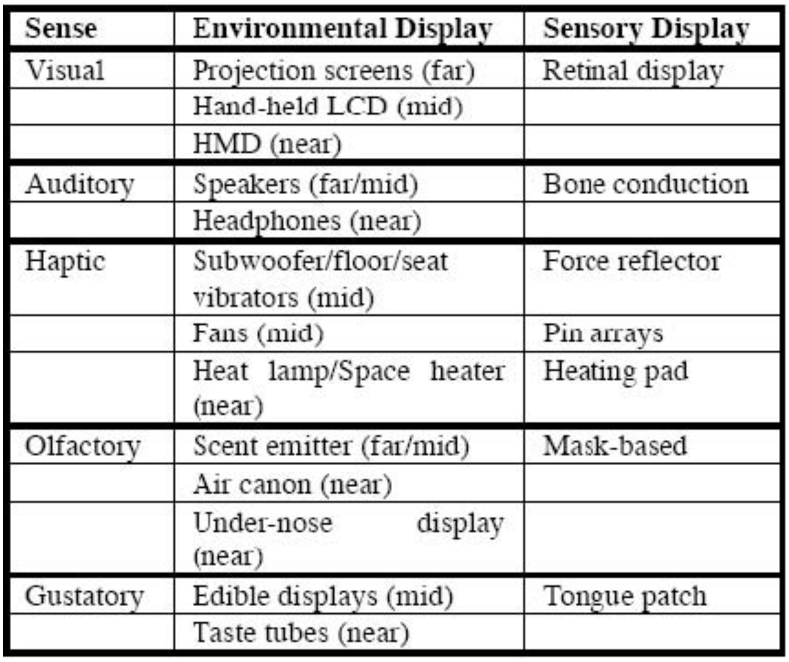
Two major groupings of HMD's have emerged: retinal displays and head-mounted displays. Both provide direct luminance to the eyes. Retinal displays, as their name suggests, project an image directly on the retina of the user while head-mounted displays are small LCD screens mounted in front of the eyes. Common issues with both retinal and head-mounted projection displays include increased user discomfort due to motion sickness induced due to mismatch between visual and display planes, limited field of view and lack of resolution. While head mounted projectors provide a larger field of view to the user they suffer from being relatively low resolution and being cumbersome. Their lack of mobility reduces their applicability to mobile Augmented Reality scenarios.

Handheld Displays

Handheld displays have been described as a "window-on-the-world" [36] and an "eye-in-hand" [22]. Due to limitations with LCD displays, hand-held displays typically use Video-see through as the merging approach for virtual object and the real environment. The use of non tethered handheld devices (PDAs, mobile phones) has been limited due to the historical lack of processing power available on such devices. However, the more recent increases in processing power available from mobile devices has allowed more recent systems to be realised [17]. Additionally commercial applications are available for major mobile operating systems [56] For interacting with virtual objects finger-tracking techniques allowed the system to recognise rotation and translation gestures by the user.

2.2 Non-Visual Based Augmented Reality

Most augmented reality systems have concentrated on visual overlays as the main modaility for information transfer about virtual objects. However there has been work carried out to access the potential of using other senses. By either combining 3D audio with a visual overlay using audio only or as in the case of navigation aids, tactile feedback. Lindeman *et al.* [30] have suggested a selection of different displays, some more plausible than others, that could be used for augmented reality systems 2.2.



Sense	Environmental Display	Sensory Display
Visual	Projection screens (far)	Retinal display
	Hand-held LCD (mid)	
	HMD (near)	
Auditory	Speakers (far/mid)	Bone conduction
	Headphones (near)	
Haptic	Subwoofer/floor/seat vibrators (mid)	Force reflector
	Fans (mid)	Pin arrays
	Heat lamp/Space heater (near)	Heating pad
Olfactory	Scent emitter (far/mid)	Mask-based
	Air canon (near)	
	Under-nose display (near)	
Gustatory	Edible displays (mid)	Tongue patch
	Taste tubes (near)	

Figure 2.2: An overview of possible displays for augmented reality systems .Taken from [30]

2.2.1 Audio Based Augmented Reality

Audio Aura [38] was designed for the office environment. Users would gain awareness of the number of emails they currently had to deal with and the activities and locations of co-workers. Rather than have the majority of this information being produced at the user's command, audio feedback was given when users walked into certain areas, for example over time you went to the coffee house, an audio reminder of your current inbox condition was played. Mynatt *et al.* [38] envisaged the following three scenarios for their system, checking email, remote group awareness and local person awareness.

Email Checking

It was commented that individuals change their coffee break habits depending on how many emails, hence work that they have to read. One possible issue that occurs when visiting a coffee shop is how long can be spent lingering talking to colleagues before the need to rush back to the office to check emails. It was imagined that every time you enter a coffee shop that you were given an auditory reminder of how many emails you have in you inbox and whom they are from. The designers hoped that it could help people make the decision on how long to stay for.

Remote Group Awareness

Collaborating groups are not forced to be located in the same area; either in the same building, town or country. By providing a 'group pulse' an auditory cue with the features of the cue being determined by the location of group members and their activities, it was hoped that by each member of the group being aware of the remote groups actions, if they are in their offices or not that a shared experience could be express, thus helping group cohesion.

Local Person Awareness

Similar to remote group scenarios, audio cues conveying qualitative information on the person of interest's whereabouts are conveyed when someone tries to find them and they are not present in their office. The idea that when entering an empty office an audio cue indicating if you have just missed the person of interest, if they have been in today is interesting as it encapsulates the notion of users leaving marks on the system and then leaving the area in which the mark is located.

Audio Cues

Local person awareness and email checking are event and location driven. When the user goes for a coffee or searches for someone cues are given to augment what the user is already experiencing. Group awareness cues are however a constant presence; thus becoming a backdrop for the other cues.

Music

By using 'earcons'; non verbal structured abstract sounds [34][51] Mynatt *et al.* [38] mapped different aspects of the information obtained by the system to the number of notes contained, rhythm and pitches of the notes making up of the earcons. For example; when the user has no new emails a high short bell melody with a rising pitch is played, this is in contrast to a long melody with the pitch falling at the end for when the user has a lot of emails.

While acknowledging the problem of overloading frequency bands, the designers sought to combine the previous worlds in a "rich, multi-layered layered environment". Unfortunately a strong evaluation of the system contrasting the different types of audio feedback discussed was not undertaken. Seven volunteers used the system with sound effect cues being used. While user impressions were favourable, users felt that the sounds played for too long and that it was difficult to remember what all the sounds meant.

These scenarios suggested by Mynatt *et al.* give an insight into the possibilities of using non visual augmented reality and interaction with virtual objects. However, they placed a limitation on the interaction afforded by their system, by forcing the majority of interactions based on the location of the user e.g. it is not possible to point to an office and find out if the owner is currently there. When designing outdoor mobile system for interacting with virtual objects, the majority of the interaction will be based on the users pointing the locations of interest.

2.3 Tactile AR

Gallace *et al.* [24] suggest that tactile interfaces have been used to support users in, but not limited to, the following scenarios; resolving spatial disorientation, waypoint based route finding and manipulation of visual attention.

2.3.1 Resolving spatial disorientation

As identified by Rupert [46] when flying in clouds, pilots often become disorientated due to the lack of visual references. As a result of this disorientation pilots are at risk from enemy activities and from their own mistakes. In an attempt to reduce disorientation, aircraft engineers developed artificial horizon systems in the belief that by providing pilots with a natural reference point and their current pitch and roll. Initial problems with these systems seemed to be based on either pilots not trusting their instruments or pilots finding it difficult to monitor multiple sources of visual information. Rupert noted that even with the introduction of Head Mounted

Displays the rate of spatial disorientation (SD) mishaps did not fall. Using the Tactile Situation Awareness System (TSAS) Rupert [46] aimed to use a vibrotactile vest (a vest that contains a mesh of actuators). TSAS activated actuators on the vest to give pilots an indication of where the gravity vector was. For example; if the plan was level and in correct orientation the pilot would experience vibrations in the centre of their abdomen, and if the aircraft was inverted actuators on the shoulder would be activated. In a similar method if the aircraft banked to the left the system would activate the actuators on the lower left side of the vest and as the aircraft continued to increase the tilt the system would activate actuators going further up the vest. Rupert experienced technical issues with actuators failing and aircraft vibrations masking the vibrations the actuators were producing. However, the system still proved successful at assisting pilots with visual instruments and cockpit windows covered complete aerial manoeuvres correctly.

2.3.2 Waypoint Based Route Finding

While Rupert's system [46] focuses on a unique situation not faced by the majority of people Van Erp *et al.* [18] have studied the navigation uses of a vibrotactile-enabled waist belt. By using belts containing multiple actuators, researchers have designed a system to present directional information to users who are either pedestrians, helicopter pilots or people driving fast boats. As identified by Burnett *et al.*, [10] direction and distance are important parameters for waypoint navigation. Direction information was encoded into 8 actuators placed in a vibrotactile belt that was placed around the user. If the next waypoint to reach was right behind the user actuator 5 would be activated (see Figure 2.3). Encoding distance into the stimulus did not have a similar natural mapping. Therefore the rhythm, or rate, of the stimulus was presented, depended on the distance of the waypoint, i.e. if the way point is far away the rhythm is slower than if the waypoint was closer. Different methods for mapping the rhythm to distance of waypoint were tested, however none performed significantly better or worse than the simple case previously illustrated.

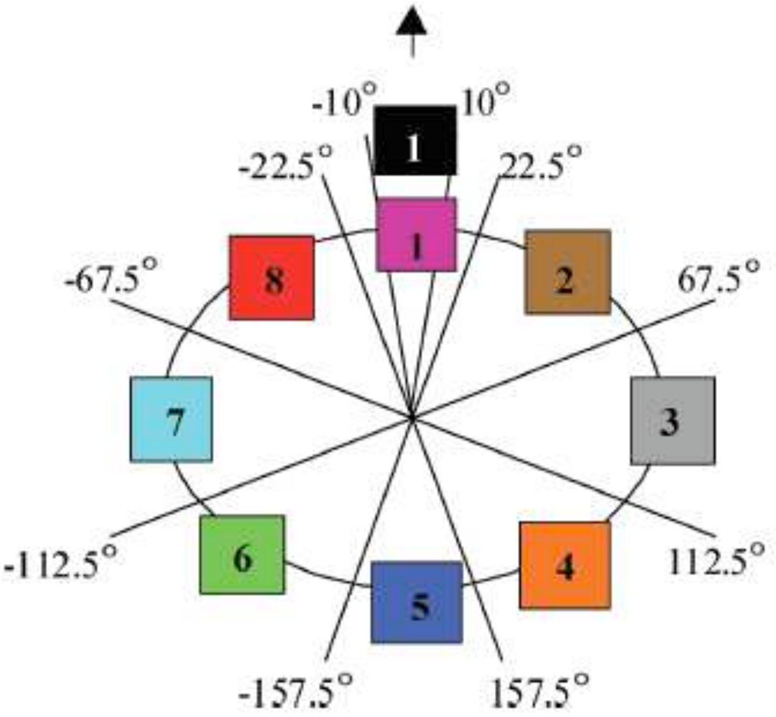


Figure 2.3: Numbering and bearing placement of 8 actuators placed around a subjects waist .Taken from [18]
.

As a result, van Erp *et al.* decided to test the pilots and boat drivers without any distance information as it was suggested that for the metrics used for testing, (speed of subject travelling to waypoint) distance from waypoint is not as important as the bearing. The experimenters also made another change to allow for finer feedback when the user is close to the correct bearing. As shown in Figure 2.3 the black box containing 1 at +/-10 degrees represents the faster rhythm rate of the stimulus presented when the user is within +/- 10 degrees of the correct bearing for the waypoint. Subjects for both scenarios had no problems with following the waypoint, with the helicopter pilot showing signs of learning the system easily. Unfortunately the results did not conclusively show that the rhythm enhancement for correct bearing selection for pilots and boat drivers increased their navigation performance. While the results published are not entirely positive they are indicative of an interesting research areas to explore. This system, along with TSAS shows the potential for using multiple actuator based displays for representing spatial information of external objects, in these cases, either waypoints or the gravity vector were represented by a virtual object generating stimuli over the areas of actuator placement.

2.3.3 Manipulation of Visual Attention

Vibrotactile stimuli have been found to have effects on other modalities; i.e. cross modal effects. As reviewed by Ho *et al.* [25] there is a large resource of studies which suggest that by giving users vibrotactile cues they can direct visual attention to the area suggested to by the vibrotactile stimuli. In a simulated driving study, Ho *et al.* [25] asked subjects to watch a display that contained three video streams; one for the front windshield of the car, another for the rear view mirror and a final one for an attention required task. During the study subjects were instructed to do two actions; pick out digits from a stream of digits and letters for the attention task and avoid being hit by a car by either pressing the accelerator or brake. During the driving simulation cars would either be getting closer to or further away from the user's car. In critical sections, when a collision would be eminent, a vibration would be produced from one of two actuators placed on the back and front of the subject's torso. When the subjects felt a vibration on their back they needed to check their rear view mirror and similarly for a vibration on the front of their body check the main display. Appropriate responses from the subjects required them to press the accelerator when a car was going to collide with the rear of their car, or press the brake when they were going to hit a car in front of them. The direction of the vibrotactile stimulus did not always suggest the correct direction of danger. To provide a comparison, a no-cue experiment was conducted with no vibrotactile stimuli being presented. This comparison suggested that even if the spatial cueing of the vibrotactile stimuli was sometimes wrong, they still assisted subjects in responding faster and more accurately in critical situations.

2.4 Virtual Objects

Regardless of the sensory modalities that we use to convey where virtual objects are and their characteristics, we need methods for manipulating them and either creating, moving or changing their physical characteristics. The following systems embody interactions with virtual objects placed in the physical environment.

2.4.1 Situated information spaces

Using a small, high resolution and portable visual display unit, Fitzmaurice [22] created a device to act as an "eye-in-hand" window into a 3D-situated, synthesised information space. This portal to a synthesised space allowed the designers to position virtual sources of information onto the physical objects by overlaying the virtual information onto the physical object. The user then interacted with this information according to the location it was found on the physical device.

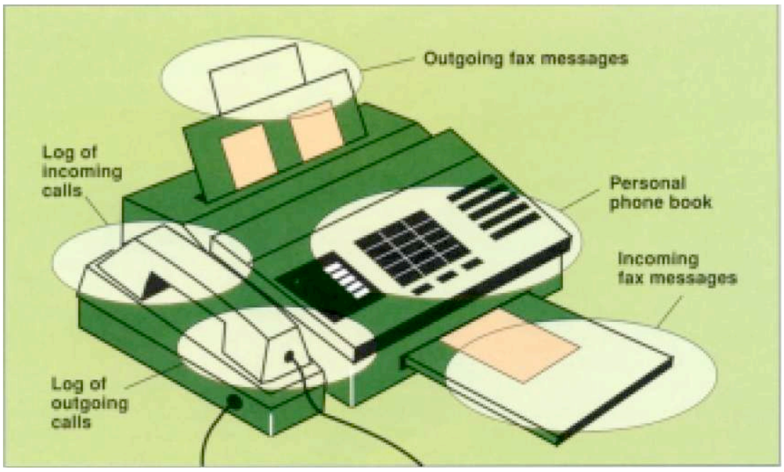


Figure 2.4: A fax machine augmented with digitally stored information. Taken from [22]

.

A fax machine, Figure 2.4, was used as an example of a physical object being augmented in this way, using the keypad on the device to situate the virtual phone book and similarly the earphone on the handset being the position of the incoming call log. Fitzmaurice suggests a key for the user to be able to move the virtual objects to "provide a logical means of partitioning and organising the associated information space and serves as a retrieval tool for users". While exploring this porthole metaphor for augmented reality scenarios, Fitzmaurice felt that by merging both the real world and the synthesised virtual space, a composite optimal medium could be produced that had the strengths of both environments.

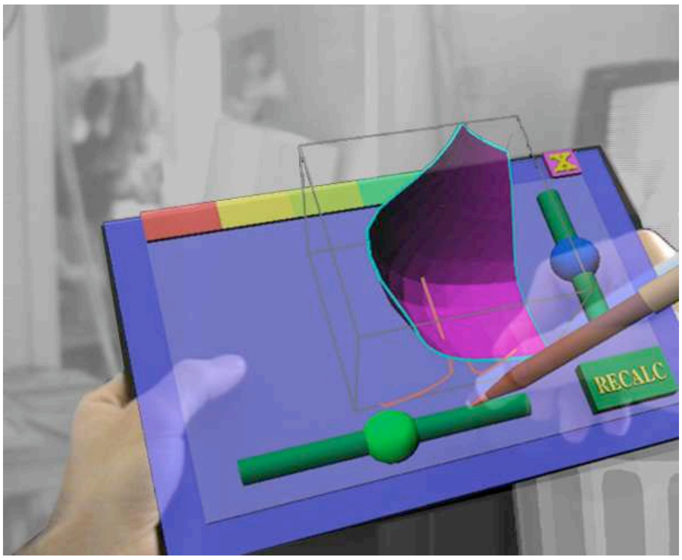


Figure 2.5: Object view zoom tools. Taken from [51]

2.4.2 Personal Interaction Panel

The majority of Augmented reality papers are focused on the technical aspects of visual overlays, as part of the Studierstube Augmented Reality Project [51] , Szalavari *et al.* in their Personal Interaction Panel (PIP) focused on how people interact with and manipulate 3D virtual objects in augmented reality systems [15]. It is suggested that 3D widgets such as virtual menus and buttons are difficult for people to use as there is no tactile feedback. As such it is further suggested that; "rather than offering virtual devices for manipulation tasks we propose extended devices" extending the real world tools by added virtual shape and functionality [59][51]. Investigation of the user of extended devices involved the construction of a head mounted display. The display presented virtual objects and extensions to the physical tool, acting as the extended device that the user could directly control. The extended device comprised of a graphics tablet and pen that allowed the head mounted display, worn by the user, to make controls appear on the tablet that the user could manipulate as shown in Figure 2.5.

Interactions using the PIP Interface

Using their extended device, a selection of object manipulation tasks was explored: selection of objects, moving objects from the PIP into the augmented space, rotation of virtual objects, changing camera position and general object customization (colour, size, etc.). In the case of object rotation the currently selected device would appear on the pad (shown bottom left of Figure 2.6) alongside the controls that were appropriate for the intended operation.

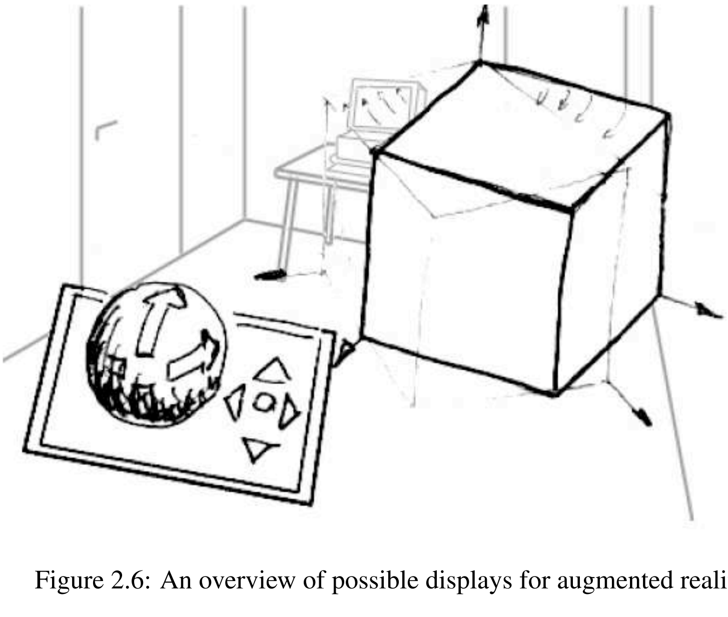


Figure 2.6: An overview of possible displays for augmented reality systems .Taken from [30]

2.5 Augmented Reality Mobile Systems

While the previous systems described could be moved around in a room, it was not until later that augmented reality systems were built expressly for use in mobile scenarios. One of the first outdoor AR systems, specifically developed for assisting in navigation was the Touring Machine [20]

2.5.1 Touring Machine

Feiner *et al.* [21] took advantage of a variety of mobile tracking technologies, GPS, magnetometer and inclinometer to allow for tracking of both the location and current direction of the user. Similar to the PIP, they used a HMD and a 2D display with pen and track pad input. However, only the head was tracked for this system and instead of using the HMD to fill in the tablet display the device overlaid the user's vision with its images. A battery belt powered the mobile system. The system was designed to give information about the buildings around a university campus and allow access to web pages about those buildings that would be shown on the display in the form of labels. Areas that had additional virtual content were represented as grey text overlaid using the HMD on the users view. The intensity of the label's brightness was determined by their position relative to the centre of the screen and the label closest to the centre changed to yellow. When the system is in 'gaze-directed selection-mode' if the same label is highlighted yellow for longer than a second, the label is selected and the user was able to access a menu containing the information about that building. It was also possible to get navigation cues from a pointer on the bottom of the HMD as the pointer always points in the direction of the currently selected building. In a similar system Hollerer ***et al.*** [52] represented information objects as flags. Users could create a path through their campus on a desktop machine and have a path projected as a pipe accented by flags. When using the mobile version the HMD would display the pipe for the user to follow. The augmented stroll [43] is another example of this type of system.

2.5.2 Human Pacman

An interesting system was developed by Cheok ***et al.*** [11] mirroring the game Pacman in an augmented reality scenario. Real people take on the roles of Pacman and ghosts, characters in the game, and to gain points have to pick up 'cookies' placed on paths. Users wear a HMD, like previous systems mentioned, but unlike the previous mobile systems some of the virtual objects have corresponding physical objects as well. Normal 'cookies' are only shown to the Pacman and the player collects them by walking through the area they are placed in by the HMD. 'Special cookies' are in fact Bluetooth enabled boxes that the user has to physically pick up. Similarly a ghost can only 'eat' the Pacman by touching the player on the shoulder. Other augmented reality games with less mobility include ARQuake [41] and AR2 Hockey [39].

The described research attempts at using virtual reality in mobile scenarios are typical of the approach taken by many designers to build systems of this nature. Either by augmenting physical features, i.e. buildings, public spaces or even individual parts of office furniture, with virtual objects embodying digital information, or by situating completely virtual information on top of the physical environment. Information spaces were initially explored by Fitzmaurice [22] and the majority of augmented reality projects can find their roots in his work and the work of others in Virtual Reality scenarios. The initial use cases for information spaces focused on augmented physical objects indoors, typical office tools and library furniture. Personal mobile scenarios in augmented reality have traditionally focused on environmental discovery, either for navigation or information placement (i.e. building and landmark identification) and games. Identification of these virtual objects has mostly relied on visual cues, either from Head Mounted Displays or palm sized displays.

2.6 Non Visual Feedback

Users in mobile situations need to be able to concentrate their visual attention on their primary tasks: walking, cycling or even driving. This is in contrast to desktop systems where the user can be focused solely on any interactions with the computer. Visually dominated interaction techniques are difficult to use in mobile scenarios, either using a separate visual display or Head Mounted Displays; this is due to the visual cues distracting the users from their primary tasks. Mobile devices that rely on a multimodal approach to feedback mechanisms have been shown to increase user's ability to use the device for a given task [32] [48]. Using audio and/or tactile feedback in this manner allows interaction designers to construct multimodal techniques that allow users to engage in meaningful interactions while some senses are fully utilised in other activities.

Researchers have made the following observations of the interactions made by users in mobile contexts:

- Mobile devices reserve our physical and attention capabilities that can lead to changes in how the user moves through their environment [40];
- Interactions and mobility are often partially exclusive (for example it is difficult to text on your mobile phone and walk at the same time);
- When there is competition for cognitive resources our ability to navigate through the environment safely is compromised [29];
- Users need to be allowed to visually assess their environment, interaction design based solely on visualisation is difficult to do in this setting [32].

Simple messages can be presented to the user by using audio and/or vibrotactile cues [7]. Embodying information via audio feedback , for example Earcons, has been explored previously as has displaying complex information via vibrotactile stimuli. The following sections will explore these.

2.6.1 Audio Feedback

Auditory notifications in the form of ringtones have been used for many years in the notification of SMS messages and incoming phone calls. The use of audio cues is not constrained to providing binary notification of events. Interaction designers can use speech synthesis and 3D audio techniques to present a large amount of information to the user. While the use of auditory displays is useful it does encroach into the social context of the user. In some scenarios it is not appropriate for the mobile device to be giving information to the user in a public fashion. While headphones can be used, and indeed the popularity of listening to music while on the move makes it a reasonable to expect such usage, it can act as a barrier between the user and the environment.

The encoding of information within audio cues has been the focus of sustained HCI research over the last two decades. Hoggan [26] and McGookin [35] provide an excellent review of this work, from which the following summary is derived.

Two main methods for audio information encoding are Auditory Icons and Earcons. While Auditory Icons are natural sounds with semantic links with the action/event they are tied with, Earcons are abstract sounds, whose meaning the user has to learn. Mcgookin[35] notes that Auditory Icons have one crucial weakness, in that semantic links are not always possible. It is difficult to discover a natural sound for abstract actions, such as renaming a file. Earcons can be parameterised by changing the musical qualities of the sound, such as the pitch, rhythm and timbre. This allows more information to be encoded with the presented audio feedback, once the user has learned the musical quality mapping.

Pitch

The pitch of a musical note can be used to distinguish separate cues from a group of similar events. Using pitch by itself has been shown to perform poorly. However in combination with other musical qualities it can be useful to extend the number of possible mappings.

Rhythm

Rhythm is the musical quality that makes a piece of music move through its natural timing. Care must be taken when designing earcons that take advantage of different rhythms. The length of time the earcon will have to be presented, as to be distinguishable, can dramatically slow down interactions between the user and system.

Timbre

Timbre is the quality of a sound that makes it distinguishable from other sounds produced with a similar pitch. The sound produced by trumpet playing a middle C, for example, is distinguishable of that from a clarinet producing the same note. The ability of human hearing to tell apart these different sounds allows the use of timbre as a parameter in earcon design. In combination with rhythm and pitch McGookin [35] achieved 90 % identification rates with three different instrument types.

2.6.2 Vibrotactile Feedback

Human Computer Interaction continues to focus on visual feedback mechanisms with attempts at supplementing that interaction with tactile and audio feedback. However, there have been attempts to use vibrotactile feedback by itself to convey information. This is understandable as it affords a rich capacity communication channel to users due to the large surface area for possible interaction. For all mammals the largest organ, by surface area, is the skin. The sensitivity and range of receptors on the skin are highest of any sense [24]. However, using the skin as a medium for feedback is still rare. Visually impaired people have used a variety of techniques to gain information about their environment through their sense of touch. Electronic Braille systems [53] have been popular to help blind users interact with books and digitally stored information while not having to wait for a Braille edition of the literature of interest to be published. Moving from digital substitution of Braille books, systems have been developed to create tactile representations of the visual view of the subject's current environment[54][27]. While these systems were exploring possible uses for the technology it became quite clear that the ability of the skin as a sensory medium to give clear reflections of sight was not possible. While the technology used was substantial, users could only recognize simple object outlines. Complex objects like faces could not be separated from the multitude of objects in the environment.

Using vibrotactile techniques rich feedback with analogies with audio can be used. The interaction afforded can be discrete, for example if the mobile device is in a pocket it does not interrupt a meeting while alerting the user of an event (phone call or text message etc). As identified by [53] the use of vibrotactile stimulus to convey information to the user has been researched for almost 50 years however; due to technology restraints applied use of this research has been limited.

As previously identified, mobile interactions are difficult due to the increased visual demand placed on the user due to navigation of the physical environment. While it is possible to increase the utility of mobile interfaces by making use of multimodal feedback mechanisms, non-visual feedback techniques suffer from having limited communication bandwidth. Current approaches to the design of non-visual feedback do not fully take into account the resulting increase in cognitive demand.

Ambient touch is an example of such a system that has been developed for the exploring the usefulness of vibrotactile feedback for mobile devices. Popyrev et al [42] identified that tactile feedback is an excellent attention grabbing technique that requires little focused attention to notice. Mobile interfaces designers need to be concerned with not only reducing visual demand, but overall cognitive demand. Popyrev [42] suggests that cognitive load for vibrotactile stimuli is highest for precise control and is lowest when it is a simulation of real-world tactile feeling (as summarised in Figure 2.7). Therefore, while non-visual interfaces may allow 'eyes-free' interaction, if the feedback design requires a significant increase in cognitive resources, it may result in users performing poorer than expected in real world tests.

In a similar way to Earcons, vibrotactile feedback can be used to create structured abstract vibrations to convey information to the user. Vibration based Earcons can

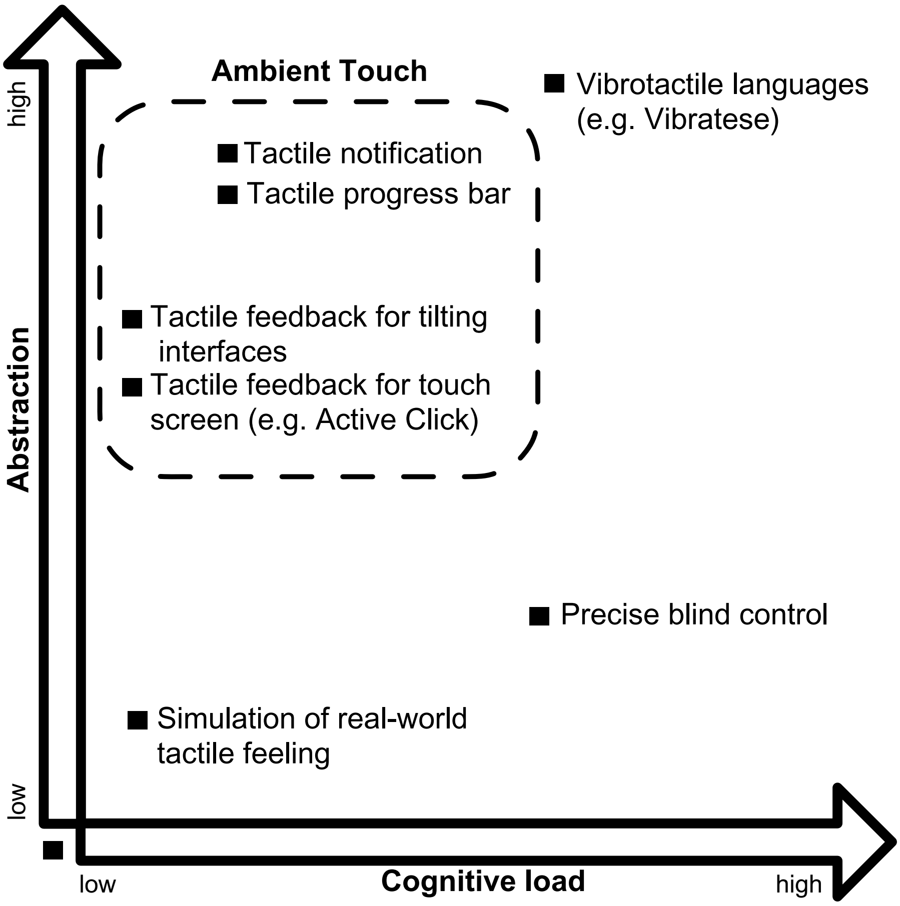


Figure 2.7: Cognitive load for vibrotactile feedback taken from [42]

been termed Tactons [8]. As identified by [17] information can coded by 4 different features of a stimulus, subjective magnitude, frequency, location and temporal patterns of stimuli.

Subjective Magnitude

While the human body has a highly variable sensitivity to a stimulus being applied on different body locations, the perceived magnitude of stimulus (a 'non linear function of amplitude') can be used to encode information. To increase the subjective magnitude of the stimuli presented to the subject, designers can either increase the area of skin being stimulated, or increase the amplitude of stimuli near the detection thresholds.

Frequency

The wide range of sensory receptors in the skin are effective at frequencies ranging from 60hz to 400Hz [13]. It is possible to encode information as being different frequencies, for example in a navigation scenario it would be possible to encode the distance you are from a way point by changing the frequency e.g. the closer you are the higher the frequency used. However user's can only perceive five to seven levels of frequency [26] and choice of actuator can limit the range of frequencies used.

Location

The location of a stimulus can also be used to convey information to the user. Applications using vibrotactile vests that contain multiple actuators have been popular for navigation aids and resolving spatial disorientation [18][46]. Utilising location in this fashion requires actuators to be placed around the body to be able to present the spatial information. Temporal perception illusions can be used to create apparent movement and phantom sensations to cause subjects to sense vibrations placed in between where actuators are actually located.

Temporal Patterns

Temporal integration of multi point vibrotactile sensations is well understood. Cholewiak *et al.* [12] Loomis *et al.* [31] and Bekesy [2] have a consensus of data to show the stimuli parameters required for two stimuli to be perceived as being separate. For two stimuli to be perceived as two separate events, the temporal separation has to be more that 5 ms. The parameters used are also influenced by the response times of the actuators used. While reported parameters vary slightly, there is still a consensus for the duration of stimuli and the gap needed between them for users to sense two stimuli rather than a single merged percept. For example, if two vibrotactile stimuli are placed on the forearm it is possible for subjects to feel a single stimulus location placed in between the stimuli points. With subtle changes in the relative temporal components of the stimulus it is possible to create a sensation that moves from one actuator point to the other. It is important to note the differences between apparent motion and phantom sensation. Apparent motion is the illusion of a stimulus moving from one stimulator site to another, while phantom sensation places a merged percept in-between the two stimulated sites. Evidence exists for suggesting that apparent motion and phantom sensation are linked illusions [2].

2.6.3 Actuators

Commonly used actuators used for the generation vibrotactile stimuli include speaker-based, pager motors and piezo electric. Speaker-based actuators, such as the C2 (as used in [26]), have the ability to change the frequency of stimulus presented while maintaining high levels of vibration amplitude. While speaker-based feedback provides feedback designers more parameters of which to base their feedback, the high power consumption of such actuators have reduced the scope for inclusion in commercial mobile devices. Pager motors are currently the most commonly used actuator in mobile phones available today. They can be easily controlled with a limited amount of additional electronics. However, pager motors are tuned to a particular frequency and cannot be changed on-the-fly. Piezo electric actuators allow for the production of vibrations in a range of different frequencies. Such actuators have lower power requirements than speaker based actuators. However, the additional electronics to produce the high voltages required have reduced their implementation on commercial devices.

2.7 Conclusion

This chapter has provided an overview of augmented reality systems using visual, audio and vibrotactile feedback. The technical capabilities of visual-based augmented reality systems have improved since the first prototypes. There are, however, concerns about using augmented reality for mobile scenarios. Competing visual demand between cues provide by the user's environment and cues presented by the system can decrease the user's performance of their main task, i.e. driving a car [19]. Vibrotactile cues have been utilised to assist in navigation tasks and are less intrusive than audio based cues. The following chapter will explore, in more detail, the presentation of simple vibrotacile cues to provide navigation aids for mobile users.

Chapter 3

Mobile Exploration

The term mobile interaction has been used as a blanket term for defining interactions that are conducted while the user is moving, in a mobile context (such as sitting on a bus) and more simply all interactions with a mobile device, regardless of context. For clarity the use of mobile interaction in this chapter denotes the use of mobile devices while the user is standing and has full freedom of movement to rotate around a static location.

The design of mobile interactions is challenging due to a marked increase in variability for almost all aspects of the interaction. During interaction the communication bandwidth (the amount of information conveyed through the interaction) can be influenced by multiple factors, including: the amount of control the user can exert over the system, the availability of cognitive resources and the suitability/utility of the system's feedback. Some of the limits placed on this interaction bandwidth are fixed and are as a result of the limits of human attention, particularly on concurrent presentation of feedback [35], or are variable depending on the context of the user [40] and the user's walking gait [16] [37]. Due to the variable amount of attention the user can afford to give the device and the error-prone nature of mobile interactions [60], development of low attention interfaces is important.

Simple interactions such as querying the device "which direction should I travel in?" may only require a small amount of feedback to provide the answer. Thus, the feedback modalities that can be used to provide the answer can have low bandwidth capabilities, such as non-speech audio and vibrotacile. Such non-visual feedback has the benefit of not distracting the user's visual attention from the physical environment.

The modalities visual, audio and vibrotactile have been successfully used to assist in mobile navigation.The use of vibrotactile feedback gained popularity as a means to discretely present notifications to the user. This is of particular importance in social contexts that forbid the use of auditory cues. While the amount of information that can be conveyed by such feedback can be limited by available actuators, simple cues can form the basis of mobile interactions. The vibrotactile modality is an attractive approach for presenting feedback for when the device is in the user's hand or pocket. While the level of detail that can be detected by the user will be variable depending on the location of the device on the body, vibrotactile feedback remains a useful attention grabbing cue and, with careful design choices, can lead to engaging interactions.

3.1 Heading Acquisition System

Current location-based map applications rely heavily on the overhead view of a map, and may re-orientate the map based on compass heading. However, the users still has to translate current knowledge of the surroundings to be able to make sense of the map. Being able to select the direction of interest is important to make the first step in integrating map information of items of interest.

3.1.1 Scanning movements

Previous systems have used vibrotactile feedback in conjunction with the user making scanning movements with the device in their hand [44][59]. When the device is pointing in the target heading a vibrotactile pulse is produced to alert the user that the device is pointing in the direction of the target/ waypoint. The design of the vibrotactile pulse, alongside the technical constraints of such systems creates a challenging environment for the implementation of such systems. As identified by Robinson *et al.* [44], pointing -based interaction have been explored by Frohlich *et al.* [23] and Rukzio *et al.* [45] and bearing based selection by Strachan et al [50]. Such scanning behaviour can be used to link the heading in which the device is pointing at to a target heading. Selection of the range of angles to accept as being on the target is important and is known as target width. Williamson *et al.*[58] used an agent navigation simulation to assess the effect of target heading width on time to arrive at a defined destination. Williamson's *et al.* analysis showed that a target width of +/- 30 degrees was good enough for navigation purposes, this is supported by experiment data produced by Magnusson *et al.* [33]. Crucially reducing the target width did not lead to greater navigation performance. The selection of target width is an important parameter for the design of such systems. If the target heading angle is too narrow users will find it harder to discover and track the target heading. In such circumstances users can spend an inordinate time trying to precisely track changes in the target heading.

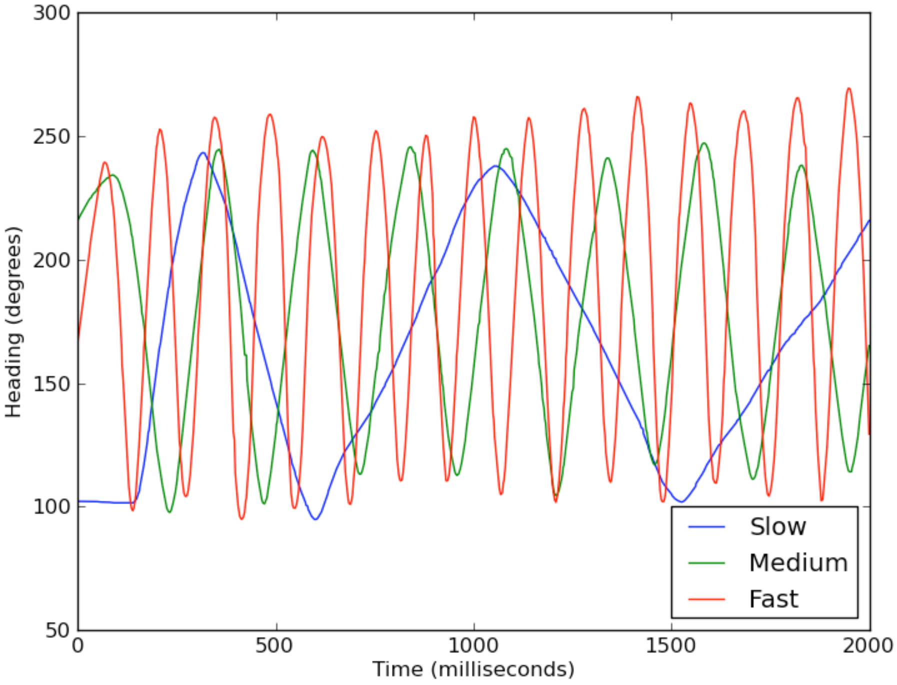


Figure 3.1: Scanning Movements at three movement speeds.

Slow	1.25 hz
Medium	3.75 hz
Fast	4.75 hz

Table 3.1: Approximate frequency of scanning movements, derived from pilot data 3.1

A SHAKE SK6 [58] sensor pack was used to provide 3D accelerometer, magnetometer and gyroscope sensor data. The SHAKE can also provided vibrotactile feedback via a pager motor and a user accessible switch coupled together with a Bluetooth module. This allowed the direction the SHAKE is pointed in to be sent to a host system and vibration profile playback commands to be received from the host system.

One of the key features of the scanning movements is their ability to be performed quickly. To estimate the speeds involved a test user holding the SHAKE (as shown in 3.1) and performed the scanning movement at three distinct speeds (Slow, Medium and Fast). Table 3.1 shows the approximate frequency of movement to move through approximilty 140 degrees.

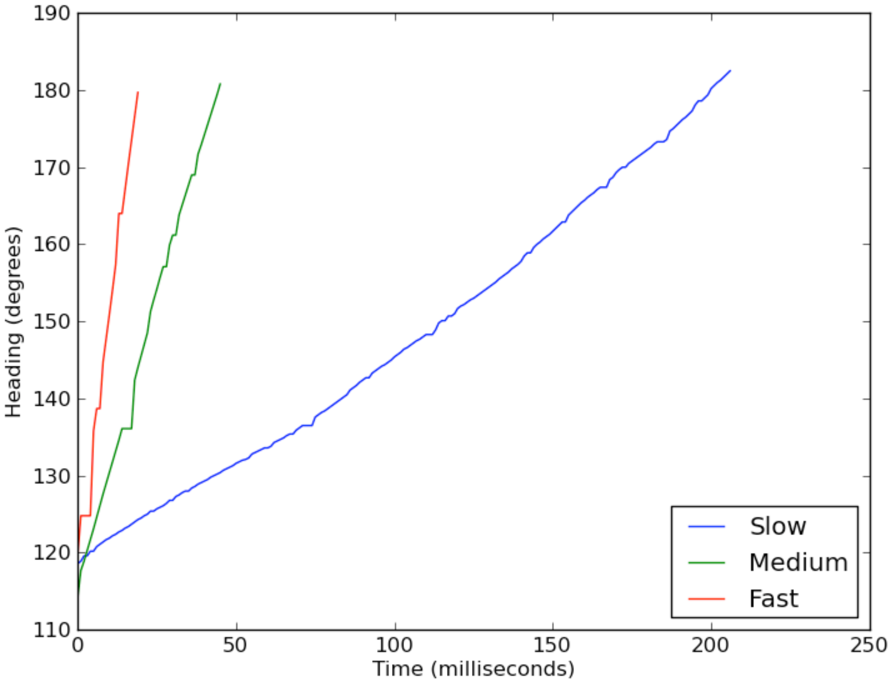
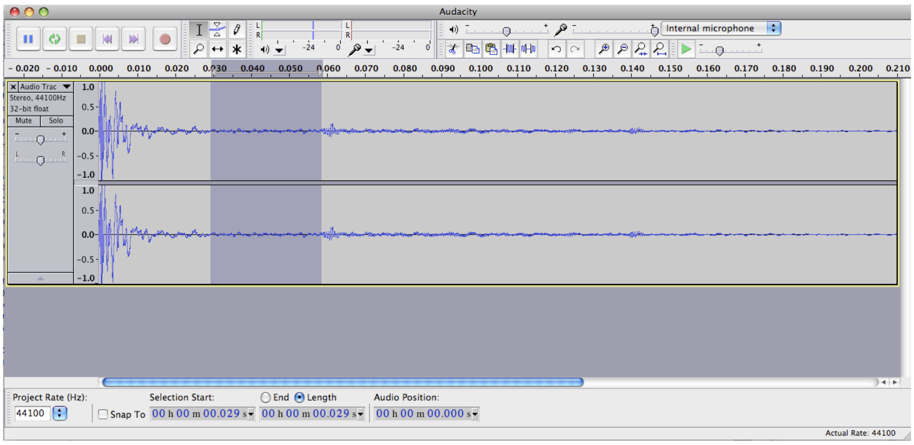


Figure 3.2: Heading vs Amount of time to pass through target heading range (60 degrees)

Figure 3.2 shows the amount of time each movement speed would take to go through the target heading range (60 degrees). Slow scanning movements could take up to 200 milliseconds, while medium and fast movements take less that 50 milliseconds. While it is possible to use sample rates high enough to capture even high speed movements the overall latency of moving through a target heading range and the presentation of a vibrotactile stimulus will impact on the usability of feedback provided. To give an indication of the latency between sending a vibration profile playback command and the vibration occurring a low cost methodology was used. The trackpad on post 2009 Macbook Pros has an audible click for button down and button up events. The internal mic can be used to record the button down event and with the SHAKE being held on the trackpad, record the low frequency vibrations (around 200 hz) from the SHAKE SK6 pager motor, as shown in Figure 3.3.

Figure 3.3: Low cost latency measuring setup. Allows for an estimation of the overall delay in triggering vibrotactile feedback.

Using a popular audio recording software package (Audacity) it is easy to identify the time difference between the click of the track pad and the resulting vibration of the SK6. Figure shows the audio recording from a button down event (0.0 seconds), the pager motor being activated (0.06 seconds) and the button up event (0.140 seconds). The selected part of the audio recording is the approximate delay, 29 milliseconds, before the vibration is played. This method was performed 10 times to ascertain the variability of the delay (Mean 30.52, Std 1.2) .While there will be a delay in the audio recording, relative measurement of events is still possible to provide an estimation of the system's delay.



Project Rate (Hz): 44100

☐ Snap To

Selection Start: 00 h 00 m 00.029 s

○ End ● Length: 00 h 00 m 00.029 s

Audio Position: 00 h 00 m 00.000 s

Actual Rate: 44100

Figure 3.4: Visual representation of the recording of mousepad clicking and subsequent vibration playing on a SHAKE SK6.

Bluetooth	60 ms
Sample Delay	15 ms
Processing Delay	15 ms
Total	90 ms

Table 3.2: Worst-case latency calculations for system operating at 60 hz. All units in milliseconds.

In the worst case scenario an upper bound on latency for the full system is equivalent to the two way Bluetooth delay + sample rate + processing delay. The Bluetooth delay is the amount of time it takes sensor data to be sent from the

SHAKE to the host system and the system to issue a play vibration command and then the vibration to be played. Given the approximate delays in Table 3.2 the total system latency is 90 ms. The latency encountered by the system places a further restriction on the production of timely feedback. In the case of slow and deliberate scanning movements the latency of system could make a +/-20 degrees error between the presentation of a vibrotactile stimulus and the current heading of the device. The problem is more acute for faster scanning movements, where the device could be moved through the target width without feedback being produced.

3.1.2 Feedback Design

The design of vibrotactile feedback with pager motors is difficult. The frequency of the vibrations is fixed, while designers can control the duration and intensity of vibration. The SHAKE SK6 has the ability to trigger predefined vibration profiles. Such that the vibration motor can play a sequence of vibrations with different durations and intensity (one such profile could be full power for 20 milliseconds then half power for 30 milliseconds). The ability increases the range of different vibration cues that we can produce. Care must be taken with the design of vibrotactile feedback as users can become desensitised by vibration when presented with long periods of continuous vibration [26]. The intensity of vibration may also cause users to have a negative reaction, becoming startled and wanting to drop the vibrating device.

Inspiration for the design of the vibration feedback came from purring cats. Many people enjoy the sensation of stroking a cat and the low pitch purring the cat produces. This seemed to be a good vibration to emulate for on-target heading feedback. To accomplish this, a vibration profile with an exponentially decaying power, with fixed pulse widths (Figure 3.5) was used. This design has the benefit of 'blurring' the percept of the target heading. Thus boundary conditions where the the device is slipping in and out of the target angle, do not become abrupt and startling.

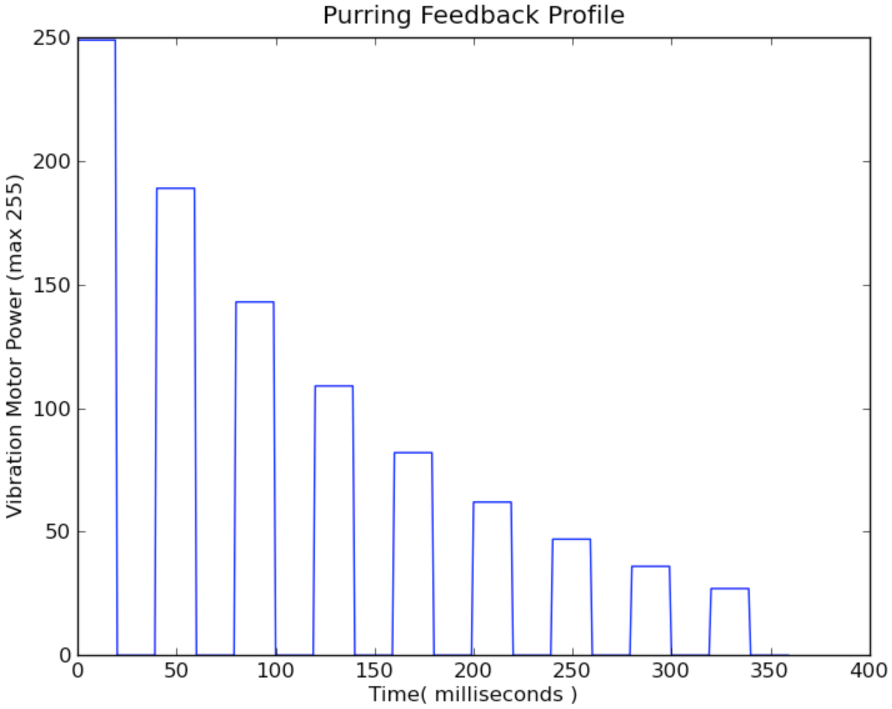


Figure 3.5: SK6 vibration motor profile, decaying power profile with fixed pulse widths of 20 ms

3.1.3 Interaction Design

Three different heading acquisition systems were implemented to study the effects of: 1. providing feedback for when the device is being moved too quickly (Limiting) and 2. providing feedback of where the target heading range is located in relation to the current heading (Flick) and 3. a traditional approach acting as a control

(Basic). For all conditions the vibrotactile profile previously discussed was used to indicate to the user that the device was pointing towards the target.

Limit

The movements of the user's hand while scanning can have a drastic effect on the accuracy of the given vibrotactile feedback. Even with a high sample rate (60hz) there are still issues with the SHAKE being moved faster than we can give feedback. The issue is further compounded by the length of time it takes to present a vibrotactile pattern. This length of time can be upwards of 90 milliseconds. This issue is prevalent on boundary conditions, for example when the user moves the SHAKE through the target heading at high speed. To reduce the speed of hand movements a sharp vibrotactile pulse (half power for 20 ms) when the user is moving too fast. The idea behind this is to alert the user that the scanning movement is too fast for reliable detection of target heading. We expect this improvement will lead to an improvement in both time to select the target and overall accuracy. This approach, while providing more information to the user about the limitations of the system, overloads the vibration channel as vibrations now convey different cues. This is mitigated by the cue for moving too fast feeling significantly different to the on-target cue.

Flick

In the first two conditions, there is no directional information provided by the feedback given. The user has no indication wither the target is closer to user's right or left hand side and so has to guess what direction to start to search. By choosing the wrong direction this can lead to the user turing 329 degrees before reaching the target heading presented (-31 degrees from their starting point). To assist users in this issue a method to provide direction information has been developed. When a small flick of the wrist is performed to either the right or left of the user the width of the active scanning by 120 degrees is extended in the direction of the flick. Thus providing a preview of the feedback given if the subject started to scan in the direction of the flick. The flick gesture can be performed quickly and gives the user an indication of the direction to start scanning.

Basic

This system will form the baseline that the proposed improvements will be tested against. On target feedback is given when the SHAKE is in the direction of the target angle +/- 30 degrees.

3.2 Experiment

3.2.1 Hypothesis

The experiment hypothesis are as follows:

- Users will be able to discover the target heading (H1)
- Indicating to the user that their movement speed is too high for reliable feedback to be conveyed will reduce errors in selection of target heading (H2)

- Flick condition will result in faster selection times (H3).

3.2.2 Apparatus

The system used to compare the three feedback conditions comprised of a SHAKE SK6 sensor pack held in the dominant hand of the subject. This was used to stream 3D accelerometer, magnetometer and gyroscope data at 60hz and was also used to provide vibrotactile feedback, via the built-in pager motor. The host system (dual core laptop running Windows XP) processed the inertia sensor data into tilt compensated heading which was then analysed and if found within the target heading, the purring vibrotactile profile was triggered. In the case a user moves outside the target heading from being within, a null vibration request is sent to the SHAKE SK6 is stop the vibration.

3.2.3 Procedure

12 participants, (8 male, 4 female) with average age of 24.6 (std 5.6) were recruited from University of Glasgow students and performed all three conditions in a counterbalanced order. None of the participants had experienced the system before the experiment. Six target headings (0, 60,120,180,240,300 degrees) were presented 10 times in a randomised order and the user selected the heading by activating the switch on the right hand side of the SHAKE SK6.

Participants were asked to select one target at a time, initially starting facing north and continuing from the previously selected heading. Once a selection had been made the next target angle could then be discovered. All selections were made while standing up holding the device in the user's dominant hand.

3.2.4 Results

Ratio of Correct Selections within Target Heading

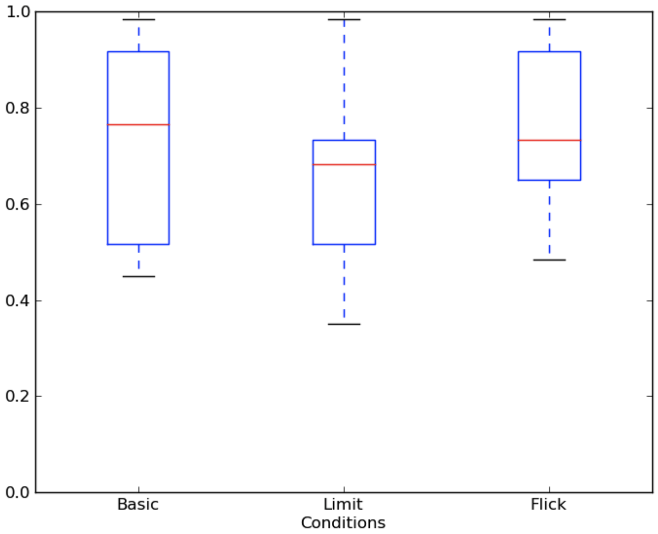
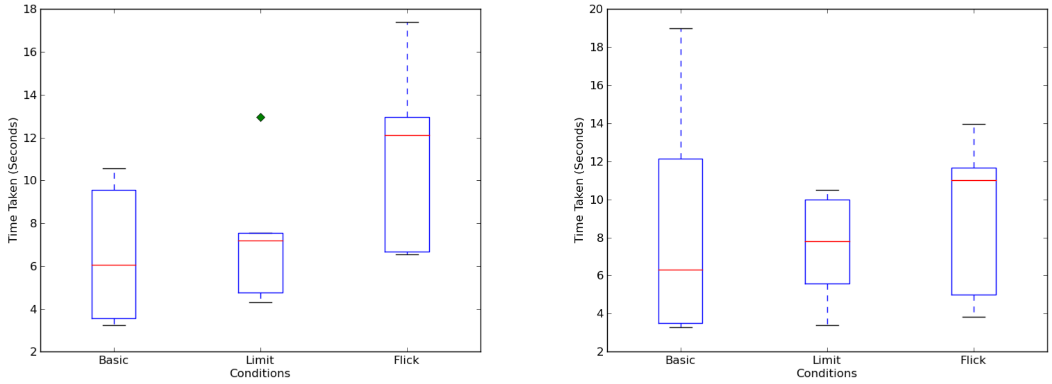


Figure 3.6: Normalised ratio (0.0 to 1.0) of heading selections within target heading range (diamonds represent outliers)

Overall, the results are positive with participants being able to select the correct target heading 71 % of the trials across all conditions. However, as shown in Figure 3.6 there is little difference of medians between the conditions. A one way ANOVA test confirms the lack of a statistical difference between the conditions ($F = 0.353$, $p > 0.7$). Thus Hypothesis H1 is accepted, however H2 is rejected as there is no evidence that error rates decrease with the limit condition.



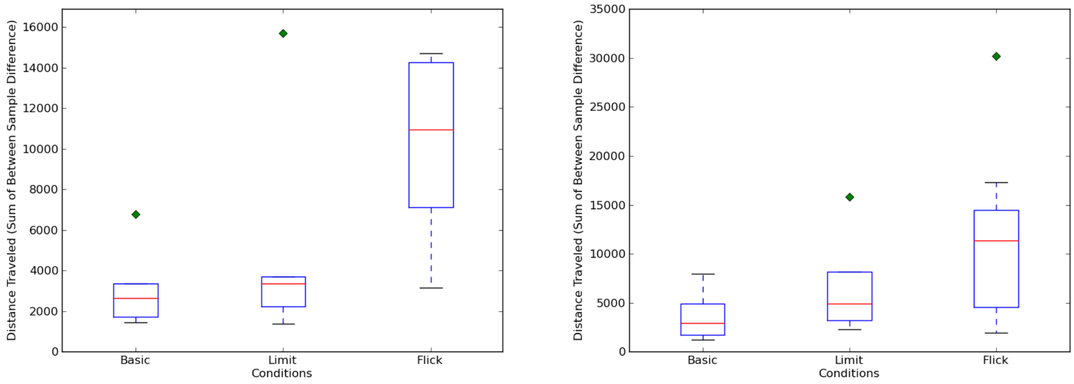
(a) Correct Selection

(b) Incorrect Selection

Figure 3.7: Average time taken for selections (diamonds represent outliers)

Time Taken

The average time to select a target heading is largely unaffected by the heading selection being correct or incorrect (showing in Figure 3.7).Correct and incorrect selections are separated in the following analysis to ensure that an bias in the data between correct and incorrect selections were exposed. In the basic feedback condition the mean of selection times was 6.45 seconds (std 2.55) for correct selections as compared to 7.82 seconds (std 4.91) for incorrect. On average the flick condition took longer for both correct selections at 10.52 seconds (std 3.56) and incorrect selections at 8.87 seconds (std 3.54). The flick condition did not perform well with subjects taking on average, 4.07 seconds longer to complete a correct selection. While the time taken for incorrect selections was not found to be significant ($F = 0.49$, $p > 0.6$), while correct selections was ($F = 5.239$, $p < 0.015$).



(a) Correct Selection

(b) Incorrect Selection

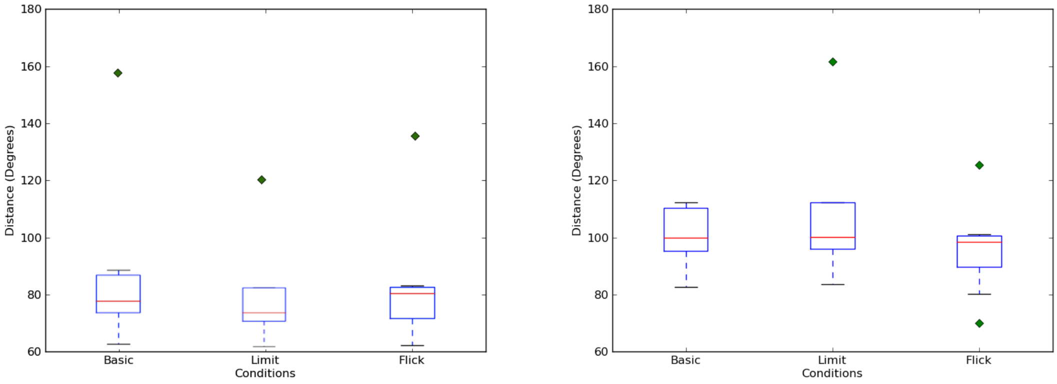
Figure 3.8: Total distance travelled

Distance Travelled

A sum of between sample differences can be used as a rough comparison measure of the amount of movement / distance travelled for selections, as shown in Figure 3.8. The flick condition resulted in a significantly higher distance travelled ($F=9.634$, $p<0.0017$). Thus indicating that the users were indeed using the the flick gesture to gain previews of direction to travel towards the target heading. A significance was found for incorrect selections ($F = 4.6.04$, $p < 0.02$).

Optimal Distance

Due to the pseudo random selection of target headings the minimum distance between the start heading and the the target heading is variable. The results are biased due to this, such as correct selections are more likely if the minimum distance is smaller than it is for incorrect selections as shown in Figure 3.9. There is no significant difference between conditions for either correct ($F=2.462$, $p > 0.1$) or



(a) Correct Selection

(b) Incorrect Selction

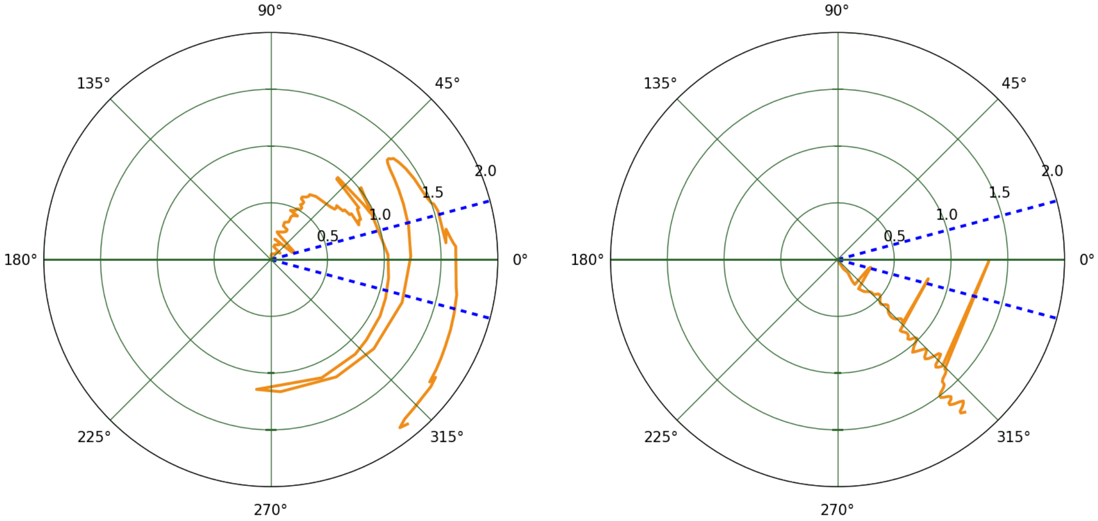
Figure 3.9: Optimal distance to travel

incorrect selections (F$=$ 0.984 , p $>$ 0.3).

3.2.5 Discussion

Overall the results are not as expected. Neither the speed reduction nor the flick-to-preview strategies had the intended effects on accuracy of selection or the time taken to make a selection. However hypothesis H1 can be accepted at participants managed to select the correct heading, on average 71% of all selections. It could be suggested that the accuracy of the the system is more dependent on the overall latency in producing feedback. Unfortunately, providing additional feedback for when users are moving too fast did not improve accuracy. These two unexpected results suggest that the trigger for presenting the speed limiting feedback was set too high and that participants moved at a speed just below the trigger threshold.

The flick condition was designed to give a preview of the next 160 degrees in the direction the flick gesture was performed. It was hoped that this additional information would allow users to pick the shortest route to the target heading. Some subjects commented after the experiment that they stopped using the flick gesture during the trial as they did not feel it helped them. Results show that the additional information provided by the flick gesture did not improve selection time, nor error rates. Thus hypothesis H3 is rejected.



(a) Fast Movements

(b) Confusing Feedback

Figure 3.10: Sample heading trajectories from experiment

Some trajectories of participants movements towards the target heading are illustrated in Figures 3.10 and 3.11. Figures 3.10a shows the result of a user moving too fast through the target region. The effects of the latency on the feedback can be clearly shown as quick sweeps are quickly corrected when the feedback is eventually produced, as shown in Figure 3.11. Figure 3.10b shows the user being on the edge of receiving feedback and due to sharp movements receives the on target feedback while staying outside the target range.

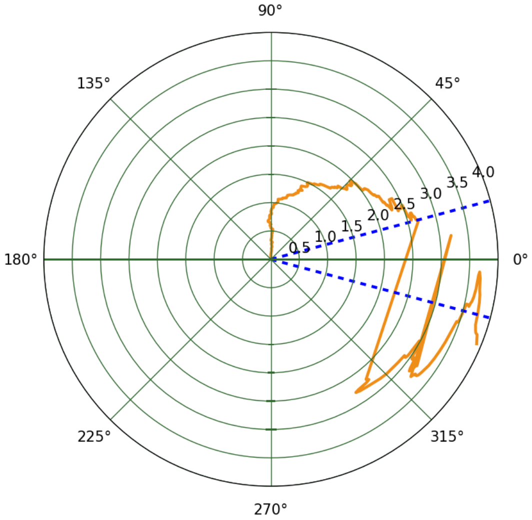


Figure 3.11: Sample heading trajectories, recorded

The expected improvements to accuracy of selection for the speed limiting condition and the time to select for the flick-to-preview were not found in the analysis of the experimental data. It is comforting to note that participants were able to correctly identify the target heading 75 % of the time. The time to select a target heading is also encouraging at only 7.14 seconds (std 3.98). It is hoped that reductions in both the noise in the heading data and the latency within the system would yield higher accuracy rates and a reduction in time taken to select targets.

3.2.6 Further Work

With the introduction of the SHAKE SK7 sensor pack it is possible to create vibrotactile cues that are triggered on both heading, pitch and roll of the device without the need for a host system. This eliminates the latency of the previously used sensor pack, by not having heading data being streamed to the host and the host evoking vibration profile play requests. By using the SHAKE SK7 it is possible to re-examine the potential benefits of the feedback improvements proposed with the compounding effect of high latency.

The flick-to-preview feedback has potential that was not properly explored. One of the main difficulties of presenting the direction in which the user should move the device to meet the target heading is the lack of direction in the feedback provided. In the previous experiment users were required to flick the device left or right to find out if they should move in that direction. Clearly this is a rather laborious process in which the user may resolve to simply move the device until feedback is presented rather than waste effort in finding in which direction they should be moving in. Vibrotactile apparent motion could be used to give the feeling of a vibration moving left to right (or right to left). This would give a cue to the user in which direction they should move the device without having to state explicitly in which direction they are searching in. It would be interesting to explore if users could tell the difference between left to right and right to left cues. If a high percentage of users can achieve high recognition rates, a similar study as detailed previously could be used to examine the validity of such a feedback design for target heading acquisition.

Chapter 4

Future Mobile Navigation Aids

The previous chapter identified potential interactions for heading acquisition supported by gestures and vibrotactile feedback. As detailed in 2.3.2 we can encode both direction and distance information in the vibrotactile navigation cues presented. This was made possible by the use of multiple actuators. The placement of the actuator enabled corresponded to the direction of the target heading. The distance to the destination/waypoint can be denoted by the frequency of the vibrotactile pulses. While this feedback design has been shown to work, the use of multiple actuators placed around the body of the user is not feasible for widespread use. Creating such feedback on mobile devices in a non visual fashion is still to be explored. Distance and direction cues could be presented concurrently. However, feedback designers may have issues with presenting concurrent information or the created feedback mechanisms may take longer to complete. While these challenges could perhaps be overcome with further research there is a fundamental flaw, in that the user may/may not be interested on both components of the feedback at all times.

The amount of cognitive resources the mobile user has at a given moment is dependent on their surroundings and other tasks they are trying to accomplish. Rather than trying to solve the problem with one interaction approach or combined feedback technique, it could be suggested that multiple one-use approaches would be a better approach to design mobile interactions. Allowing the user to choose the interaction approach could result in selections being based on their suitability as measured by the available cognitive resources. This would also result in reducing the amount of information that needs to be encoded within the non-visual feedback. There is little point in having non-visual interfaces, that while providing all the benefits of non-visual interfaces, vastly increase the cognitive demand placed on the user. Designing interactions that allow the user to vary the amount of information/cognitive demand required for control is fundamental to creating mobile interfaces that are robust to usage in mobile scenarios.

While the design of non-visual interfaces has gained much attention, there is little research on the merits of designing interactions that have multiple input possibilities. Should our approach be to invest in complex non-visual interactions or have multiple simple non-visual interactions that the user can decide the utility of use for themselves? While the idea of creating simpler and imperfect interactions is counter intuitive, it puts the emphasis of control back on to the user. The ability to restrict the 'richness' of an interaction results in control over the amount of attention the user has to give to the device. Further work would need to be undertaken to examine the benefits of such an approach, as the cognitive load is not removed but rather transformed so that the user has to decide what interaction inout methods to use for a given task.

Having multiple interaction possibilities creates problems for mobile system developers. There would be a significant need for devices to change modes easily and with minimum effort, without such, the benefits of this alternative approach could be negated. One such way of the user defining the application context, or changing modes, is by analysising the way the device is being held. Taylor et al [55] used capacitive sensors placed around a mobile device to create a map of hand placement around the device. Inference techniques were then applied to give the system a belief of what context/application the user wanted to use the device in.

In Section 3.2.6 it was proposed that the by rolling the device to the side, further information could be presented to the user of which direction they should move in to reach the target heading in the shortest time possible. In a similar fashion the way the device is being held, or particular movements around the casing of the device, could be used for the user to query the device for distance information. A potential approach for providing distance information is by using a physical groove situated on the side, or back of device. The physical form affords movement of the finger along it and helps guide the movement in a manner expected by the system. The distance travelled by the finger can be scaled to a distance in meters. As the finger approaches the correct point on the scale, a short vibrotactile pulse will alert the user that they have reached the point where the destination exists.

The distance query interaction is not best suited for absolute distance and accurate measurements. However, it does give the user useful information for making comparisons between different destinations. When users wish to know the distance to a high accuracy, this interaction approach does not inhibit the use of a visual display to present such information. By designing mobile devices with a selection of interaction techniques this will allow the user to be in control of the interaction with the device and can pick the most appropriate one and change when the current method becomes either unusable or less optimal than other choices. Such segregation of distance and direction cues may allow the feedback presented to the user to be simple and therefore effective in mobile scenarios.

4.1 Conclusion

This thesis has explored a new interaction possibility for the presentation and control of vibrotactile navigation cues. Vibrotactile feedback has been found to be adequate for directional information to be conveyed to the user. Interesting vibrotactile feedback patterns have been explored with patterns mimicking that of a purring cat. The utility of such feedback design and interaction techniques could be improved upon with the availability of low latency feedback. This would form a significant part of future work in this area of research.

Chapter 5

Bibliography

Bibliography

[1] Azuma R.; Baillot Y.; Behringer R.; Feiner S.; Julier S.; MacIntyre B. Recent advances in augmented reality. Computer Graphics and Applications, IEEE, 21(6):34 – 47, Jan 2001.

[2] Georg Von Békésy. Sensory Inhibition. Princeton University Press, Jan 1967.

[3] Mark Billinghurst and Hirokazu Kato. Collaborative augmented reality. Commun. ACM, 45(7):64–70, July 2002.

[4] Oliver Bimber and Ramesh Raskar. Modern approaches to augmented reality. In ACM SIGGRAPH 2006 Courses, SIGGRAPH '06, New York, NY, USA, 2006. ACM.

[5] Frank Biocca and Ben Delaney. Communication in the age of virtual reality. chapter Immersive virtual reality technology, pages 57–124. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1995.

[6] Mike Blackwell, Constantinos Nikou, Anthony M DiGioia, and Takeo Kanade. An image overlay system for medical data visualization. Medical Image Analysis, 4(1):67 – 72, 2000.

[7] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. Multimodal 'eyes-free' interaction techniques for wearable devices. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '03, pages 473–480, New York, NY, USA, 2003. ACM.

[8] Lorna M. Brown, Stephen A. Brewster, and Helen C. Purchase. Multidimensional tactons for non-visual information presentation in mobile devices. In Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, MobileHCI '06, pages 231–238, New York, NY, USA, 2006. ACM.

[9] Buchmann, Volkert and Violich, Stephen and Billinghurst, Mark and Cockburn, Andy. Fingartips: gesture based direct manipulation in augmented reality. Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia, Jan 2004.

[10] G. E. BURNETT and J. M. 2002 PORTER. An empirical comparison of the use of distance versus landmark information within the human-machine interface for vehicle navigation systems. Human Factors in Transportation, Communication, Health, and the Workplace., pages (pp. 49 – 64), Jan 2002.

[11] Adrian David Cheok, Kok Hwee Goh, Wei Liu, Farzam Farbiz, Siew Wan Fong, Sze Lee Teo, Yu Li, and Xubo Yang. Human pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. Personal Ubiquitous Comput., 8(2):71–81, May 2004.

[12] Roger Cholewiak and Amy Collins. The generation of vibrotactile patterns on a linear array: Influences of body site, time, and presentation mode. Attention, Perception and Psychophysics, 62:1220–1235, 2000.

[13] Justin Cohen, Masataka Niwa, Robert W. Lindeman, Haruo Noma, Yasuyuki Yanagida, and Kenichi Hosaka. A closed-loop tactor frequency control system for vibrotactile feedback. In CHI '05 extended abstracts on Human factors in computing systems, CHI EA '05, pages 1296–1299, New York, NY, USA, 2005. ACM.

[14] C. Conn, J. Lanier, M. Minsky, S. Fisher, and A. Druin. Virtual environments and interactivity: windows to the future. SIGGRAPH Comput. Graph., 23(5):7–18, July 1989.

[15] Brookshire D. Conner, Scott S. Snibbe, Kenneth P. Herndon, Daniel C. Robbins, Robert C. Zeleznik, and Andries van Dam. Three-dimensional widgets. In Proceedings of the 1992 symposium on Interactive 3D graphics, I3D '92, pages 183–188, New York, NY, USA, 1992. ACM.

[16] Andrew Crossan, Roderick Murray-Smith, Stephen Brewster, James Kelly, and Bojan Musizza. Gait phase effects in mobile interaction. In CHI '05 extended abstracts on Human factors in computing systems, CHI EA '05, pages 1312–1315, New York, NY, USA, 2005. ACM.

[17] Jan B F Van Erp. Guidelines for the use of vibro-tactile displays in human computer interaction. Proceedings of Eurohaptics, 2002(2):18–22.

[18] Jan B. F. Van Erp, Hendrik A. H. C. Van Veen, Chris Jansen, and Trevor Dobbins. Waypoint navigation with a vibrotactile waist belt. ACM Trans. Appl. Percept., 2(2):106–117, April 2005.

[19] Marie Falahee, Kezzy Latham, and Erik Geelhoed. Safety and comfort of eyeglass displays. In Peter Thomas and Hans-W. Gellersen, editors, Handheld and Ubiquitous Computing, volume 1927 of Lecture Notes in Computer Science, pages 236–247. Springer Berlin / Heidelberg, 2000.

[20] Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Anthony Webster. A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. Personal Technologies, 1:208–217, 1997. 10.1007/BF01682023.

[21] Steven Feiner, Blair Macintyre, and Dorée Seligmann. Knowledge-based augmented reality. Commun. ACM, 36(7):53–62, July 1993.

[22] George W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. Commun. ACM, 36(7):39–49, July 1993.

[23] Peter Fröhlich, Rainer Simon, Lynne Baillie, and Hermann Anegg. Comparing conceptual designs for mobile access to geo-spatial information. In Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, MobileHCI '06, pages 109–112, New York, NY, USA, 2006. ACM.

[24] Alberto Gallace, Hong Z. Tan, and Charles Spence. The body surface as a communication system: The state of the art after 50 years. Presence: Teleoper. Virtual Environ., 16(6):655–676, December 2007.

[25] Cristy Ho, Hong Z. Tan, and Charles Spence. Using spatial vibrotactile cues to direct visual attention in driving scenes. Transportation Research Part F: Traffic Psychology and Behaviour, 8(6):397 – 412, 2005.

[26] Eve Hoggan. Crossmodal Audio and Tactile Interaction with Mobile Touchscreens. , University of Glasgow, UK (2010). PhD thesis, University of Glasgow, 2010.

[27] C.M. Johnson, L.A.; Higgins. A navigation aid for the blind using tactile-visual sensory substitution. Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE, pages 6289 – 6292, 2006.

[28] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top ar environment. In Augmented Reality, 2000. (ISAR 2000). Proceedings. IEEE and ACM International Symposium on, pages 111 –119, 2000.

[29] Kristoffersen, Steinar and Ljungberg, Fredrik. "making place" to make it work: empirical explorations of hci for mobile cscw. Proceedings of the international ACM SIGGROUP conference on Supporting group work, Jan 1999.

[30] Robert W. Lindeman and Haruo Noma. A classification scheme for multi-sensory augmented reality. In Proceedings of the 2007 ACM symposium on

Virtual reality software and technology, VRST '07, pages 175–178, New York, NY, USA, 2007. ACM.

[31] J Loomis. Tactile pattern perception. Perception, 10(1):5–27, Jan 1981.

[32] Joanna Lumsden and S. Brewster. A paradigm shift: alternative interaction techniques for use with mobile and wearable devices. The 13th Annual IBM Centers for Advanced Studies Conference (CASCON'2003), Jan 2010.

[33] Charlotte Magnusson, Kirsten Rassmus-Gröhn, and Delphine Szymczak. Scanning angles for directional pointing. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, MobileHCI '10, pages 399–400, New York, NY, USA, 2010. ACM.

[34] D. K. McGookin. Providing a structured method for integrating non-speech audio into human-computer interfaces. PhD thesis, University of York, UK, 1994.

[35] D. K. McGookin. Understanding and Improving the Identification of Concurrently Presented Earcons. PhD thesis, University of Glasgow, 2004.

[36] P Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. IEICE Transactions on Information Systems, 5(12):1–5, Jan 1994.

[37] Roderick Murray-Smith, Andrew Ramsay, Simon Garrod, Melissa Jackson, and Bojan Musizza. Gait alignment in mobile phone conversations. In Proceedings of the 9th international conference on Human computer interaction with mobile

devices and services, MobileHCI '07, pages 214–221, New York, NY, USA, 2007. ACM.

[38] Elizabeth D. Mynatt, Maribeth Back, Roy Want, Michael Baer, and Jason B. Ellis. Designing audio aura. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '98, pages 566–573, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[39] T. Ohshima, K. Satoh, H. Yamamoto, and H. Tamura. Ar2hockey: a case study of collaborative augmented reality. In Virtual Reality Annual International Symposium, 1998. Proceedings., IEEE 1998, pages 268 –275, 18-18 1998.

[40] Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '05, pages 919–928, New York, NY, USA, 2005. ACM.

[41] Wayne Piekarski and Bruce Thomas. Arquake: the outdoor augmented reality gaming system. Commun. ACM, 45(1):36–38, January 2002.

[42] Ivan Poupyrev, Shigeaki Maruyama, and Jun Rekimoto. Ambient touch: designing tactile interfaces for handheld devices. In Proceedings of the 15th annual ACM symposium on User interface software and technology, UIST '02, pages 51–60, New York, NY, USA, 2002. ACM.

[43] Philippe Renevier and Laurence Nigay. Mobile collaborative augmented reality: The augmented stroll. In Proceedings of the 8th IFIP International Conference

on Engineering for Human-Computer Interaction, EHCI '01, pages 299–316, London, UK, UK, 2001. Springer-Verlag.

[44] Simon Robinson, Parisa Eslambolchilar, and Matt Jones. Sweep-shake: finding digital resources in physical environments. In Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '09, pages 12:1–12:10, New York, NY, USA, 2009. ACM.

[45] Enrico Rukzio, Karin Leichtenstern, Vic Callaghan, Paul Holleis, Albrecht Schmidt, and Jeannette Chin. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. In Paul Dourish and Adrian Friday, editors, UbiComp 2006: Ubiquitous Computing, volume 4206 of Lecture Notes in Computer Science, pages 87–104. Springer Berlin / Heidelberg, 2006.

[46] A.H. Rupert. An instrumentation solution for reducing spatial disorientation mishaps. Engineering in Medicine and Biology Magazine, IEEE, 19(2):71 –80, mar/apr 2000.

[47] J. Kenneth Salisbury and Mandayam A. Srinivasan. Phantom-based haptic interaction with virtual objects. IEEE Comput. Graph. Appl., 17(5):6–10, September 1997.

[48] Nitin Sawhney and Chris Schmandt. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. ACM Trans. Comput.-Hum. Interact., 7(3):353–383, September 2000.

[49] Mandayam A. Srinivasan and Cagatay Basdogan. Haptics in virtual environments: Taxonomy, research status, and challenges. Computers and Graphics, 21(4):393 – 404, 1997. Haptic Displays in Virtual Environments and Computer Graphics in Korea.

[50] Steven Strachan and Roderick Murray-Smith. Bearing-based selection in mobile spatial interaction. Personal Ubiquitous Computing, 13(4):265–280, May 2009.

[51] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. “studierstube”: An environment for collaboration in augmented reality. Virtual Reality, 3:37–48, 1998.

[52] Hollerer T., Feiner S., Terauchi T., Rashid G., and Hallaway D. Exploring mars: developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers and Graphics, 23(6):779–785, 1999.

[53] Hong Z. Tan and Alex Pentland. Tactual displays for wearable computing. In Proceedings of the 1st IEEE International Symposium on Wearable Computers, ISWC ’97, pages 84–90, Washington, DC, USA, 1997. IEEE Computer Society.

[54] Hui Tang and D.J. Beebe. An oral tactile interface for blind navigation. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 14(1):116 –123, march 2006.

[55] Brandon T. Taylor and V Michael Bove. The bar of soap: a grasp recognition system implemented in a multi-functional handheld device. In CHI ’08 extended

abstracts on Human factors in computing systems, CHI EA '08, pages 3459–3464, New York, NY, USA, 2008. ACM.

[56] techsplurge (viewed 31/05/12). http://techsplurge.com/3214/mega-list-33-awesome-augmented-reality-apps-games-android/.

[57] Layar: (viewed 31/05/12). www.layar.com.

[58] John Williamson, Simon Robinson, Craig Stewart, Roderick Murray-Smith, Matt Jones, and Stephen Brewster. Social gravity: a virtual elastic tether for casual, privacy-preserving pedestrian rendezvous. CHI '10: Proceedings of the 28th international conference on Human factors in computing systems, Apr 2010.

[59] Matthias M. Wloka and Eliot Greenfield. The virtual tricorder: a uniform interface for virtual reality. In Proceedings of the 8th annual ACM symposium on User interface and software technology, UIST '95, pages 39–40, New York, NY, USA, 1995. ACM.

[60] Koji Yatani and Khai N. Truong. An evaluation of stylus-based text entry methods on handheld devices in stationary and mobile settings. In Proceedings of the 9th international conference on Human computer interaction with mobile devices and services, MobileHCI '07, pages 487–494, New York, NY, USA, 2007. ACM.

Chapter 6

Appendices