# INTERACTION TECHNIQUES WITH NOVEL MULTIMODAL FEEDBACK FOR ADDRESSING GESTURE-SENSING SYSTEMS

EUAN FREEMAN

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
*Doctor of Philosophy*

SCHOOL OF COMPUTING SCIENCE

COLLEGE OF SCIENCE AND ENGINEERING
UNIVERSITY OF GLASGOW

FEBRUARY 2016

## Abstract

Users need to be able to *address* in-air gesture systems, which means finding where to perform gestures and how to direct them towards the intended system. This is necessary for input to be sensed correctly and without unintentionally affecting other systems. This thesis investigates novel interaction techniques which allow users to address gesture systems properly, helping them find where and how to gesture. It also investigates audio, tactile and interactive light displays for multimodal gesture feedback; these can be used by gesture systems with limited output capabilities (like mobile phones and small household controls), allowing the interaction techniques to be used by a variety of device types. It investigates tactile and interactive light displays in greater detail, as these are not as well understood as audio displays.

Experiments 1 and 2 explored tactile feedback for gesture systems, comparing an ultrasound haptic display to wearable tactile displays at different body locations and investigating feedback designs. These experiments found that tactile feedback improves the user experience of gesturing by reassuring users that their movements are being sensed. Experiment 3 investigated interactive light displays for gesture systems, finding this novel display type effective for giving feedback and presenting information. It also found that interactive light feedback is enhanced by audio and tactile feedback.

These feedback modalities were then used alongside audio feedback in two interaction techniques for addressing gesture systems: *sensor strength feedback* and *rhythmic gestures*. Sensor strength feedback is multimodal feedback that tells users how well they can be sensed, encouraging them to find where to gesture through active exploration. Experiment 4 found that they can do this with 51mm accuracy, with combinations of audio and interactive light feedback leading to the best performance. Rhythmic gestures are continuously repeated gesture movements which can be used to direct input. Experiment 5 investigated the usability of this technique, finding that users can match rhythmic gestures well and with ease.

Finally, these interaction techniques were combined, resulting in a new single interaction for addressing gesture systems. Using this interaction, users could direct their input with rhythmic gestures while using the sensor strength feedback to find a good location for addressing the system. Experiment 6 studied the effectiveness and usability of this technique, as well as the design space for combining the two types of feedback. It found that this interaction was successful, with users matching 99.9% of rhythmic gestures, with 80mm accuracy from target points. The findings show that gesture systems could successfully use this interaction technique to allow users to address them. Novel design recommendations for using rhythmic gestures and sensor strength feedback were created, informed by the experiment findings.

## Acknowledgements

The research in this thesis was partly funded by Nokia Technologies, Finland.

First, I would like to thank my supervisor, Stephen Brewster, for giving me the opportunity to do this PhD and for keeping me on the right track. His help, encouragement and seemingly endless knowledge made this all possible. I would also like to thank Vuokko Lantz, my supervisor at Nokia, for all of her helpful input and for the interesting discussions we had every couple of weeks. Finally, I would like to thank my second supervisor at Glasgow, Matthew Chalmers, who challenged me during my first and second year exams with a different perspective, which ultimately improved this research.

I first discovered my interest in HCI through working with Helen Purchase, who taught me a lot of good research habits and methods, and whose advice has stuck with me to this day. I would not be half as productive as I am today were it not for what I learned from her. Thank you, Helen!

I would like to thank my thesis committee — Kenton O'Hara, Mary Ellen Foster and Craig Macdonald — for the interesting discussions during the viva and for giving feedback which has improved this thesis. Thanks, also, for being the only people who will ever have read this work in full!

I would also like to thank everyone, past and present, in the Multimodal Interaction Group. You're all great, especially those in F131 who endure me. Special thanks to Dong-Bach Vo, whose baking and choice of restaurants has introduced me to some great food.

Finally, I would like to thank those closest to me: Allie, for convincing me to do this and for her endless patience, love and support over the past few years; my family[1] and friends, for keeping me sane and helping me through it all; and my bicycles, for being there for me.

[1] Especially my mum and dad who play a starring role in this thesis as the protagonists in Section 2.3.1.

## Declaration

The research presented in this thesis is entirely the author's own work. Research in this thesis has been published at the following venues, using only the parts of these papers that are directly attributable to the author:

- The research in Chapter 3 has been published at ICMI 2014: Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. *Tactile Feedback for Above-Device Gesture Interfaces: Adding Touch to Touchless Interactions*. Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14, ACM Press, 419–426.
- The research in Chapter 4 has been partly published at INTERACT 2015: Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2015. *Towards In-Air Gesture Control of Household Appliances with Limited Displays*. In Proceedings of INTERACT '15 in LNCS 9299. IFIP, 611–615.
- The research in Chapter 5 has been partly published at INTERACT 2015: Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2015. *Interactive Light Feedback: Illuminating Above-Device Gesture Interfaces*. In Proceedings of INTERACT '15 in LNCS 9299. IFIP, 478–481.
- The research in Chapter 4 and Chapter 5 has been partly published at the NordiCHI 2014 workshop on Interaction with Light: Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. *Illuminating Gesture Interfaces with Interactive Light Feedback*. Proceedings of the Beyond the Switch: Explicit and Implicit Interaction with Light workshop at NordiCHI '14.
- The research in Chapter 5, Chapter 6 and Chapter 7 will be published at CHI 2016: Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2016. *Do That, There: An Interaction Technique for Addressing In-Air Gesture Systems*. Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI '16, ACM Press, to appear.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Motivation

In-air gesture systems allow people to interact with technology using hand movements or hand poses performed in mid-air. These types of gesture have received considerable interest of late, as improvements in technology have made them a feasible input modality. Smaller and more powerful sensors, with lower energy and computation requirements, mean that gesture interfaces can be added to increasingly smaller devices, such as mobile phones or household objects. In-air gesture sensing approaches are also improving, with greater computational resources available to support sophisticated gesture-recognition algorithms. Together, these advances are driving interest in using gestures for human-computer interaction. This thesis focuses on in-air gestures, rather than other types of gesture, for example those performed while holding a device or those performed on a touchscreen; when the word "gesture" is used, it specifically means in-air gestures.

Many gesture systems have been studied in the human-computer interaction literature over the past few decades, including the seminal *Put-That-There* [16] system, which combined in-air gestures with speech input. It was another three decades before Xbox Kinect[1] appeared and became the first successful consumer in-air gesture system, allowing people to interact with their televisions and games consoles from across the living room, using just their hands in mid-air. Research has since investigated ways of introducing these capabilities to smaller devices, like mobile phones and household appliances. Consumer versions of these products have also launched with in-air gesture systems; for example, Samsung[2] and Microsoft[3] mobile phones and the Nest smoke alarm[4] allowed users to interact using gestures.

[1]Xbox Kinect: `www.xbox.com/en-GB/Xbox360/Accessories/kinect` Accessed 05/05/15
[2]Samsung Galaxy S4: `www.samsung.com/global/microsite/galaxys4` Accessed 05/05/15
[3]Microsoft Lumia: `www.microsoft.com/en-gb/mobile/phones/lumia` Accessed 05/05/15
[4]Nest Protect: `www.nest.com/protect` Accessed 05/05/15

### 1.1.1 Gesture Usability Problems

Advances in gesture-sensing technology have lead to gesture systems which fit in our pockets, yet the human-computer interaction community still lacks solutions to important gesture usability problems, which were discussed in detail by Bellotti *et al.* [11]. These usability issues may be why gesture-sensing capabilities have not made much of an impact in the market, outside of entertainment systems like Xbox Kinect. Research is needed to investigate solutions to them, especially now that the technology is ready for gesture systems to be integrated into a wide variety of everyday devices.

This thesis focuses on the problem of *addressing* a gesture system, which is the way users direct input towards it to initiate interaction. This problem is "*so fundamental that it is often taken for granted in UI design*" [11, p147]. While this is simple for some input modalities (reaching out and touching a screen or pressing a key on a keyboard, for example), addressing an in-air gesture system presents two challenges: (1) users need to know where to gesture, so that their hand movements can be sensed; and (2) they need to know how to direct their input towards one system, so that their actions do not have unintended effects on other systems.

These challenges have received little attention in commercial gesture systems so far. For example, the mobile phones discussed before do not give feedback during gesture interaction and users only know they have addressed the system properly when, if at all, they see the effects of their actions. Xbox Kinect uses its display capabilities to mirror its camera image, or show a stylised version of it, but this leaves users to infer if they are addressing it properly or not. The human-computer interaction literature also lacks a good solution to the challenges of addressing a gesture system, as existing ones are impractical and have limitations, as discussed in the following Literature Review. This thesis investigates ways of helping users address gesture-sensing systems, taking a first step towards making them easier to use.

Addressing a gesture system happens *before* users provide input to control a gesture system, occurring when they initiate interaction. It may be an explicit part of the interaction, requiring action from the user (to direct their input, for example), or it may be implicit, with users assuming that their input is being treated as intended. The following chapter will look at the problem of addressing a gesture system in greater detail.

### 1.1.2 Limited-Display Devices

Many have argued that gesture systems need to give good feedback to help users overcome some of their usability problems. Norman [97] argued that gesture interfaces are so unnatural that they need good feedback to be usable and understandable. Wensveen *et al.* [143] argued that they lack inherent couplings between action and function, unlike direct manipulation interfaces, making additional feedback from other sources necessary. In terms used by

Reeves *et al.* [111], system manipulations are visible but the effects of those manipulations on the system are not, without feedback. Finally, Zamborlin *et al.* [156] identified "meaningful feedback" as a requirement for modern gesture systems, calling it necessary for users to learn and improve interactions with the system.

However, many of the research prototypes and consumer products discussed so far are limited in their ability to give feedback. Mobile phones and household controls do not have large-sized screens for visual feedback, like gesture systems for televisions, laptops and desktop computers do. Instead, they have small screens or no screens at all. This thesis calls these devices **limited-display devices**, as they are limited in their ability to show feedback. Those without screens are unable to present information visually, meaning other display types are required. Those with small screens may present feedback that is difficult to see when gesturing from more than an arm's length away, meaning users may miss it or not find it useful. On small screens, feedback about gestures also takes away from the already limited space available for content. Other approaches are therefore needed for limited-display devices to present feedback. This thesis investigates the use of three alternative display types—audio, tactile and interactive light—which are discussed further in the Literature Review. Using these displays means gesture feedback may be more salient and screens—if devices have them—can be kept free for showing the content users are interacting with.

The research in this thesis focuses on gesture interaction with limited-display devices, like mobile phones and 'smart' household controls. These devices are growing in popularity and there is already research studying how they may use and benefit from gesture-sensing capabilities (as discussed in Section 2.5). The interaction techniques developed as part of this research will focus on limited-display devices, but will also benefit gesture systems without the same display limitations. For example, gesture systems for large displays, such as Xbox Kinect or Leap Motion[5], will also be able to use the novel interaction techniques to allow users to address them properly. They will also be able to use the methods investigated here to present multimodal feedback as well as, or instead of, feedback on their screens.

This thesis focuses on using feedback to improve usability and to help users address in-air gestures systems. Other approaches to improving gesture usability include relaxing gesture recognition requirements (so that users can be less precise in their interactions) and creating interactions which have a stronger coupling between gesture and action (so that feedback is inherent in the action's effects). These approaches do not consider the challenges of addressing a gesture system, however, which happen *before* input to control a system. A different approach is necessary. This thesis investigates interaction techniques which provide a different approach, with feedback playing a central role in those interactions. Feedback is important because it can inform users of what is happening during the address stage of

[5]Leap Motion: `www.leapmotion.com` Accessed 05/05/15

interaction, before the effects of any subsequent actions are clear.

## 1.2 Thesis Statement

This thesis asserts that multimodal feedback can help users address and interact with in-air gesture systems, even when their devices have limited or no screens. Tactile, audio and interactive light displays offer gesture systems a variety of ways of presenting information during gestural interactions, overcoming the need for visual feedback on screens. This thesis presents novel interaction techniques which use these modalities to help users address in-air interfaces, showing them where and how to gesture.

## 1.3 Contributions

This thesis makes novel contributions in three areas. First, it contributes a study of the use of tactile feedback for in-air gesture interactions. While tactile feedback has been the subject of much research in human-computer interaction, it has received little attention in the gesture interaction literature. In-air gestures are non-contact interactions so delivering tactile information is challenging; however, recent technologies have made the research in this thesis possible. Wearable devices and ultrasound haptic technology are two ways in which gesture interfaces can deliver tactile feedback to users' bodies. Two experiments focused specifically on the presentation and design of tactile feedback for in-air gestures, while a further three used tactile feedback as part of new interaction techniques. The first experiment in this thesis is also the first comparison of feedback from an ultrasound haptic display with feedback from conventional vibrotactile displays.

Second, this thesis contributes a study into interactive light displays, a novel output type for limited-display devices. While interactive lighting is an active area of research, the work presented here is the first detailed study of its use as gesture feedback. One experiment evaluated the feasibility of using interactive light displays for gesture feedback and a further three used them in a variety of gestural interaction techniques.

Finally, this thesis also contributes novel gesture interaction techniques for addressing gesture interfaces. While past research has investigated certain aspects of addressing a gesture interface, these have shortcomings which limit their use in environments with more than one gesture-sensing system. The interaction techniques studied within this thesis overcome these shortcomings and allow users to address gesture systems successfully. While these interaction techniques and their experimental studies focus on limited-display devices, they are appropriate for all types of gesture-sensing system.

## 1.4 Research Questions

This thesis aims to answer the following questions:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

**RQ2:** Can interactive light be used to present gesture feedback to users?

**RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

**RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

**RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

## 1.5 Thesis Structure

**Chapter 2**, *Literature Review*, begins by discussing problems which users must overcome when interacting with gesture-sensing systems. It then reviews interaction techniques for addressing gesture-sensing systems, one of the problems identified at the beginning of the chapter. The literature review then discusses gesture interaction with limited-display devices, including mobile phones and small household devices, showing the need for good feedback but the difficulty in providing it from such systems. Finally, this chapter discusses three output modalities—audio, tactile and interactive light—which can enhance the feedback capabilities of limited-display devices, so they can help users address them properly. This chapter provides a background for the work which follows in this thesis and motivates the research questions identified previously.

**Chapter 3**, *Tactile Feedback for Gesture-Sensing Systems*, investigates tactile feedback for gesture systems, addressing **RQ1**. Two experiments are described which compare methods of delivering tactile feedback from a distance and look at how information about gestures can be encoded using vibration. Later chapters build on this one by using tactile feedback in interaction techniques for addressing gesture systems.

**Chapter 4**, *Interactive Light Feedback for Gesture-Sensing Systems*, investigates interactive light feedback for gesture systems. It describes the *Gesture Thermostat*, a limited-display device which uses an interactive light display for feedback. This chapter also presents an experiment which evaluates the thermostat and its feedback, contributing an answer to **RQ2**. Its findings inform the later use of interactive light displays for feedback in interaction techniques for addressing gesture systems.

**Chapter 5**, *Showing Users Where to Gesture*, discusses an interaction technique called *sensor strength feedback*, which uses multimodal feedback to help users find where to gesture when addressing a gesture system. This chapter describes an experiment which investigates the effectiveness of this interaction technique, contributing a partial answer to **RQ3**.

**Chapter 6**, *Directing Input Towards A Gesture-Sensing System*, investigates the problem of directing input towards a system when addressing it and describes *rhythmic gestures*, an interaction technique which allows users to do this. This chapter describes an experiment which studies the use of rhythmic gestures, resulting in design recommendations for the effective use of this technique and providing an answer to **RQ4**.

**Chapter 7**, *An Interaction Technique for Addressing Gesture-Sensing Systems*, describes the combination of the interaction techniques evaluated in Chapters 5 and 6, resulting in an approach which helps users address gesture systems by showing them how and where to interact. This chapter presents the final experiment of this thesis, which investigates the effectiveness of this interaction technique; findings from this experiment contribute answers to **RQ3** and **RQ5**.

**Chapter 8**, *Conclusions*, summarises the research within this thesis and reflects upon how it answered the research questions identified in the introduction. This chapter also identifies the main contributions of this thesis, outlines its limitations and discusses possibilities for future work.

### 1.5.1 Overview of Experiments

Six experimental studies are discussed in this thesis, each contributing to the research questions outlined in Section 1.4. Three of these experiments focus on novel feedback modalities for limited-display devices, while the other three focus on interaction techniques for addressing gesture interfaces. Table 1.1 provides a brief overview of each experiment and shows which research questions they contribute answers to.

As this research is motivated by the limited visual feedback capabilities of mobile phones and small household appliances—limited-display devices—the experiments described in this thesis use both. Three experiments involve gesture interaction with mobile phones; the other three involve gesture-controlled household controls. Table 1.1 also shows which experiments use which type of limited-display device.

| Topic | Experiment | Research Questions | Experiment Context and Purpose |
|---|---|---|---|
| Feedback Modalities | Experiment 1 | RQ1 | • *Mobile phones*<br>• Compare tactile displays for in-air interfaces<br>• Compare wearable tactile display locations |
| | Experiment 2 | RQ1 | • *Mobile phones*<br>• Understand effective feedback design<br>• Study effects of more complex encodings |
| | Experiment 3 | RQ2 | • *Small household appliances*<br>• Evaluate effectiveness of interactive light feedback |
| Address Interactions | Experiment 4 | RQ3 | • *Mobile phones*<br>• Evaluate *sensor strength* interaction technique<br>• Compare feedback modality combinations |
| | Experiment 5 | RQ4 | • *Small household appliances*<br>• Evaluate *rhythmic gestures* interaction technique<br>• Identify design guidelines for *rhythmic gestures*<br>• Evaluate effectiveness of rhythmic gesture feedback |
| | Experiment 6 | RQ3, RQ5 | • *Small household appliances*<br>• Evaluate address interaction technique<br>• Compare feedback combinations<br>• Evaluate *sensor strength* for household controls |

**Table 1.1:** Summary of experiments presented within this thesis, showing which research questions they contribute answers to and which type of device was used (in italics).

# Chapter 2

# Literature Review

## 2.1 Introduction

The opening chapter introduced the idea that despite recent advances with in-air gesture sensing technologies, the human-computer interaction literature still lacks solutions to well-known usability problems with gesture interfaces. With increasing numbers of limited-display devices being enriched with sensing capabilities, these usability problems become more challenging. This literature review introduces these problems (in Section 2.3), illustrated using two scenarios. Each of these scenarios demonstrates usability issues when using gesture-sensing systems, especially when they have limited output capabilities and when they are in environments where there is more than just one gesture-sensing system.

This review begins by taking a quick look at the appeal of gestures and identifies the types of in-air gesture used in human-computer interaction. This is used to provide context for the following section, which looks at usability issues in gesture interaction. These usability problems are discussed in the context of *Making Sense of Sensing Systems* [11]. In that paper, Bellotti *et al.* drew inspiration from how people interacted with each other and identified five challenges which designers of sensing systems must consider. They argue that while traditional user interfaces have well-understood solutions to these problems, sensing-based user interfaces do not. This thesis focuses on the *address* problem in particular, which concerns how users direct their input towards a system. As this review will show, research is needed to find an effective way of addressing in-air gesture systems.

Following that introduction to gesture usability problems, this review then discusses interaction techniques for addressing gesture systems (in Section 2.4); this builds on the overview of the *address* problem given in Section 2.3.2. A critique of these state-of-the-art techniques identifies limitations which makes them impractical, as they do not consider the possibility of more than one gesture-sensing system within an environment.

This chapter then looks at gesture interaction with limited-display devices (in Section 2.5), the interaction context on which this thesis focuses. In particular, the review covers gesture interaction with household appliances and with mobile phones (which are often called *around-device interfaces* in the mobile human-computer interaction community). When discussing work in these areas, specific focus is given to the types of feedback given to users. As will be shown, feedback and usability is often an afterthought, with most work focusing on gesture-sensing techniques or other aspects of interaction design, like choosing which gestures do which actions. The work also often identifies the need for better feedback techniques and many of the usability problems discussed in the previous section are observed.

Finally, this review looks at how limited-display devices may give feedback during gesture interaction (in Section 2.6). Three output modalities for gesture interfaces are reviewed: audio feedback, tactile feedback, and interactive light feedback. Audio and tactile feedback are well understood in other interaction areas; however, their use in gesture interaction is more limited. For tactile feedback, this is often the case because delivering tactile cues is difficult when users are not in contact with the device they are interacting with. Recent technologies can overcome this problem. This section of the review discusses how tactile feedback can be delivered from a distance, using technologies such as ultrasound haptics and wearable tactile displays.

Recent developments in interactive lighting technology mean that small and simple light sources can be used to communicate with users visually, rather than doing so using screens. Interactive light, then, could be used by limited-display devices as a means of giving visual feedback, by illuminating the area around the device. Examples in commercial products are given which show that many limited-display devices have already been enriched with such lighting sources, although their use is typically non-interactive or for novelty. This review considers how others have used interactive light in interaction design. It defines *interactive light feedback* as a novel type of feedback, which can be used alongside audio and tactile displays to create multimodal gesture feedback from limited-display devices.

## 2.2 Gestures in HCI

Gestures are appealing in human-computer interaction because they have the potential to allow us to communicate with technology in the same ways that we might communicate with other people and because they can allow us to manipulate virtual objects in the same ways that we might manipulate physical objects in the real world: by grasping, pushing, pressing, etc. McNeill [87] argues that gesture is an integral part of language, providing the imagery in what he calls an "imagery-language dialectic". He makes the point that gesture does not merely ornament the use of language but is an important part of the language itself, found in around 90% of spoken utterances.

This strong coupling between speech and gesture has been explored in HCI. One of the earliest gesture systems was *Put-That-There* [16], which allowed users to provide instructions using combinations of speech and gesture. Gesture was an integral part of the system, providing context and extra information about the commands that were giving using speech: gesture told the system what "that" was, and where "there" should be, so that the system could put *that* over *there*. In other gesture-sensing systems, gesture has been used as the *only* means of communication, with hand and body movements providing all of the information for the interaction. In such systems, gesture is typically used to present commands (for example, performing the "next chapter" gesture in *Charade* [9]) or to manipulate some aspect of interactive content (for example, rotating an object by turning the hand in *PalmSpace* [74]).

Others have developed taxonomies which describe the varied use of gesture in HCI, providing a more formal way to think about the types of gestures used in gesture systems. Karam and Schraefel [67] reviewed the use of gestures in the HCI literature and identified five types of gesture: deictic, manipulative, semaphoric, gesticulation and language gestures (e.g. sign language). Deictic gestures are those which use pointing to provide context for spoken information, like those in *Put-That-There* [16]. Manipulative gestures are those which have a strong coupling between virtual entities and the gesture movements, like in *PalmSpace* [74]. Semaphoric gestures are those which issue commands to a system, like in *Charade* [9]. Gesticulation gestures are those which accompany speech in conversation. Finally, language gestures are those used for sign languages, strongly grounded in linguistics. This range of classifications shows the many ways in which gesture can be used to interact with technology.

Something which is common to all types of gesture is the need for good feedback, an issue which was raised in the introduction. Even when there is a close relationship between a gesture and its effects on the system (e.g. Karam and Schraefel's "manipulative" gestures [67]), feedback is needed to provide important information which makes the interactions usable and understandable (as argued by Norman [97]). This means that the research in this thesis, specifically that investigating non-visual gesture feedback, could be applied to all types of gesture system, as the need for feedback is universal. The other main focus of this thesis research is the problem of addressing in-air gesture systems, something which happens before users provide input to control the system. As will be discussed in the following section, the address problem is a general usability issue, so this aspect of the thesis is also relevant to all gesture systems.

## 2.3 Making Sense of Gesture-Sensing Systems

### 2.3.1 Scenarios Illustrating Gesture Usability Problems

Two scenarios have been developed as part of the thesis to illustrate usability problems with in-air gestures:

> **Scenario 1:** Katie is in her living room watching television. Her mobile phone is on the coffee table a few feet in front of her. Her phone starts ringing; after glancing at its screen and seeing an unrecognised number, she gestures towards it to dismiss the call. As she gestures, her television also recognises her hand movements and changes channel, whilst her phone continues to ring.

> **Scenario 2:** Bobby is getting ready to leave the house. As he walks past the thermostat in the hallway, he gestures — waving, as though saying goodbye — to turn off the heating. He gestures once more towards the hallway lights as he opens the door and the lights turn off. Although switching the lights off had an immediately noticeable effect, Bobby is unsure if the heating turned off.

In the first scenario, Katie encounters two problems: her gesture did not silence her mobile phone, but it did affect her television. Her mobile phone may have continued to ring for many reasons: it may have been unable to sense her gesture, it may have sensed her gesture but not recognised it, or it may have recognised her gesture but it was not the gesture to dismiss a phone call. More formally, Katie's gesture may have been *not sensed*, *sensed* but *not desired*, or *desired* but incorrect (using terminology defined in Benford *et al.'s* framework [13]). These are outcomes which users must deal with when interacting with gesture systems; users need to know where to gesture, which gestures are available, and how to perform them so that they will be recognised. Katie's second problem was that she unintentionally changed television channel with her gesture. This illustrates the *Midas Touch* problem [70] (also called *immersion syndrome* [9] or the *live mic* problem [144]), where sensing interfaces continuously look for input and consider any recognised signal as intentional interaction.

Bobby's scenario illustrates a further gesture interaction problem. While his gesture to switch off the lights produced an immediately noticeable effect, indicating its success, Bobby does not know if his first gesture was recognised and had its desired effect. His first gesture may have been recognised by the thermostat, but the effect of the gesture will not be immediately noticeable. In this case, Bobby does not know if his gesture was recognised and, if it was, had its expected effect on the system.

### 2.3.2 Making Sense of Sensing Systems

Bellotti *et al.* [11] identified more general versions of these usability problems, which they called *Address*, *Attention*, *Action*, *Alignment* and *Accident*. Address concerns how users direct their actions towards the systems they wish to interact with, while not affecting others. Katie encountered this problem when her gesture affected her television; she needs a way to *address* her mobile phone which makes it clear she does not intend to interact with other systems which may also be sensing her actions. Attention concerns how users become aware that a system is ready and is paying attention to their actions. Bobby's thermostat did not show attention to his movements, making him unsure if his gesture was sensed or not. Action concerns how users bridge the *Gulf of Execution* [96], by determining what action is required to have the intended effect on the system. Alignment concerns how users understand the effect that their actions have had — or are having — on the system; or how users overcome the *Gulf of Evaluation* [96]. Katie and Bobby both experienced issues with action and alignment; Katie did not know if she chose an incorrect gesture or if her gesture was just not recognised, while Bobby was not aware what effect — if any — his gesture had on his thermostat. Finally, Accident concerns how users avoid or overcome mistakes during input. Katie experienced accidental input when her television also interpreted her hand movements, something which could have been avoided had she been able to address her gestures towards her mobile phone.

Many of the problems illustrated in Katie and Bobby's scenarios have been investigated before, although their solutions are not always appropriate in practice. Gesture systems are being introduced to ever smaller devices, which lack the display capabilities for easy visual feedback. Increasing numbers of gesture-sensing devices also mean that prior solutions, assuming users are only being sensed by one system, are impractical. This section examines existing solutions to some of these usability problems, identifying why they may be unsuitable for the next generation of gesture-sensing devices.

#### Address: Initiating Interaction

The way in which users address a user interface is "*so fundamental that it is often taken for granted in UI design*" [11, p147]. Yet, such a fundamental part of interaction can cause a variety of challenges for users. Bellotti *et al.* [11] identify two problems present when addressing sensing systems: unwanted response, from unintentional input; and no response from being unable to initiate interaction. This section discusses these problems in turn and identifies how others have tried to overcome them. It should be remembered that address happens *before* input to control a gesture system and that any interaction to address a system happens separately from other interaction techniques.

### *Avoiding unintentional gestures*

In Katie's scenario, her gesture was unintentionally recognised and acted upon by her television. Gesture sensors often sense all movements within range, whether users intend for these movements to be gestures for them or not [9]. Bellotti *et al.* [11] identified the importance of being able to *address* a sensing interface, so that users can specify which, of the potentially many available, they wish to interact with. Gesture interfaces, likewise, need ways of knowing when they should, or should not, be sensing and interpreting input. While this helps systems avoid interpreting gestures meant for other systems, it also means that everyday actions — such as reaching out and lifting a coffee mug — do not get treated as gesture input.

Unintentionally gesturing is known as the *Midas Touch* problem [70], after King Midas from Greek mythology, who had the — often unfortunate — ability to turn anything he touched to gold. Others have called this *immersion syndrome* [9], as users are always immersed in interaction, even if they do not wish to be. Wigdor and Wixon [144] called unintentional gestures *false-positive recognitions*, noting that these mis-recognitions must be avoided if users are to feel confident and skilled. Preventing unintentional gestures, then, is important if users are to feel comfortable using in-air interactions.

Many solutions have been proposed for overcoming the *Midas Touch* problem and these will be reviewed in Section 2.4. As will be shown, these solutions can help users avoid unintentional input, but they each have limitations which make them impractical when there is more than one gesture-sensing system. As our homes and personal devices continue to expand their sensing repertoires, we need new ways of addressing gesture systems. Addressing gesture systems explicitly allows users to direct their movements towards the one they wish to interact with, while also avoiding unintentional input through an explicit action of intent. This thesis will investigate novel interaction techniques for addressing gesture systems, which overcome the limitations with existing approaches.

### *Knowing* where *to gesture*

Another aspect of addressing an interface is knowing *where* to interact. Users must gesture within range of the input sensors for their movements to be sensed and understood. Movements outside of sensor range will be undetected, but users may have no way of knowing if their movements were not sensed or just unrecognised; Katie experienced similar uncertainty when trying to interact with her mobile phone in her scenario. Gesture-sensing systems have ambiguous input areas and it is not always clear where sensors are located and what they can and cannot see. A common approach for helping users find where to gesture is to give them feedback when they are within sensor range. This is called showing system attention, which is discussed in the following section.

As well as showing system attention, interfaces can also tell users how well they can be seen. Interfaces with high-resolution displays often show users what the gesture sensors can see; for example, *StrikeAPose* [141] displayed a stylised version of sensor data, allowing users to see their position in the sensor field of view. Xbox Kinect games often use similar visualisations. In *Proxemic Flow* [139], a large floor display presented visualisations under users' feet as they approached an interactive screen; they used a traffic light metaphor to show users how well they could be seen, with green light around their feet meaning sensing quality was high. Users can use this information to adapt how they gesture, for example moving closer to the sensor or towards the centre of the field of view.

Not all gesture interfaces have these display capabilities, so other output modalities are required. Morrison *et al*. [90] looked at how audio feedback could help users understand how well they can be seen. They found that telling users which body parts could not be seen offered no benefit over simply saying that the body was not in full view; it was also more complex. Although feedback saying the body could not be fully seen was more ambiguous than naming unseen body parts, such feedback could encourage users to explore the input space and develop their own understanding about how they need to gesture. Active exploration is an important part of *scaffolding* [144], where users are given incremental cues and situations to help develop their skills with a user interface. With scaffolding, users learn by exploring an interface's possibilities and limitations, which are revealed through affordances and feedback.

#### *Summary of addressing gesture-sensing systems*

Addressing gesture-sensing systems presents two problems: avoiding unintentional input and being able to initiate interaction. While others have studied these problems separately, little research has considered them together. Research is needed to develop interactions which allow users to direct their input towards a system, while also helping them find where to interact, especially as addressing sensing systems is "*so fundamental*" [11, on p417]. Solutions to this challenge also need to support limited-display devices, which are unable to give the same rich visual feedback that approaches like *StrikeAPose* [141], *Proxemic Flow* [139] and Xbox Kinect use to help users discover where to interact. Later sections of this review will provide more background on this problem area, including state-of-the-art techniques for addressing gesture systems (in Section 2.4), gesture interaction with limited-display devices (in Section 2.5) and ways in which these devices could communicate with users without screens (in Section 2.6).

#### *Addressing gesture systems in collaborative contexts*

There are an increasing number of gesture systems that support multiple users providing input at the same time. For example, O'Hara *et al.* [101] describe a gesture system for surgical settings where there may be multiple users who need to take control of the system and provide input. In these situations, users also need to know who is in control of the system, if *they* are in control of the system, and what actions they need to take to either take or relinquish control. While the research in this thesis focuses on single-user gesture interactions, these issues are worth considering and will be revisited in the Conclusions chapter.

### Attention: Showing Response to Action

When Bobby gestured at the thermostat and light fitting in his hallway, neither acknowledged his actions as he began gesturing. The light fitting eventually gave a response through *functional feedback* [143] as the light switched off; however, had the interface not understood his gesture then no feedback would have been given. Katie also unintentionally interacted with her television as it gave no indication that it was sensing her hand movements. Sensing interfaces should show their users whether and when they are paying attention to them; Bellotti *et al*. [11] call this showing *system attention*, noting that it is missing in many sensing systems. Had Katie's phone shown system attention, she would have known that it could not sense her through its lack of feedback as she began gesturing.

Users need feedback about system attention when gesturing to know if their movements are likely to be sensed. If gesture interfaces do not show system attention — or *"signs of life"* [144, p50] — then users do not know if lack of a response after a gesture means their movements were not sensed, or if they were sensed but not recognised as a valid gesture. One of the design challenges in showing system attention is making sure users will notice it. Golod *et al*. [37] suggest using as many available output modalities as possible to show system attention, making it more likely that users will notice it. Visual feedback alone is insufficient as users may not always be looking at a device as they interact with it [11]; this was the case in Katie's scenario, as she gestured towards her ringing phone while continuing to watch television. Gesture interfaces may also have limited or no display capabilities, like the simple household devices in Bobby's scenario, meaning other types of feedback are necessary.

### Action: Bridging the Gulf of Execution

Katie's gesture to silence her mobile phone did not have its desired effect, as her phone continued to ring. There are many possible reasons for this: her gesture was not sensed, her gesture was sensed but not a valid gesture, or her gesture was sensed and valid but was mapped to a different action (for example, increasing the ringer volume). Katie was unable to bridge the *Gulf of Execution* [96]. For her gesture to be successful, she needed to know where to gesture, which gesture to perform, and how to perform it so that it would be recognised; however, she received no feedback to help her make these decisions. Gesture systems need to give users sufficient information about interaction so that they can gesture successfully.

As well as knowing where to gesture (as part of *addressing* an interface), users also need to know which gestures to use to accomplish their goals. Gestures can be *self-revealing* [144], meaning they become apparent to users through affordances and feedback. Revealing all available gestures and their actions is infeasible, however; especially for devices with limited displays. Instead, users can be introduced to fewer gestures at a time, a type of *scaffolding*. Scaffolding [144, p53] is a design approach where users are given just enough assistance to help them gradually develop skills and experience using a system. Over time, users would become less reliant on assistance about gestures. Others have proposed selecting gestures which users already know instead. These include using user-defined gestures, where users map their own gestures to actions [95, 78], or guessable gestures [152], which users are likely to associate with certain actions (although others have argued that this is rarely the case [144]).

Users also need to know *how* to perform gestures, so that they can be reliably sensed. Solutions to this problem typically include giving textual and iconic hints (for example, *just-in-time chrome* [144], *LightGuide* [124] and *StrikeAPose* [141]), which may be accompanied by further feedback during gestures (for example, *Glissando* [91] and *Gestu-Wan* [118]) or after gestures (for example, *Recognizer Feedback* [66]). Gesture feedback is discussed more in the following section, which looks at how users align their conceptual models of interaction with the feedback they receive.

### Alignment: Bridging the Gulf of Evaluation

Another challenge in gesture interaction is helping users *align* [11] their understanding of how their actions are affecting the interface with the information they are being given. Users need feedback to know if their gestures are being recognised successfully and the effects they are having. Feedback can help users gesture by giving insights into how their movements are being sensed by the gesture-sensing system, allowing them to adapt how they gesture (for example, moving closer to sensors) and informing their future interactions with the system. Feedback is also needed about the effects gestures are having on the system, to tell users how, if at all, they have affected it and if they have accomplished their goals. This section discusses each of these feedback types — sensing and effects — separately.

#### *Feedback about sensing*

Neither Katie or Bobby received feedback *during* their gestures; Katie did not know if her mobile phone could sense her hand movements at all and Bobby did not know if the thermostat responded to his gesture. Users need feedback about how their movements are being interpreted so that they can adapt how they gesture, to be more accurate in future. For example, users may discover that their gestures will be more reliably sensed if they move their hands at a slower speed. Feedback during gestures may be *seamful* [25], revealing raw sensor data so users can interpret its ambiguity for themselves. Seamful feedback has also been called *echo feedback* [144, p83], as user interfaces return a representation of their incoming sensor data. Whilst raw sensing data is ambiguous to users, it may encourage active exploration of how the gesture sensors work. For example, a user may discover that data appears less noisy when they move closer to the sensor.

An alternative way of presenting sensor data is using *seamless feedback* [72], where raw data is presented in an idealised form. For example, a seamless design may filter noise in hand movements to visualise gestures using straight trajectories, whereas a seamful design may present unfiltered hand trajectories. While exposing raw sensor data may provide users with additional insight about gesture sensing, it can also be overly complex. A comparison of seamful and seamless gesture feedback (using the designs in the previous example) found that seamless feedback was more effective [72]. Users were distracted by the seamful design as it clashed with their mental models of how their gestures were being recognised; they thought their hand movements were being sensed as smooth paths, rather than the fuzzy, imprecise paths their hands actually moved in. Wigdor and Wixon [144, on p83] discussed the importance of understanding associations between cause and effect in natural user interfaces; providing sensor data on its own — fully seamless feedback — may give insufficient information about these associations.

Feedback about gesture sensing may not always be necessary if there is a clear correspondence between the input gestures and their noticeable effects on the system, as users would know if their gestures have been recognised correctly and had the desired effect. This type of feedback is known as *functional feedback* [143] and will be discussed more in the following section. Even when this is the case, users may find feedback about sensing useful as it can provide additional insight into the quality of the interaction. For example, it might let users know how well they performed a gesture, so that they can decide if they need to change how they interact in future. Feedback about sensing also lets users know about the interaction when it does *not* work as expected; for example, showing them that the system is responding but does not understand their gestures.

### *Feedback about effects*

Gesture-sensing systems also need to give feedback about the effects gestures are having on system state, which Wigdor and Wixon call *semantic feedback* [144, on p83]. If providing sensor data as feedback shows what an interface *sees*, semantic feedback is what an interface *knows*. For example, gesture interfaces which let users make selections may give semantic feedback about which item is currently selected. Bobby's thermostat could have given semantic feedback showing that his gesture resulted in the heating being turned off, for example.

Gestures will often give *functional feedback* [143], where effects of their action become apparent through its function. Music stopping after a 'pause music' gesture is recognised is an example of functional feedback, as users will hear that their gesture produced the desired effect. However, functional feedback may not always be available or practical. Some actions may have no immediately perceivable effect, for example, so other feedback must also be considered. In Bobby's scenario, functional feedback from the thermostat would be the eventual change in temperature; however, this is not instant and may not even be perceived. Bobby would need additional feedback to know his gesture was recognised and caused its desired effect: turning the heating off.

Reeves *et al.* [111] discussed issues surrounding the visibility of system *manipulations*, which are actions performed by a user on the system, and system *effects*, which are the results of those manipulations. In a mid-air gesture system, manipulations are visible to the user performing them and to others nearby, who see the user gesturing. However, their effects are not always visible, as was the case in Bobby's scenario. As well as causing uncertainty in the user about the success of their manipulations, this also raises an interesting social issue. Users know their intentions when gesturing and may assume that their input was recognised successfully and had its intended effect, but others nearby may not know what the user had intended. Feedback about effects, then, also serves to inform "spectators" [111] about what the user of a gesture system is doing. Appropriate feedback about gestures and their effects would make the system "expressive" [111], where manipulations and effects are visible to all, rather than "suspenseful" [111], where manipulations are visible but their effects are not.

#### *Summary of feedback for alignment*

Users need feedback from gesture-sensing systems so that they can align their mental models of system state with the actual system state. There are two types of feedback which can help users do this. Feedback about how gestures are being sensed can help users understand how to gesture more effectively in future, for example by giving insight into why a movement may not have been recognised as a gesture. Feedback about the outcomes of gestures lets users know if their movements are having their intended effects, or not (i.e. making the effects of manipulations clear [111]). While changes in system state may become apparent through functional feedback, this is not always appropriate and more explicit types of feedback should also be given. This thesis does not investigate the alignment problem of gesture interaction, as there is a large body of work which already addresses this area.

### Accident: Avoiding and Correcting Mistakes

Being able to avoid and correct mistakes is a fundamental design principle in human-computer interaction. Norman [97] notes that gesture interfaces will need an undo mechanism, as users may wish to reverse unintended gestures or unintended effects of gestures. Bellotti *et al*. [11] discuss the need to also correct errors as they happen, for example cancelling a gesture mid-course. Wigdor and Wixon [144, on p193] recommend that gestures have a negative action, cancelling their effects, and a reciprocal action, which produces the opposite effect. An alternative to reciprocal actions is to have a single 'undo' gesture, which reverses the effect of the previous gesture; for example, *Wear Ur World* [88] let users undo gestures by drawing an "x" symbol in mid-air. In Katie's scenario, she could correct the unintentional gesture to her television by performing a reciprocal gesture to change the channel, or by performing an undo gesture to reverse the most recent gesture effect.

Existing solutions for avoiding and correcting mistakes when gesturing — cancelling gestures midcourse, performing a reciprocal gesture, and performing an 'undo' gesture — may be needed less frequently if gesture interface designers address the previously discussed problems. For example, Katie would not need to undo the effects of an unintended gesture if she was able to address the intended interface in the first place. As gesture interaction becomes more widespread, it is likely that a standard 'undo' gesture will emerge, similar to the widely-adopted Ctrl+Z/Cmd+Z keyboard shortcut; users will be able to use such a gesture with any gesture-sensing system to correct mistakes.

### 2.3.3 Summary

The scenarios at the beginning of this section illustrate the usability problems users often encounter when interacting with gesture-sensing systems. In these scenarios, Katie and Bobby both experienced difficulty using gestures: Katie's gestures had unintended effects on another system whilst producing no response from the one she wanted to interact with; and Bobby was unsure if one of his gestures was recognised and acted upon. These problems have been discussed by others, most notably by Bellotti *et al*. [11], who framed sensing interaction problems by comparing human-computer interaction to human-human interaction. They, along with others (as discussed in Section 1.1 of the Introduction chapter), have highlighted the importance of giving feedback during sensing interactions, to help users interact confidently and effectively.

Addressing a gesture-sensing system is perhaps the most important part of gesture interaction. Without first addressing an interface, users would be unable to accomplish their goals through using it. Despite its importance, the literature is missing solutions which allow users to effectively address gesture-sensing systems. Users must overcome two problems when addressing a gesture-sensing system: they need to know where to gesture and they need to know how to direct their gestures towards one system in particular. This thesis investigates ways in which gesture-sensing systems can help users overcome these problems.

Feedback is important when addressing an interface, as it is throughout an entire gesture interaction. However, many gesture-sensing devices have limited output capabilities, which restricts their ability to give good feedback about gestures. The Introduction chapter introduced *limited-display devices*, devices which are increasingly being enhanced with new input modalities — like gesture input — but which have limited, or no, visual displays. Gesture-sensing systems typically rely on visual feedback, shown on large, high-resolution displays. Alternative feedback mechanisms are required for limited-display devices. This thesis also investigates how limited-display devices — or any device, for that matter — can use other output modalities for effective feedback.

The other sensing-system problems (Attention, Action, Alignment, and Accident) are not focused upon in this thesis. They have received more interest than the Address problem, which has notable gaps in the literature. However, this thesis still makes contributions from which these problems will benefit. Improving the output capabilities of gesture-sensing systems using multimodal feedback means that the interaction techniques discussed throughout this chapter will have more ways of communicating with their users. Future work in these areas could also build on this thesis by enhancing the interaction techniques used for addressing in-air gesture interfaces.

## 2.4 Addressing Gesture Interfaces

So far, this review has discussed usability problems which users may encounter when using in-air gestures. Five usability problems were discussed in the previous section and the *address* problem, in particular, was found to require more research. This section reviews the literature for techniques which allow users to address devices using in-air interactions, building on the earlier introduction to this problem. Four types of technique are discussed: pointing gestures, activation gestures, active zones and gaze sensing. As will be shown, there are limitations for each of these techniques which mean they will be impractical, requiring further research to develop an alternative.

### 2.4.1 Pointing Gestures

Many have explored pointing at things (Figure 2.1) — with handheld devices or with fingers — as a means of addressing them and interacting with them. One reason this method is popular is that it is direct: it allows users to address "*that one there*" [133]. Despite the direct nature of pointing at things, it is often problematic. Many have reported issues with ambiguity about what users are pointing at and others have designed interaction techniques which attempt to resolve this ambiguity. Some of these pointing interactions require hand-held devices; however they are discussed here because their findings can inform the design of in-air pointing techniques too.

*GesturePen* [133] was a stylus which allowed users to select between smart objects by pointing at them in mid-air. Users pointed at objects ("that one there") and then confirmed their selection by pressing a button on the stylus. *DopLink* [7] also let users address objects by pointing a device at them, although they used gestures rather than button presses as a means of selection. Users could flick or push the device towards an object to select it. *PICOntrol* [120] allowed users to interact with distant objects by pointing a handheld pico projector at them. Projectors were also used to show user interface controls in the space surrounding objects. Like *GesturePen*, they used buttons on the projector to confirm selection. Users could also move the projector for simple input gestures, like rotating it to turn a virtual dial.

None of these interaction techniques gave users ways to overcome ambiguous selections. If users pointed towards two or more objects which were close to each other, they had no way of disambiguating which one they wanted to address. *Point & Control* [22] addressed this issue by asking users to make a further selection if the system was not sure which object was being pointed at. Users pointed a smartphone towards an object to address it and, if selection was ambiguous, a list of possible targets was displayed on the touchscreen. Upon selecting the intended object, users could then interact with it using the smartphone touchscreen.

Others have used in-air pointing gestures, without a handheld accessory, for addressing interfaces. *MISO* [34] let users interact with devices by pointing at them, clicking their fingers, and then performing a short gesture. Clicking was used to show intention to interact, avoiding unintentional input. If users gestured correctly, a short confirmation sound was given. However, users received no feedback about their gestures before and during a gesture performance. *MISO* also lacked a method for resolving ambiguity over which device users were pointing at. Delamare *et al.* [28, 29] addressed ambiguity with in-air pointing by also using wrist-rotation to select between objects. Users first pointed in the general area of the device they wanted to address; then, they rotated their wrist to select between objects in that area. They evaluated their technique and found that users could reliably select between densely-populated objects.

Few of these interaction techniques gave users any feedback as they addressed objects, which may explain the problems with selection ambiguity. *PICOntrol* [120] gave users plenty of feedback about their interactions, as its projections made system state visible. Users could see when they were successfully addressing a device in the environment as its projected interface would appear around it. When Delamare *et al.* [28] applied their point-and-rotate gesture to interactive light sources (lights placed within small balls, in this case), they were able to use the lights themselves to give users feedback as they gestured. When users were

**Figure 2.1:** Pointing gestures allow users to address devices directly by pointing at them.

pointing in the general area of a light source, it illuminated at medium brightness; when they used wrist-rotation to select a particular light, it illuminated at maximum brightness.

### 2.4.2 Activation Gestures

An alternative to using pointing gestures to address in-air interfaces is to use a *clutch mechanism*, which tells the gesture system when the user is providing input. Hinckley *et al.* [50] discussed examples of clutch mechanisms for spatial input devices, noting that a poorly designed clutch can be the cause of difficult usability problems. The examples they discuss use push buttons or foot pedals, which are to be held continuously during the clutch operation. Such an approach would be inappropriate for mid-air gesture systems, as pressing a button during the interaction would take away some of the benefits of gesture input; for example, users would not be able to gesture freely from anywhere in the room if they had to first locate a physical clutch mechanism.

The in-air gesture solution to this problem was to use an arbitrary gesture as the clutch mechanism, to *activate* gesture input. Such *activation gestures* are typically modal: users perform a discrete gesture to activate the system, after which they perform other gestures. By performing an activation gesture (or *gating gesture* [149]) when addressing an interface, users show their intent to interact. Such gestures must be uncommon so that users do not perform them accidentally [70, 37]. Once gesture interaction is active, sensing continues until users end interaction, either by performing another arbitrary gesture (a *closure* gesture [37]) or by leaving the sensor space. Activation gestures include dynamic hand movements, such as the

Xbox 360 "Wave to Kinect" gesture[1], static hand poses, as with the pointing interactions described previously, and finger clicking, as in *MISO* [34]. Users could also perform full-body gestures and poses to address a gesture interface. In *StrikeAPose* [141], users had to perform the *Teapot* gesture — placing one hand on their hip — to engage with a large public display.

When using activation gestures, it is important to select an appropriate body pose or gesture. Hinckley *et al.* [50] discussed the implications of a poorly designed clutch mechanism and many of the same issues are likely to apply to in-air gesture systems with activation gestures, too. O'Hara *et al*. [101] found that their users often activated interaction unintentionally when using their activation gesture, which they thought would be unlikely to be performed accidentally. Although activation gestures allow users to show intent to interact, they are not ideal for specifying *which* gesture-sensing system a user wishes to address. Each interface would require its own unique gesture, which must then be learned in advance or communicated to users. Communicating activation gestures can be difficult, especially for devices with limited output capabilities. Even with an appropriate display, showing users which activation gesture to use can be difficult as they are unnatural poses or movements, often with strict sensing requirements, so that they are less likely to be performed accidentally.

### 2.4.3 Active Zones

In *Charade* [9], users addressed the interface by placing their hand within an *active zone*, an area of the sensor space in which all sensed movements were treated as gestures (illustrated in Figure 2.2). Using an active zone for address means that users are less likely to accidentally gesture with ordinary hand movements, although this remains possible. While this may be a successful approach for a single gesture system, it could be problematic in environments with multiple gesture-sensing devices. Multiple active zones may overlap, leading to uncertainty over which device a user is gesturing at. More active zones also means more space in which users may accidentally provide input. Another problem with active zones is that their position and extent are unclear; users may gesture accidentally because they are not aware that their hands are within an active zone for a gesture interface. Likewise, users may have difficulty finding an active zone when they want to interact, especially if the zone is conservatively sized to try to reduce the *Midas Touch* problem.

Others have combined active zones with activation gestures, meaning users must perform a certain gesture within a certain area. For example, Golod *et al*. [37] and Sørensen *et al*. [126] looked for a particular in-air gesture performed above a table surface. However, active zones with activation gestures suffer many of the same problems as each technique on its own.

[1] `support.xbox.com/en-US/xbox-360/kinect/body-controller` Accessed 09/03/15


**Figure 2.2:** Active zones are subsets of the space covered by a gesture sensor. Gestures within these zones will be treated as input; hand movements outside these zones will be ignored by gesture-sensing systems.

### 2.4.4 Gaze

Schwarz *et al*. [121] combined body pose, gaze and movement information to estimate how willing users were to engage with a gesture interface. Their approach meant devices could infer when they were being addressed without users having to gesture first. Much like the pointing gestures discussed earlier, gaze may be ambiguous. Users may look in the direction of two or more sensors at once or may not look at sensors at all. When investigating multimodal speech and gaze input to attentive user interfaces — those which anticipate users' needs — Maglio *et al*. [81] found that users did not always look at devices before addressing them. There were also occasions when users did not look at devices at all during interaction. If sensors are embedded and hidden within devices then users may not know where to look in the first place, adding further ambiguity about where they are looking and what they are addressing.

### 2.4.5 Summary of Addressing Gesture Interfaces

Users must be able to address a gesture-sensing system. Addressing a gesture-sensing system involves finding where to gesture, so that movements can be sensed, showing intent to gesture, so that ordinary hand movements are not accidentally interpreted as input, and directing input towards one system in particular, so that gestures do not unintentionally affect other systems. This section discussed a variety of interaction techniques for addressing interfaces, including pointing at them, performing an activation gesture, interacting within a certain area and looking at them. These interaction techniques allowed users to show intent to interact and some also helped users identify which system they wanted to interact with; however, each had disadvantages which would make them impractical, especially when there is more than just one gesture-sensing system.

The literature paid little attention to an important aspect of addressing gesture interfaces: finding *where* to gesture. Users need to know where to gesture so that their movements can be sensed by the system. Of the techniques discussed, only those which mirrored sensor views on-screen (like *StrikeAPose* [141] and the Xbox 360 “Wave to Kinect” interaction) gave users information which could help them find where to gesture. As discussed earlier in this thesis, limited-display devices often lack the capabilities to give users this feedback. Other feedback approaches are needed to help users of these devices discover where their gestures will be recognised.

This thesis addresses gaps in the literature discussed in this section in two ways: (1) it studies a way of helping users find where to address a gesture-sensing system; and (2) it investigates address techniques which work with more than one gesture-sensing system. This research also considers how limited-display devices can achieve these aims, using alternatives to on-screen feedback. The following research questions of this thesis address these gaps:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?
>
> **RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?
>
> **RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

Note that **RQ3–5** appear before **RQ1&2**, which will be introduced later in this chapter; this is because the research questions are numbered in the order in which they will be addressed in the remainder of the thesis. Although **RQ3–5** motivate **RQ1&2**, it is necessary to provide answers to **RQ1&2** before the remaining questions can be investigated.

## 2.5 Gestures with Limited-Display Devices

This section discusses gesture interaction with two types of limited-display device: simple household appliances and smartphones. It looks at each of these separately, discussing why there is growing interest in gesture interaction with such devices and identifying where more research is needed to improve their usability.

### 2.5.1 Simple Household Appliances

New home technologies and the greater availability of "smart" appliances have led to increased research interest in interaction with household devices. These devices give their users greater control over their homes, typically through dedicated controllers, smartphone applications and automation services. Two types of household appliance in particular have seen great interest: thermostats and light fittings. Intelligent, connected thermostats, such as the *Nest Thermostat*[2] (Figure 2.3), became popular as they offered greater efficiency and energy savings, by learning from homeowners' habits and allowing them greater control. Interactive household lighting has also grown in popularity, through products such as the *Philips Hue* range[3] (Figure 2.4). Users can interact with these light sources — which include ordinary light fittings, desk lamps and strips of lights — to control the lighting and ambience within their homes.

This section of the literature review looks at interaction with such systems. It focuses on simple appliances — like light switches, thermostats, window blinds — rather than more complex ones, like televisions and videogame systems. Such systems benefit from having greater display capabilities, which allows them to give users rich visual feedback during interaction. Simple appliances, in contrast, cannot give such detailed feedback. Garzotto and Valoriani [36] noted a similar problem when designing gesture interfaces for household appliances. They called using gestures to interact with such devices "gestures in the small", referring to their small screen sizes.

This review begins with an overview of why in-air gestures are ideal for interacting with simple home appliances. It compares gestures to a common alternative — using smartphones


**Figure 2.3:** *Nest Thermostat* is a 'smart' thermostat which users can program and interact with using their smartphones[4].

[2]Nest Thermostat: `www.nest.com/thermostat` Accessed 07/05/15
[3]Philips Hue: `www.meethue.com` Accessed 07/05/15
[4]© Nest. Image from `www.nest.com/uk/press` Accessed 29/05/15
[5]© Philips. Image from *Philips Newscenter*, free for editorial use.

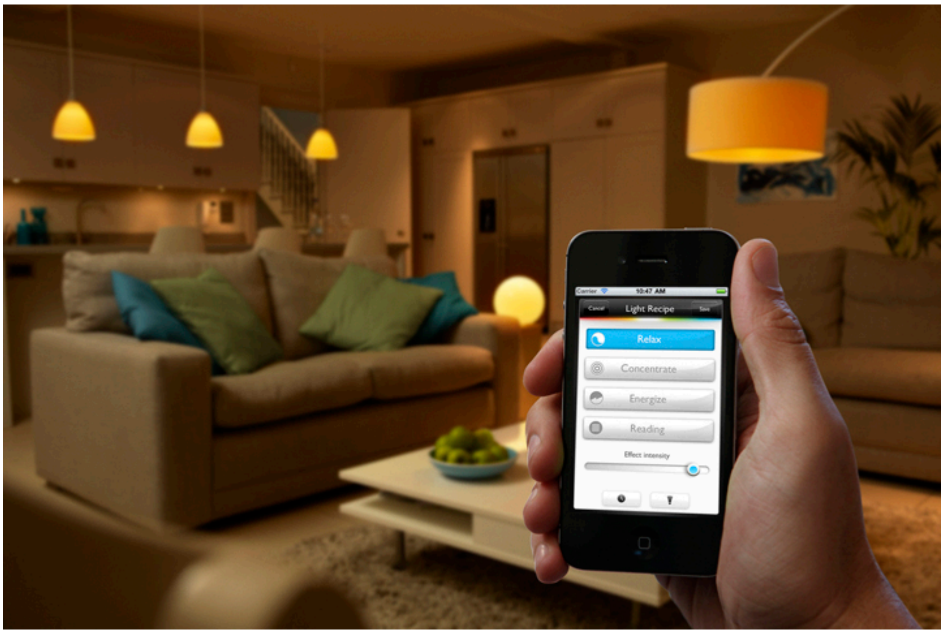



**Figure 2.4:** *Philips Hue* lighting gives users control over their household lighting, through their smartphones and through dedicated controllers[5].

as remote controls — and investigates what others have learned about using gestures in this context. As will be shown, users need effective ways of addressing these systems and they need more ways to give feedback to their users.

### Why In-Air Gestures?

Mobile phones have emerged as the primary means of interacting with simple household appliances. People typically have their mobile phones with them, or near them, making them a readily available and accessible interface. Although some initial effort is required — retrieving the phone, unlocking it, and then launching an application — users seem willing to interact using them as it means not having to approach the physical controls of an appliance.

Greater availability means mobile phones allow quick access to smart appliances. Koskela *et al.* [71] performed a long-term ethnographic study of three user interfaces within a smart home environment, comparing mobile phones to PC and television-based interfaces. They found two main activity patterns for interacting with household devices using these three user interfaces: *pattern control*, where tasks are planned and automated; and *instant control*, where users want to control a device immediately. They found that mobile phones were used most often and were particularly effective for instant control, since they were more easily accessed than the other interfaces. Their findings show the importance of allowing immediate control over a household device, which in-air gestures can also provide. Gestures are also available when mobile phones may not be or when interaction with a mobile phone is inconvenient — when cooking, for example.

Offermans *et al.* [99] looked at interaction with new lighting technologies and focused on understanding homeowners' motivations for interacting with lighting controls. They found that control location and availability had a particularly large influence on willingness to interact, with users more motivated to adjust their lighting when controls were easily accessible. They noted that mobile interfaces — remote controls and smartphone applications — have disadvantages, such as higher effort and inconsistent availability. In-air gestures could be a more accessible input modality, because users do not first have to locate and pick up a device before giving input. There are still challenges — such as addressing the system correctly — which must be overcome for gestures to be useful, however. The following section discusses two examples of in-air gesture interfaces for lighting controls.

The findings discussed by Offermans *et al.* [99] about lighting will also apply to other household appliances, such as temperature controls and music systems. They found that simply having to approach a control to use it dissuaded people from interacting with it. Koskela *et al.* [71] described similar needs for "instant control" over household appliances. In-air gesture interaction could let users interact with devices from a short distance away, providing them with immediate and spontaneous control. Gestures may not always be appropriate, such as when fine-grained control is required, for example; however, when they are suitable for the task users wish to complete, they could allow convenient interaction. With appropriate sensors, users could gesture to interact from wherever they are in the home, whether unoccupied on the couch or engaged in other activities, like preparing a meal.

Finally, users may also prefer gestures to other modes of input in certain circumstances. Valkkynen *et al.* [136] found that users preferred to interact with objects using pointing gestures, rather than touching them, when more than a step away. Even when on their feet or walking past objects, some users would rather gesture than go out of their way to touch them. Rukzio *et al.* [119] also reported preference of distal interactions instead of approaching objects to interact with them, although this would be dependent on the task and the user's goals.

In-air gestures also support the variable degrees of freedom which Offermans *et al.* [99] found desirable in an interactive household system. Users could use simple and imprecise gestures for low-effort needs, such as turning lights off or switching the air conditioning on. More complex and precise gestures could be used for high-effort needs, like adjusting brightness of lights or adjusting thermostat temperature. Using gestures would also give users a more direct way of interacting with appliances; users could gesture while focusing on the devices they are controlling, rather than abstractions of them within a remote control or smartphone [120].

Gestures would not replace existing ways of interacting with household devices; instead, they would offer an alternative means of control which users may find more convenient.

Not all appliance functionality would be appropriate for a gesture interface, however. More complex needs like programming a thermostat or mapping notifications to interactive lights — what Koskela *et al.* [71] called "pattern control" — would be better suited for other user interfaces, such as smartphone applications.

### Gesture Interaction with the Home

In-air gestures are an appealing alternative to other input modalities for controlling simple household appliances and many research prototypes have demonstrated the feasibility of using gestures to interact with devices in the home. *Gesture Pendant* [127] was a wearable gesture-sensing device which allowed users to interact with appliances through in-air gestures. Its wearers could gesture in front of their body to adjust lighting levels or music volume, for example. The pendant was created as an always-available and easily accessible alternative to traditional household controls, which elderly or disabled people may have had difficulty using. Some of their motivations for using gestures, such as easy control from a distance, are also desirable for users of all abilities, however.

Others have also used in-air gestures for interacting with lighting systems. Djajadiningrat *et al*. [31] presented a gesture-controlled wake-up light, *Grace*, which users could interact with using mid-air gestures and on-device touch input. By moving their hands up or down, or from left to right, users could adjust *Grace*'s brightness and alarm volume, respectively. Sørensen *et al.* [126] also used in-air gestures as a means of controlling interactive lighting, this time light fittings over a dinner table. Users could reach out over the table and "grab" the light to manipulate its appearance. Through gestures, users could enlarge or shrink the light beam and move its location on the table, as well as adding and removing new light sources. They were motivated to use in-air gestures to interact with lighting as gestures are more readily available than smartphone input and they may also be more socially acceptable, as others in the room can see users perform actions which cause changes in the environment.

As interactive lighting becomes more pervasive, users may have to deal with increasingly complex lighting scenes. Delamare *et al.* [28] used pointing gestures combined with wrist orientation as a way of selecting one light source from many. Their technique allowed precise selection of a single light through gesture interaction, letting users adjust interactive lighting remotely. They used the lights themselves for giving feedback; lights displayed a medium brightness when they were in the roughly selected area and the currently selected light displayed its maximum brightness. While appropriate for lighting systems, other types of feedback need to be considered for other appliances.

In-air gestures could be especially useful in the kitchen, where users are often engaged in messy, hands-on activities. *Kinect in the Kitchen* [103] was an in-air gesture system for the kitchen which let users follow recipes, set timers and control music playback. In-air gestures were chosen because they let users interact with technology when their hands were messy and while they were engaged in other cooking tasks. They also explored the possibility of using other body parts to gesture, when hands were unavailable. In-the-wild evaluation found that users were generally successful when using gestures while cooking. However, most users had problems with accidental input, as hand movements during cooking activities were interpreted as input. Panger concludes that while *Kinect in the Kitchen* showed the potential of in-air gestures within the home, users would need plenty of feedback and ways of avoiding unintentional input; that is, they need ways of addressing the gesture system.

Neßelrath *et al*. [94] also considered using gesture interaction in the kitchen. They presented two concepts which could be used to make gesture vocabularies easier to learn: inferring which devices users want to interact with, allowing gestures to be reused across devices; and allowing users to control many devices at once through shortcut gestures. Although they implemented a simple prototype of their concepts, they noted that an unsolved challenge was determining contexts for gestures — that is, which device users wanted to *address*.

#### Summary of Simple Household Appliances

Most of the research discussed in this section focused on the technical challenges of using gestures for interacting with household appliances. Little attention was given to the usability of these interfaces, although some of the problems discussed earlier in the review were identified. Users need ways of *addressing* gesture interfaces: *Kinect in the Kitchen* [103] reported unintentional gestures as users performed cooking activities; and Neßelrath *et al.* [94] noted that establishing which appliance users gestured at would allow gestures to be reused, although determining which interface was being addressed was an unsolved problem.

Users need good feedback about their gestures; however, most of the interfaces described only gave *functional feedback* [143] — users observed the effects of their gestures through their effects, such as lights turning off or the sound of an air-conditioning system coming on. As discussed in Section 2.3, functional feedback is not always appropriate, as not all systems produce an immediately noticeable effect. One of the reasons these simple household appliances did not give users other feedback about their gestures was that they lacked the means to do so — they are limited-display devices and other output modalities have not been considered.

### 2.5.2 Smartphones

Gesture interaction with mobile phones, often called *around-device interaction* in the mobile human-computer interaction literature, has received a lot of research interest in recent years. Smartphones have a variety of sensors available which can be used to sense many types of input, including in-air gesture input. This section of the literature review discusses in-air gesture interaction with mobile devices. It begins by looking at the benefits of using gestures to interact with smartphones, before looking how these interfaces have been used in commercial products and in research.

### Why In-Air Gestures?

Smartphones are limited-display devices because of their small screens. With touchscreens being the dominant way of interacting with smartphones and other mobile devices, these small screens also restrict how users provide input. Small targets can be difficult to select using touch, especially when fingers occlude the screen during input. In-air gestures, however, can be performed in the wider space surrounding a device, which prevents occlusion of the screen and allows users to make more precise movements within a larger space. *SideSight* [23] used proximity sensors to detect finger movements in the space beside a smartphone, allowing users to point and make selections without occluding targets. As users moved their fingers, a small cursor on the screen showed where they were pointing, allowing precise selections. *Abracadabra* [43] was a similar interaction technique for smart-watches, although it used a magnetic sensor to track a magnetic object for input. As users pointed in the space above the watch-face, a cursor showed what they were pointing at. *AD-Binning* [46] used similar in-air pointing gestures for information storage and retrieval, allowing users to access content stored in "bins" surrounding a smartphone.

Gesture sensing could also be used to enhance other types of interactions with smartphones. *Around-device devices* [109] explored using tangible interactions with objects near a device, to allow precise input which leveraged the affordances of those other objects; for example, using a coffee mug as a volume dial. *Air+Touch* [27], *MagPen* [60] and *BeyondTouch* [157] used gesture sensing to enhance touchscreen interactions with smartphones. In *Air+Touch*, gestures over the touchscreen were used before, during and after touchscreen input to allow easy access to functionality which would be more complex to use with touch alone. *MagPen* combined touchscreen stylus input with in-air stylus gestures, allowing users to quickly access extra functionality through gestures. In *BeyondTouch*, gestures were used for input beside the device (like in *SideSight* [23]), giving users an alternative input style.

In-air gestures can also allow users to interact when other forms of touch input are unavailable or inconvenient. Users could gesture at their smartphones from across the room for quick input, like with *Surround-See* [155], meaning they do not have to first approach their device. Gestures could also be used for *casual* interactions when users are disinterested and do not wish to engage more; for example, waving to dismiss an unwanted phone call when busy with another task [108]. Users may wish to use gestures when their hands are wet or messy; when in the kitchen, for example. In such situations, hand movements near the device avoid getting the screen messy and allow quick input; users could gesture to move to the next step of a recipe, for example.

### Gesture Interaction with Smartphones

Using gestures to interact with smartphones offers many benefits, as discussed in the previous section: they avoid occlusion and allow precise control within a larger space; they can extend touch interactions and allow convenient access to functionality; they can be used from a distance; and they can be used when other types of interaction are less convenient or unavailable. Smartphone manufacturers have now started exploring how to use in-air gesture sensing to provide their users with these benefits. Nokia introduced a "peek" gesture to many of their *Lumia* smartphones[6], which allowed users to check for notifications by holding their hand above the device. Samsung introduced "air gestures" and "air view" to their *Galaxy* smartphones, beginning with the *Galaxy S4*[7]. These features let users interact without touching their phones and provided fast access to extra functionality which was unavailable using touch. For example, users could hover their fingers over a photograph thumbnail to see a larger preview of it without selecting it from the gallery application. Google have also been exploring ways of enhancing smartphone sensing with *Project Tango*[8]. Their hardware prototypes feature a depth-sensing camera and powerful computer-vision processors, which could be used for sensing in-air gestures.

The interaction techniques in these products are in their infancy so little is known about how usable they are. Users only receive functional feedback when using the Nokia and Samsung gesture interactions, so users may experience the usability problems discussed in Section 2.3. Not much research has looked at the usability of smartphone in-air gesture interfaces, either, with most attention being given to developing sensing techniques. While some have studied aspects of usability, no research considers how smartphone in-air gesture interfaces can overcome the sensing problems discussed throughout this thesis. Ahlström *et al.* [1] investigated gesture usability; however, they focused on the social acceptability of performing gestures near mobile phones. Their work presents recommendations for designing socially-acceptable interaction techniques, although does not consider how to create usable ones. Hincapié-Ramos *et al.* [49] also looked at gesture usability, although they focused on understanding how to minimise arm-fatigue during interaction. Hasan *et al.* [46] studied ergonomic aspects of in-air gesture performance and give suggestions for using the space around devices effectively. They presented *AD-Binning*, an interaction technique for storing and accessing information through pointing gestures. They discussed the importance

[6] Microsoft Lumia: `www.microsoft.com/en-gb/mobile/phones/lumia` Accessed 12/05/15
[7] Samsung Galaxy S4: `www.samsung.com/global/microsite/galaxys4` Accessed 12/05/15
[8] Project Tango: `www.google.com/atap/project-tango` Accessed 12/05/15

of giving visual feedback during input, although only considered feedback which shows the effects of gestures.

Others have also highlighted the importance of giving feedback from smartphone gesture interfaces. Kratz *et al.* [73] argued that since there is no direct manipulation metaphor when using gestures, users need feedback to help them interact "more effectively" [73, p7]. They gave suggestions for how their *HoverFlow* technique may help users understand how their hand movements are being sensed, although this was unstudied future work. Jones *et al.* [63] suggested that feedback could help users gesture and would inform them of tracking errors, however they did not look at these issues in more detail. Their work also noted the difficulties of giving visual feedback on limited-display devices, saying that this would "[take] away valuable screen real-estate" [63, p92].

#### Summary of Smartphones

Most research in this area has focused on the technical challenges of sensing gestures using smartphones, with little attention paid to the usability of in-air gestures with such small devices. These gesture systems have also been studied in isolation, meaning the practical issue of addressing them has not been investigated. Understanding how to address them and improve their ease of use is timely as the technology has already been introduced to commercial products. Research is needed to develop solutions to the usability problems discussed in Section 2.3, so that users can confidently and effectively gesture to interact with their smartphones.

### 2.5.3 Summary of Gestures with Limited-Display Devices

This section of the literature review discussed gesture interaction with two types of limited-display device: simple household appliances and mobile phones. Simple appliances, like light switches and thermostats, have small displays or no display on which to give users visual feedback. Mobile phones do have screens, although their small size means that feedback may not be perceived from a distance and it takes away already-limited space available for content. Other output modalities could be used for feedback, overcoming these issues. The next section of this review considers three alternative ways for presenting feedback in gesture-sensing systems.

## 2.6 In-Air Gesture Feedback

Users need feedback when interacting with gesture-sensing systems, as has been discussed throughout this thesis. What little research has considered the usability of addressing and using gesture-sensing systems has largely focused on giving users visual feedback. However, many devices which are being enriched with gesture interfaces lack the display capabilities necessary to give users plentiful and useful visual feedback. This section of the literature review considers three other ways in which limited-display devices can communicate with their users: (1) using audio feedback; (2) using tactile feedback; and (3) using interactive light feedback. Each of these feedback types is now discussed in turn.

### 2.6.1 Audio Feedback

Audio feedback can be given with, or instead of, visual feedback in gesture systems. It does not require visual attention so can help users gesture while they focus on other activities or when a visual display is not available. Audio can also be used to convey additional information during interaction, reducing the complexity of visual feedback. This section discusses a variety of ways in which sound has been used to give feedback during gesture interactions, including abstract sonification of sensor data, speech feedback about selections, and musical techniques affected by hand and body movements. Although some of these feedback techniques were designed for touch or motion gestures while holding mobile devices, they are discussed here as they could also be used for in-air hand gestures. There is limited research exploring audio feedback techniques for in-air gestures, although there has been growing interest due to the availability of new commodity gesture sensors.

#### Speech Audio Feedback

*earPod* [158] combined a circular touch menu with spatial audio feedback to allow eyes-free menu selection using touch gestures. As users explored the *earPod* menu, selected items were spoken aloud. When moving to a new target, current feedback was interrupted, a short click sound was played, and then the newly selected item name was spoken. Once users confirmed their selection, a camera shutter sound was given. *Nenya* [6], a ring which could be rotated around the finger to select from circular menus, also spoke selected item names aloud. Similar speech feedback was used by *Imaginary Phone* [41], where users selected menu items by touching the palm of their hand with their other index finger. By memorising the layout of their smartphone menus, users could gesture with *Imaginary Phone* to access applications non-visually, without having to take their phone out of their pockets. Upon tapping their palm, users heard the name of the selected application read aloud.

Kajastila and Lokki [65] used spatial speech feedback in their in-air circular menus. Users selected items positioned in a flat circle centered around their hand, when held in front of their body. As users explored the menu, they received visual feedback on a screen or audio feedback through headphones, where selected item names were spoken aloud. Headphones allowed spatial cues positioned around the head, allowing users to hear where their hand was positioned, relative to the centre of the circular menu. They found that gesture performance with auditory circular menus was comparable to visual circular menus, with some users favouring audio because it did not require visual attention. These examples used speech feedback to give information about the effects of gestures, although gave little information about the gestures themselves. In the absence of feedback, users would not know if it was because they could not be sensed (not interacting in the right place, for example) or because their gestures were not recognised.

Oh *et al.* [100] compared two audio techniques for teaching visually impaired users how to perform touchscreen gestures. *Corrective verbal feedback* described gestures to users and then gave suggestions on how they could perform them more accurately. *Gesture sonification* represented finger movement through sound, allowing users to hear how their gesture performance compares to the reference gesture. They found that pitch (y-axis) and stereo panning (x-axis) were the most effective combination of non-speech audio parameters for gesture sonification. An evaluation of the two techniques found that visually-impaired users preferred speech feedback about their gestures, although sonification was especially useful for conveying time-based characteristics, like speed and duration.

### Non-Speech Audio Feedback

Unlike speech, non-speech audio can be continuously presented and modulated in response to gestures. A number of design parameters mean that non-speech audio can be used to encode several types of information, often simultaneously. It has mostly been used to guide users towards performing a gesture correctly, often by sonifying sensor data or gesture features (as in *gesture sonification* feedback [100], discussed previously). This section of the review discusses use of non-speech audio about gestures. It focuses on informative non-speech audio, which goes beyond simple confirmation sounds (for example, a short tone to show a gesture was detected, analogous to a button “click” sound).

Williamson and Murray-Smith [148] proposed using sound to represent system ambiguity during gesture recognition. As users gestured, the probability of different gestures within the gesture vocabulary being recognised were sonified with different sounds. When system ambiguity was high (for example, once users started a gesture and it was unclear which they were performing), audio feedback sounded incoherent. As performance continued and the system became more confident about which gesture users were performing, feedback became more distinct. In earlier work [146] they presented dissonance feedback, a similar idea which used inharmonious tones for each gesture. Audio feedback sounded discordant when system ambiguity was high, becoming more distinct and pleasant sounding as the gesture recognition system became more confident about which gesture was being performed.

Audio feedback has also been used to guide users through movements by telling them when their motions are (or are not) as expected. Charbonneau *et al.* [26] used multimodal feedback to teach body poses in body-controlled video games. They used non-speech audio and vibrotactile cues to guide movements, giving positive reinforcement when a limb was in the correct pose and negative reinforcement when not. Audio and vibrotactile feedback were delivered independently to each limb, informing users which parts of their body were not in the correct position. They found that gamers performed correct poses fastest when given visual feedback, but did not prefer visual feedback any more than non-visual feedback. Their findings also suggest that audio feedback should be used for positive reinforcement whereas vibrotactile feedback should be used for negative reinforcement.

Similar types of feedback were used by Morrison-Smith and Ruiz [91] to reveal device motion gestures. They compared two techniques, Silenzio and Glissando. Silenzio gave speech feedback after gesture attempts, telling users which gesture was recognised. Glissando gave continuous audio feedback during gestures, mapping tones to desired device movements and giving users negative reinforcement when they gestured incorrectly. A different tone represented each of three accelerometer axes and changes in their properties (pitch or volume) were used to show deviations from desired movement. A variety of sonification approaches were evaluated. Additive Pitch was the most effective; it used continuous tones to indicate desired movement, with undesired movement around a particular axis causing extra tones. Their feedback design is similar to Williamson and Murray-Smith's dissonance feedback, except undesired movement produces pleasant sounding tones rather than discordant ones. In their evaluation, users strongly preferred Glissando, finding continuous feedback more helpful than speech while gesturing.

Tahiroğlu *et al.* [134] used audio feedback to guide users towards target shapes when manipulating deformable interfaces. They used two continuous audio feedback designs to guide deformations: one used real-life sounds, like cracking and twanging, while the other used musical sounds. In their prototype evaluation, users were given feedback about three target shapes, with the volume of the corresponding feedback increasing as users reached the target deformation. They found that audio feedback enhanced users experience with the interface, adding a sense of affordance which was lacking when no feedback was present. Audio feedback encouraged users to explore the capabilities of the prototype and allowed users to attribute meaning to the deformations and their feedback. In-air gesture interfaces may likewise benefit from continuous non-speech audio feedback.

These examples all used abstract audio feedback to guide users towards certain movements. However, users also need other types of information when gesturing. Morrison *et al.* [90] looked at using sound to tell users which of their body parts a gesture-sensing system could and could not "see". They compared feedback which informed users a body part was not visible to feedback which told users *which* body part was not visible. They found that identifying body parts added complexity but offered no benefit. Some participants in their study even ignored audio feedback during tasks because it was not helpful. Part of the reason their audio feedback was ineffective was that it did not convey *why* body parts could not be seen; users understood that there were sensing issues but did not know how to resolve them. Their findings suggest that audio feedback must convey sufficient information to help users, otherwise it may distract users instead of help them. With tuition and experience, users may eventually benefit from such feedback when they can interpret it more effectively.

Whereas Kajastila and Lokki [65] used spatial speech feedback as users interacted with circular menus, Park *et al*. [104] used abstract audio instead. Users selected items from a vertical circle in front of their body, moving their hands forward to confirm a selection. Some feedback designs gave discrete cues about item crossings whilst others continuously sonified position within menu items. They found that continuous feedback about hand position generally outperformed other feedback designs, especially when feedback was given about the confirmation gesture.

### Summary of Audio Feedback

Speech and non-speech audio have both been used to give users feedback about touch, deformation, motion and in-air hand gestures. Descriptive speech feedback has typically been used to overcome lack of visual output, reading item names aloud as users make selections and explore gesture-based menus. However, Oh *et al.* [100] demonstrated how speech can also be used to give users verbal feedback about how to improve their touchscreen gesture performances. Continuous non-speech audio has been used to guide movements in a variety of interaction modalities, to tell users how their movements were being interpreted and how they could adapt their movement patterns to match expected behaviour. Morrison *et al.* [90] also looked at how audio could be used to inform users what body parts gesture sensors can and can not see.

The audio feedback techniques reviewed in this section could be used by limited-display devices to give users feedback while they gesture. Audio feedback could be used to enhance visual feedback when a screen is available — when gesturing towards a mobile phone, for example — or to overcome the lack of a screen when not available. The research in this thesis will build on this body of work through its use of audio feedback about in-air gestures. This modality is readily available to a variety of limited-display devices — for example, loudspeakers are common in smartphones and many household controls — so is ideal for giving feedback. Whereas audio feedback is a more used and understood type of output in gesture-sensing systems, tactile feedback is less well studied, as the next section will discuss.

### 2.6.2 Tactile Feedback

Tactile displays have been used in human-computer interaction to communicate with users non-visually, allowing them to experience information through their sense of touch. In the human-computer interaction literature, "tactile feedback" generally means "vibrotactile feedback": information presented as vibration. This thesis, likewise, uses "tactile" as a synonym for "vibrotactile", focusing on information encoded, presented and perceived as vibration. Vibration shares many of the same dynamic properties of sound [17, 19], giving interaction designers a rich and expressive design space for tactile communication. This makes tactile feedback an ideal output modality for limited-display devices, as a complement to visual feedback or as a replacement when visual feedback is not available.

This review now discusses the literature on tactile feedback for in-air gestures. Unlike audio feedback, discussed in the previous section, there is little research on tactile feedback about gestures. Giving tactile feedback about in-air gestures is difficult, as users do not physically touch the devices they interact with. Any vibrations direct from these devices will go unnoticed. However, recent technologies mean this problem can now be overcome. This section discusses two ways in which gesture-sensing systems can present tactile feedback from a distance: non-contact, mid-air tactile displays; and wearable tactile displays, which give *distal tactile feedback* [84, 85]. As will be shown, little is known about how to use these devices for feedback from gesture-sensing systems. While others have demonstrated the feasibility of tactile feedback about gestures, more research is needed to understand how to use this output modality effectively.

#### Mid-Air Tactile Feedback

##### *Ultrasound Haptics*

Recent technologies have made mid-air perception of tactile information possible. These include using focused ultrasound to create areas of acoustic radiation pressure and using air vortex generation to create moving fields of air pressure. Using focused ultrasound to create perceivable tactile sensations was first demonstrated by Iwamoto *et al.* [62]. They used a two-dimensional array of ultrasound transducers[9] (like that shown in Figure 2.5) to focus

[9]Ultrasound haptic displays typically use 40 kHz speakers — those used in car parking sensors — because they are small, readily available, and the ultrasound strength remains strong over a moderate distance, up to

sound upon a fixed point above the array. As an ultrasound focal point came into contact with skin, almost all of the ultrasound was reflected, imparting a small area of pressure while doing so. By modulating the ultrasound at a perceptible frequency (between 200 Hz and 300 Hz [79]), sensations of vibration are created.

Later work improved the capabilities of these ultrasound haptic devices. Hoshi *et al.* [56] developed a system which could dynamically move an ultrasound focal point in three dimensions above the transducer array. Their prototype produced a focal point with 20mm diameter, offering a high resolution and flexible tactile display. A further version [47] added the ability to change the dynamic properties of the ultrasound feedback, allowing designers to vary more than just the modulation frequency (effectively allowing designers to use the same design space as Tactons [17]). Subsequent iterations of Iwamoto and Hoshi's technology focused on creating smaller hardware prototypes [54], understanding how to increase its power and range [135, 47], investigating the quality of perceived tactile sensations [55] and using ultrasound haptics with virtual screens [123, 53, 89].

These early ultrasound haptic devices had several limitations: they could only produce a single point of vibration, their spatial resolution of 20mm was less than theoretically possible, and quality of perception was affected by surrounding areas of pressure. Alexander *et al.* [2] began considering how multiple independent points of feedback could be created using ultrasound haptics. They proposed two solutions: spatial and temporal multiplexing. Spatial multiplexing dedicated separate areas of the ultrasound array to separate points, whilst temporal multiplexing rendered alternating points in sequence, creating the perceptual illusion of both appearing simultaneously.

Carter *et al.* [24] noted limitations with these approaches and presented a new method for generating multiple focal points simultaneously. As well as producing multiple focal points, their approach had a spatial resolution of 10mm and suppressed unintentional areas of pressure, which would have a negative effect on the quality of perceived tactile sensations. They paired their ultrasound haptics technology, *UltraHaptics*[10], with an acoustically-transparent projection surface and a gesture sensor, allowing users to interact with and "feel" virtual objects projected onto the display surface. A later iteration of the technology [79] was used to create volumetric haptic shapes. Users experience these haptic shapes through active exploration, feeling a cross-section of the shape outline as their hand moves through it. Their user study found that users could accurately identify most of these ultrasound haptic shapes (80% identification of five shapes). Wilson *et al.* [151] investigated perception of *UltraHaptics* in detail. They found that users could reliably perceive one ultrasound haptic pixel with resolution of 20mm and also found the technology capable of creating noticeable feelings of movement. Vo *et al.* [140] investigated active exploration of ultrasound haptic feedback,

approximately 40cm.

[10]UltraHaptics: `www.ultrahaptics.com` Accessed 08/05/15

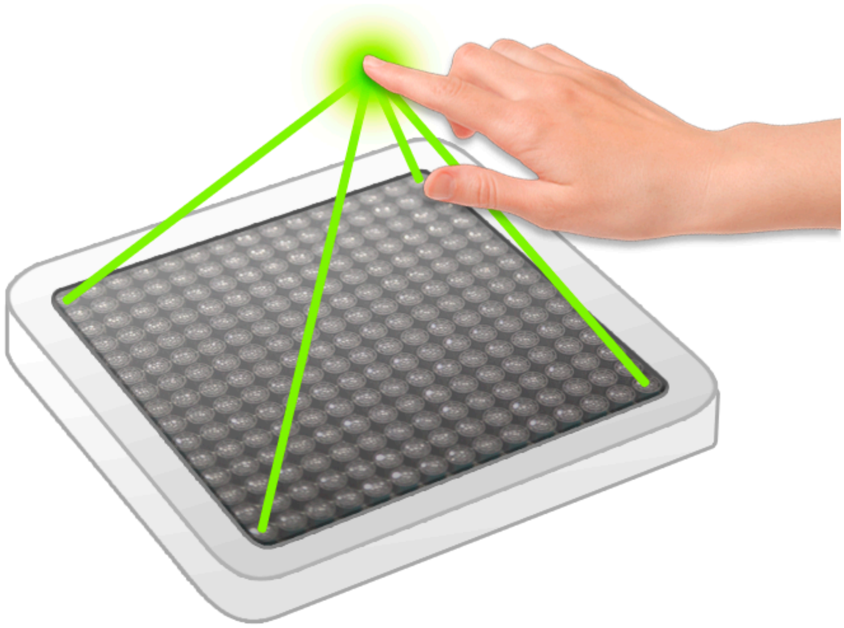


**Figure 2.5:** Ultrasound haptic displays use an array of ultrasound speakers to create tactile sensations in mid-air.

demonstrating its potential for helping users interact with in-air user interface widgets. They found that such feedback could help users locate widgets more easily than visual feedback.

### *Air Pressure*

An alternative to using ultrasound acoustic pressure for creating mid-air tactile sensations is to use air pressure instead. *AIREAL* [125] and *AirWave* [40] both investigated using air vortex generation to create moving air pressure fields for tactile communication. As a vortex — a compressed air pressure field — comes into contact with an object — a hand, for example — it imparts a small but perceptible force. Each of these systems generated an air vortex which held its shape over distances of up to 2.5 metres away, with a spatial resolution of around 85mm. A limitation of their approach is that vortexes take time to reach their target; for example, *AirWave* had an upper bound of 470ms when users were operating at its furthest distance [40]. Furthermore, air vortex generation is unable to produce continuous tactile feedback. Ultrasound, on the other hand, moves at the speed of sound (340 m/s) and can be rendered quickly enough for perceptually continuous feedback.

Air-jets have also been used to create mid-air tactile sensations. Suzuki and Kobayashi [131] developed a tangible in-air interface where users interacted with virtual objects through a hand-held paddle. A large array of air-jets were used to produce pressure against the paddle, giving users the impression that the paddle was encountering resistance from a real object. Another air-jet display [5] produced stimulus directly on the skin, although only from a distance of 5mm. Mid-air tactile feedback using air-jets is inaccurate and difficult to control [79], however, so has received little research interest.

### Distal Tactile Feedback from Wearables

While the previously discussed technologies can produce tactile experiences in mid-air, users are limited to interacting within their output range. Output range is often short, especially in the case of ultrasound haptic displays[11]. Tactile feedback from accessories worn on the body, however, allow users to interact anywhere and still experience feedback. Wearable tactile displays have been explored in human-computer interaction research, although many of the devices used are bulky which encumber users and are unlikely to be worn often. One of the advantages of gesture interaction is that it allows users spontaneous, hands-free interaction without further engagement; for example, gesturing inattentively to dismiss an unwanted notification from a mobile phone or gesturing to turn lights off when leaving a room. An ideal wearable tactile display would be always available, socially acceptable, unobtrusive when not being used, and worn where its feedback will be easily perceived.

Emerging wearable accessories like smart-watches and fitness trackers (such as those in Figure 2.6) satisfy many of these criteria. Such devices typically interface with a mobile phone and act as a secondary information display, delivering notifications and information to users' wrists. Many modern wearables use vibration to draw attention to notifications, giving simple tactile cues just like a mobile phone would. These capabilities could also be used for delivering gesture feedback. Accessories like smart-watches are likely to be worn throughout the day, meaning they would be available for giving feedback when required. Such accessories could also be used in gesture sensing, for example, their motion sensors could provide additional sensor data for gesture-sensing systems.

This review now investigates how others, motivated by these advantages, have used wearable tactile displays for feedback. First, it considers perception of vibrotactile information from such devices, which can inform feedback design. Then, it discusses examples of wearable accessories as tactile displays in human-computer interaction.

#### *Tactile Patterns on the Wrist*

While much is known about perceptual aspects of tactile stimulus on our hands and arms, little is known about how vibration is perceived on the wrist. Oakley *et al.* [98] investigated perception and localisation of vibration from a tactile display worn atop the wrist. They found that vibration is more accurately localised when there are spatial landmarks, in this case the left and right sides of the forearm. They also found that localisation is better across the wrist than along it, suggesting vibration from a watch strap could be accurately localised.

[11] Ultrasound haptic displays typically have a range of 40cm from the device, after which the strength of the tactile sensation reduces.


**Figure 2.6:** Many wearable devices, like the *Jawbone Up* activity tracker (left) and the *Moto 360* watch (right) have tactile displays for delivering vibration notifications[12].

A further study found that a dense tactile display can create more intense vibrations which still feel like a single stimulus. Their findings suggest that wearable tactile displays could create feelings of movement which feel continuous by using multiple nearby actuators in synchrony. Findings about strong localisation across the wrist are also positive for wearable tactile displays as these suggest watch-straps and bracelets could be used to deliver tactile sensations which can be reliably localised.

Matscheko *et al.* [83] investigated the information capacity of wrist-worn tactile displays. They compared two actuator placements: four placed in a circle atop the wrist under a watch-face; and four distributed around the whole wrist, using the entire wrist-band of a watch. They found that actuators placed around the wrist had greater information capacity and resulted in higher recognition of their vocabulary of eight tactile patterns.

More research was needed to understand the design space of tactile feedback from wrist-worn devices. Paneels *et al.* [102] evaluated identification of tactile patterns on a wrist-top tactile display. They found static tactile patterns difficult to identify as users could not localise feedback, instead perceiving the device vibrating as a whole; similar findings were noted in *AirTouch* [77], whose users did not notice localised feedback about their hand position. Instead, they found dynamic tactile patterns more appropriate, as patterns which changed over time were easier to perceive and identify. A limitation of their prototype device was its small size and actuator placement; they suggested that users had difficulty localising vibration because actuators were too close together.

Similar research by Lee *et al.* [76] also explored how a wrist-top tactile display could be used to present spatial and temporal tactile patterns. Using a 3x3 grid of actuators, they conducted a series of exploratory studies into tactile pattern design and delivery. They found that

[12] © Jawbone and © Motorola. Product images freely available from Motorola Press Box and Jawbone Press `www.jawbone.com/productshots`.

actuators with a small tip were more effective than actuators with a flat surface, suggesting the smaller point of contact may have been easier to perceive. They also found that repeating pulses in tactile patterns (sensory saltation) led to slightly better perception and tactile motion patterns were more easily perceived on the outsides of the wrist; the latter was also found by Oakley *et al.* [98] in earlier work.

The research described here shows the strengths and weaknesses of tactile feedback from wearable devices on the arm. Spatial landmarks, like the sides of the forearm, help users localise vibrotactile sensations, meaning multiple actuators can be used for spatial patterns. These could also be effective if delivered around the entire wrist, rather than just on one side. Static vibrotactile messages were more difficult to identify, with users generally just perceiving the presence of vibration. However, dynamic vibrotactile patterns were easier to identify as users perceived how they changed over time. Although these studies investigated the perception of tactile information from wearables, little research has studied how such information could be used as feedback about interaction. The following section reviews the literature on use of wearable tactile displays as feedback.

#### *Distal Tactile Feedback for Tabletop Interaction*

Mobile phones have been used as *ad hoc* wearable tactile displays, as they have vibrotactile output capabilities and are often near the body when not in use (in pockets, for example). McAdam *et al.* [84, 85] called this *distal tactile feedback* and they investigated its use with tabletop computers, which typically lack tactile feedback due to their large size. In two studies, they found that feedback from mobile phones placed in a trouser pocket [85] and worn on the wrist [84, 85] and upper arm [84] improved text entry performance on tabletop computers. Their work shows that tactile feedback on the body can improve human-computer interaction and the success of feedback on the wrist suggests that wearable accessories in this location may be similarly effective as mobile phones.

#### *Wearable Tactile Displays in Gesture Interaction*

*AirTouch* [77] was a gesture-sensing watch which had four proximity sensors on the watch-face and four corresponding vibration motors on its underside, arranged in a square. As users gestured in mid-air over the watch, it provided vibration feedback to show which of its sensors were tracking their hand. In their user-study, participants reported that tactile feedback was helpful, although none recognised that feedback was localised to show hand position over input sensors. Their participants found the presence of tactile feedback useful because it confirmed that sensors could see their gestures, even if they did not notice its spatial encoding of information.

Pasquero *et al.* [105] also developed a watch with a tactile display, called *Haptic Wristwatch*. They used a piezoelectric transducer that allowed them to create rich tactile cues by varying the dynamic properties of the vibrations it produced. Despite these capabilities they only used it for generating a sequence of identical pulses, contributing a demonstration of the technology.

Wearable tactile displays offer an alternative to mid-air tactile displays, allowing users to feel feedback without the constraints of gesturing in range of a fixed-position device like *UltraHaptics* [24]. As shown in this review, the wrist is a popular site for wearable tactile displays as wrist-worn devices are already widely accepted. There are also many commodity wearables — like activity trackers and smart-watches — which already have vibrotactile output capabilities, which could in future be used as tactile displays. Few have considered using wearable tactile displays for gesture feedback; those who have, focused on a technical contribution rather than the interaction design. Research is needed to understand if and how wearable tactile displays can be used effectively in gesture interfaces.

### Summary of Tactile Feedback

This section of the literature review examined two approaches for giving tactile feedback in gesture interfaces: (1) mid-air tactile displays; and (2) wearable tactile displays. Research has demonstrated the feasibility of using ultrasound pressure and air pressure to create mid-air tactile displays, allowing users to experience vibrotactile sensations without physically touching anything. These technologies each have their advantages. Ultrasound haptics offers a high resolution with low latency, which means users can experience continuous tactile sensations, even experiencing motion. Air pressure can be used over larger distances but has a lower resolution of feedback than ultrasound haptics. These differences suggest that they may be appropriate for different types of gesture-sensing system: ultrasound haptics when users are close to a device (when gesturing over a mobile phone or desktop device, for example); and air pressure when users are further away from a device (when gesturing across the room at a household appliance or television, for example).

Unlike mid-air tactile displays, wearable tactile displays require users to wear an accessory. However, many everyday devices — like activity trackers and smart-watches — have vibrotactile output capabilities and would be ideal tactile displays. Such devices are worn throughout the day, meaning they would be available as a tactile display when needed. They also allow users to gesture anywhere, rather than within range of a mid-air tactile display. While most research focuses on wrist-worn tactile displays, accessories in other locations could also be used to provide feedback. Rings are one compelling example, as finger-worn objects are common and socially acceptable; feedback from a ring would also be delivered directly to fingers as users gesture.

Research is needed to understand how mid-air tactile displays compare to wearable tactile displays, in the context of tactile feedback about gestures. Work also needs to explore other locations for wearable tactile displays, extending the literature by moving beyond devices worn on the wrist. Other body sites were effective in tabletop interaction although it is unknown if this is the case for in-air gesture interaction. Finally, research is needed to understand if, and how, these technologies can be used to create effective gesture feedback, to allow limited-display devices to communicate with users non-visually. This thesis addresses these gaps in the literature through study of the following research question:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

There was not a similar research question for audio feedback (discussed in Section 2.6.1), as that is a much better understood and more widely used type of feedback for in-air gesture interfaces. Tactile feedback, however, presents many unsolved challenges and it is unknown how recent technologies can be used by gesture-sensing systems for feedback.

### 2.6.3 Interactive Light Feedback

Limited-display devices are limited in the amount of visual feedback they can give users. If devices have a small screen, visual feedback restricts the space available for content and may not be easily perceived from a short distance as users gesture. If devices have no screen then visual feedback cannot be given. This section of the literature review introduces a novel type of feedback which allows devices to extend their visual feedback capabilities by illuminating surrounding areas: *interactive light feedback*. Small interactive light sources, like LEDs (light emitting diodes), placed around devices can be used to create a simple but easily noticeable display by illuminating nearby surfaces. Wall-mounted gesture-sensing devices could then illuminate the wall surrounding them for feedback, for example. This section of the literature review discusses the use of interactive light sources for presenting information. It begins by discussing research on the expressive capabilities of LED displays and then looks at how these have been used in gesture interfaces. While others have used LED displays as output in gesture-sensing systems, research is needed to evaluate their effectiveness and to understand how to use them effectively for gesture feedback.

#### Expressive Capabilities of LEDs

LEDs have long been used in product design as simple status lights which can communicate information through their state (on or off) and how this changes over time (flickering to show activity, for example). Many LEDs can now adjust their brightness and hue on demand, allowing them to be used for more complex and expressive information display. Research has explored how these properties of LED lights can be used to communicate information. Harrison *et al.* [42] explored the design potential of a single, fixed-colour LED whose brightness they animated over time. They found that animated changes in brightness, which they call "light behaviours", were very expressive and effectively conveyed certain meanings. Xu *et al.* [154] built on their work by also considering the use of hue in conveying information. They explored the design space of screen-less smart-watches, using LEDs on the watch-face to represent information. As little as four LEDs were able to create informative smart-watch displays, communicating rich information with users through changes in hue and brightness.

As well as having rich, expressive capabilities, LEDs are small and have low power requirements. This makes them ideal for integrating into small devices or for providing simple displays where a screen is not necessary. Some Sony *Xperia* smartphones featured an *Illumination Bar*[13] along the bottom edge of the phone (as in Figure 2.7), which used LEDs for presenting notifications or for showing visualisations during music playback. Samsung also used illumination as a secondary display, although instead of LEDs they used the curved edge of their *Galaxy S6 Edge*[14] screen to present call notifications when the phone was placed face-down on a table. Similar LED displays have also been added to "smart" household appliances. For example, the Honeywell Lyric[15] thermostat has LEDs behind its dial, which illuminate the surrounding wall when users approach it and use it.

### Interactive Light in User Interfaces

Research has extended the idea of light displays in mobile phones, by illuminating the entire space around the device instead of one side. Qin *et al.* [110] embedded LEDs in the edges of a smartphone (shown in Figure 2.7, left), creating a low-resolution extension of the screen. They presented two example uses of their technology: rendering off-screen areas of interest and showing users when gestures are available. The latter example allowed users to respond to incoming phone calls by touching the table beside the phone. Red lights illuminated the table on the left of the device and green lights illuminated the right; users could reject or accept calls by touching the left or right side of the phone, respectively. *Sparkle* [93] also used lights around a mobile device to show off-screen points of interest. Rather than place lights around the device, they embedded them in an enlarged transparent device bezel, allowing users to still see light when holding the device.

Little work has explored the use of interactive light as feedback during human-computer interaction. Qin *et al.* [110] used light as static feedforward, giving users hints as to how they can use the surface surrounding the phone for input; however, their approach did not give users feedback as they interacted. *Grace* [31], a gesture-controlled wake-up lamp, used its light for feedback to show the effects of gestures. As users moved their hands — up

[13]Sony Illumination Bar Developer API: `developer.sonymobile.com/knowledge-base/experimental-apis/illumination-bar-api` Accessed 11/05/15

[14]Samsung Galaxy S6: `www.samsung.com/uk/galaxys6` Accessed 11/05/15

[15]Honeywell Lyric: `lyric.honeywell.com` Accessed 11/05/15

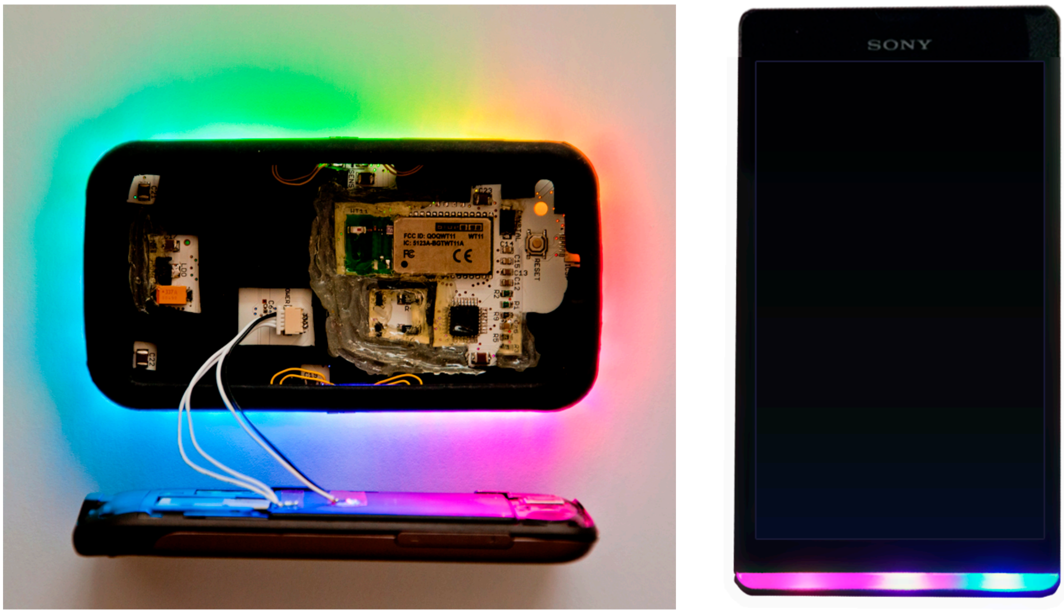



**Figure 2.7:** Lights embedded in devices can illuminate surrounding areas for feedback. Interactive lighting has been used in prototype smartphones, like that developed by Qin *et al.* [110] (left) and commercial smartphones, like the Sony *Xperia SP* (right)[16].

or down to adjust brightness, left or right to adjust alarm volume — light from the lamp reflected their movements. Its feedback and interactions were not evaluated and they only used brightness (full brightness or off completely) in their designs.

*Rainbowfish* [38], a gesture-sensing surface, used a rectangular LED display to give users feedback as they gestured. Light animations were used to show movement directions; for example, showing users which direction to swipe over the display. They also used light as feedback, glowing green, or red, to show if gestures were accepted, or not. A simple user study found that using interactive light for feedback helped novice users and provided insight into errors during interaction. However, a more thorough investigation is needed to evaluate the strengths and weaknesses of this type of visual feedback.

### Summary of Interactive Light Feedback

This section of the literature review introduced the idea of interactive light feedback: using interactive light sources for feedback from gesture-sensing systems. Like audio displays and tactile displays, discussed previously, interactive light displays can also be used by limited-display devices for output. Small devices, like mobile phones and simple household appliances, could use lights embedded in their edges to illuminate surrounding surfaces, allowing them to give visual feedback. This visual feedback does not affect content on the screen, if the device has one, and may be more easily noticed from a short distance away.

[16]Image on left provided by Qian Qin under a CC-BY license: `www.qianqin.de`. Image on right is modified from a Wikimedia Commons image, provided under a CC-BY-SA license: `commons.wikimedia.`

Despite the simplicity of such LED displays, others have found that they can be informative and expressive, communicating information through changes in brightness and hue. Interactive light displays have been found in commercial products and in research prototypes, although more research is needed to understand the effectiveness of this type of feedback and to understand how it can be used effectively in helping users to address and interact with gesture-sensing systems. This thesis addresses these research needs through the following research question:

**RQ2:** Can interactive light be used to present gesture feedback to users?

## 2.7 Summary of Literature Review

This chapter reviewed research on four topics: (1) usability problems with gesture-sensing systems; (2) interaction techniques for addressing in-air gesture interfaces; (3) gesture-sensing systems with limited visual feedback capabilities; and (4) alternative feedback types for limited-display devices. As discussed in Section 2.4, there is a need for more research into how users can address gesture-sensing systems. Limitations with existing interaction techniques and a lack of research into helping users find where to perform gestures have motivated the following research questions[17]:

**RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

**RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

**RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

Gesture-sensing systems can help their users interact by giving them feedback before, during and after interaction. However, as discussed in Section 2.5, many gesture-sensing systems are unable to give their users sufficient feedback. More research is needed to understand how such systems can support their users. Section 2.6 discussed research on three types of feedback which could be used by limited-display devices: (1) audio feedback; (2) tactile feedback; and (3) interactive light feedback. While audio feedback is better understood, research is needed to investigate the effectiveness and effective design of tactile and interactive light feedback. These three output types could be used by gesture-sensing systems to complement on-screen visual feedback, reducing the amount which needs to be given, or could

org/wiki/File:Sony_Xperia_SQ_lightbar.jpeg.

[17] Section 2.4.5 (p24) explains why the research questions appear out of order in the literature review.

be used in the absence of visual feedback, when devices have no screen or are too far to see clearly. These needs have motivated the following research questions:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

**RQ2:** Can interactive light be used to present gesture feedback to users?

This thesis now addresses these research questions, starting with **RQ1** and **RQ2**. Research in Chapters 3 and 4 will investigate how tactile feedback and interactive light feedback, respectively, can be presented by gesture-sensing systems. A better understanding of how to present and use these modalities in in-air gesture interfaces will inform their use in research discussed later in this thesis, which addresses the remaining questions.

# Chapter 3

# Tactile Feedback for Gesture-Sensing Systems

## 3.1 Introduction

Users need feedback when addressing and interacting with gesture-sensing systems. Without feedback, they have no way of knowing if they addressed the system properly, if their gestures were recognised and acted upon, and if they had their intended effects. However, many systems are unable to provide feedback effectively. Small devices with limited display capabilities need alternatives to visual feedback, the predominant type of feedback used in human-computer interaction. They may have no display on which to show information, or their display may be small, limiting the amount of information which can be given and making it difficult to notice when gesturing from a short distance away.

Tactile feedback could be given during interaction, replacing visual feedback when it is unavailable or complementing it when it is. When used together, visual and tactile feedback could be more salient than visual alone. Additional feedback may also provide extra reassurance that systems are responding and showing system attention, which would help users as they address them. However, presenting tactile information from an in-air gesture interface is a challenge, as users may not be touching a device while they gesture towards it. The first research question in this thesis, therefore, considers how gesture-sensing systems can overcome this challenge:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

This chapter describes two experiments — Experiment 1 and Experiment 2 — which investigate this question. Experiment 1 compares tactile display technologies to better understand how in-air gesture interfaces can deliver tactile cues from a distance. Experiment 2 looks at one of these approaches in particular and investigates feedback design, to see how presenting more complex feedback affects gesture performance and user experience. These experiments focus on gesture interaction with smartphones, one of the limited-display devices considered throughout this thesis. As discussed in the following section, the experiments investigate tactile feedback for in-air selection gestures, in particular.

### 3.1.1 Chapter Structure

Section 3.2 describes the design of two selection gesture interactions for mobile phones. These techniques are used in the experiments discussed in this chapter. Section 3.3 and Section 3.4 discuss Experiments 1 and 2, respectively. Section 3.5 discusses limitations of these experiments. Conclusions are given in Section 3.6, which also revisits the research question discussed earlier in the introduction.

## 3.2 Interaction Design for Experiments 1 and 2

The two experiments within this chapter investigate tactile feedback for in-air selection gestures with smartphones. Selection was chosen as the input context because it is a fundamental smartphone interaction: users select from icons, list items, hyperlinks, etc, by tapping them on the touchscreen, with sequences of selections helping them accomplish more complex tasks (like emailing a friend or finding directions to a restaurant). Selection gestures are gestures which allow users to make a selection from one of many possibilities shown on-screen.

Selection gestures have also been the focus of other gesture-sensing systems for mobile devices (like *SideSight* [23], *Abracadabra* [43], and *AD-Binning* [46], for example) so findings from these experiments could make a contribution to improving these existing interaction techniques. Selection is also a continuous and *focused* [108] interaction, requiring more engagement from users than other gesture techniques. Such interactions will benefit more from extra multimodal feedback than *casual* [108] ones, like inattentively waving over a device to dismiss interruptions. This is because focused interactions require more precise control over a longer duration, so users may be able to make greater use of information presented to them. Finally, continuous selection is also similar to the interaction techniques which will be explored later in this thesis for addressing gesture systems, so these findings will inform work discussed in later chapters. Findings from these experiments will also have implications for the design of simpler interaction techniques, however: understanding how to effectively deliver tactile feedback in gesture-sensing systems benefits all types of in-air gesture.

This section describes the design and implementation of two interaction techniques which are used in these experiments. These interactions combine in-air gestures with visual feedback, presented on a smartphone screen, and tactile feedback, presented using an ultrasound haptic display or a wearable tactile display.

### 3.2.1 Gesture Design

Two selection gestures, *Count* and *Point*, were chosen for these experiments. These gesture designs come from an earlier study [35] which collected and evaluated gestures for smartphones; both gestures performed well, with participants in that study rating them highly. Two gestures were used in these experiments, rather than just one, to see if findings about tactile feedback depended on gesture choice. This section now describes each of these gestures, before discussing the design of accompanying visual and tactile feedback.

#### Count

*Count* allows users to select from numbered targets by counting with their fingers. For example, if a user wished to select the second numbered target, they would extend two of their fingers (as in Figure 3.1). To make selections, users have to hold a counting posture for 1000ms. Dwelling for one second means users have a chance to see the effect of their gestures, giving them a chance to correct if necessary. This also simplifies the issue of knowing when to accept input from users. When developing this technique, pilot evaluations compared different dwell times (ranging from 500ms to 1500ms). Findings from these pilot studies suggested that 1000ms was ideal, giving users enough time to react to feedback but not being so long that interaction was cumbersome.

A limitation of counting using one hand is that users can only select from up to five targets. One solution would be to use both hands to extend the selection range to ten, or for more complex counting combinations like those used by Bailly *et al.* [8]. They used two hands to select from up to 25 targets, however their evaluation found that this was a mentally demanding selection technique. Requiring two hands also restricts users from interacting when one hand is unavailable; one of the benefits of gesture interaction is that it allows input when other types of interaction are inconvenient, which may be because users are holding something or are doing something else while they want to interact with their device. In this study, *Count*, instead, lets users select from groups of up to five targets, with groups being selected based on hand position, relative to the input device. Figure 3.1 illustrates how selection targets may be divided into groups and shows how a group may be selected based on palm position. Groups of selection targets are faded out when not active, so users can identify which set of targets they are selecting from.

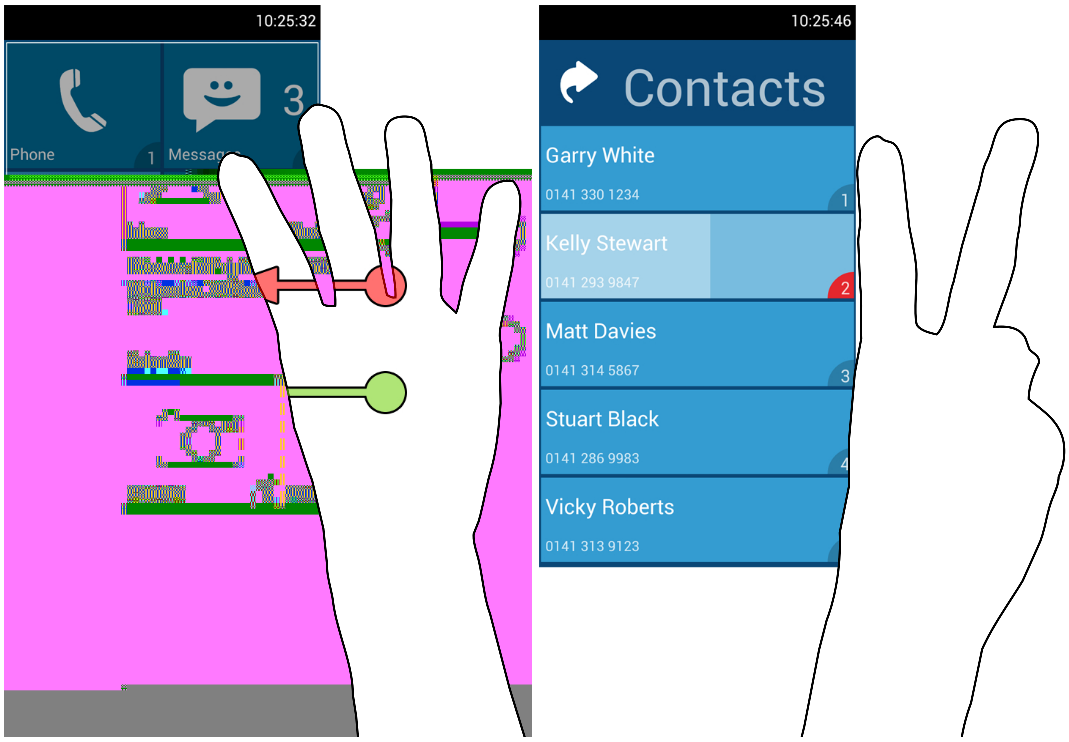



**Figure 3.1:** *Count:* users select from numbered targets by extending an appropriate number of fingers. The left image shows how palm position (green/lower arrow) determines which group of targets is active; if the palm was closer to the top half of the screen (red/upper arrow), the top four targets could be chosen from. In the right image there is only one group of targets, so users can gesture anywhere over the device.

### Point

*Point* is similar to the selection gestures used by other gesture interfaces for mobile devices, like *Abracadabra* [43] and *SideSight* [23]. Users point at targets using a virtual cursor which is mapped to the position of their extended index finger. As users move their hands around the space beside the screen, the cursor moves with their index finger. Figure 3.2 shows how cursor position is mapped to finger position within the space beside the screen. An extended finger was used because the fingertip can be reliably sensed by vision-based sensors and because users are already familiar with the idea of pointing at targets on a touchscreen; only, this time, they are pointing remotely.

As with *Count*, users can select targets by dwelling over them with the cursor for 1000ms. An alternative 'tapping' gesture was considered, allowing users to select targets without waiting; after pointing at a target, a fast downwards finger movement, as though tapping a keyboard button, confirmed the selection. However, pilot testing found that inadvertent finger movement during the 'tap' sometimes caused the wrong target to be selected. Dwelling over targets avoided this issue and also meant that this technique was similar to *Count*, allowing them to share feedback designs.

When using both gestures, users can navigate to the previous menu by swiping their hand from right-to-left over the device. A 'back' button at the top left of the screen could also be selected when using *Point* (Figure 3.2, right).

## 3.2.2 Feedback Design

### Visual Feedback

Figures 3.1 and 3.2 show the graphical user interface design for these selection techniques. Inspired by the *Windows Phone*[1] tile-based user interface, it has large rectangular targets to make selection with *Point* easier. When users are interacting with *Point*, a white, circular cursor represents the finger position in the interaction space. As users point at selection targets, the target background is highlighted and the cursor fills using a clock timer animation, showing the dwell progress (see call-out in Figure 3.2).

When users are interacting with *Count*, each target has its number in the bottom-right corner. Groups of targets are enclosed within a rectangular outline and groups are faded out when not being selected from. As users make a selection, their chosen target is highlighted. Its background fills from left to right, showing dwell progress. This animation was also tried with the *Point* gesture, although pilot participants found that the changing background distracted them from the cursor position.

[1]Windows Phone: `www.windowsphone.com` Accessed 13/05/15

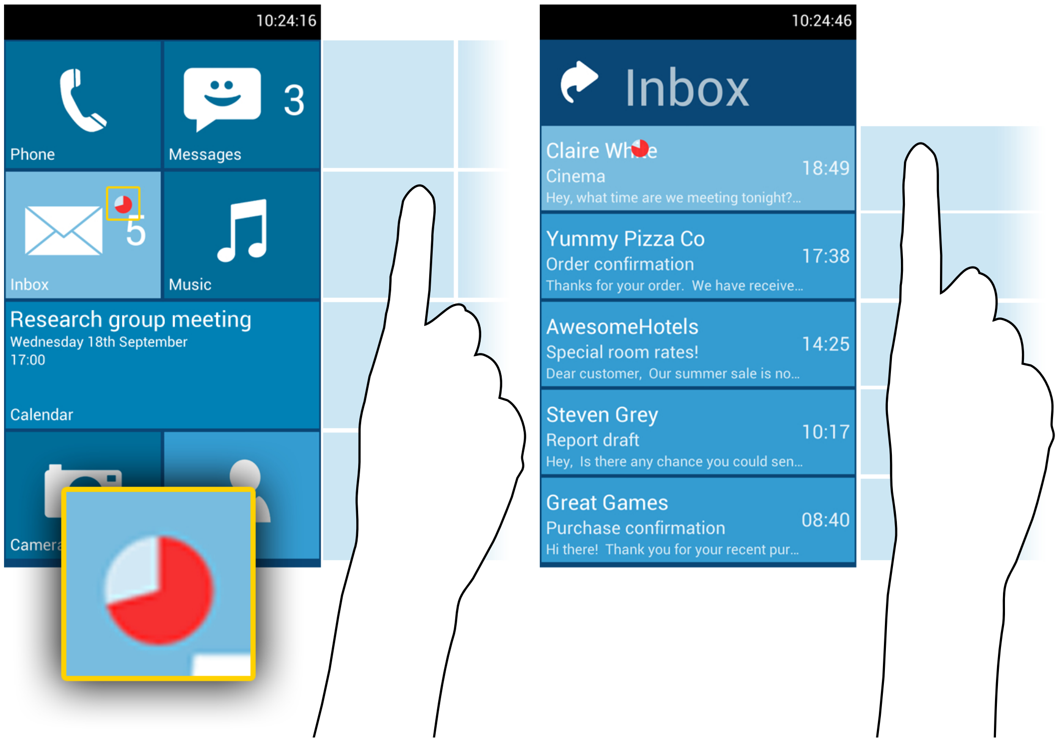



**Figure 3.2:** *Point:* a circular cursor (close-up shown in call-out on left) is mapped to finger position in the space beside the device. These images visualise how the space is divided between selection targets. Users gesture beside the screen, rather than above it, to avoid occluding targets.

### Tactile Feedback

Two types of tactile feedback were initially created for these selection gestures: *Continuous* and *Discrete*. These designs were intended for use with two types of tactile display: (1) an ultrasound haptic display; and (2) a vibrotactile actuator, which could be used in wearable accessory prototypes.

*Continuous* presented constant vibration which changed as users gestured. When selecting a target, users felt smooth vibration (a 175 Hz sine wave[2]); when gesturing but not selecting a target, users experienced a rougher sensation (a 175 Hz sine wave modulated with a 20 Hz sine wave, a technique used by Brown *et al.* [20]). No feedback was given if users were not gesturing. The aim of *Continuous* feedback was to show system attention through constant feedback, with changes in vibration reflecting user interface events. Changes in feedback let users know when they: (1) started making a selection (e.g. when moving over a button using *Point*, feedback felt smoother); (2) finished making a selection (e.g. after selection, feedback returned to feeling rough); and (3) were gesturing incorrectly, or were not being tracked (e.g. feedback stopped entirely when a hand stopped being recognised).

[2]This frequency was chosen as the vibrotactile actuator used to deliver tactile feedback, discussed in Section 3.2.4, had an optimal resonant frequency of 175 Hz.

*Discrete* feedback used short Tactons [17] (tactile icons), mapping feedback to the same user interface events that *Continuous* feedback identified. The selection start and selection complete Tactons were 150ms and 300ms smooth vibrations, respectively (both 175 Hz sine waves). The 'tracking error' Tacton was a 300ms rough vibration (a 175 Hz sine wave modulated with a 20 Hz sine wave). *Discrete* feedback aimed to communicate information about selection gestures in a less obtrusive way than *Continuous*, using short tactile messages rather than continuous vibration. New feedback designs were created for Experiment 2, based on findings from Experiment 1; these designs will be discussed in Section 3.3.5.

### 3.2.3 Tactile Displays

One of the aims of the research in this chapter was to compare different ways of presenting tactile feedback for in-air gestures. Two types of tactile display were chosen for comparison: an ultrasound haptic display and a vibrotactile actuator, which can be placed in different locations as a wearable tactile display. These technologies were chosen because they are appropriate for use with smartphones. Ultrasound haptic displays have a low latency and high resolution [24, 151], ideal for the small hand and finger movements required for the selection gestures discussed previously. Wearable tactile displays allow users to feel tactile cues anywhere on their bodies as they gesture, offering more freedom of movement than an ultrasound device. Wearable devices could also be used by many gesture-sensing systems, giving users familiar and consistent feedback from all devices they gesture with.

Tactile displays could be worn anywhere on the body, however two locations were chosen for evaluation in these experiments: the index finger and the wrist of the hand which users gesture with. Wearing objects in these locations is already commonplace and acceptable, and an increasing amount of jewellery, watches, etc., are being enriched with interactive capabilities. Many smart-watches and activity trackers, for example, have a vibrotactile display for providing notifications; these capabilities could also be used for tactile feedback. Interactive rings have received interest for their input potential (e.g. discreet eyes-free input with *Nenya* [6]) but could also be used as output devices, delivering feedback directly to the fingers which users interact with. The following section describes the prototype wearable devices used in these experiments, as well as the ultrasound haptic display.

### 3.2.4 Implementation and Apparatus

#### Gesture Sensing

A Leap Motion[3] sensor was used to track users' hands and fingers for input. The sensor field of view is 150 degrees and offers reliable finger tracking from short distance, making it more appropriate than alternative sensors, such as the Xbox Kinect[4], which are designed to track body movements from a greater distance. A gesture detector ran on a desktop computer using Leap Motion's C# library. Information about hand movements was sent to an Android smartphone via a wireless network, allowing the phone to provide visual feedback during interaction (demonstrated in Figure 3.3).

#### Ultrasound Haptic Display

A prototype ultrasound haptic display was used to provide non-contact tactile feedback. The ultrasound display (shown in Figure 3.4, top left) was the same as Wilson *et al.* [151] used in their study of the perception of ultrasound haptic feedback. It has sixty-four 40 kHz transducers arranged in an 8 x 8 grid. Each transducer has a diameter of 10mm; at 80mm x 80mm, the device is slightly wider than a smartphone. Focal points could be created on a flat plane 100mm above the display (a limitation of the experimental prototype; ideally focal points could be presented at varying heights). As the human hand cannot detect vibration at ultrasound frequencies, ultrasound was modulated at 200 Hz to create a perceivable sensation (as explained in Section 2.6.2 of the Literature Review). Modulation frequency was fixed in the prototype so it was unable to create different types of vibration (e.g. to distinguish between targeting and not targeting a selection target for *Continuous* feedback). Instead, a focal point of constant vibration (at 200 Hz) followed users' fingertips for *Continuous* feedback. This implementation effectively only showed system attention to gestures; it was unable to represent selection state through change in texture.

During development, pilot users had difficulty perceiving the ultrasound versions of Tactons used by the *Discrete* design, due to the subtle sensation produced by the ultrasound haptic display. As a result of this difficulty, the *Discrete* feedback design was not provided for the ultrasound haptic display. This was not an issue for the wearable tactile displays, as their vibrotactile actuator produced a stronger sensation.

[3] Leap Motion: `www.leapmotion.com` Accessed 05/05/15
[4] Xbox Kinect: `www.xbox.com/en-GB/Xbox360/Accessories/kinect` Accessed 05/05/15

**Figure 3.3:** A user performing the *Point* gesture. A Leap Motion sensor tracks finger movements, while an Android smartphone application provides visual feedback. A velcro ring, worn on the index finger, could have a vibrotactile actuator attached to it for tactile feedback.

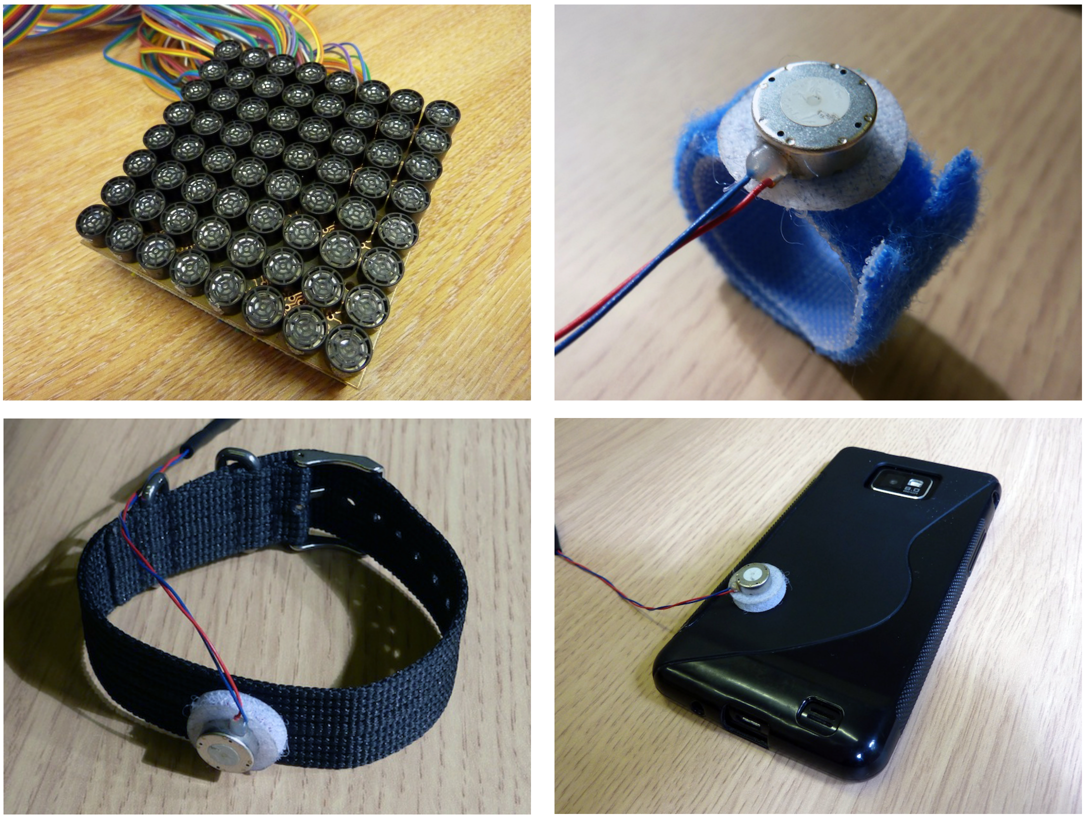


**Figure 3.4:** Tactile display prototypes used in Experiment 1 (clockwise, from top left): ultrasound haptic display; vibrotactile ring, smartphone; vibrotactile watch.

### Wearable Tactile Display Prototypes

Two types of wearable tactile display were prototyped for these experiments: a ring, worn on the index finger of the pointing hand in *Point*; and a watch, worn on the wrist of the gesturing hand (top-right and bottom-left of Figure 3.4). Participants were asked to wear the ring prototype on their index finger, as that was the one they gestured with when using *Point*. These prototypes used a Precision Microdrives[5] *C10-100* linear resonant actuator. This actuator was chosen as its small size (10mm diameter) and light weight meant the ring prototype was not cumbersome. A control signal for the vibrotactile actuator was synthesised in real-time using Pure Data[6], which generated an audio signal. A portable amplifier was used to amplify the audio output from the desktop computer before it reached the actuator, increasing the resulting strength of the vibration to an easily perceptible level.

For Experiment 1, the actuator was also attached to a smartphone (Figure 3.4, bottom right), allowing users to experience tactile feedback direct from the phone if they were holding it while gesturing towards it. This actuator was used rather than the rotational motor in the phone for consistency. It also allowed greater control over the produced vibration than is currently possible using typical smartphones, which limit developers to turning vibration on and off for set periods of time.

## 3.3 Experiment 1

### 3.3.1 Research Aims

Experiment 1 investigated how different tactile display technologies could be used to present feedback from in-air gesture-sensing systems. It aimed to evaluate the feedback designs discussed in the previous section of this chapter and compare the four tactile displays discussed previously: an ultrasound haptic display, a vibrotactile ring, a vibrotactile watch, and a smartphone. One aim of this experiment was to compare ultrasound haptic feedback to vibrotactile feedback from wearable devices. These displays have their limitations — ultrasound haptic displays have limited range and wearables do not; but wearable devices require an accessory to be worn and ultrasound haptics does not — and research is needed to understand what effects these limitations have on gesture interaction and acceptance of the technology.

Another aim of this experiment was to investigate the locus of tactile feedback. These tactile displays deliver feedback to different locations, which may or may not affect their usability and usefulness. Ultrasound haptics presents feedback directly to a fingertip or to a point on

[5]Precision Microdrives: `www.precisionmicrodrives.com` Accessed 13/05/15
[6]Pure Data: `www.puredata.info` Accessed 13/05/15

the hand during a gesture, whereas the wearable tactile displays deliver feedback to the base of a finger (using the ring) or to the wrist (using the watch). In this experiment, feedback was also given directly from the phone, held by the non-gesturing hand. Tactile feedback directly from a handheld device is familiar to many users, due to its use in touchscreen interactions, so this experiment compares this familiar feedback to distal tactile feedback.

Finally, this research aimed to compare the *Continuous* and *Discrete* tactile feedback designs, described in Section 3.2.2, to understand what types of information, and how much information, users need during interaction. These designs contrast in the amount of tactile feedback they present: *Continuous* constantly presents information, whereas *Discrete* delivers information in short messages (Tactons). Research is needed to understand how to effectively and acceptably present information with tactile displays during gesture interaction.

The three main aims of this experiment are to: (1) compare tactile display technologies to understand their effectiveness; (2) investigate effects of the locus of tactile feedback; and (3) evaluate the *Continuous* and *Discrete* feedback designs. These aims begin to contribute an answer to the first research question of this thesis:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

### 3.3.2 Experiment Design

In this experiment, participants completed selection tasks using the user interface described in Section 3.2. Only the *Point* gesture was used, to reduce the number of experimental conditions; *Point* and *Count* are compared in Experiment 2. Experimental conditions comprised combinations of tactile display (ultrasound haptics, ring, watch, phone) and feedback design (continuous and discrete), with a final condition in which no tactile feedback was given; the eight conditions are shown in Table 3.1. This was a within-subjects design, with participants experiencing all conditions.

| | Name | Tactile Display | Feedback Design |
|---|---|---|---|
| 1 | Ultrasound-Continuous (UC) | Ultrasound haptic display | Continuous |
| 2 | Ring-Continuous (RC) | Actuator attached to ring | Continuous |
| 3 | Ring-Discrete (RD) | Actuator attached to ring | Discrete |
| 4 | Watch-Continuous (WC) | Actuator attached to wrist | Continuous |
| 5 | Watch-Discrete (WD) | Actuator attached to wrist | Discrete |
| 6 | Phone-Continuous (PC) | Actuator attached to phone | Continuous |
| 7 | Phone-Discrete (PD) | Actuator attached to phone | Discrete |
| 8 | None (None) | None | None |

**Table 3.1:** Conditions in Experiment 1.

There was no Ultrasound-Discrete condition because pilot test users had difficulty perceiving Tactons presented using ultrasound haptics, as discussed in Section 3.2.4. Participants completed a block of 14 tasks for every condition and condition order was balanced using a Latin square.

### Task Design

For each task, participants had to make three consecutive selections using the *Point* gesture. Task instructions were presented on the smartphone screen and remained there until the gesture sensor detected their hand. Figure 3.5 describes an example task instruction and shows which selections would be required. Tasks used three selections as this required an active engagement with the interface and exposed participants to more tactile feedback. These tasks were also intended to be representative of the actions users may perform with their smartphones. Participants were asked to complete two types of task: selecting an action for an inbox message and selecting an action for a person in the contacts list. Of the 14 tasks in each block, seven were inbox-based tasks and seven were contacts list-based tasks. Task order was randomised for each block.

### Measures

For each task, the time taken to complete the final selection (*Time*) was measured. This measurement started when the *Point* gesture was initially recognised at the beginning of a task. Participants were asked to rest their hand on the table between tasks, which limited fatigue and meant that measurements for *Time* would be consistent.

For each completed NASA-TLX survey, an overall *Workload* value was calculated. This value was the mean of the six TLX ratings (mental demand, physical demand, temporal demand, perceived performance, effort, and frustration). *Workload* ranged from 0 to 100, where higher values mean a greater task workload.

At the end of the experiment, participants ranked feedback locations (finger, wrist, phone) and feedback design (continuous, discrete) from favourite to least favourite. Participants were also asked if they would prefer to receive tactile feedback during gestures, or not.

### Procedure

Participants were given a short tutorial at the start of the experiment session, which demonstrated how to use the *Point* interaction to make selections and gave them a chance to practice using it. No tactile feedback was presented during this part of the tutorial. Next, participants

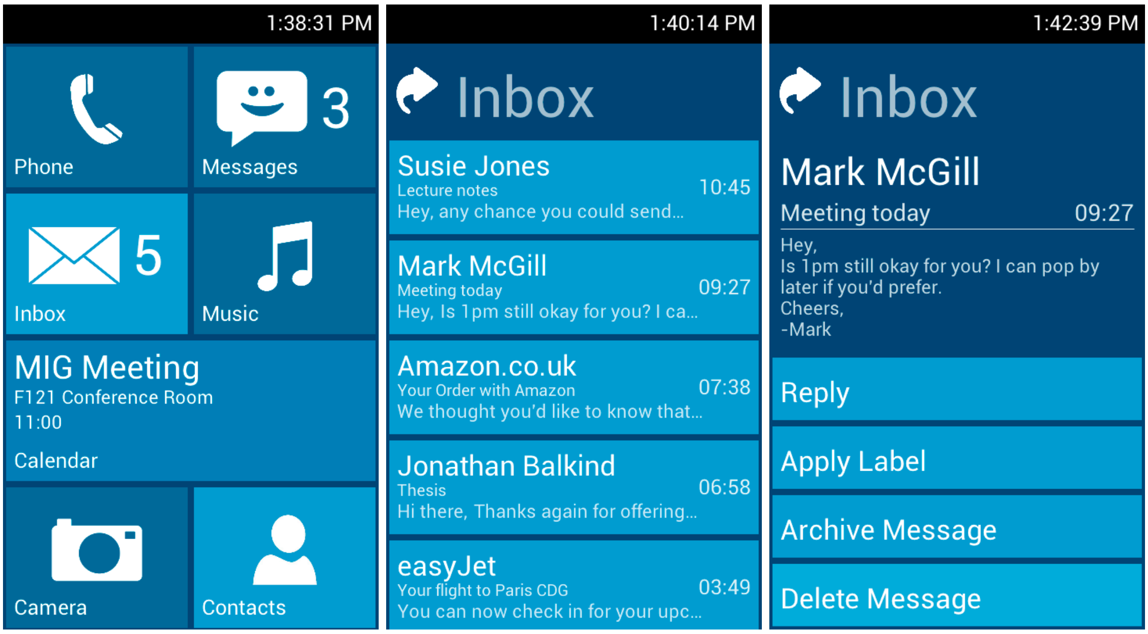



**Figure 3.5:** As an example of an experiment task, "*Reply to the second message in the inbox*" would require three selections: (1) select the Inbox item (left); (2) select the second message (middle); and (3) select the Reply item (right). An in-air gesture interface for smartphones may then allow users to speak message content aloud, for example. Participants only needed to select menu items; their actions had no effect on the experiment user interface and required no further input.

were given a demonstration of each tactile display and the types of feedback they would experience during the experiment. They were given a chance to interact with the system again, this time with tactile feedback.

During experiment tasks, participants were asked to hold the smartphone in their non-dominant hand, which they rested on a table in front of them. Although in-air gestures do not require holding a device — that is one of their benefits — this meant that holding the smartphone would not be a confounding factor for the two conditions where feedback was given directly from the device (Phone-Continuous and Phone-Discrete). Figure 3.6 shows how the apparatus used in this experiment was arranged. Right-handed participants held the device in their left hand and gestured over the ultrasound device with their right; this was reversed for left-handed participants. The ultrasound haptics device was affixed to the table, separate from the smartphone, because it had several cables that may have restricted movement if it was attached to the phone and held in hand.

After each block of tasks, participants were asked to complete a NASA-TLX (Task Load Index) [45] survey. This survey provides an estimate of task workload and is often used in human-computer interaction studies [44]. Participants were interviewed at the end of the experiment session, to better understand their preferences for feedback and what they liked and disliked about the types of feedback they experienced. During the interview, participants


**Figure 3.6:** Experiment apparatus: participants held the smartphone in their non-dominant hand beside the Leap Motion sensor and ultrasound haptic display. By gesturing over the ultrasound device, participants would experience tactile feedback during the Ultrasound-Continuous condition.

were asked to rank feedback location (Phone, Finger, Wrist) and feedback design (Continuous, Discrete). Although the Ultrasound and Ring devices delivered feedback to different parts of the finger (the fingertip and the base of the finger, respectively), these locations were combined in the interview to encourage participants to discuss feedback location rather than the way it was delivered. Participants were also asked if they would prefer to receive tactile feedback when gesturing, or not. The interview was unstructured, with preference rankings used as prompts for discussion. Interviews were recorded for later transcription and analysis.

### Hypotheses

- **H1:** *Time* will be lower for None than for the other conditions;
- **H2:** *Workload* scores will be higher for UC than the other conditions;
- **H3:** Participants will prefer tactile feedback to no tactile feedback;
- **H4:** Tactile feedback will be more preferred on the finger than elsewhere;
- **H5:** Participants will prefer the *Continuous* tactile feedback design.

**H1** anticipates participants being more effective when tactile feedback is given because that additional feedback will let them interact more confidently. Tactile feedback is also expected to make participants more aware of mistakes — like accidentally slipping off selection targets, for example — which will allow them to correct those mistakes faster. This hypothesis arose from previous experiments involving touch; for example, Brewster *et al.* [18] found that mobile text entry was improved with the addition of tactile feedback about selection. They concluded that this benefit could apply to all types of button selection on the screen so this hypothesis investigates if this may also be true for selection with in-air gestures.

Ultrasound haptic feedback is less noticeable than vibration directly on the skin so it may place more demands on users during interaction. **H2** predicts that this will be the case, with higher *Workload* scores showing those increased demands.

Tactile feedback is intended to improve in-air gesture interaction by giving users more feedback about their actions, which has been found to be beneficial for input on the screen (as discussed before). **H3** expects tactile feedback to be preferred by participants as a result of this. **H4** states that the finger will be the preferred location for feedback because it is closest to the point where users control the gesture interface (their fingertip); feedback on the wrist, in contrast, is expected to be less preferred as the feedback is not directly connected to the locus of interaction. Finally, **H5** predicts that *Continuous* will be the most preferred feedback design because it gives users more feedback during interaction, which shows continued system attention and responsiveness to movement.

#### Participants

Sixteen people took part in this study. Of these sixteen participants, five were female and three were left-handed. Participants were recruited using university email lists and were mostly undergraduate and postgraduate students. Each experiment session lasted one hour and participants were paid £6 for taking part.

### 3.3.3 Results

#### Performance

Mean *Time* was 7225ms (sd 1504ms), which includes at least 3000ms spent dwelling over selection targets. Figure 3.7 shows mean selection times for each condition. A repeated-measures ANOVA found that **Condition** had a significant effect on *Time*: $F(7, 105) = 2.99$, $p = 0.007$. *Post hoc* pairwise comparisons found that selection times were significantly higher for UC than for None ($t(106) = 3.68$, $p = 0.009$) and PC ($t(106) = 3.24$, $p = 0.03$). No other comparisons were significant. This result means the null hypothesis for **H1** cannot be rejected, as selection times were not significantly slower for the None condition.

#### Workload

Mean *Workload* was 38.9 (sd 16.7); Figure 3.8 shows mean workload for each condition. *Workload* was not normally-distributed (Shapiro-Wilk: $W = 0.96$, $p = 0.001$), so the Aligned-Rank Transform [153] was applied to workload scores prior to analysis. This approach

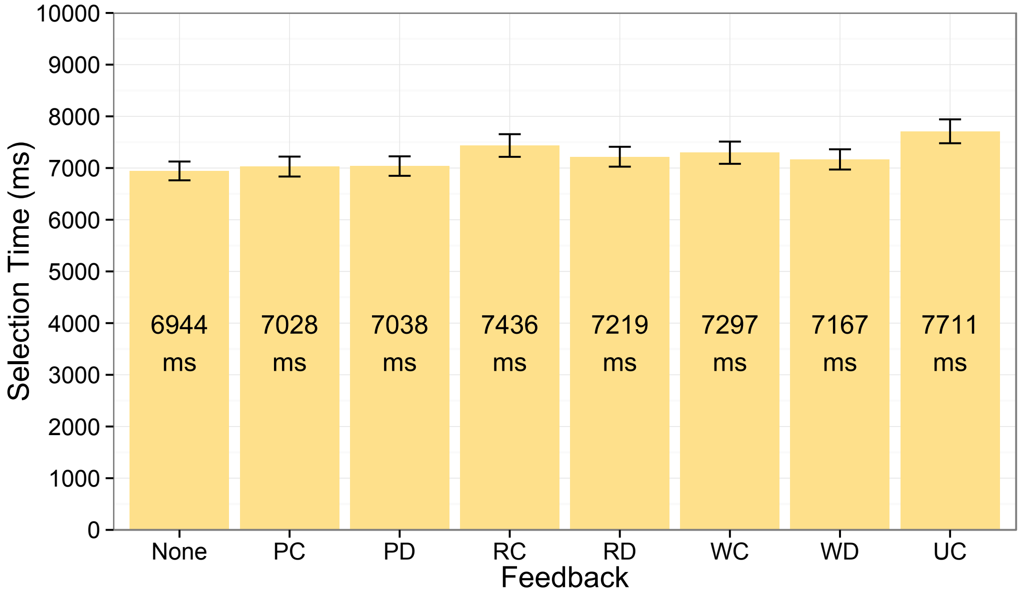



**Figure 3.7:** Mean *Time* for each **Condition**. Error bars show 95% CIs.

transforms non-parametric data into a form which can be analysed using parametric statistical tests, such as ANOVA with *post hoc* t-test comparisons [153]. A repeated-measures ANOVA on the transformed data found that **Condition** had no significant effect on *Workload*: F(7, 105) = 1.82, p = 0.09. This result fails to reject the null hypothesis for **H2**, as task workload was not higher for UC than for the other conditions.

### Preference

Fourteen participants indicated that they would prefer to receive tactile feedback, with two preferring no tactile feedback. A one-sided t-test shows that this proportion is significantly greater: t(15) = 10.25, $p < 0.001$, supporting acceptance of **H3**.

Median ranks for feedback location and design are shown in Table 3.2, where a rank of ‘1’ was the most preferred option. Friedman’s rank sum test was used to analyse ranked data for feedback locations. Ranks for **Location** were not significantly different: $\chi^2(2) = 4.1$, p = 0.13; this fails to reject the null hypothesis for **H4**, as participants did not prefer feedback on their finger.

| Location | | | Design | |
|---|---|---|---|---|
| Phone | Finger | Wrist | Continuous | Discrete |
| 1 | 2 | 2 | 1 | 2 |

**Table 3.2:** Median ranks for feedback location and design.

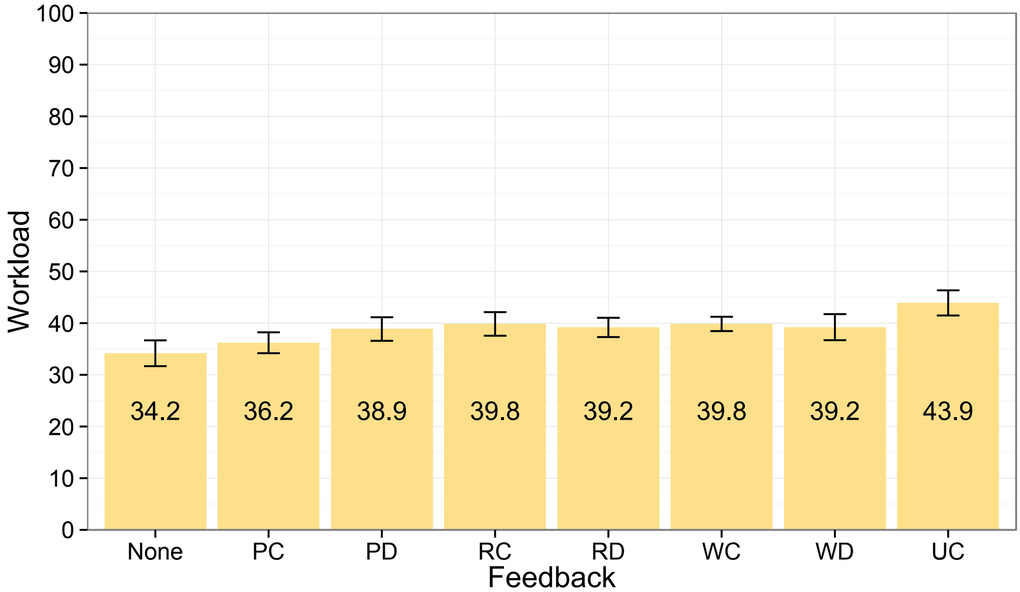



**Figure 3.8:** Mean *Workload* for each **Condition**. Error bars show 95% CIs.

Wilcoxon's signed-rank test was used to compare ranks for **Design**, finding that these were also not significantly different: Z = -0.5, p = 0.80; this result does not support **H5** being accepted, as *Continuous* was not more preferred.

### 3.3.4 Discussion

#### Benefits of Tactile Feedback

There was little difference in selection time between the experiment conditions. **H1** could not be accepted because time without tactile feedback was not worse than conditions with it. Comparing conditions where tactile feedback was given, there was only one significant difference: participants took longer with Ultrasound-Continuous than with Phone-Continuous. Workload was also comparable across conditions, meaning **H2** could not be accepted. **H2** predicted higher task workload when the ultrasound display was used for feedback, because it creates a more subtle sensation of vibration than a device in contact with the skin; it was expected that perceiving this feedback would be more demanding than the other tactile displays. While some participants said in the interview that ultrasound feedback was more difficult to notice than vibration from wearables, workload data did not reflect this.

In this experiment, tactile feedback did not affect task performance (in terms of time) or workload. However, most participants said they would prefer to receive it, supporting **H3** being accepted. This finding suggests that gesture-sensing systems may still benefit from tactile feedback, although through improved user experience. In the interviews, participants gave many reasons for liking tactile feedback. Some liked that it reduced their reliance on visual feedback by letting them "feel" that the interface was responding to them; for example, changes in *Continuous* feedback made users aware when they "slipped" off of a selection target or when there were tracking issues. Others felt that tactile feedback made the interface more engaging, as it interacted with them using more than just visual information. Getting multimodal feedback also gave users more assurance that the system was responding to their input, through increased system attention. This was most beneficial when they started each task, as it let them know they were "doing the right thing" (that is, they were addressing it correctly and their gestures were being recognised).

### Feedback Location

Preference for feedback location was divided, with no significant differences in rankings. This meant **H4** could not be accepted, as participants did not prefer tactile feedback on their fingers. **H4** expected the finger to be the preferred location for feedback because users were controlling the *Point* interaction with their fingertip, whereas the wrist and the hand holding the phone were further away from the point of control. Analysis of the interview data suggests that participants considered which hand received feedback more than they considered where on the hand received feedback. Some preferred feedback in the hand holding the phone, as this was familiar to them from their own smartphones. One participant explained that it made sense to receive feedback from the phone since he was already holding it, rather than use an alternative technology to receive feedback. Another preferred feedback from the phone because she also received visual feedback on its screen, so she found it simpler to get them both from the same device. Others preferred feedback on their gesturing hand because the feedback was giving information about their gestures and hand movements. One participant said they were focusing on giving input with their right hand, so having feedback presented to her left hand was confusing.

While participants in this experiment were divided over where tactile feedback should be given, there are situations where feedback directly from the device will not be possible. Users are more likely to interact with gesture interfaces without first approaching them and picking them up [119, 136]. In these situations, feedback given to the gesturing hand (either by wearable displays or ultrasound displays) can be perceived during interaction. There may be situations where both types of feedback are used together; for example, if a device knows it is being held, it may choose to deliver feedback to the holding hand as well as to a wearable accessory on the other wrist. Some types of information may also be more appropriate for presenting to certain hands; for example, feedback about an application error may be more appropriate from the device, whereas feedback about a recognised gesture may be more appropriate for the hand.

### Feedback for Fingers: Ultrasound vs Wearables

In the interviews, some participants said they liked ultrasound feedback because it felt connected to their use of the *Point* gesture. Ultrasound feedback was presented to the fingertip, which controlled the on-screen cursor, whereas the wearable tactile displays delivered vibration to the base of the finger or to the wrist. One participant described ultrasound feedback as being at the "correct" part of the finger, unlike the ring: the fingertip was correct because that was the point of control, discussed previously. Vibration from the ring was also described as "intrusive" and another participant felt that it interfered with their gestures. Ultrasound feedback, being more subtle, was more acceptable in these cases.

When delivering tactile feedback to fingers, ultrasound displays have advantages over wearable accessories, as discussed. However, there are limitations to using ultrasound displays which also need to be considered. Their size means that integrating them with small devices — like smartphones, in this instance — requires novel solutions to avoid changing their form factor significantly. Carter *et al.* [24] demonstrated how an ultrasound display could be placed beneath an acoustically-transparent projection surface, allowing feedback to pass through the screen. Similar approaches could be used here, positioning the ultrasound display within a smartphone rather than as an extension beside it. Smaller ultrasound haptic displays could also be used, although this would reduce the strength of the feedback. The strength of feedback in this experiment was subtle so reducing the size of the ultrasound display further would make it more difficult to perceive.

Although ultrasound feedback could be experienced without any wearable accessories, users are limited to interacting within range of the ultrasound display. Wearables, however, would allow users to receive tactile feedback over a wider range; feedback from wearables would be as noticeable up-close as it would if users were gesturing from across the room. Gesture-sensing systems, if delivering tactile feedback directly to users' fingers, therefore have to consider expected input range when selecting which tactile display technology to use. If users are to interact within close range of the device, ultrasound haptic displays are ideal. If users will be gesturing from greater distances, however, wearable tactile displays are more appropriate.

### Feedback Design

There were no significant differences in feedback design rankings, meaning **H5** could not be accepted. Participants who liked *Continuous* feedback felt it made them more aware of how the system was responding to their gestures. The presence of continuous feedback assured them that they were being sensed (that is, system attention let them know they were addressing the system properly) and subtle changes in vibration reflected changes in interface state. However, some participants thought constant vibration was obtrusive and distracting. They preferred *Discrete*, as feedback was given as short Tactons instead. Some participants thought the short messages were more useful than changes in constant vibration for conveying information, as the onset of vibration for a Tacton was easy to notice.

Ultrasound haptic feedback was more acceptable for the *Continuous* design because it produced a more subtle sensation than the wearable displays. Gesture-sensing systems could give feedback using a combination of ultrasound haptic and wearable devices, using ultrasound displays for constant feedback while wearables give more discrete information, like Tactons. Some participants suggested that a combination of the feedback designs would be more appropriate, as they found that feedback all the time was too much, but discrete feedback did not provide as much reassurance or tell them enough about the interaction.

### Summary

Experiment 1 was an initial investigation of tactile feedback for in-air gesture systems. Of the hypotheses identified in Section 3.3.2, only one was accepted:

> **H3:** Participants will prefer tactile feedback to no tactile feedback.

Participants preferred tactile feedback because it improved their awareness of how the system was responding to their actions. Little consensus was shown about how tactile feedback should be presented, however. A variety of tactile displays was used in this experiment and each has its strengths and weaknesses. Ultrasound displays are ideal for gesture-sensing systems because they do not require users to wear additional accessories; however, they limit where users can gesture and require novel solutions if they are to be integrated in small devices like smartphones. Wearable devices allow users to interact in a wider area and can deliver stronger feedback, although this feedback needs to be designed so it is not obtrusive and distracting.

Initial feedback designs were also compared in this experiment, with little agreement over which was best. Users liked that *Continuous* feedback showed system attention and kept them aware of how the interface was responding to their movements. However, some thought that feedback all of the time was obtrusive. They, instead, preferred the *Discrete* design, which communicated using short Tactons. While these Tactons were less distracting than constant vibration, they did little to reassure users that their gestures were being recognised and responded to. Experiment 2, described in the next section, investigates new tactile feedback designs informed by these findings. These refined feedback designs are discussed in the following section.

### 3.3.5 Refined Feedback Designs

A new feedback design which combined aspects of *Continuous* and *Discrete* was created, informed by suggestions from participants. Rather than present tactile information all of the time (like *Continuous*), this new design only gave constant vibration (a 175 Hz sine wave, as before) while a selection gesture was taking place. Giving constant feedback during gestures is intended to reassure users that their gestures are being recognised and responded to, one of the benefits of *Continuous* feedback found in this experiment. Tactons were presented when hands entered and left the sensor view, showing system attention when sensing begins and confirming that gesture sensing has stopped. These Tactons were 300ms-long smooth vibrations (at 175 Hz). Users receive no feedback while their hands are being tracked but they are not performing a gesture; this situation may arise while users are reading on-screen content or thinking about their next actions, for example. Feedback in such situations was considered obtrusive so this new design aims to be more acceptable.

*Continuous* and *Discrete* only presented information about gesture state: *Continuous* changed vibrotactile roughness as state changed (for example, when a selection begins); and *Discrete* indicated state change using Tactons (for example, when a gesture ends after making a selection). Vibration has several properties which can be perceived (for detailed investigation of these, see work by Brewster and Brown [17, 19]) and these could be used to communicate more information about gestures than just user interface state. For the *Point* and *Count* selection gestures, for example, vibration could also encode selection progress.

Two variations on the already described feedback design (named *Static*) were created, which dynamically encoded selection progress using properties of vibration: *Amplitude* and *Roughness*. *Amplitude* mapped the amplitude of the vibration signal to selection progress, increasing from 0% to 100% as a gesture took place. As a gesture progressed, the vibration became more intense; this change occurred over 1000ms, the time needed to complete a selection gesture. A 175 Hz sine wave was used, as with *Static*. *Roughness* modulated how smooth the vibration felt, moving from 'rough' to 'smooth' as selection progressed. This effect was achieved by modulating a 175 Hz sine wave with another sine wave, whose frequency increased from 0 Hz to 75 Hz. As the secondary frequency increases, the resulting vibration feels less 'rough'. Modulating sine waves was the approach used to change roughness in the *Continuous* feedback type and is also recommended by Brown [19]. A third property of vibration—frequency—was also considered; however, the limited frequency range of conventional vibrotactile actuators makes this property of vibration unsuitable for conveying information [20, 19].

To summarise this section, three new tactile feedback designs were created:

1. *Static*: short Tactons when sensing starts and ends, continuous vibration during a selection gesture;

2. *Amplitude*: short Tactons when sensing starts and ends, continuous vibration which increases in amplitude during a selection gesture;
3. *Roughness*: short Tactons when sensing starts and ends, continuous vibration which decreases in roughness during a selection gesture.

## 3.4 Experiment 2

### 3.4.1 Research Aims

Experiment 2 investigates how presenting extra information tactually affects the performance and usability of an in-air gesture system. Two of the refined feedback designs discussed previously (*Amplitude* and *Roughness*) encode additional information about gestures. However, presenting extra information may come at a cost. Greater feedback complexity may make it less effective or make interaction more challenging. Research is needed to understand how more complex feedback affects its use and if users find additional tactile information useful.

Experiment 2 also evaluates these designs, which try to bridge the divide between users who liked and disliked *Continuous* feedback. Participants in Experiment 1 showed divided preference for these designs, although all but two said they would prefer tactile feedback to no tactile feedback. Some liked *Continuous* as it was informative, showed continuous response to input, and let them know how their movements were being tracked. Others disliked receiving constant tactile information and instead preferred *Discrete* feedback, as they found it informative without being intrusive and distracting.

Only the *Point* gesture was used in Experiment 1, to reduce the complexity of the experiment. Experiment 2 also used the *Count* selection gesture (described in Section 3.2.1) to study what effects, if any, gesture type has on the effectiveness of feedback. Feedback would ideally be just as effective regardless of which gesture is used.

The main aims of this experiment are to: (1) investigate the effects of presenting extra information about interaction using tactile feedback; and (2) evaluate the acceptability of the refined feedback designs. These, along with knowledge from Experiment 1, will contribute towards the first research question of this thesis:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

### 3.4.2 Experiment Design

As in Experiment 1, participants completed selection tasks using the user interface described in Section 3.2. Unlike that experiment, however, both *Point* and *Count* were used. Experimental conditions comprised combinations of two within-subjects factors: **Feedback** (None, Static, Amplitude, Roughness) and **Gesture** (Point and Count). Table 3.3 shows the eight conditions in this experiment.

Participants only received tactile feedback from a wrist-based wearable tactile display (the same device used in Experiment 1). This reduced the complexity of the experiment and meant participants had greater chance to focus on the feedback itself rather than how or where it was presented. The wrist was chosen as it performed similarly to the other approaches in Experiment 1 and it allowed feedback for the *Count* gesture to be given to the whole hand, rather than an individual finger which may not always be involved in a finger-count pose.

Participants completed a block of 14 tasks for every condition and condition order was balanced using a Latin square. Task design was identical to Experiment 1; see Section 3.3.2.

### Measures

*Time* and NASA-TLX *Workload* were measured as they were in Experiment 1. Participants were asked to rank feedback type (Static and Dynamic, as discussed previously) and gestures (Point and Count), from favourite to least favourite. They were also asked if they would prefer to receive tactile feedback, or not.

### Procedure

Participants were given a short tutorial at the start of the experiment session, like in Experiment 1. This tutorial demonstrated the *Point* and *Count* gestures, introduced the types of tactile feedback they would receive, and gave them a chance to practice making selections.

All tactile feedback was delivered to the wrist in this experiment, unlike in Experiment 1 where many tactile displays were compared. Since no feedback was given directly from

| | **Name** | **Feedback Design** | **Gesture** |
|---|---|---|---|
| 1 | None-Point (NP) | No feedback | Point |
| 2 | None-Count (NC) | No feedback | Count |
| 3 | Static-Point (SP) | Constant vibration (Static) | Point |
| 4 | Static-Count (SC) | Constant vibration (Static) | Count |
| 5 | Amplitude-Point (AP) | Dynamic change in amplitude (Amplitude) | Point |
| 6 | Amplitude-Count (AC) | Dynamic change in amplitude (Amplitude) | Count |
| 7 | Roughness-Point (RP) | Dynamic change in roughness (Roughness) | Point |
| 8 | Roughness-Count (RC) | Dynamic change in roughness (Roughness) | Count |

**Table 3.3:** Conditions in Experiment 2.

the smartphone, participants were not asked to hold it in their non-dominant hand during tasks. The same wrist-based tactile display as Experiment 1 was used (shown in Figure 3.4). Gesture sensor placement was the same as the last experiment, with right-handed participants performing the *Point* gesture above and to the right of the smartphone (as in Figure 3.3). This was reversed for left-handed participants, so that they gestured on the left side of the screen using their left hand.

After each block of tasks, participants completed a NASA-TLX survey. They were also interviewed at the end of the experiment session, as in the last experiment. Preference rankings were gathered during the interview for gesture (Point and Count) and feedback type (Static and Dynamic). Rather than ask participants to rank *Amplitude* and *Roughness* separately, these were considered together as 'Dynamic' feedback. Dynamic feedback designs were grouped together so that the interview could investigate preference for feedback which conveys increased amounts of information, rather than the properties of vibration which encode that information. Another reason these were grouped together is that participants were not expected to be able to identify the different designs during the experiment. Pilot testing found that the designs were perceptually similar; participants were able to identify the change in vibration but often perceived it as an overall increase in the strength of the vibration. Interviews were unstructured, with preference rankings used as prompts for discussion. They were recorded for later transcription and analysis.

### Hypotheses

- **H1:** *Workload* will be significantly higher for Dynamic feedback than for Static.
- **H2:** Dynamic feedback will be the most preferred feedback type.

**H1** predicts that Dynamic feedback will have significantly greater *Workload* than Static feedback. This is due to their increased complexity, as both designs feature continuously changing vibration that encodes more information (i.e. selection progress) than is given by the Static design. Understanding this additional information is therefore expected to place greater demand on users. If this were the case, it is not necessarily a negative outcome, as users may find the additional information useful. The feedback is not considered to be too complex as research on vibrotactile perception has found people capable of accurately perceiving concurrent properties of vibration [21]; however, gesture interaction also places demands on users and it is not known how performing gestures affects vibrotactile perception. Experiment 1 found that tactile feedback had no effect on *Workload*, but its feedback designs were less complex than the Dynamic designs used here.

Despite the expected increase in *Workload* for Dynamic feedback, **H2** predicts that participants will prefer these designs. In Experiment 1, participants preferred receiving tactile feedback and found it useful when changes in feedback let them know what was happening with the interface. Since Dynamic feedback is designed to produce these changes in vibration, they are expected to be more preferred than the Static design.

### Participants

Sixteen people took part in this study. Six were female, three were left-handed and five participated in the previous study. Participants were recruited using university email lists and were mostly undergraduate and postgraduate students.

## 3.4.3 Results

### Performance

Mean *Time* was 8323ms (sd 2428ms); Figure 3.9 shows mean times for each type of **Feedback** and **Gesture**. A repeated-measures ANOVA found that **Feedback** had no significant effect on *Time*: $F(3, 45) = 0.34$, $p = 0.80$. **Gesture** did have a significant effect: $F(1, 60) = 64.31$, $p < 0.001$. *Post hoc* pairwise comparisons show that *Time* was significantly lower for *Point* than *Count*: $t(60) = -8.02$, $p < 0.001$. The interaction effect between **Feedback** and **Gesture** was not significant: $F(3, 60) = 0.10$, $p = 0.96$.

Five of the sixteen people who took part in this experiment also participated in Experiment 1. A repeated-measures ANOVA was used to investigate what effect, if any, their prior experience with the *Point* gesture had on task performance in this experiment. No significant effect was found: $F(1, 49) = 0.005$, $p = 0.95$.

### Workload

Mean *Workload* was 36.2 (sd 17.5); Figure 3.10 shows mean task workloads for each type of **Feedback** and **Gesture**. Workload scores were not normally-distributed (Shapiro-Wilk $W = 0.95$, $p < 0.001$), so the Aligned-Rank Transform [153] was applied to *Workload* values before further analysis. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Workload*: $F(3, 105) = 2.72$, $p = 0.04$. *Post hoc* pairwise comparisons found no significant differences between types of **Feedback**, however (all $p \geq 0.09$). This finding does not support accepting **H1**, as the Dynamic feedback conditions did not have higher *Workload* than Static.

**Gesture** also had a significant effect on *Workload*: $F(1, 105) = 50.47$, $p < 0.001$. A *post hoc* comparison of both gestures found that workload was significantly lower for *Point* than for *Count*: $t(105) = -7.1$, $p < 0.001$. The interaction effect between **Feedback** and **Gesture** was not significant: $F(3, 105) = 0.30$, $p = 0.82$.

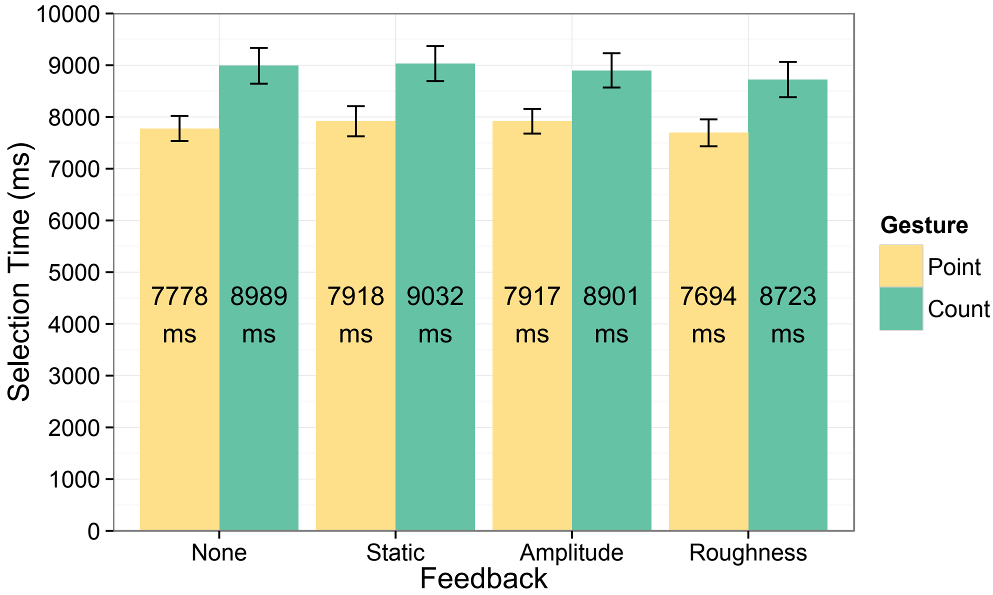



**Figure 3.9:** Mean *Time* for each **Feedback** and **Gesture**. Error bars show 95% CIs.

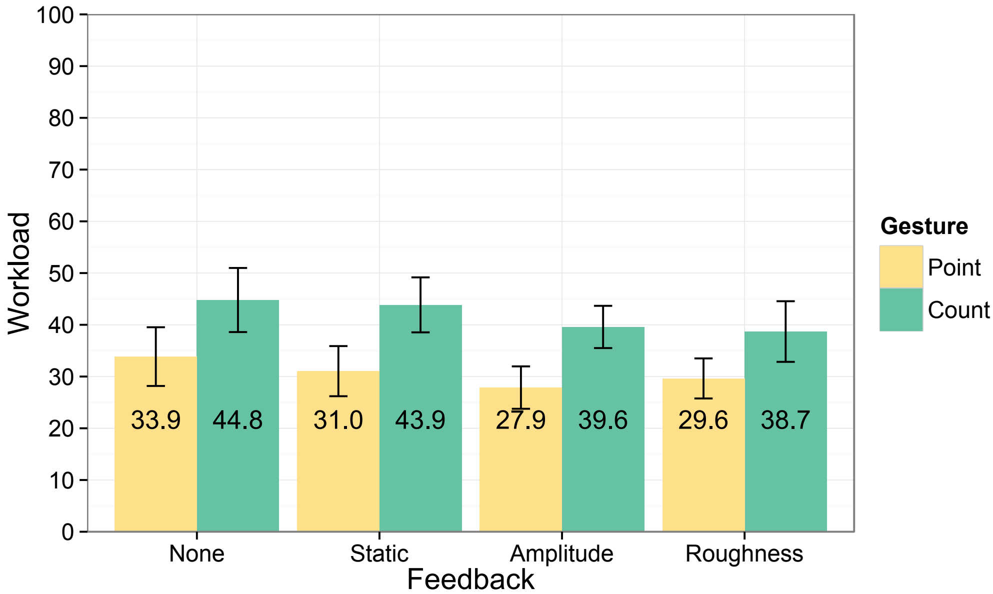



**Figure 3.10:** Mean *Workload* for each **Feedback** and **Gesture**. Error bars show 95% CIs.

A further ANOVA was used to investigate if the lower workload estimations for *Point* were influenced by the five participants with prior experience using that gesture. There was no significant difference in *Workload* between those who took part in Experiment 1 and those who did not: F(1, 14) = 0.78, p = 0.39.

#### Preference

Median ranks for feedback type and gesture are shown in Table 3.4, where a rank of '1' was the most preferred option. Wilcoxon's signed-rank test was used to compare preference rankings for **Feedback Type** and **Gesture**. Ranks for **Feedback Type** were not significantly different: Z = 1.94, p = 0.09; this result does not support accepting **H2**. Ranks for **Gesture** were significantly different: Z = -3, p = 0.004, with *Point* being more preferred than *Count*.

| **Feedback Type** | | **Gesture** | |
|---|---|---|---|
| Static | Dynamic | Point | Count |
| 2 | 1 | 1 | 2 |

**Table 3.4:** Median ranks for feedback type and gesture.

Fourteen participants indicated that they would prefer to receive tactile feedback, while two said they would prefer none. A one-sided t-test shows that this proportion is significantly greater: t(15) = 10.25, $p < 0.001$.

### 3.4.4 Comparison of Experiment 1 and Experiment 2

#### Performance

Selection time data for the *Point* gesture was also compared with data from Experiment 1. Welch's two-sample t-test was used to compare mean *Time* from Experiment 1 with Experiment 2 (excluding blocks of tasks using the *Count* gesture). This test was used because it is appropriate for comparing data from two independent samples where equal variance cannot be assumed. Welch's t-test found no significant difference in *Time*: t(23) = -2.07, p = 0.05.

#### Workload

*Workload* ratings for *Point* were compared with ratings from Experiment 1, where that gesture was also used. *Workload* was aggregated for each participant in each experiment, excluding the *Count* gesture in Experiment 2. The Mann-Whitney U test was used to compare *Workload* between experiments, as it is appropriate for comparing non-parametric data

(*Workload* was not normally-distributed) from independent samples (different sets of participants in each experiment). The test found no significant difference in *Workload* between Experiment 1 and Experiment 2: U = 170, p = 0.12.

### 3.4.5 Discussion

#### Encoding Extra Information about Gestures

One of the aims of this experiment was to investigate the effects of presenting more information about interaction using other properties of vibration. Two 'Dynamic' feedback designs encoded selection gesture progress, using changes in intensity (*Amplitude*) and roughness (*Roughness*). These designs were more complex, so it was hypothesised (**H1**) that their workload would be higher than feedback which just represented interface state (*Static*). This was not the case, however, as **Feedback** had no significant effect on *Workload*.

Presenting additional information as tactile feedback did not improve task times, although this was expected based on findings from Experiment 1. In Experiment 1, tactile feedback improved user experience — mostly through improved awareness — but did not affect performance, so similar outcomes were expected here. Interviews after this experiment revealed similar trends, with participants reporting that tactile feedback improved their awareness of how their gestures were being sensed and the effects they were having. The dynamic feedback designs were especially effective at providing this awareness, as their continuously changing vibration informed users of how their gestures were progressing. This resulted in no objective difference in performance, however. Many found this awareness useful as it reduced their reliance on visual cues about progress.

Changes — expected or unexpected — in dynamic feedback were also considered helpful. One participant described how unexpected changes in feedback informed him of problems; for example, feedback changed when he 'slipped' off targets using *Point* and when his *Count* gestures were being misinterpreted. Recognising the subtle changes in vibration allowed him to correct his gestures before making an unintended selection. This participant's experience shows one of the benefits of giving consistent feedback about interactions: users can form an expectation of what feedback their actions will create and when it is *not* as expected, it catches them off guard and helps them to resolve the interaction problem. Feedback could also be purposefully manipulated to make it unexpected, as a means of teaching users about their interactions with the system. For example, if a user's gesture is recognised but it only just matches the recognition requirements, the system could provide some unexpected feedback to make the user think about their gesture and to try something different in future. Similar ideas were discussed by Rogers and Muller [117]; they discussed how unexpected outcomes may cause users to reflect and try to understand *why* the unexpected happened, noting that this is central to how we learn about things.

Despite these benefits of presenting extra gesture information, preference rankings for feedback type were divided. This meant that **H2** was rejected, as participants did not prefer the Dynamic designs. Some participants found the changes in vibration for the Dynamic feedback difficult to perceive and reported that they focused on the presence (or absence) of vibration instead. Those participants preferred the *Static* design as it was easier to perceive. One participant felt that redundantly presenting information using vibration was unnecessary, since the visual feedback already informed her of selection progress. However, redundantly encoding information tactually may benefit interaction in other ways; for example, the previously discussed case where sensing issues became apparent through unexpected changes in vibration.

### Feedback Designs

Another aim of this experiment was to evaluate the acceptability of the tactile feedback designs, after the divided preferences observed in Experiment 1. Some participants in that experiment found continuous vibration obtrusive while others found it informative. The designs in this experiment were a compromise between the *Continuous* and *Discrete* designs, using bursts of continuous vibration during gestures, with short Tactons keeping users informed in between gestures. In the interviews, only two participants indicated that they disliked the feedback in this experiment. Both found it irritating and would prefer to receive no tactile feedback instead.

### Summary

Experiment 2 investigated tactile feedback for in-air gestures, focusing on feedback design rather than tactile display choice (which was the focus of Experiment 1). Like the first experiment, tactile feedback led to improved user awareness but had no effect on task time or gesture workload. Participants found dynamic tactile feedback designs especially useful. Not necessarily because they encoded more information, but because encoding constantly changing information meant that changes in vibration could be informative. Expected changes in vibration — for example, feedback increasing in intensity as a selection progresses — assured users that their gestures were being responded to as they expected. Unexpected changes in vibration — for example, when intensity decreased suddenly — informed users that something was wrong, like a misinterpreted gesture or accidentally switching to a new selection target.

## 3.5 Limitations

### 3.5.1 Ultrasound Haptic Display Prototype

The ultrasound haptic display used in Experiment 1 (described in Section 3.2.4) was a prototype device with limited functionality. While ultrasound haptics can be used to create points of stimulus anywhere in 3D space over the display, this prototype was restricted to creating points on a flat plane with a fixed height of 100mm. As a result, ultrasound feedback followed the fingertip position in only two dimensions. If participants gestured too far above or below this plane, the feedback would be difficult to perceive. Another limitation of this hardware was that the vibration frequency was fixed to 200 Hz, meaning the *Continuous* feedback design could not be fully implemented. Despite these limitations, ultrasound haptics performed well in Experiment 1 and participants generated lots of useful discussion comparing it to the wearable tactile displays. The same hardware has also been used successfully in perceptual research [151], where these limitations did not affect the detailed study of ultrasound haptic perception. Experiment 1 was a successful initial investigation of tactile feedback for in-air gestures and the answer given to **RQ1** is not affected by these prototype limitations.

### 3.5.2 Gesturing at a Handheld Device

One of the aims of the research in this chapter was to investigate ways for gesture-sensing systems to deliver tactile feedback, since users will not always be holding a device while they gesture towards it. However, in Experiment 1, participants held the smartphone while they gestured. They were asked to hold it so that they could feel tactile feedback in the Phone-Continuous and Phone-Discrete conditions, where vibration was given directly from the phone. They were asked to hold it during all conditions so that holding the phone would not be a confounding factor. During the experiment, participants rested their arm on a table surface so could not benefit from lifting or moving the device during interaction. In Experiment 2, feedback was no longer given from the phone, so participants gestured at it without touching it. A comparison of data from these experiments (see Section 3.4.4) found no significant differences in performance and workload measurements, suggesting that participants were not necessarily advantaged or disadvantaged by holding the phone in Experiment 1.

### 3.5.3 Feedback Locations

Whereas Experiment 1 investigated tactile feedback presented at a variety of locations, only the wrist was used for feedback in Experiment 2. This should be considered when interpreting the findings of these studies, as the *Count* gesture was only evaluated with feedback from a wrist-worn device. Using only one feedback location reduced the complexity of Experiment 2. The wrist was selected rather than the index finger because it meant feedback about the *Count* gesture would be given to the whole hand rather than to a single finger, which may not necessarily be involved in finger-counting. For the *Point* gesture in Experiment 1, there were no significant differences between feedback at the wrist and at the finger, so using the wrist was not considered a disadvantage in Experiment 2.

## 3.6 Conclusions

This chapter investigated tactile feedback for in-air gesture systems through two experiments. Experiment 1 focused on the delivery of feedback. It compared an ultrasound haptic display with wearable tactile displays and also explored different locations for tactile feedback. Participants in that experiment reported benefiting from receiving tactile feedback, despite it having no significant effects on task performance and workload. These benefits, which include an improved awareness of system response and reassurance that they were addressing the system properly, led to a majority of participants saying they would prefer to receive tactile feedback about their in-air gestures. Similar benefits were found in an evaluation of *AirTouch* [77]. There was little consensus in this experiment over where to present tactile information, with participants showing divided preferences for feedback location. Wearable devices were well received and users found their feedback useful, even when it was presented away from the point of interaction. Feedback on the wrist about finger movements was acceptable and no-one expressed difficulty understanding how that distal feedback related to their gestures.

Experiment 2 focused on evaluating new feedback designs, also investigating the effects of tactile presentation of more complex information about gestures. Users reported many of the same benefits as those in Experiment 1, with tactile feedback keeping them informed about how they were being sensed. Although this information was also available visually, its multimodal presentation was more salient and almost all participants preferred to receive it tactually than to not. More complex information presentation did not benefit task performance or workload, but participants found it useful. The information which was presented (selection progress, in this case) was not as useful as the way its encoding changed the dynamics of the vibration. These dynamic changes in feedback acted as helpful cues. Expected changes reassured users that their gestures were being recognised and having their desired effect, while unexpected changes informed them that something was wrong. A sudden, unexpected change in dynamic feedback might suggest that a gesture was misinterpreted (like detecting two fingers instead of three, for the *Count* gesture) or that users made a mistake (like moving off the intended target with the *Point* gesture).

### 3.6.1 Research Question 1

These findings suggest that tactile feedback about in-air gestures is beneficial. Gesture-sensing systems should present tactile feedback and the research presented in this chapter suggests how this may be done. The outcomes of these experiments are now summarised as recommendations in response to the following research question:

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

Ultrasound haptic displays and wearable tactile displays are two ways of delivering tactile feedback as users gesture in mid-air, allowing them to receive tactile information from a device they may not be touching. Ultrasound haptic displays are ideal for systems where users interact from a short distance away, as their range is limited but they do not require any other accessories for users to experience feedback. Wearable devices produce a stronger sensation and allow users to interact from a greater distance. These would be more appropriate when users are gesturing from further away, like waving at their phone from across the room (as in *Surround-See* [155]) or gesturing at interactive lighting controls (as in *Grace* [31]).

Feedback itself should be dynamic, as change in vibration can create rich and useful cues. This may be achieved by encoding more information about the state of the gesture-sensing system, as in Experiment 2, for example. Users benefit from feedback throughout a gesture interaction, from the moment their hands are first sensed as they address the system, to the moment they stop interacting with it. Continuous vibration throughout an entire interaction would be obtrusive, however. Instead, discrete types of information — such as Tactons [17] — should be used when continuous tactile information is not necessary. These shorter messages show system attention and keep users informed, without being obtrusive.

### 3.6.2 Contributions

The research in this chapter makes the following contributions:

- It presents a comparison of an ultrasound haptic display with wearable tactile displays for gesture feedback;
- It investigated different body locations for wearable tactile displays, finding no difference between wrist and finger for remote tactile cues about gestures;
- It found that tactile feedback improves the user experience of in-air gesture input; especially continuous, dynamic feedback which creates rich and informative cues.

# Chapter 4

# Interactive Light Feedback for Gesture-Sensing Systems

## 4.1 Introduction

Limited-display devices can illuminate the space around themselves using interactive light displays, an approach discussed in Section 2.6.3 of the Literature Review. This allows them to give visual feedback when they would be otherwise unable to. Presenting feedback in this way can benefit these devices in three ways:

- Gesture feedback does not constrain the limited screen space available for content, because it does not have to be fit on-screen;
- Users may be able to see the feedback from a greater distance, meaning they do not have to gesture close by;
- Feedback can be given from devices which have no screen.

Other research has demonstrated the potential of using interactive light sources for feedback. For example, interactive light displays have been used to extend mobile phone screens for off-screen content [93, 110] and have been used for output when screens are unavailable [31, 38]. However, little is known about how such displays should be used in gesture systems. While others have suggesting using light for feedback about gestures [110, 31], this has not been evaluated in detail. The need to better understand the use and usability of interactive light feedback motivated the following research question:

**RQ2:** Can interactive light be used to present gesture feedback to users?

If the answer to this research question is ‘yes’ and interactive light feedback is effective, then this thesis would be able to make informed use of interactive light in interaction techniques for addressing gesture systems. If the answer is 'no', then alternative feedback would have to be used instead. This chapter describes research which investigates an answer to this question. It describes the *Gesture Thermostat*, an example of a simple household control enriched with a gesture interface. The thermostat features an interactive light display which it uses to give visual feedback. Experiment 3 evaluates the thermostat and its use of interactive light feedback, contributing an answer to **RQ2** as well as a study of the effectiveness of interactive light feedback. This experiment also considers the use of audio and tactile feedback alongside interactive light, to see if these modalities can improve the experience of gesturing at an interactive light display.

### 4.1.1 Chapter Structure

Section 4.2 describes the design and implementation of the Gesture Thermostat, an example of a limited-display device which gives interactive light feedback. Section 4.3 discusses the design and outcomes of Experiment 3, which investigated the usability of interactive light feedback and the Gesture Thermostat. Section 4.4 identifies limitations of the experiment described here. Finally, Section 4.5 provides conclusions from Experiment 3 and an answer to **RQ2**.

## 4.2 Gesture Thermostat

The experiment presented in this chapter investigates the use of interactive light feedback about gestures. A gesture-sensing system, called the *Gesture Thermostat*, was created for use in this experiment. As its name suggests, the Gesture Thermostat is a household thermostat which can be controlled using in-air gestures. It is an example of a limited-display device because, like many household controls which are being enhanced with new input modalities, thermostats typically lack displays (Figure 4.1, left) or have small screens which may be difficult to see from more than a couple of metres away (Figure 4.1, right). In order for gesture input to be a feasible option for such devices, they need to be able to give good feedback. A thermostat was chosen for this study because they are common and simple devices which experiment participants may be familiar with. They are also one of the most successful 'smart' devices being increasingly introduced into homes, as discussed in the Literature Review.

The rest of this section now describes the design and implementation of the Gesture Thermostat. It begins with an overview of the hardware design, describing its interactive light display and its output capabilities, then finishes with a description of the gestures it supports and the feedback given about them.


**Figure 4.1:** Honeywell RLV210A (left), an example of a screen-less thermostat control; and the Honeywell Lyric (right), an example of a thermostat with a small screen.

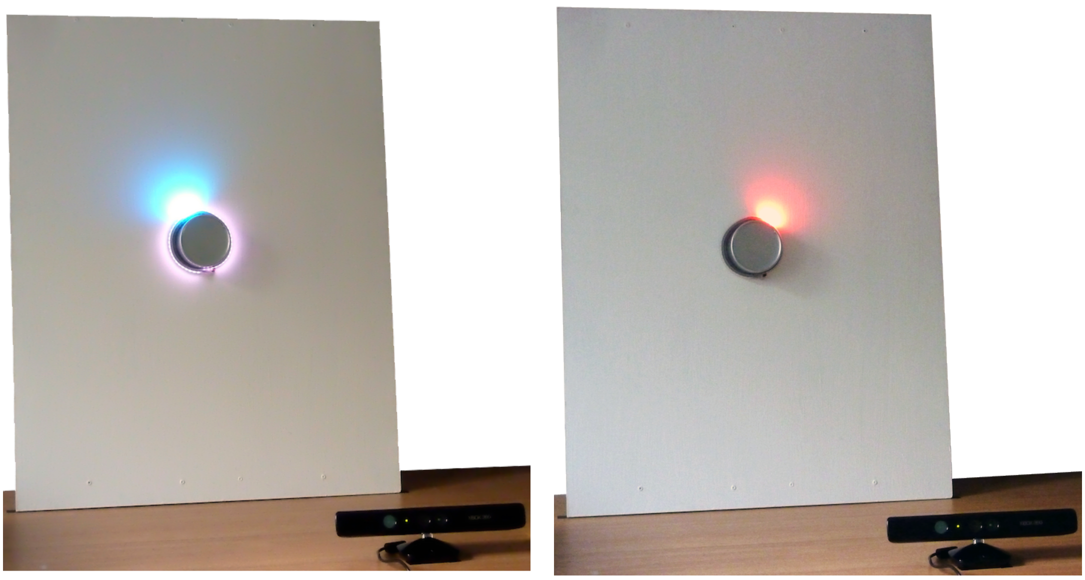


**Figure 4.2:** The Gesture Thermostat: showing a temperature setting on the colder side of the dial, while the rest of the lights gently pulse white (left); and showing a temperature setting on the warmer side of the dial as it gets adjusted, with the rest of the lights off (right). The Xbox Kinect was used for sensing gestures to control the Thermostat.

### 4.2.1 Hardware Design

The Gesture Thermostat is a round dial, like the Honeywell Lyric (Figure 4.1) and the Nest Thermostat (Figure 2.3). This form factor was chosen because it is similar to popular contemporary thermostats. A circular cardboard dial prototype (⌀10cm) was mounted on a large wooden panel which wires and electronic components could be hidden behind, so that they do not obscure the interactive light feedback around the dial.

Interactive light feedback from the thermostat was created by LEDs placed around the dial, which illuminate the surrounding wooden panel when turned on; Figure 4.2 shows the Gesture Thermostat dial and two examples of its interactive light output. The thermostat used a flexible strip of 46 Adafruit NeoPixel LEDs[1] (Figure 4.3), wrapped around its edge. Each of these LEDs can be controlled independently of the others and their hue and brightness can be adjusted. The LEDs are controlled by an Arduino microcontroller[2] and can be updated hundreds of times per second, allowing the LEDs to be used for smooth animations.

An Xbox Kinect[3] depth sensor was used for tracking hands and gestures. This sensor was chosen because it can track hand movements from across the room, making it appropriate for the range over which users of the Gesture Thermostat will be performing gestures. In future, similar sensing technology could be integrated into the actual thermostat; there are already mobile devices[4] with this technology, demonstrating the small size of contemporary depth-sensing cameras.

The Gesture Thermostat also gave audio and tactile feedback about gestures. Audio cues were delivered using laptop speakers; this kept the prototype simple, although speakers could be added to the thermostat in future. Tactile cues were given using the wrist-based tactile display from Experiments 1 and 2. This tactile display was chosen because it was successful in the earlier studies. A laptop computer was used for sensing gestures, maintaining thermostat state and producing its multimodal feedback. Interactive light was controlled by sending commands to the Arduino microcontroller, while audio and tactile feedback were synthesised using Pure Data[5] (as in Experiments 1 and 2 for tactile feedback).

### 4.2.2 Interaction Design

Two thermostat functions were selected for the Gesture Thermostat: (1) checking the temperature, which users would normally do by looking at the dial or display after approaching it; and (2) adjusting the temperature, which users would normally do using physical controls

[1]Adafruit NeoPixel Strip: `www.adafruit.com/products/1506` Accessed 07/07/15
[2]Arduino: `www.arduino.cc` Accessed 07/07/15
[3]Xbox Kinect: `www.xbox.com/en-GB/Xbox360/Accessories/kinect` Accessed 05/05/15
[4]Google Project Tango: `www.google.com/atap/project-tango` Accessed 03/08/15
[5]Pure Data: `www.puredata.info` Accessed 13/05/15

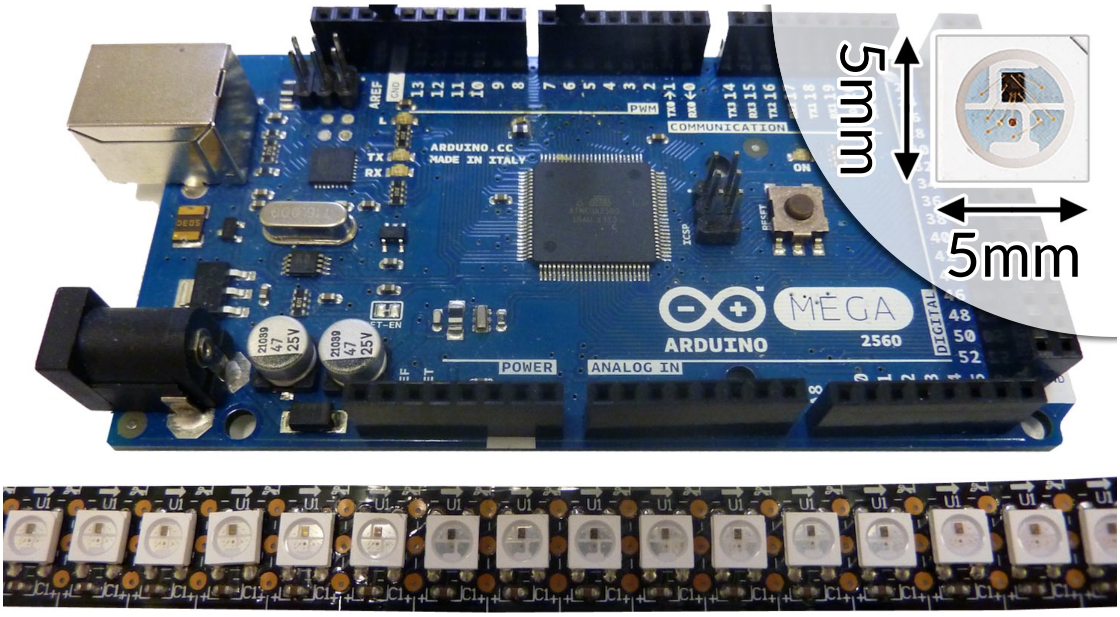



**Figure 4.3:** The LED strip used in the Gesture Thermostat prototype. Its high density of LEDs (144 per metre) meant 46 LEDs could be fit around the thermostat dial (⌀10cm). The LEDs were controlled by an Arduino microcontroller. A laptop computer sent instructions to the Arduino using its USB serial port.

or a smartphone application (for modern 'smart' thermostats). These functions provide users with what Koskela *et al.* [71] call "instant control" (see Section 2.5.1 in the Literature Review). More complex functions, like programming the thermostat, are less appropriate for gesture interaction so were not provided.

When users are first sensed by the thermostat, all of its lights turn on at 25% brightness, showing white light. Here, interactive light is used to show system attention [11, 37] so that users are aware that the thermostat can be interacted with and so that they know it is sensing their movements. This feedback was intended to reveal the thermostat and show attention, which could help users address it as they know they are sensed by it. This design was also intended to be unobtrusive when users are not interested in interacting with the thermostat. If they show no intention of gesturing, or if they leave its sensor view, the lights fade out.

### Checking the Thermostat Setting

Users address the Gesture Thermostat by raising their hand with a closed fist (Figure 4.4). This is the *Query* gesture and it has two purposes: it allows them to address the thermostat and, by doing so, check its temperature setting. This gesture design was chosen for this implementation because it could be easily and reliably recognised. Upon performing the *Query* gesture, the interactive light display increases to full brightness and then begins to pulse gently, fading between 100% and 50% brightness. This animation was used because it shows continuous response, even when users are not performing gestures. Experiments 1 and 2 found that this type of continuous feedback was important as it reassures users that they are addressing the system correctly and are still being sensed. The increased brightness from when users are first detected provides confirmation that they are addressing the thermostat properly. Earcons and Tactons, described in Table 4.1 are also given when the *Query* gesture is recognised, providing additional feedback that the gesture was detected.

During the *Query* gesture, the thermostat temperature setting is displayed as an area of coloured light at an appropriate location around the dial. This light has a fixed brightness (at 100%), while the rest of the lights display the pulsing white animation. If the thermostat setting is on the cooler left side of the dial then blue light is used; if it is on the warmer right side of the dial then red light is used, as shown in Figure 4.2[6]. Two LEDs are used to represent each discrete step around the dial, meaning the thermostat can display a temperature range of 23 degrees; this is comparable to the analogue thermostat dial in Figure 4.1. A limitation of interactive light displays like the one used here is that they are less able to precisely represent quantities — such as the actual temperature — when they are illuminating a surface. Speech output could accompany the interactive light feedback, reading aloud the temperature (for example, "twenty degrees Celsius"); however, this was not necessary for the experiment in this chapter.

### Changing the Thermostat Setting

Two types of gesture were created for changing the thermostat setting: *Quick Change* and *Precise Change*. These interactions differ in the amount of adjustment they allow. Whereas *Precise Change* is a continuous gesture which allows users to change the setting in one degree increments, *Quick Change* is a discrete one which uses three degree increments. These gestures support different interaction goals; for example, users may prefer to use *Quick Change* when they just want to turn the thermostat down and would use *Precise Change* when they want to set it to a specific temperature. These gestures will allow the following

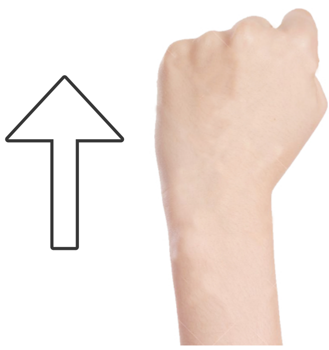


**Figure 4.4:** *Query*: raise hand with a closed fist.

[6]The photographs make the area of coloured light appear larger than it actually is.

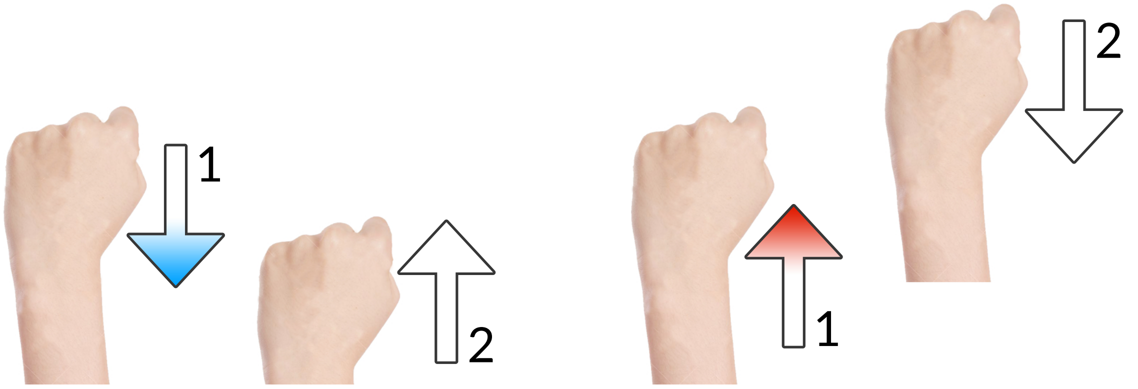



**Figure 4.5:** *Quick Change*: move hand down-then-up (left) or up-then-down (right) to lower or raise temperature, respectively.

**Figure 4.6:** *Precise Change*: turn hand to the left or right to lower or raise temperature, respectively; the same as turning the thermostat dial.

| Event | Description | Earcon Notes | Pattern |
|---|---|---|---|
| *Start of interaction (Query gesture)* | | | |
| | Four tones with rising pitch | C6, D6, E6, F6 | 300ms tones with 50ms gaps |
| *End of interaction (hand lowered)* | | | |
| | Four tones with falling pitch | F6, E6, D6, C6 | 300ms tones with 50ms gaps |
| *Start of Precise Change gesture* | | | |
| | Two tones with rising pitch | G#4, C5 | 300ms tones with 100ms gaps |
| *End of Precise Change gesture* | | | |
| | Two tones with falling pitch | C5, G#4 | 300ms tones with 100ms gaps |
| *Temperature changed (Quick Change or Precise Change)* | | | |
| | Single tone, pitch increasing with temperature | E5–G#9 | 100ms tone |

**Table 4.1:** Earcon designs for the Gesture Thermostat. Tactons used the same patterns but had a fixed frequency of 150 Hz, the resonant frequency of the actuator. Earcons were paired so that rising pitch and falling pitch indicated the beginning and end of an interaction mode, respectively. For example, four tones of increasing pitch signalled the start of interaction and four tones of decreasing pitch signalled the end of interaction.

experiment to investigate how interactive light feedback can meet the demands of interactions which offer different levels of control to users.

*Quick Change* and *Precise Change* both begin from the closed-fist pose of the *Query* gesture, described previously. For *Quick Change*, users can increase or decrease the temperature by moving their hand up then down again, or down then up again, respectively; see Figure 4.5. When this gesture is detected, a short Earcon and Tacton are presented (described in Table 4.1). The interactive light feedback from the *Query* gesture continues uninterrupted, although the new temperature setting is shown instead. Feedback is synchronised so that the interactive light display changes in time with the Earcon and Tacton. Since *Quick Change* is a discrete gesture, this feedback only shows the effects it had on the Gesture Thermostat.

*Precise Change*, in contrast, is a continuous gesture which allows greater control over the thermostat setting. Users begin this gesture by opening their hand (Figure 4.6), an action which is easily sensed and lets the thermostat know that users have started gesturing. When this action is detected, an Earcon and Tacton are presented and the interactive light feedback changes. The thermostat setting is shown as before, using coloured light, although this pulses gently between 100% and 50% brightness. The rest of the lights are switched off. This difference makes it clear that the thermostat is in a different mode and the pulsing of the thermostat setting reassures users that their gesture is being sensed and responded to.

Figure 4.7 shows examples of how the temperature setting appears during this gesture.

From the open hand position, users can increase or decrease the setting by turning their hand to the right or to the left, respectively (Figure 4.6), as though they were physically turning the thermostat dial. This gesture uses fixed rate-based sensing, where the temperature increases or decreases by a fixed amount over time, rather than position-based sensing, where the temperature changes based on the hand angle. Fixed rate-based sensing was chosen because it can be reliably sensed; noise in the sensor data and hand instability means position-based gestures are more difficult to sense accurately [9]. As the temperature changes, the area of coloured light moves to reflect the new setting. This is accompanied by audio and tactile feedback, described in Table 4.1. Once users have finished adjusting the temperature to the desired setting, they can end the gesture by closing their hand again. When this happens, an Earcon and Tacton are presented (Table 4.1) and the interactive light display returns to its previous state, showing the temperature setting and the pulsing white lights.

## 4.3 Experiment 3

### 4.3.1 Research Aims

Experiment 3 investigates the use of interactive light feedback in a gesture-sensing system. While others have used interactive light displays for gesture feedback before, there is little work evaluating the usefulness and usability of this feedback. This experiment addresses this gap and also investigates audio and tactile feedback from the Gesture Thermostat and how these affect the interactive light feedback.

The main aims of this research are: (1) investigate the effectiveness of interactive light feedback; (2) study if and how interactive light is affected by multimodal presentation with audio and tactile feedback; and (3) evaluate the feedback designs for the Gesture Thermostat. This research contributes an answer to the following research question:

**Figure 4.7:** Examples of three temperatures shown during the *Precise Change* gesture. Here, the temperature is shown in the middle setting (centre) and offset by three increments in each direction (left and right).

**RQ2:** Can interactive light be used to present gesture feedback to users?

### 4.3.2 Experiment Design

In this experiment, participants were asked to complete three types of task, each using one of the three gesture interactions described in Section 4.2; task design will be described in the following section. There were two blocks of tasks for each gesture: first, participants were only given interactive light feedback; then, they were given audio and tactile feedback as well as interactive light. **Feedback** was a within-subjects factor and participants always experienced interactive light feedback on its own before getting multimodal feedback. This was so that the study design could also investigate how audio and tactile cues complemented interactive light feedback; by presenting light on its own first, participants could make a comparison between the unimodal and multimodal feedback. This has implications for the experiment findings, however, as there may be a learning effect present in the multimodal results.

The experiment did not present audio and tactile feedback without interactive light, as the main research aim was to investigate interactive light feedback in order to answer **RQ2**. Audio and tactile feedback were not separately paired with interactive light, as this was not necessary to answer the research questions and would have increased the length and complexity of the experiment. There was also no baseline condition where no feedback (in any modalities) was given, as interaction would be too difficult without feedback to tell users what was happening. In many cases, feedback revealed information about the thermostat (e.g. its temperature) and was therefore necessary to complete the tasks.

The order in which participants used each gesture was balanced using a Latin square. There were six task blocks, shown in Table 4.2.

| | **Name** | **Task Type** | **Feedback Type** |
|---|---|---|---|
| 1 | Query-Light | Query | Interactive light only |
| 2 | Query-Multimodal (NC) | Query | All modalities |
| 3 | QuickChange-Light | Quick Change | Interactive light only |
| 4 | QuickChange-Multimodal | Quick Change | All modalities |
| 5 | PreciseChange-Light | Precise Change | Interactive light only |
| 6 | PreciseChange-Multimodal | Precise Change | All modalities |

**Table 4.2:** Conditions in Experiment 3.

### Task Design

This experiment had three types of task, one for each gesture. For the *Query* gesture, participants were shown a picture of the thermostat showing its temperature and were asked if the current thermostat setting was higher, the same, or lower than in the picture. Figure A.1 in Appendix A shows the task user interface, including the three buttons used to provide an answer for this task. Tasks ended once participants provided an answer. This task design was chosen because it evaluates how easily participants can interpret information communicated using the interactive light display. While such displays are not intended for presenting precise quantities, they could be useful for providing a glanceable overview of information; such as giving a rough indication of thermostat setting, in this case.

For the *Quick Change* gesture, participants were shown a picture of the thermostat displaying its temperature and were asked to use *Quick Change* to adjust the current thermostat setting towards that shown in the picture. They were required to raise or lower the temperature for each task, although the number of times they did this did not matter; they only had to perform the correct gesture. Figure A.2 in Appendix A shows the user interface instruction for this task. Tasks started when participants first raised their hand and ended once they lowered it. This task design requires participants to perform discrete gestures and evaluates the feedback given about them.

Finally, for the *Precise Change* gesture, participants were instructed to adjust the current thermostat setting by a particular amount. They were asked to raise or lower the temperature by a random number, with a minimum change of three degrees and a maximum of five. This minimum amount meant participants had to gesture for a prolonged period of time, meaning they were given more feedback. Figure A.3 in Appendix A shows the user interface instruction for this task. Tasks started when participants first raised their hand and ended once they lowered it again. This task design requires participants to perform continuous gestures and evaluates how effective their feedback was. All tasks started with participants addressing the thermostat with the *Query* gesture. This means these tasks can also provide insight into how effective interactive light displays are for presenting feedback as users address them.

There were 20 *Query* and *Quick Change* tasks in each block of tasks. There were only 15 *Precise Change* tasks in a block, as these took more time to complete. These numbers of tasks were chosen so that the experiment would not last more than one hour, including time at the beginning and end of the session for discussion.

### Measures

#### *Task Performance*

Task time was recorded (*Time*). Timing began when participants raised their hand at the start of each task and ended upon task completion: for the *Query* tasks, this was when they used the laptop to give their response; for the other tasks, this was when they lowered their hand to end a gesture.

The proportion of correctly completed tasks was also recorded for each block of tasks (*Correct*). For the *Query* tasks, participants had to select the correct response on the laptop. For the *Quick Change* tasks, they had to perform the right gesture. For the *Precise Change* tasks, they had to set the thermostat to within one degree of the target value; for example, if they had to lower the temperature by five degrees then lowering it by four or six was also accepted. Because *Precise Change* was a continuous gesture, its tasks were more difficult so this tolerance allowed small mistakes.

#### *Surveys*

After each block of tasks, participants were asked to complete a survey. Two surveys were used: one asked questions about interactive light feedback, given after the first block for each gesture; and the other asked questions focusing on audio and tactile feedback, given after the second block for each gesture. Both are shown in Appendix B and are described in this section. Using separate survey designs meant that specific questions could be asked to address the research aims of this study. For example, many of the questions in the second survey focused on how well audio and tactile feedback worked alongside interactive light.

These surveys asked participants to rate their agreement with a series of statements, using a five-point scale from “Strongly Disagree” (1) to “Strongly Agree” (5). The statements for the first survey were:

S1Q1: Interactive light feedback was useful when I was gesturing.
S1Q2: Interactive light feedback was relevant to my gestures.
S1Q3: I was given enough interactive light feedback about my gestures.
S1Q4: Light made me aware of when the thermostat was responding to my input.
S1Q5: I understood how the light feedback changed in response to my hand movements.
S1Q6: Light feedback made me aware of when my movements were being tracked.
S1Q7: Interactive light feedback was annoying to look at.
S1Q8: Interactive light feedback was pleasant to look at.

The statements for the second survey are listed below; participants gave two responses for S2Q1–S2Q3, one for each type of feedback.

S2Q1: Audio / Tactile feedback was useful when I was gesturing.

S2Q2: Audio / Tactile feedback was relevant to my gestures.
S2Q3: I was given enough audio / tactile feedback about my gestures.
S2Q4: I would prefer audio feedback was well as light feedback.
S2Q5: Audio feedback went well with the light feedback.
S2Q6: Audio feedback would let me complete tasks without light feedback.
S2Q7: I would prefer tactile feedback as well as light feedback.
S2Q8: Tactile feedback went well with the light feedback.
S2Q9: Tactile feedback went well with the audio feedback.
S2Q10: Tactile feedback would let me complete tasks without light feedback.

These surveys were designed to investigate how helpful each of the three types of feedback were (S1Q1–S1Q3 and S2Q1–S2Q3), how effective interactive light was as a feedback modality (S1Q4–S1Q6), how pleasant, or not, it was (S1Q7 and S1Q8), and how well audio and tactile feedback complemented interactive light (S2Q4–S2Q10).

### Procedure

At the beginning of the experiment session, participants were introduced to the Gesture Thermostat system. They were given a tutorial which showed them each of the three gestures and the feedback which accompanied them. They were also told about the experimental tasks and were given a few minutes to try gestures themselves (with and without the non-visual feedback). They were asked to complete a block of practice tasks. Once participants were familiar with the interactions and tasks, the experiment started.

Participants were seated 3m from the wall-mounted thermostat prototype, shown in Figure 4.2. The laptop computer which displayed task instructions was placed on a desk in front of them; the laptop was positioned on the side of their non-dominant hand, so that it did not obscure the Kinect sensor's view of their other hand as they performed gestures. After each block of tasks, they were asked to complete a short survey. At the end of the experiment, they were also interviewed about their experience. This was an unstructured interview, with their survey responses being used to drive the discussion. The experimenter took notes during the interviews, for later analysis.

### Hypotheses

- **H1:** *Time* will be lower for *Precise Change* tasks with multimodal feedback, but not for the other tasks;
- **H2:** *Correct* will be higher for *Precise Change* tasks with multimodal feedback, but not for the other tasks;
- **H3:** Median agreement for S2Q4 and S2Q7 will be greater than, or equal to, four.

This experiment investigates the usability of three interactions with the Gesture Thermostat. Of these three, the *Precise Change* gesture demands the most precision and engagement from participants. Additional feedback is expected to make those tasks easier, as discrete audio and tactile feedback will make changes in light during the continuous gesture more noticeable. As such, **H1** and **H2** anticipate lower task times and higher task success, as participants will be more aware of the rate at which the thermostat changes. Task performance for the other gestures is not expected to benefit as much from the audio and tactile feedback. However, a better user experience is expected, as was found in Experiments 1 and 2: users found tactile feedback beneficial because it improved their awareness of how the system was responding to their input and this is expected to be the case here. The potential benefits of extra feedback about users' actions means that they are expected to prefer to receive audio and tactile feedback, as indicated by agreement with S2Q4 and S2Q7 (**H3**).

#### Participants

Sixteen people participated in this experiment (five females; mean age of participants was 27.2 years, sd 10 years). They were recruited using university email lists and social media. All were paid £6 for participating in this experiment.

### 4.3.3 Results

This section presents results from this experiment, one task at a time. However, survey responses for some questions (S1Q4–S1Q8 and S2Q4–S2Q10) will be discussed together at the end of this section.

#### Query Tasks

##### *Task Performance*

Mean *Time* for the *Query* tasks was 3645ms (sd 460ms); see Figure 4.8. The effect of **Feedback** was examined using a repeated-measures t-test. It found that *Time* was significantly lower when multimodal feedback was given: t = 3.95, p = 0.001.

Mean *Correct* for the *Query* tasks was 83.3% (sd 8.7%); see Figure 4.8. A t-test was also used to examine the effect of **Feedback**. It found no significant difference: t = -0.48, p = 0.64.

##### *Survey Responses*

Table 4.3 shows median agreement ratings for Q1–3 in each of the surveys for the *Query* tasks. Friedman's rank sum tests were used to investigate the effect of feedback modality on these ratings. Ratings for Q1 were significantly different: $\chi^2(2) = 16.53$, $p < 0.001$. *Post hoc*

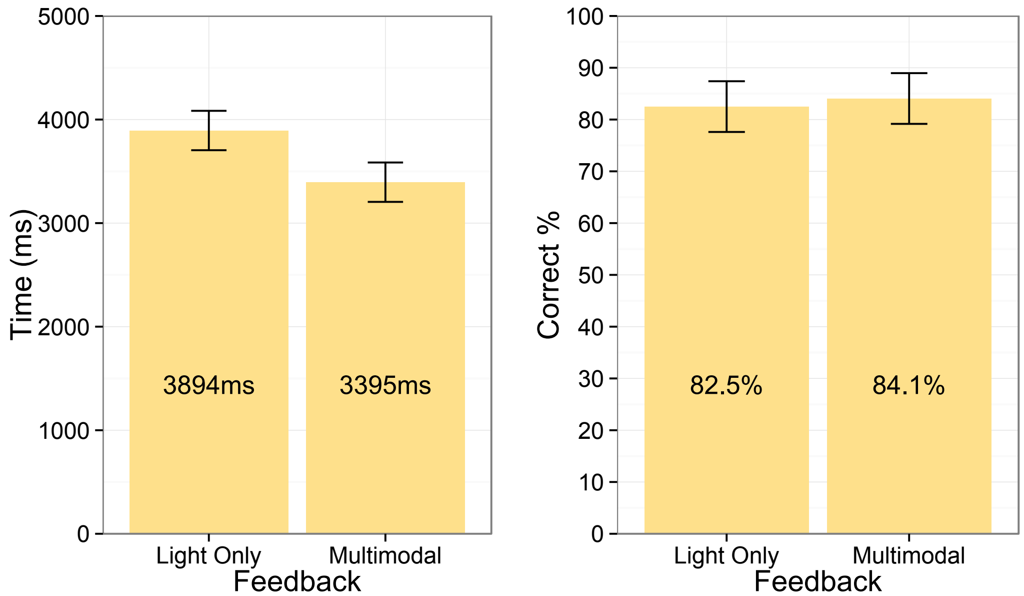



**Figure 4.8:** Mean *Time* and *Correct* for each type of **Feedback**, for the query tasks. Error bars show 95% CIs.

Wilcoxon's tests found that Light and Audio were rated more useful than Tactile feedback (both $p \leq 0.01$). There was no difference between Light and Audio feedback ($p = 0.7$).

Ratings for Q2 were significantly different: $\chi^2(2) = 9.72$, $p = 0.01$. *Post hoc* Wilcoxon's tests found that Audio was rated more relevant than Tactile ($p = 0.04$). No other comparisons were significant ($p \geq 0.17$). Ratings for Q3 were also significantly different: $\chi^2(2) = 9.7$, $p = 0.01$. However, *post hoc* Wilcoxon's tests found no significant differences (all $p \geq 0.1$).

### Quick Change Tasks

#### *Task Performance*

Mean *Time* for the *Quick Change* tasks was 3744ms (sd 693ms); see Figure 4.9. *Time* did not have a normal distribution (Shapiro-Wilk: $W = 0.84$, $p < 0.001$) so Wilcoxon's signed-rank test was used instead of a t-test. It found no significant difference: $Z = -0.94$, $p = 0.35$.

| **Statement** | **Feedback** Light | Audio | Tactile |
|---|---|---|---|
| *S1Q1 & S2Q1: Feedback was useful* | 4 | 4.5 | 4 |
| *S1Q2 & S2Q2: Feedback was relevant* | 4.5 | 5 | 4 |
| *S1Q3 & S2Q3: Enough feedback was given* | 4 | 5 | 4 |

**Table 4.3:** Median agreement ratings for the *Query* survey statements.

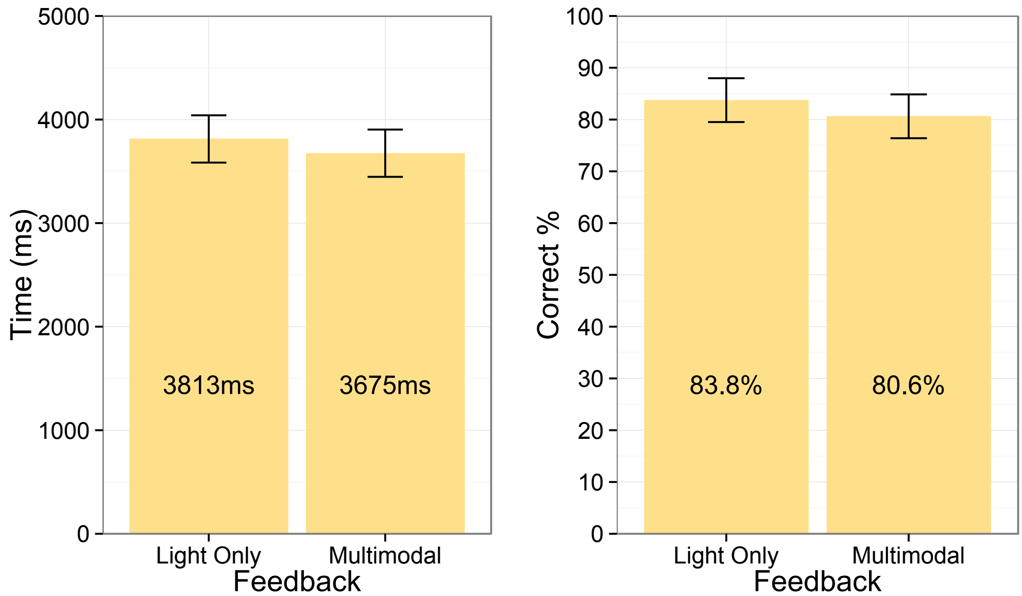



**Figure 4.9:** Mean *Time* and *Correct* for each type of **Feedback**, for the quick change tasks. Error bars show 95% CIs.

Mean *Correct* for the *Quick Change* tasks was 82.2% (sd 13.6%); see Figure 4.9. *Correct* did not have a normal distribution (Shapiro-Wilk: W = 0.85, $p < 0.001$) so Wilcoxon's signed-rank test was used instead of a t-test. It found no significant difference: Z = -0.91, p = 0.36.

### *Survey Responses*

Table 4.4 shows median agreement ratings for Q1–3 in each of the surveys for the *Quick Change* tasks. Friedman's rank sum tests were used to investigate the effect of feedback modality on these ratings. Ratings for Q1 were significantly different: $\chi^2(2) = 12.32$, $p < 0.001$. *Post hoc* Wilcoxon's tests found that Light and Audio were rated more useful than Tactile feedback (both $p \leq 0.03$). There was no difference between Light and Audio feedback (p = 1.0).

Ratings for Q2 were significantly different: $\chi^2(2) = 8.49$, p = 0.01. *Post hoc* Wilcoxon's tests found that Light was rated more relevant than Tactile (p = 0.01). No other comparisons were

| **Statement** | **Feedback** Light | Audio | Tactile |
|---|---|---|---|
| *S1Q1 & S2Q1: Feedback was useful* | 4 | 4 | 4 |
| *S1Q2 & S2Q2: Feedback was relevant* | 4.5 | 4 | 4 |
| *S1Q3 & S2Q3: Enough feedback was given* | 4 | 4 | 3.5 |

**Table 4.4:** Median agreement ratings for the *Quick Change* survey statements.

significant ($p \geq 0.55$). Ratings for Q3 were not significantly different: $\chi^2(2) = 4.6$, $p = 0.1$.

### Precise Change Tasks

#### *Task Performance*

Mean *Time* for the *Precise Change* tasks was 8116ms (sd 2219ms); see Figure 4.10. A repeated-measures t-test examined the effect of **Feedback** on *Time*. It found that *Time* was significantly lower when multimodal feedback was given: $t = 4.78$, $p < 0.001$.

Mean *Correct* for the *Precise Change* tasks was 62.9% (sd 21.4%); see Figure 4.10. *Correct* did not have a normal distribution (Shapiro-Wilk: $W = 0.92$, $p = 0.02$) so Wilcoxon's signed-rank test was used instead of a t-test. It found that participants completed more tasks correctly when given multimodal feedback: $Z = -2.7$, $p = 0.007$.

#### *Survey Responses*

Table 4.5 shows median agreement ratings for Q1–3 in each of the surveys for the *Precise Change* tasks. Friedman's rank sum tests were used to investigate the effect of feedback modality on these ratings. Ratings for Q1 were significantly different: $\chi^2(2) = 14.82$, $p < 0.001$. *Post hoc* Wilcoxon's tests found that Audio was rated more useful than Tactile feedback: $p = 0.003$. No other comparisons were significant (both $p \geq 0.06$).

Ratings for Q2 were significantly different: $\chi^2(2) = 7.95$, $p = 0.02$. However, *post hoc* Wilcoxon's tests found no significant differences (all $p \geq 0.05$). Ratings for Q3 were also significantly different: $\chi^2(2) = 6.43$, $p = 0.04$. However, *post hoc* Wilcoxon's tests found no significant differences (all $p \geq 0.07$).

### Other Survey Responses

The previous sections did not present responses to S1Q4–S1Q8, from the first survey. This is because Friedman's tests for each question found that responses were not significantly different for each type of task: all $\chi^2(2) \leq 2.67$, $p \geq 0.26$. As such, these responses are presented here as a whole, in Table 4.6.

| **Statement** | **Feedback** | | |
|---|---|---|---|
| | Light | Audio | Tactile |
| *S1Q1 & S2Q1: Feedback was useful* | 5 | 5 | 4 |
| *S1Q2 & S2Q2: Feedback was relevant* | 4.5 | 4 | 4 |
| *S1Q3 & S2Q3: Enough feedback was given* | 4 | 4 | 3 |

**Table 4.5:** Median agreement ratings for the *Precise Change* survey statements.

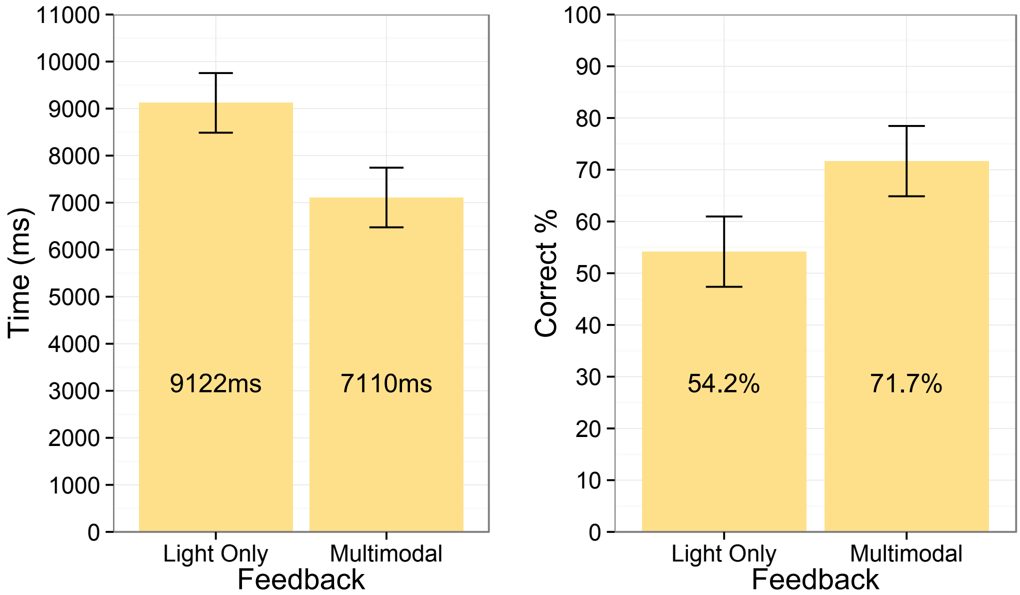



**Figure 4.10:** Mean *Time* and *Correct* for each type of **Feedback**, for the precise change tasks. Error bars show 95% CIs.

| Statement | Median | Deviation |
|---|---|---|
| *S1Q4: Light made me aware of response to my input* | 5 | 0 |
| *S1Q5: I understood how light changed with my movements* | 4 | 0 |
| *S1Q6: Light made me aware of when I was being tracked* | 4 | 0.74 |
| *S1Q7: Light was annoying to look at* | 1 | 0 |
| *S1Q8: Light was pleasant to look at* | 5 | 0 |

**Table 4.6:** Median agreement, with median absolute deviation, for S1Q4–S1Q8.

Responses to S2Q4–S2Q10, from the second survey, were also not significantly different for each task: all $\chi^2(2) \leq 3.5$, $p \geq 0.17$. These responses have also been aggregated for presentation, shown in Table 4.7. Wilcoxon's signed-rank tests were used for pairwise comparisons between: S2Q4 and S2Q7; S2Q5 and S2Q8; and S2Q6 and S2Q10. These pairs asked participants to rate their agreement with the same statements, for audio and tactile feedback respectively. Participants agreed more with the statements about audio feedback than they did for tactile feedback: all $Z \geq 4.37$, $p < 0.001$.

### 4.3.4 Discussion

#### Effectiveness of Interactive Light

Interactive light feedback allowed participants to interact and complete tasks with the Gesture Thermostat, even though it lacked a screen for displaying information and feedback. Participants gave positive ratings about the effectiveness of the feedback (S1Q1–S1Q3); especially for the more difficult *Precise Change* tasks, where they strongly agreed that it was useful (S1Q1). In interviews after the experiment, they were asked why they found the light feedback useful. Many explained how it supported their awareness of when and how the thermostat was responding to their actions. For example, they knew they were being sensed when the lights increased in brightness at the start and the pulsing animations let them know they were still being sensed, even if they were not actively performing a gesture. Survey responses to S1Q4 and S1Q6 support these findings. This awareness of being sensed is important when addressing and interacting with a gesture system, because it lets users know they are gesturing in a way which can be seen by the system.

A few participants said they would have liked more feedback when something went wrong because the light feedback was often ambiguous in those situations. They knew that a tracking error had occurred (for example, when the Kinect sensor had difficulty recognising their gestures or when it lost sight of their hands) because the feedback changed in unexpected

| Statement | Median | Deviation |
|---|---|---|
| *S2Q4: I prefer audio feedback as well as light* | 5 | 0 |
| *S2Q5: Audio feedback went well with light* | 4 | 0 |
| *S2Q6: I could complete tasks with just audio feedback* | 3 | 1.48 |
| *S2Q7: I prefer tactile feedback as well as light* | 4 | 0.74 |
| *S2Q8: Tactile feedback went well with light* | 4 | 0 |
| *S2Q9: Tactile feedback went well with audio* | 4 | 0 |
| *S2Q10: I could complete tasks with just tactile feedback* | 1.5 | 0.74 |

**Table 4.7:** Median agreement, with median absolute deviation, for S2Q4–S2Q10.

ways; however, they did not know what the error was. More explicit feedback about sensing errors, or how to resolve them, would help users overcome these issues.

Participants generally understood how interactive light feedback changed in response to their hand movements (S1Q5). Mode changes (for example, when starting and ending the *Precise Change* gesture) were especially salient, as change in brightness strongly affected the appearance of the thermostat. The use of colour was also effective, as this supported the thermostat metaphor used in the visual feedback design. Participants said that change in colour was easily noticed when adjusting the temperature setting between the hot and cold sides of the dial. They also rated the aesthetic appeal of the interactive light feedback highly, suggesting it was pleasant (S1Q8) and unobtrusive (S1Q7).

### Multimodal Feedback

Findings from this experiment suggest that audio and tactile feedback might have improved gesture interaction with the thermostat. Performance (in terms of task time and success rate) was better for *Precise Change* when multimodal feedback was given, as was expected, although there may have been a learning effect as participants always experienced multimodal feedback second. Task times and success rates were not better for the other tasks, although task times were lower for *Query* ones. These results support accepting **H2** and partly accepting **H1**. In the interviews after the experiment, many participants described how the discrete sound and vibration of the multimodal feedback made the system feel more responsive to their actions. This, in turn, made it easier for them to exert continuous control over the thermostat in the *Precise Change* tasks, as they knew when to stop gesturing because the desired setting was reached. The non-visual feedback did not present information that was unavailable visually, it just enhanced their perception of it. Because of these benefits, participants showed a preference for getting audio and tactile feedback as well as light (S2Q4 and S2Q7), supporting the acceptance of **H3**.

Responses to S2Q5, S2Q8 and S2Q9 were positive (all had a median of 4) suggesting that the feedback designs went well together. This may partly explain the positive effects audio and tactile cues had on interactive light feedback. For example, the sound and vibration occurring at the same time that the light changed may have made that change more noticeable. Many participants said that this was the case, with the non-visual feedback providing "confirmation" of the change in light feedback. Similar findings were reported in the previous chapter; participants in Experiments 1 and 2 said that tactile feedback was helpful because it reinforced feedback which was already being shown on screen. Tactile feedback was less positive here than in those experiments, although this may have been because audio feedback was more effective.

Survey responses suggest that participants found audio feedback more helpful than tactile

(S2Q1–S2Q10), although tactile feedback also received high ratings. Audio may have been preferred because it was easier to notice and understand than the tactile cues. For example, one participant said that he found tactile feedback to be “second best” because it was not as noticeable as audio; however, he did say that he liked tactile too because it “inspired more confidence” than interactive light feedback on its own. Another participant said that he liked the audio feedback design more because it was familiar to him; he had experienced similar sounding feedback in an entertainment system menu.

This experiment presented audio and tactile feedback together, although not individually. This was to save time and reduce the complexity of the experiment. However, a limitation of this choice is that the experiment findings must be interpreted with it in mind; it is difficult to draw conclusions about the effect of each modality individually as both were evaluated together.

### Addressing the Gesture Thermostat

This experiment focused on investigating interactive light feedback and evaluating the feedback from the Gesture Thermostat. However, it also provides insight into one of the problems users encounter when addressing a gesture system — finding where to gesture — and it shows how interactive light feedback could be used to help them overcome that problem. Participants liked how the thermostat lit up and responded to them raising their hand at the start of the interaction; this feedback, along with the audio and tactile cues, let them know they were interacting “properly”. Similarly, an absence of feedback at the beginning of tasks let them know that something was wrong. However, the presence or absence of feedback may be insufficient for helping users learn how best to address systems like the thermostat. The following chapter looks at this issue in more detail, investigating an interaction technique which uses multimodal feedback to help users find where to perform gestures.

### Gesture Interaction with Simple Household Appliances

Although this experiment investigated interaction with the Gesture Thermostat, some of its findings could also be applied to simple household appliances in general. The results show that interactive light feedback about gestures can be effective for displaying the basic state of a device and for giving feedback about gesture interactions. This suggests that appliances with limited display capabilities would benefit from integrating an interactive light display, as it gives them a means of providing feedback to users interacting from a distance. Some of the feedback designs used by the thermostat could also be used by other gesture systems, such as the pulsing animations which showed responsiveness to input and the fade in and out used to show when the gesture system was engaged and disengaged, respectively.

## 4.4 Limitations

### 4.4.1 Device Form Factor

This experiment studied interactive light feedback from the Gesture Thermostat, a circular wall-mounted device. Findings from this experiment may not necessarily apply to other form factors or usage contexts. For example, it is unknown if interactive light would also be effective when given around a mobile phone laid flat on a table. This limitation will be addressed in a later experiment in this thesis.

### 4.4.2 Application-Specific Feedback

Some of the feedback from the Gesture Thermostat was application-specific, representing information about the thermostat settings and using a traditional thermostat dial as a metaphor in the feedback design. However, some of the feedback which participants found most beneficial (discussed in Section 4.3.4) was generic and could be used in other application contexts. For example, the pulsing white light animations and the changes in brightness when starting and ending interaction were both effective and are not specific to the thermostat. The experiment also found a greater need for more explicit feedback when sensor errors occur, which applies to all gesture-sensing systems using interactive light.

## 4.5 Conclusions

This chapter investigated interactive light feedback, a novel way of giving feedback from limited-display devices. It described the Gesture Thermostat — a limited-display device because it has no screen — which turned the surrounding wall surface into a display, using LEDs around its edge. The thermostat was then evaluated in Experiment 3, which studied the usability of its interactive light feedback as well as the use of audio and tactile feedback about gestures. Overall, the findings from Experiment 3 suggest that this can be an effective way of giving feedback about gestures. Interactive light supported users and was good at supporting their awareness of how the thermostat was responding to their actions. This helped when they were addressing the thermostat; when the interactive light display lit up, they knew they could be sensed. Similar benefits were noted about tactile feedback in the previous two experiments. While interactive light was useful on its own, it may have been enhanced by audio and tactile feedback. These modalities made changes in light more salient, which participants may have benefitted from during the more demanding tasks.

The first three experiments of this thesis have investigated novel ways in which limited-display devices can present gesture feedback. Tactile feedback and interactive light feedback have both been found to be effective and these experiments have contributed a greater understanding of their capabilities and how they may be delivered by gesture systems. Along with audio feedback, which is better understood and presents less technical challenges, they can be used as part of multimodal feedback about gestures. In the next chapter, this thesis starts to focus on interaction techniques for *addressing* gesture-sensing systems, building on the work so far by investigating an interaction which uses these feedback modalities.

### 4.5.1 Research Question 2

The findings from this chapter contribute an answer to the following question:

**RQ2:** Can interactive light be used to present gesture feedback to users?

Interactive light displays can be used to present visual feedback in the space surrounding limited-display devices. For devices with no screen, like the Gesture Thermostat, this allows visual feedback which would not otherwise be possible. For devices which do have screens, like mobile phones, this allows the screen to be kept free for showing other content. Designers can use properties of light, like hue and brightness, when creating interactive light feedback. These properties can also be animated, like done here, creating responsive and aesthetically appealing feedback. This feedback is effective on its own but its perception may be enhanced by cues presented in other modalities, like crossmodal audio and tactile feedback.

### 4.5.2 Contributions

The research in this chapter makes the following contributions:

- It examines the usability of interactive light feedback. It demonstrates that such feedback about gestures can be effective;
- It investigates interactive light as part of multimodal feedback. It found that interactive light displays can be enhanced by audio and tactile feedback.

# Chapter 5

# Showing Users Where to Gesture

## 5.1 Introduction

Users must be able to *address* a gesture-sensing system to be able to interact with it. As discussed earlier in this thesis, addressing a gesture system requires finding out where to gesture, so that hand movements can be sensed, and finding out how to direct input, so that gestures only affect the system which users intend to interact with. While many have investigated interaction techniques for addressing gesture-sensing systems (discussed in Section 2.4 of the Literature Review), these techniques have focused on how users can direct their actions towards a particular interface. Research is needed to understand how systems can help users find where to gesture, so that their actions will be sensed. Some gesture-sensing systems show visualisations of what sensors 'see' (for example, Figure 5.1 and Figure 5.2), which can guide users; however, this approach is unsuitable for limited-display devices. On small screens, these visualisations would take up a large amount of space and may be difficult to notice from a distance. Alternative interaction techniques are needed, which also work for screen-less devices. Therefore, the third research question of this thesis asks:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

The research in this chapter investigates how limited-display devices can help users find where to gesture, so that they can address them effectively. It describes the development and evaluation of an interaction technique — *sensor strength feedback* — which uses interactive light, audio and tactile feedback to guide users as they address a gesture system. In a later chapter, it will be used alongside another interaction technique to provide a complete solution to the problem of addressing a gesture system.

**Figure 5.1:** Gesture-sensing systems where users interact with hand and finger movements, from a short distance away, often show users a visualisation of their hand, constructed from what sensors can detect.

**Figure 5.2:** Gesture-sensing systems where users interact from further away often visualise parts of the body, or the whole body, showing what its sensors can 'see'. In this example, the user sees an annotated silhouette of himself; alternatively, the system could have used the sensor camera feed to mirror his body.

### 5.1.1 Chapter Structure

Section 5.2 discusses sensor strength feedback, a focus group study which informed feedback design, and implementation of a prototype which will be used in Experiment 4. Section 5.3 describes Experiment 4, which investigates the effectiveness of this interaction technique. Section 5.4 discusses limitations of this experiment. Section 5.5 discusses the findings of Experiment 4 and revisits the research question discussed earlier in the introduction.

## 5.2 Sensor Strength Feedback

Users need to know where to gesture, so that they interact where their hand movements can be detected by a gesture-sensing system. Gestures performed outside of a system's sensor range cannot be detected and if users are acting at the limits of this range then gesture detection will be unreliable. For example, vision-based sensors (like depth cameras) will be unable to detect enough detail if hands are too far away or too close to them, as in Figure 5.3, or if hands are at the edge of the field of view, as in Figure 5.4. Many gesture-sensing systems show visualisations of what their sensors can 'see', which allows users to move so that they can be seen clearly; Figure 5.1 and Figure 5.2 show examples of such visual feedback. As discussed in the Introduction, this is inappropriate for limited-display devices; detailed visualisations require screen space which may be limited or unavailable. Other output modalities could be used instead, delivering cues which help users find where to gesture. This thesis has identified three types of appropriate output: interactive light, audio and tactile.

These outputs could be used to tell users how well they can be sensed, allowing them to make the same adjustments they would make when using sensor visualisations. However, these modalities have a limited ability to represent spatial information; while spatial audio and tactile cues are possible, they require headphones and more complex tactile displays, respectively, which may not always be available. Interactive light displays may also lack a spatial component; for example, if a single light source is used then feedback is limited to

**Figure 5.3:** Sensors will have difficulty detecting enough detail to sense gestures performed too far from them (left), or too close to them (right).

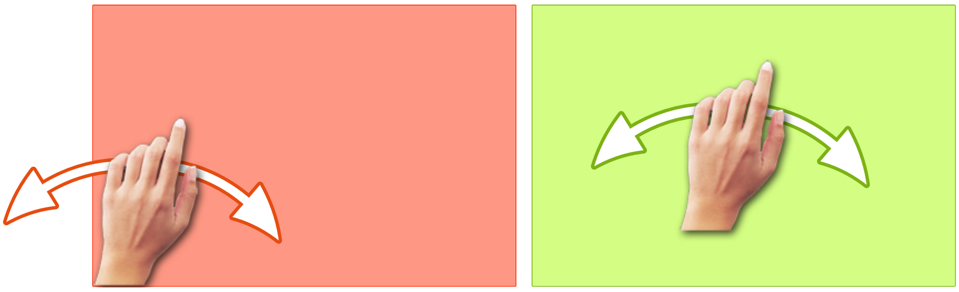


**Figure 5.4:** Although sensors may detect hands at the limits of their sensing, they may be unable to accurately track all hand movements (left). When users gesture towards the centre of the sensor range (right), hand movements are more likely to be captured by the sensor.

using hue and brightness in its design. These limitations mean an alternative information encoding is required; gesture-sensing systems cannot rely on spatial information to guide users when telling them where to gesture.

Gesture-sensing systems could tell users *how well* they can see them, allowing them to move and find a better position for input. This research suggests an interaction technique called *sensor strength feedback*, which uses a measurement called *sensor strength* to tell users how well they can be sensed. Sensor strength is an estimate of how well users and their actions can be detected, based on where they are gesturing. This measurement can then be presented as feedback (mapped to properties of sound, for example). Rather than use spatial cues to tell users where to move (closer to the sensor, towards the left, for example), this technique requires them to actively explore the space around them. Through active exploration, users would find a location where their hand movements can be better sensed; that is, a location where sensor strength is greater. This feedback would have an ambiguous meaning, as it does not tell users what actions they need to take; however, such ambiguity may be beneficial. By exploring how feedback changes with their actions, users might be able to form their own understanding of where and how they should gesture. As well as helping them find where to interact, this feedback could also help them discover the capabilities and limitations of the gesture-sensing technology.

More formally, sensor strength is a value which ranges from 0 to 1, where higher values mean users can be better sensed. Its calculation depends on the sensing technology used. In vision-based approaches, like depth-sensing cameras, it may be the Euclidean distance between hand position (centre of palm or a fingertip, for example) and a position in the sensor space where hands can be optimally sensed. This position is likely the centre of the field of view, at a distance which is not too close or too far from the sensor (as in Figure 5.5). As sensor strength has just one dimension, it can be used by a variety of sensing approaches, including those where positional information may be unavailable. For example, sensing approaches like magnetic sensing around a device [43, 69] or sensing from mobile GSM signals [68], could use signal magnitude for sensor strength.

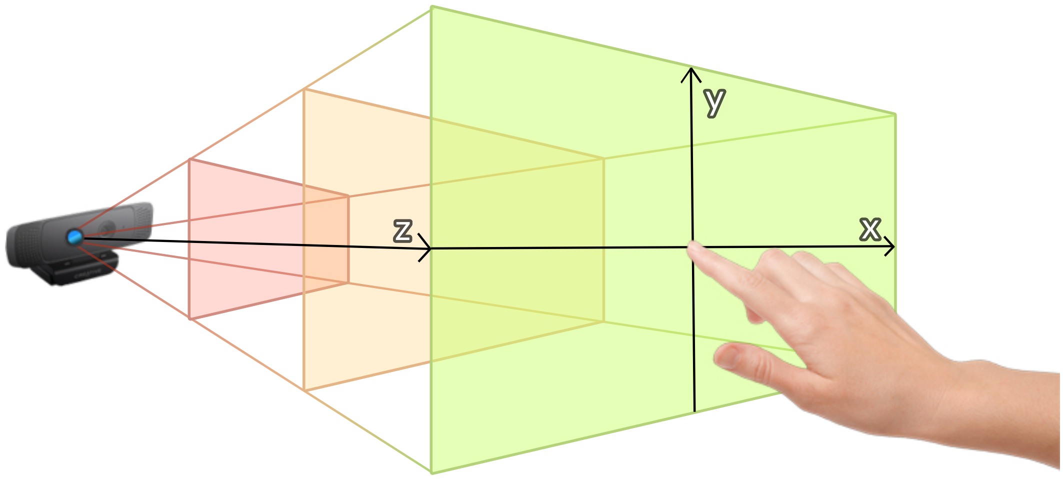



**Figure 5.5:** Vision-based sensing systems may calculate sensor strength as a function of hand position relative to a location where hands can be robustly sensed. In this example, this location (being pointed at by the index finger) is the centre of the field of view at a distance not too close or too far from the sensor.

Sensor strength feedback allows limited-display devices to help users find where to perform gestures as they address those systems. Although non-visual feedback techniques have been created to tell users *how* to perform gestures, these have not considered telling users *where* to perform them. Morrison *et al.* [90] used audio feedback to inform users when they were not fully visible to sensors. Users could use this feedback to find a position where they were completely sensed, although it would not necessarily help them find the best position. They may still be too far from the sensor or in a location where their movements can be sensed, albeit noisily and unreliably. Other feedback techniques discussed in Section 2.6.1 of the Literature Review only gave feedback about how closely gestures matched recognition templates, guiding users towards more accurate gesture performances. Sensor strength feedback has a different purpose to these techniques as it aims to help users find *where* to provide input; once they have addressed the system, these other types of feedback can then be given to help them learn *how* to perform gestures more accurately. Users may rely on sensor strength feedback less often as they gain expertise with a particular gesture system and learn how best to interact with it. In those circumstances, users could go straight to providing input and could ignore the feedback, although they may wish to reduce or disable it. This issue is discussed more in the conclusions of this thesis.

### 5.2.1 Feedback Design Space

Multimodal sensor strength feedback can be created by mapping properties of light, sound and vibration to sensor strength. However, these output types have a variety of design parameters, which results in a large design space. This section discusses this design space and identifies possible mappings between sensor strength and feedback parameters. A focus group study, described in Section 5.2.2, then investigates the designs identified here, to come up with a final feedback design.

### Interactive Light Feedback

Interactive light has two properties which could be mapped to sensor strength: brightness and hue. Spatial location of light was not considered as it would be unsuitable for systems with only a single light source. Mapping brightness to sensor strength — so that lights are brighter for higher values — uses feedback visibility as a metaphor: when feedback is more noticeable, it is because users' hands can be easily sensed; when feedback is more difficult to see, it is because users hands can not be sensed as clearly. Experiment 3 found animated changes in brightness were a good way of showing responsiveness, suggesting that mapping brightness to sensor strength could result in effective and responsive feedback here. When mapping hue to sensor strength, colours have to be chosen which make sense to users. Selecting hue from a gradient between red, yellow and green (see Figure 5.6), for example, would be appropriate, as these colours have negative and positive connotations, which make sense with respect to sensor strength. When feedback is 'good' (green), it is because users are gesturing in a good location for sensing. Both brightness and hue[1] are considered in the focus group study.

### Audio & Tactile Feedback

Sound and vibration share many continuous properties which could be mapped to sensor strength, like frequency or intensity. Mapping these properties to sensor strength would produce continuous sonifications, which users may find obtrusive. Findings from Experiment 1 suggest that continuous tactile feedback can be unacceptable, with some users preferring timely, discrete cues instead. This is likely to be the case for audio feedback as well, which is not just personal to the user but could be heard by others too. An alternative to continuous


**Figure 5.6:** Sensor strength could be mapped to a colour on a gradient from red to yellow to green, using a traffic light metaphor to tell users about their gesture location. Higher sensor strength values would be on the green side of the gradient.

[1] A notable issue with the hue design is that it would be inappropriate for people with red-green colour blindness, leading to difficulty interpreting the feedback; in those cases, an alternative design would have to be used.

audio and tactile feedback is to give short bursts of feedback instead, with temporal properties — like the duration of cues or the intervals between them — encoding information.

Varying the interval between repeating cues creates feedback which sounds or feels like a Geiger counter, with the duration between cues encoding information. This feedback metaphor has been used successfully in navigation applications. *AudioGPS* [52] and *Audio Bubbles* [86], for example, both used this style of audio feedback to encode distance to navigation targets, guiding users to nearby landmarks. As they approached their target, users heard increasingly frequent feedback. *MARSUI* [145], a deformable user interface, used similar audio feedback to guide users to pre-defined deformations; they were given increasingly frequent feedback as their deformations approached the target shapes. Tactile equivalents of this feedback metaphor have also been used for navigation; for example, Pielot *et al.* [106] used tactile patterns to guide users, with patterns repeating more often as they approached the target. Similar Geiger counter-like designs could be used for sensor strength feedback, so are investigated in the focus group study.

A disadvantage of using audio feedback in any interaction is that it may be heard by other people nearby who are not involved in the interaction. For example, officemates or family members may also hear feedback about a user's gestures. In these situations, audio feedback may be less acceptable, even if it is beneficial to the person interacting with the system. The audio and tactile 'Geiger counter' described here is a crossmodal [51] feedback design, as the properties used to encode information are shared between the audio and tactile modalities. An advantage of crossmodal feedback like this is that users could choose the most appropriate modality for certain situations. For example, when around others, users may prefer to receive the tactile version of the feedback, as it would not disturb others like audio might. Changing between modalities may not necessarily have a negative impact on the effectiveness of the interaction in these situations: research into crossmodal learning suggests that cues taught in one modality (audio or tactile) can be reliably identified in the other [51].

### 5.2.2 Focus Group Study

A focus group with six interaction design experts (three PhD candidates, two post-doctoral researchers and one member of research faculty from the School of Computing Science) investigated the feedback mappings discussed previously. These feedback mappings comprise part of the sensor strength feedback design space and this focus group study explored this space to create a final feedback design, for use in Experiment 4. All participants had experience designing and evaluating non-visual user interfaces and interaction techniques, so were familiar with at least one of the feedback types. During the focus group, they used the tool shown in Figure 5.7 to configure sensor strength feedback then try it themselves. The next part of this section describes the design space exposed by this configuration tool, followed by a description of the focus group study procedure.

### Feedback Design Parameters

In the tool shown in Figure 5.7, **Max Distance** controlled the real-world distance between the maximum and minimum value of sensor strength. Figure 5.8 shows how **Max Distance** defines a volume in which sensor strength changes. Hand movements within this volume (from a pre-defined point, **P** in Figure 5.8) would cause a change in sensor strength; hand movements beyond this distance would have a sensor strength value of zero. Restricting the distance in which feedback varies could encourage users to gesture within a smaller area, as their hand movements only elicit noticeable changes in feedback within that region.

**Mapping** let participants choose between a linear or exponential mapping of sensor strength to feedback parameters. An exponential mapping would cause greater changes in feedback as users approached higher values of sensor strength, and more subtle changes as they approached lower values. Others have suggested that exponential transfer functions make for more effective feedback designs, as users benefit from the more noticeable changes in feedback; for example, Pitt *et al*. [107] found that exponential change in volume was better than linear change for communicating target distance and Müller *et al.* [92] found that exponential colour change (between red and green) was better than linear change for communicating remaining time in their *Ambient Timer*.

**Red and Green** and **Brightness** allowed participants to compare the two interactive light feedback designs discussed previously.

For audio and tactile feedback, the **Vary Intensity** option also mapped signal intensity to sensor strength, so that lower values of sensor strength would produce quieter sounds and more subtle vibrations. This approach was used by *Audio Bubbles* [86] for their Geiger counter feedback; as users approached their destination, audio feedback increased in volume as well as rate.

**Attack** and **Decay** adjusted the audio and tactile waveform envelope, which changed the duration of cues and how they sounded/felt. **Frequency** controlled the audio waveform frequency; frequency was fixed at 175 Hz for tactile feedback. Finally, **Click Delay** set the maximum time between successive tones; a faster delay meant more frequent feedback at low sensor strength values. Sliders were not given numerical labels, to encourage focus group participants to fully explore design parameters.

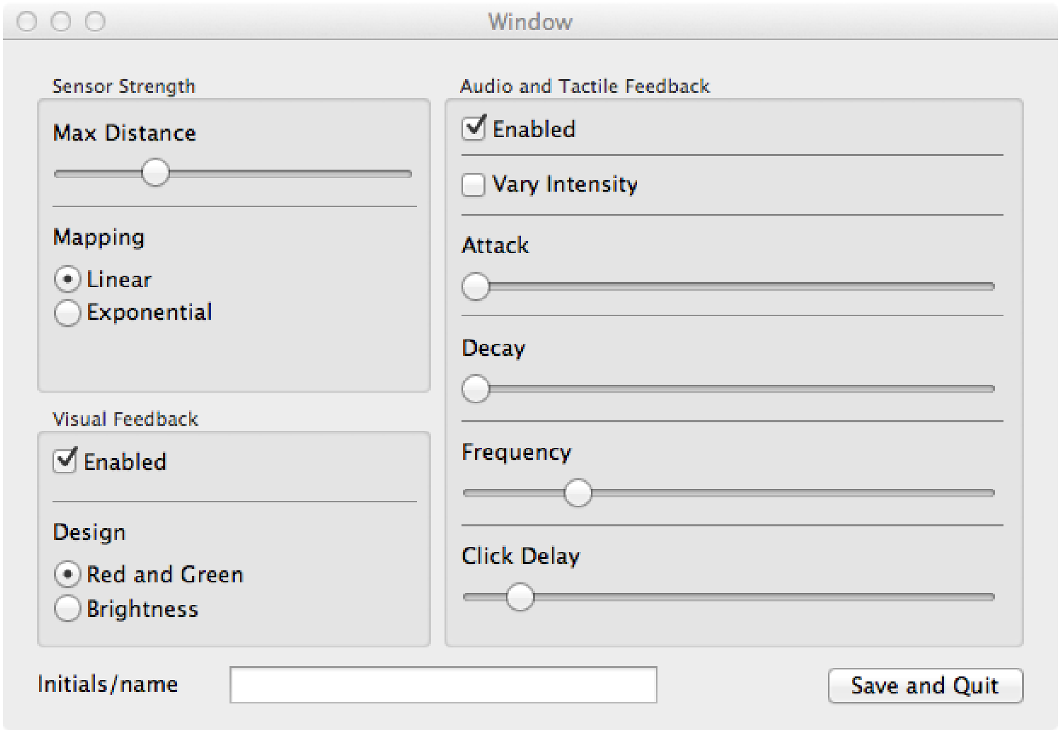



**Figure 5.7:** Configuration tool used during the sensor strength feedback focus group.

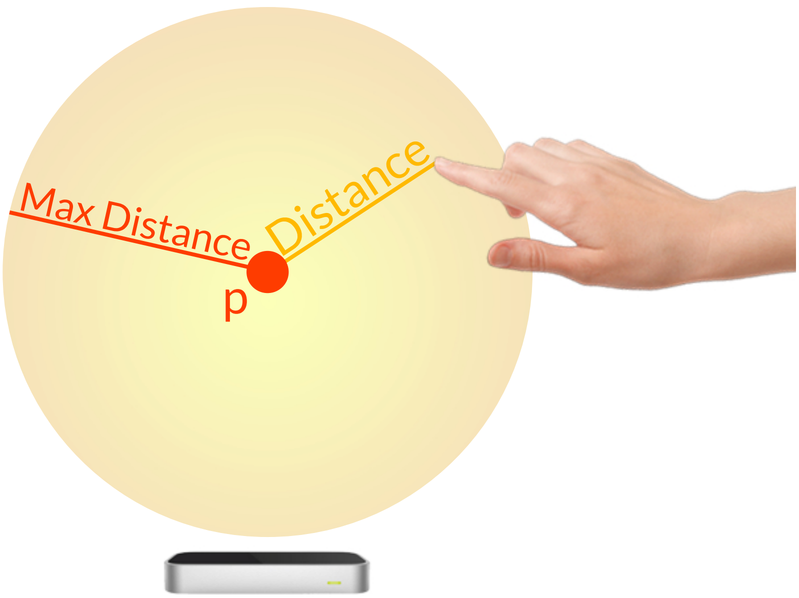



**Figure 5.8:** An example of the volume defined by **Max Distance**, relative to a central point **P**. Sensor strength varies as hands move within this volume; beyond it, sensor strength has a value of zero.

### Focus Group Procedure

Participants were instructed to explore the options available to them, discuss the feedback with each other, and configure a final design which they think will be successful for guiding users as they find where to perform gestures as they address a gesture system. They used the configuration tool, which ran on a laptop computer, to create feedback designs and then try them. A mobile phone was provided as a prop; participants were asked to gesture relative to this device. A Leap Motion sensor, placed over the phone, tracked their hand movements and feedback was given in real time using a strip of LEDs, the laptop loudspeakers and a vibrotactile actuator. They were given as much time as required to use the configuration tool, try the feedback and discuss it. The experimenter observed the focus group and took notes about what was discussed, focusing on justification for design choices. By the end of the focus group, the participants had reached agreement about the feedback design; this configuration was then saved.

## 5.2.3 Final Feedback Designs

As a result of the focus group study, the following feedback design decisions were made:

- **Max Distance:** Feedback will vary up to 300mm from the optimal sensing point. It is important to note that this was calibrated in the context of interaction above a mobile phone; other gesture-sensing systems may be able to sense users over greater distances so a more appropriate value may have to be chosen.
- **Mapping:** An exponential transfer function between sensor strength and feedback parameters will be used. This was chosen because it created more noticeable changes at higher sensor strength values, which could encourage users to find the best location for performing gestures.
- **Brightness**: Interactive light will map brightness of white light to sensor strength, becoming less visible as hands become more difficult to sense. This was considered a better choice than **Red and Green** colour change, which some thought could be too ambiguous. Although red and green have clear negative and positive connotations, participants thought the meaning of the intermediate colours (see Figure 5.6) would be unclear and users may not understand the traffic light metaphor. Figure 5.12 shows an example of how brightness changes with sensor strength.
- **Vary Intensity:** Audio and tactile feedback will not change in intensity, as users may find low intensity feedback difficult to hear and feel. Fixed intensity means that absence of feedback is because users cannot be sensed, rather than they can be sensed, but are gesturing in a less ideal location.
- **Attack, Decay, Frequency:** Sound and vibration tones will last for 50ms, with a 5ms

attack and 45ms decay; this creates a sharp “clicking” sound. Audio will use a 370 Hz frequency (equivalent to F#4 on piano); higher frequencies were considered annoying and “piercing”, while lower frequencies were not as pleasant sounding. Figure 5.9 shows how these parameters create the audio signal used for feedback.

- **Click Delay:** Audio and tactile feedback will use a range of 70ms to 370ms between tones. This range was chosen because it resulted in easily noticeable changes in sensor strength feedback over the interaction range, without unnecessarily long breaks between cues.

### 5.2.4 Implementation and Apparatus

A smartphone with an interactive light feedback display was prototyped, so that these sensor strength feedback designs could be studied in Experiment 4. This prototype, shown in Figures 5.10–5.12, uses sixty LEDs placed around the edge of a smartphone to illuminate the surrounding table surface. Each LED can be controlled independently of the others and can vary its hue and brightness. These LEDs were the same as those used in the *Gesture Thermostat* described in Chapter 4. The phone screen was not used during Experiment 4; users only received visual feedback from the LEDs. A Leap Motion sensor, centred over the phone screen, was used to detect hand movements above the smartphone. This was for simplicity; gesture sensors could be integrated into the mobile phone and there are already consumer products available with these capabilities (discussed in the Literature Review). Hand movements were tracked by a C# application running on a laptop computer; this interfaced with the LED display through an Arduino micro-controller, using its USB serial port for communication. Audio and tactile feedback were synthesised using Pure Data, which ran on the laptop computer. The laptop speakers delivered audio feedback and the wrist-worn tactile display used in Experiments 1–3 delivered tactile feedback to users’ wrists.

## 5.3 Experiment 4

### 5.3.1 Research Aims

Experiment 4 investigates the effectiveness of sensor strength feedback, using the designs created during the focus group. An interaction technique which helps users find where to gesture must be able to accurately guide them, so this experiment aims to understand the performance of this type of feedback. As well as evaluating each of the three feedback modalities, this experiment will also consider their multimodal combinations. If there are benefits to presenting this information multimodally then gesture-sensing systems should make the most of their output capabilities.

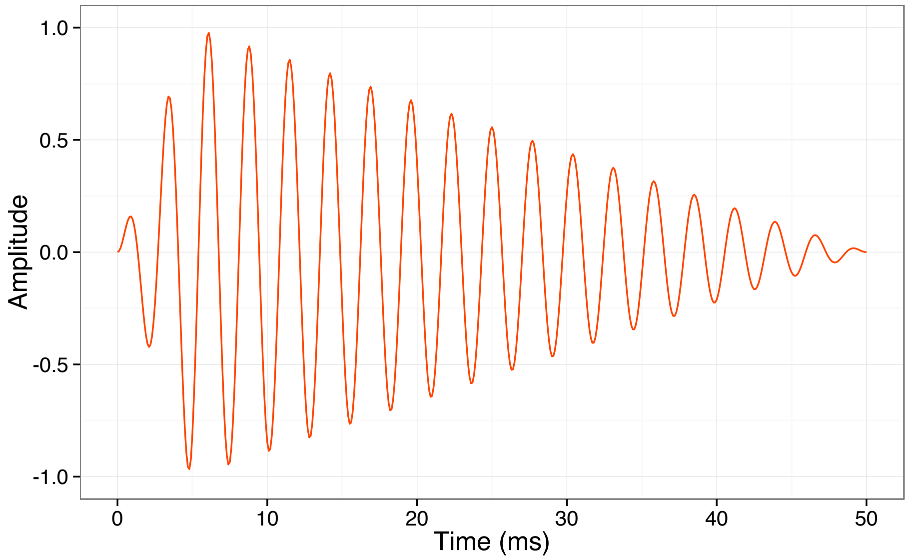



**Figure 5.9:** Audio signal for the final feedback design, showing how **Attack** (5ms), **Decay** (45ms), and **Frequency** (370 Hz) combine to create audible tones.

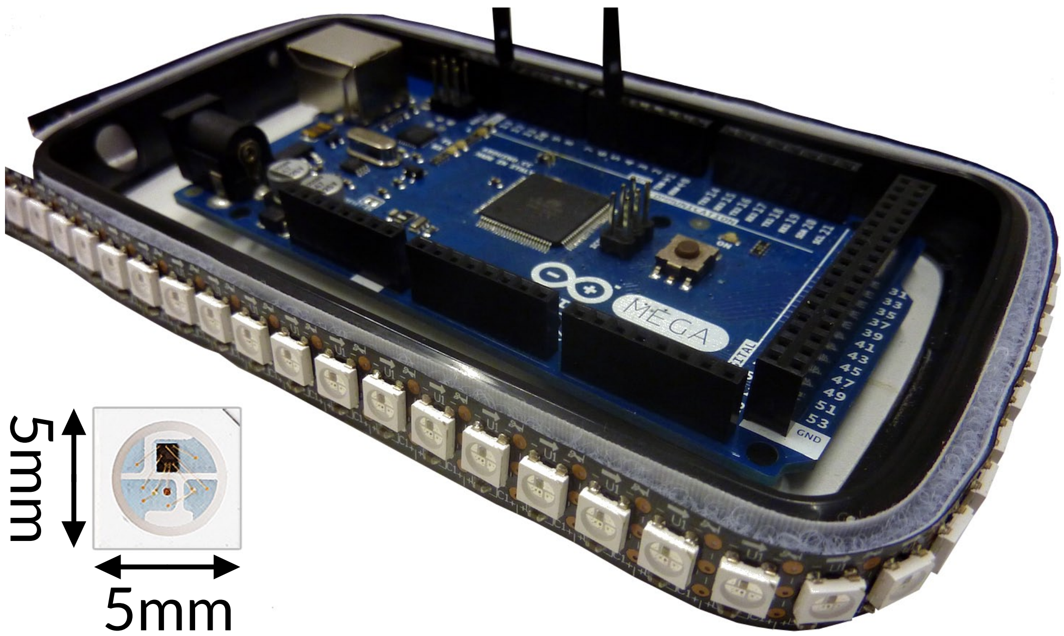



**Figure 5.10:** Interactive light display for Experiment 4. Sixty LEDs placed around the edge of a phone case illuminated the surrounding table surface for feedback. The LEDs were controlled by an Arduino microcontroller.

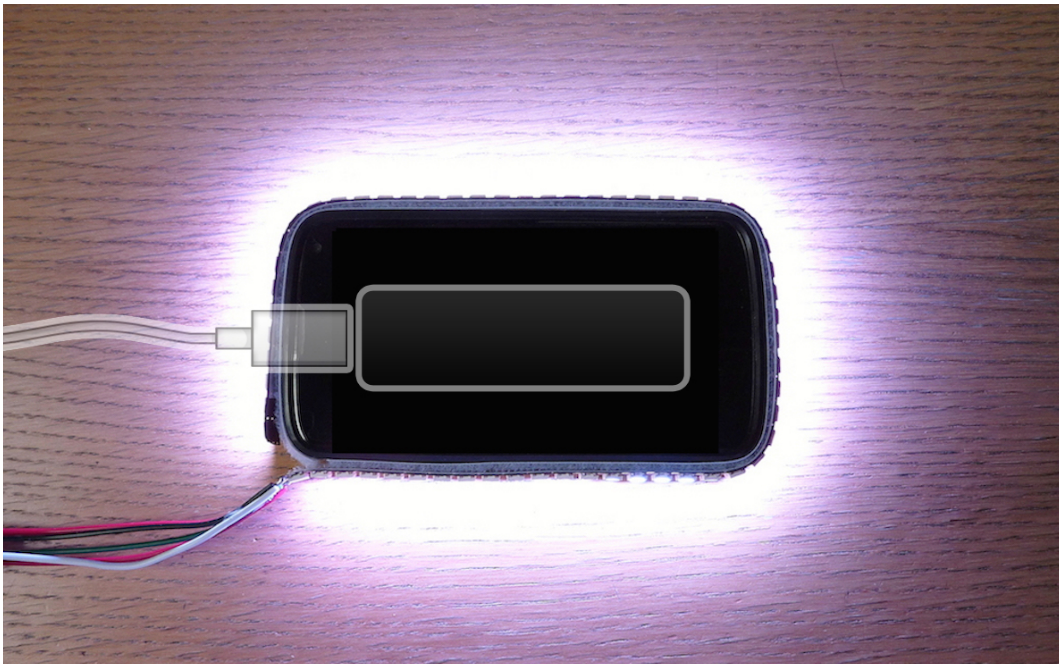


**Figure 5.11:** A mobile phone with the interactive light display from Figure 5.10. A Leap Motion sensor (overlaid in the foreground) was centred over the screen for hand-tracking.

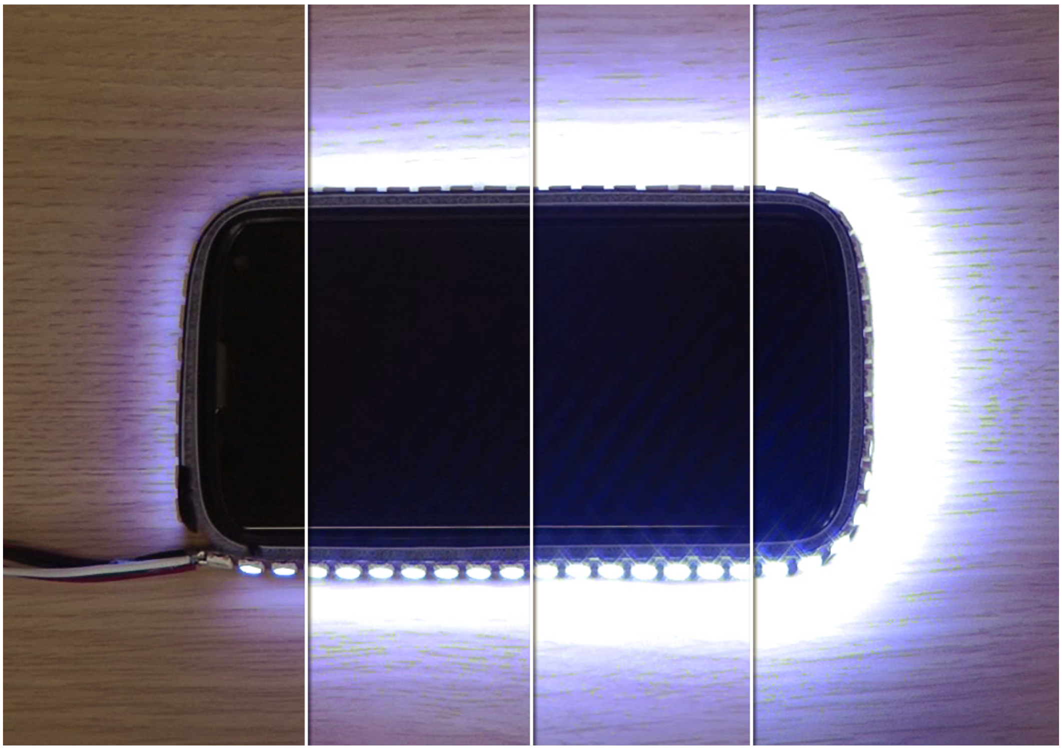


**Figure 5.12:** An example of varying brightness for sensor strength feedback.

The main aims of this experiment are to: (1) investigate the effectiveness of sensor strength feedback; (2) evaluate the feedback designs created during the focus group; and (3) understand what modalities users prefer for feedback when addressing a gesture-sensing system. These will contribute to the following research question:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

### 5.3.2 Experiment Design

In this experiment, participants were asked to complete tasks using all unimodal and multimodal combinations of sensor strength feedback, which used the designs described in Section 5.2.3. These combinations led to seven conditions for the within-subjects factor **Feedback**, listed in Table 5.1. Participants completed a block of 15 tasks for each condition and condition order was randomised using a Latin square design.

#### Task Design

Participants were asked to locate randomly-positioned target points as accurately as possible, guided by the cues given by each type of **Feedback**. Target points were placed randomly within a volume from (-100mm, 200mm, -100mm) to (100mm, 300mm, 100mm), relative to the Leap Motion sensor used for hand tracking; Figure 5.13 shows this volume relative to the sensor. This range was chosen so that hands would remain fully visible to the sensor during tasks. The sensor was placed in the centre of the prototype smartphone described in Section 5.2.4, meaning the target point range was centred over the device. Participants were asked to explore the space over the device using the extended index finger of their dominant hand; sensor strength feedback was given about the fingertip position relative to the target point. A single extended finger was used because this creates a point which can be reliably and accurately tracked by the Leap Motion sensor. Alternative landmarks for tracking, like palm position, were less accurately sensed.

| | Name | Feedback Type |
|---|---|---|
| 1 | Light Only (L) | Interactive light only |
| 2 | Audio Only (A) | Audio only |
| 3 | Tactile Only (T) | Vibration only |
| 4 | Light & Tactile (LT) | Interactive light and vibration |
| 5 | Light & Audio (LA) | Interactive light and audio |
| 6 | Audio & Tactile (AT) | Audio and vibration |
| 7 | Light & Audio & Tactile (LAT) | All modalities |

**Table 5.1:** Conditions in Experiment 4.

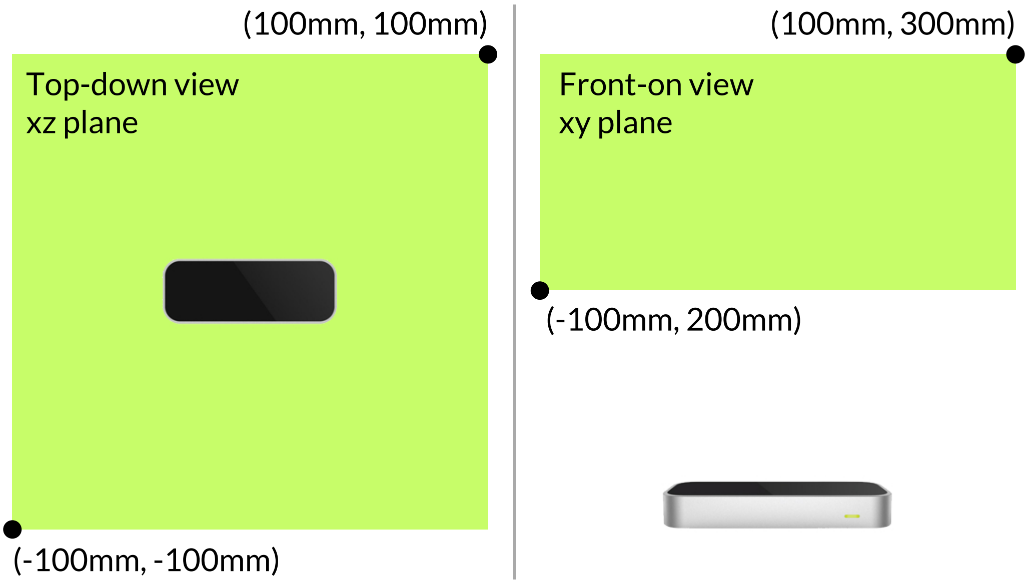



**Figure 5.13:** Task target points were positioned within a $20\text{cm}^2$ by 10cm volume.

This experiment used a targeting task so it could evaluate how accurately users could be guided by sensor strength feedback. This task design is not representative of how users would address a gesture system, as they may not need to be so precise. However, this experiment investigates how effective this feedback is and aims to understand how users make use of it, so these tasks can provide these insights. A later chapter, informed by these findings, uses sensor strength feedback alongside another gesture interaction for a more realistic usage scenario. Findings from this experiment could also inform the use of sensor strength-type feedback in other interaction techniques, where high precision is necessary. For example, the selection interface used in Experiments 1 and 2 could use similar feedback designs to help users make more accurate selections.

### Measures

For each task, the distance between the target point and the final hand position (*Distance*) was calculated, giving an indication of how accurate participants were. This was calculated as the Euclidean distance in three dimensions: $\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. Values for *Distance* are in mm, as the Leap Motion sensor provides hand position in real-world coordinates, rather than using a virtual coordinate system.

Task time (*Time*) was also measured, starting when participants raised their hand over the sensor and ending when they pressed the spacebar key to end the task. Participants were asked to rest their hand on the table surface between tasks, meaning they started each task by lifting their hand and reaching over the sensor.

At the end of the experiment session, participants were asked to give preference rankings for feedback modality (*Modality*) and feedback combination (*Feedback*), from favourite to least favourite. Participants were also asked to complete a survey after the experiment. This survey (shown in Appendix C) asked them to rate their agreement with five statements for each of the three feedback modalities. Participants responded using a five-point scale, from 'Strongly Disagree' (1) to 'Strongly Agree' (5). The statements were:

1. Changes in *light*/*audio*/*tactile* feedback were easy to notice.
2. I understood how *light*/*audio*/*tactile* feedback changed with my hand movements.
3. *Light*/*Audio*/*Tactile* feedback let me know my movements were being tracked.
4. *Light*/*Audio*/*Tactile* feedback helped me find the target point.
5. *Light*/*Audio*/*Tactile* feedback was annoying.

Participants were asked to rate their agreement with these statements because their responses could give insight into how effective each modality was for sensor strength feedback. These statements also probed how effective each type of feedback was for showing system attention to input and how acceptable the feedback was.

### Procedure

Participants were given a short tutorial at the start of the experiment, which explained the tasks they would be asked to complete and gave them a chance to become familiar with the feedback they would receive. During the experiment, participants sat at a table with the experiment apparatus and a desktop keyboard in front of them. They completed tasks using their dominant hand and wore the tactile display on that wrist. Tasks started when participants raised their hand over the gesture sensor and finished when they pressed the spacebar key on the keyboard with their non-dominant hand. Ending tasks this way meant that final hand position could be accurately recorded. They then returned both hands to the table before the next task started. At the end of the experiment, participants gave preference rankings and completed a short survey; both are described in the following section.

### Hypotheses

- **H1:** *Distance* will be higher for Light than Audio and Tactile;
- **H2:** *Time* will be lower for Light than Audio and Tactile.

**H1** expects participants to be less accurate with interactive light than with audio or tactile feedback. This hypothesis is based on speculation that subtle changes in brightness are likely to be less noticeable than the temporal changes in the non-visual cues. The audio and tactile modalities are known to have a high temporal resolution (e.g. [142]) so participants are expected to detect these changes in feedback with good accuracy. However, listening for changes in feedback over time means that participants are likely to spend more time finding target points with these modalities (as the feedback does not appear instantly, as with interactive light feedback); therefore **H2** predicts that participants will spend less time gesturing when just given interactive light feedback.

#### Participants

Sixteen people completed this study. Of these sixteen participants, six were female and two were left handed; their mean age was 26.1 years (sd 3.4 years). They were recruited using university email lists and were mostly university students. Each experiment session lasted one hour and participants were paid £6.

### 5.3.3 Results

#### Accuracy

Mean *Distance* was 51.4mm (sd 15mm); Figure 5.14 shows mean *Distance* for each type of feedback. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Distance*: $F(6, 90) = 15.83$, $p < 0.001$. *Post hoc* pairwise comparisons found that all types of **Feedback** containing audio (A, LA, AT, LAT) were significantly more accurate than all of those not containing audio (L, T, LT): all $t(90) \geq 3.47$, $p \leq 0.01$. No other pairwise comparisons were significant.

#### Time

Mean *Time* was 7174ms (sd 2819ms); Figure 5.15 shows mean task times for each type of feedback. *Time* was not normally-distributed (Shapiro-Wilk $W = 0.95$, $p < 0.001$), so data was transformed using the Aligned-Rank Transform [153] prior to analysis. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Time*: $F(6, 90) = 6.20$, $p < 0.001$. *Post hoc* comparisons found the following significant differences; no others were significant:

- L faster than A: $t(90) = -4.2$, $p = 0.001$
- L faster than AT: $t(90) = -4.69$, $p < 0.001$
- LT faster than A: $t(90) = -3.79$, $p = 0.005$
- LT faster than AT: $t(90) = -4.28$, $p < 0.001$

#### Preference

Table 5.2 shows the median ranks for feedback modality and feedback type, where a rank of ‘1’ was the most preferred option. Friedman’s rank sum tests were used to see if these

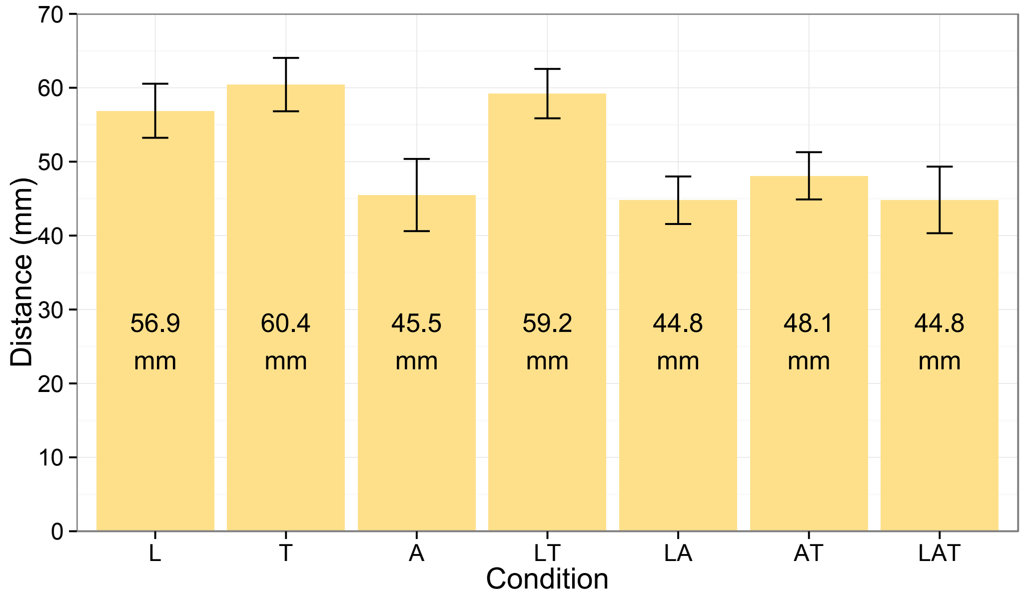



**Figure 5.14:** Mean *Distance* for each type of **Feedback**. Error bars show 95% CIs. (L = Light, A = Audio, T = Tactile)

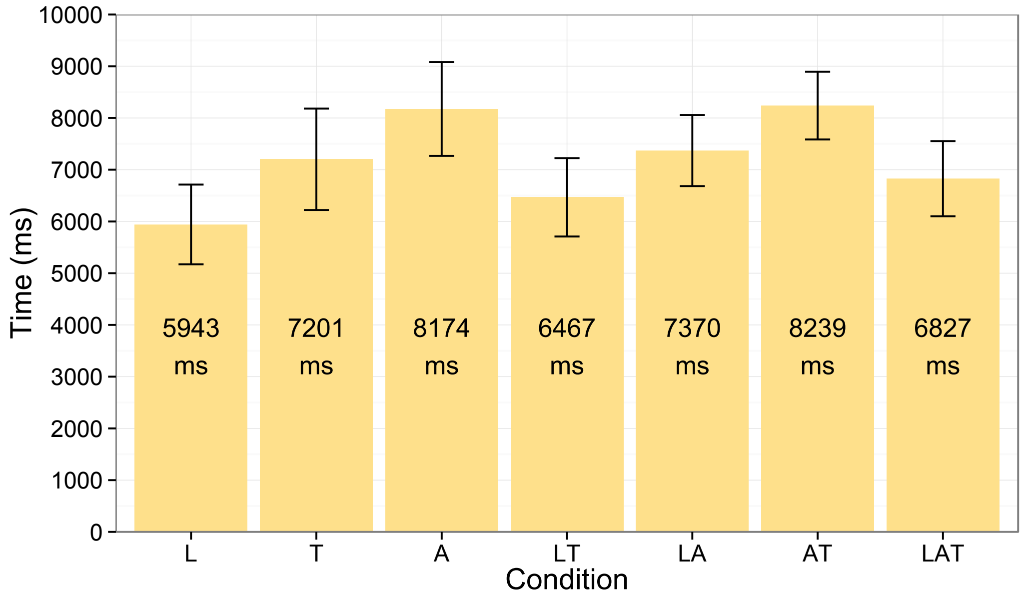



**Figure 5.15:** Mean *Time* for each type of **Feedback**. Error bars show 95% CIs.

rankings were significant or not. Friedman's test found that **Modality** rankings were significantly different: $\chi^2(2) = 11.62$, $p = 0.003$. *Post hoc* comparisons using Bonferroni-corrected Wilcoxon's signed-rank tests found that rankings for Audio were significantly higher than for Tactile ($W = 14$, $p = 0.009$). No other pairwise comparisons were significant (both $p \geq 0.05$). Although Light had a lower median rank than Audio, participants showed more disagreement in their rankings (interquartile range of 1 versus 0.25); therefore, rankings for Light and Tactile were not significantly different as well.

| **Modality Rank** | | | **Feedback Rank** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Light | Audio | Tactile | L | A | T | LT | LA | AT | LAT |
| 1 | 2 | 3 | 4 | 4 | 7 | 4 | 2 | 5 | 1.5 |

**Table 5.2:** Median ranks for feedback modalities and modality combinations.

A Friedman's test found that **Feedback** rankings were also significant: $\chi^2(6) = 42.96$, $p < 0.001$. *Post hoc* comparisons, using the same method as for **Modality**, found the following significant differences:

- LA greater than L: $W = 136$, $p = 0.009$
- LA greater than T: $W = 131$, $p = 0.02$
- LA greater than LT: $W = 4.5$, $p = 0.02$
- AT greater than T: $W = 136$, $p = 0.009$
- LAT greater than T: $W = 132$, $p = 0.02$

There was only one significant difference for LAT in comparison to three for LA, despite LA having a lower median ranking. This is because participants showed less agreement in their rankings for LAT (interquartile range = 2) than for LA (interquartile range = 1).

### Survey Ratings

This section discusses responses to the post-experiment survey questions about each feedback modality. Friedman's rank sum tests were used to compare ratings across feedback modalities. When significant differences were found, Wilcoxon's signed-rank test was used for *post hoc* pairwise comparisons, with Bonferroni-corrected p-values. Table 5.3 shows median responses to each survey statement.

#### *Q1: "Changes in feedback were easy to notice"*

Friedman's test found that ratings were significantly different between modalities: $\chi^2(2) = 13$, $p = 0.002$. *Post hoc* comparisons show that ratings for Audio were significantly higher than for Tactile: $W = 102$, $p = 0.002$. No other comparisons were significant (both $p \geq 0.16$).

| Question | Light | Audio | Tactile |
|---|---|---|---|
| *Q1: Changes in feedback were easy to notice* | 4 | 5 | 3 |
| *Q2: I understood how it changed with my hand movements* | 4.5 | 5 | 4 |
| *Q3: Feedback let me know my movements were being tracked* | 5 | 5 | 4 |
| *Q4: Feedback helped me find the target point* | 4 | 5 | 3 |
| *Q5: Feedback was annoying* | 1 | 2.5 | 2 |

**Table 5.3:** Median survey responses to each question for each type of feedback. Responses show agreement on a five-point scale, from "Strongly Disagree" (1) to "Strongly Agree" (5).

*Q2: "I understood how feedback changed with my hand movements"*

Friedman's test found that ratings were significantly different between modalities: $\chi^2(2) = 7.13$, p = 0.03. *Post hoc* comparisons found that ratings for Audio were significantly higher than for Tactile: W = 36, p = 0.04. No other comparisons were significant (both $p \geq 0.36$).

*Q3: "Feedback let me know my movements were being tracked"*

Friedman's test found that ratings were significantly different between modalities: $\chi^2(2) = 10.75$, p = 0.005. *Post hoc* comparisons found that ratings for Light were significantly higher than for Tactile: W = 51.5, p = 0.04. No other comparisons were significant (both $p \geq 0.11$).

*Q4: "Feedback helped me find the target point"*

Friedman's test found that ratings were significantly different between modalities: $\chi^2(2) = 15.69$, $p < 0.001$. *Post hoc* comparisons show that ratings for Audio were significantly higher than for Tactile: W = 78, p = 0.007. No other comparisons were significant (both $p \geq 0.05$).

*Q5: "Feedback was annoying"*

Friedman's test found that ratings were not significantly different: $\chi^2(2) = 5.78$, p = 0.06.

### 5.3.4 Discussion

Participants located target points more accurately in conditions with Audio than in those without; based on this result, **H1** cannot be accepted. **H1** expected changes in brightness to be more difficult to notice than changes in the temporal properties of the non-visual feedback, leading to lower accuracy for Light. While this was true for Audio, participants were not more accurate with Tactile feedback. Survey responses show that changes in Tactile cues were more difficult to discern than Audio cues (survey Q1 and Q2), despite them using the same temporal pattern. Two participants also said that when intervals were shorter, vibration often felt continuous to them. This difference between Audio and Tactile perception may be because temporal gap detection is better for the auditory modality than for tactile [59, 30];

the Geiger counter feedback may have been too frequent at higher sensor strength values for participants to perceive it accurately. A future implementation of sensor strength feedback, used in Experiment 6, will address this issue by increasing the minimum interval between feedback stimuli.

Accuracy with Audio often came at the expense of task time; task times were significantly longer for conditions with Audio and without Light (A and AT), than for conditions with Light and without Audio (L and LT). Based on this finding, **H2** also cannot be accepted. **H2** expected participants to take longer with Audio and Tactile because these conditions would allow them to be more precise, thus spending more time trying to locate target points. However, this was not the case for Tactile feedback. Difficulty in perceiving tactile cues, as discussed previously, may have caused participants to spend less time adjusting their hand position.

Time and accuracy data revealed two trends: participants were faster but less accurate with interactive light feedback; and they were slower but more accurate with audio feedback. Combinations of these modalities could be effective as users would benefit from faster localisation (from light) and greater accuracy (from audio cues). All but two participants ranked combinations of Light and Audio (LA and LAT) the highest, with many suggesting that they preferred those feedback designs for that reason. Seven of the sixteen participants reported that they preferred these combinations because they could use these modalities together to quickly and accurately locate targets. Although small changes in Light were difficult to perceive, participants could quickly tell when their hands were in the "right area", as the difference between low and high brightness was easily noticed. Survey responses (to Q3) support this, showing that Light was especially effective for informing users when their hand movements could be sensed. Participants then used Audio cues for more precise adjustments, "homing in" on the target points. This finding shows one of the benefits of multimodal feedback about gestures: different modalities have different strengths which users can take advantage of during interaction, e.g. using the high temporal resolution of the audio feedback for precision in this experiment.

Participants generally found Tactile feedback least helpful (survey Q4), with survey responses and post-experiment comments suggesting that this was because the feedback was more difficult to perceive and understand. However, some said that they would rather use tactile feedback than audio feedback when around others, as it would only be noticeable to them. Audio was also rated as most annoying (survey Q5; however, this rating was not significant), with one participant saying it was irritating because it was "monotonous". A later implementation, for Experiment 6, will use a new sound design which aims to be more acceptable.

## 5.4 Limitations

### 5.4.1 Interaction Time

For users to address a gesture-sensing system, they must do two things: they need to find where to gesture so their actions can be sensed and they need to direct their gestures towards the system they wish to interact with. This experiment focused on the first goal and evaluated a technique which helped users find where to gesture by telling them how well they can be seen. However, the mean task time from this experiment (seven seconds) is longer than users would be expected to spend finding where to gesture when they address a gesture system. In normal use, finding where to gesture may require less precision than that requested by the experimental tasks; users would use feedback to find the "right area" without necessarily having to "home in", as they did here.

This task design was chosen as it allowed the effectiveness of each type of sensor strength feedback to be investigated fully and for the first time. Had a more representative task been used, where participants only imprecisely interpreted feedback over a shorter period of time, less would have been learned about the strengths and weaknesses of the feedback. Findings from this experiment are used to inform the design of a new interaction technique, discussed in Chapter 7, which looks at the entire problem of addressing a gesture system. Study of that technique, in Experiment 6, presents a more realistic use of sensor strength feedback, addressing this limitation.

### 5.4.2 Interaction Over a Smartphone

In this experiment, participants gestured over a gesture-sensing smartphone from close proximity; target points were located within a $20cm^2$ area over the device, with height ranging from 20cm to 30cm. Participants explored this space using an extended index finger. Findings from this experiment may not necessarily apply to other types of gesture-sensing system, where users gesture from a greater distance and over a wider area than used here. Users gesturing at a device from across the room will likely use larger hand and arm movements than they would if gesturing at a device almost within reach. Sensor strength feedback designs were also configured to work with this type of interaction so it is unclear if these designs will also be effective when interacting from a greater distance or with larger hand movements.

As such, **RQ3** can only be partially answered by this experiment; more work is needed to understand how effectively gesture interfaces can guide users when making larger arm movements. Research in Chapter 7 addresses this limitation by investigating the use of sensor strength feedback when interacting with a wall-mounted device from a greater distance, allowing a complete answer to be given for **RQ3**.

## 5.5 Conclusions

This chapter investigated sensor strength feedback, a novel interaction technique which guides users as they address a gesture-sensing system. Multimodal feedback — comprising interactive light, sound and vibration — means limited-display devices can use this technique without requiring feedback on a screen. Participants in Experiment 4 found sensor strength feedback helpful and easy to use, especially when interactive light and audio were combined. This combination was especially effective because it allowed them to quickly find where to position their hand (using interactive light feedback) before making precise adjustments (using audio feedback).

Users may not always need help to find where to gesture. As they become more familiar with a particular gesture system, sensor strength feedback may become less useful as they know from experience where to perform input. In these cases, users would be able to go straight to providing input and could ignore (or even disable) the feedback. Finding where to gesture is also only part of addressing a gesture-sensing system. Users must also be able to direct their actions towards the system they wish to interact with, to avoid unintentionally affecting other gesture interfaces. Research is needed to investigate how users may do this and how a possible solution can work alongside sensor strength feedback.

### 5.5.1 Research Question 3

The research in this chapter contributes a partial answer to the following question:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

Findings from Experiment 4 show that limited-display devices can use multimodal feedback to accurately and effectively guide their users when they address them from a short distance away. Participants achieved a mean accuracy of 51.4mm, with the best performing feedback guiding them to within 45mm of target points. However, more research is needed to understand how effective this interaction technique is when users are gesturing from a greater distance (a limitation discussed in Section 5.4.2). Research in Chapter 7 will allow a more complete answer to be given for **RQ3**.

### 5.5.2 Contributions

The research in this chapter describes and evaluates *sensor strength feedback*, an interaction technique for guiding movement as users begin to address a gesture system.

# Chapter 6

# Directing Input Towards A Gesture-Sensing System

## 6.1 Introduction

When addressing a gesture-sensing system, users need to know where to perform gestures and how to direct those gestures towards only that system. Sensor strength feedback, an interaction technique described in Chapter 5, provided a solution to the first problem, by guiding users to gesture in a location where their hand movements can be reliably sensed. However, an accompanying interaction is needed to allow users to direct their gestures towards a particular system. Directing input is necessary because, without it, users may unintentionally affect other gesture-sensing systems; this is known as the *Midas Touch* problem [70]. By directing gestures towards a particular system, it then knows that subsequent hand movements should be treated as meaningful input; other systems also know that they should not be detecting gestures because they have not been addressed. Existing solutions to the *Midas Touch* problem have limitations which mean they would be impractical in use. Section 2.4 of the Literature Review discussed these solutions — pointing, activation gestures, active zones, and gaze — and their limitations, motivating the following research question:

> **RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

The research in this chapter investigates a novel interaction technique — *rhythmic gestures* — which allows users to direct their input towards a gesture-sensing system, overcoming the limitations found with existing approaches. This technique also considers the limited display capabilities of many gesture-sensing devices, meaning it can be used by a variety of devices, from small household controls and mobile devices, to entertainment systems with large displays.

### 6.1.1 Chapter Structure

Section 6.2 discusses rhythmic gestures, an interaction technique which allows users to direct their actions towards a particular gesture-sensing system. Section 6.3 describes Experiment 5, which investigates how well users can perform rhythmic gestures and looks at how feedback affects gesture performance. Section 6.4 identifies limitations of Experiment 5. Section 6.5 gives design recommendations for using rhythmic gestures and also discusses an answer to the research question mentioned before.

## 6.2 Rhythmic Gestures

Users must be able to direct their gestures towards a system when they wish to interact with it, so that their actions do not unintentionally affect other systems at the same time. This must also be done in a way which shows intent to interact, so that everyday movements — like reaching out and lifting a cup of coffee — are not treated as meaningful input. Other interaction techniques for addressing gesture-sensing systems, reviewed earlier in this thesis, do not meet these criteria. *Active zones* are prone to detecting ordinary hand movements, as they are unable to distinguish between interactive and non-interactive actions. Inferring input from *gaze* and *pointing* can be ambiguous, meaning users cannot reliably direct their gestures as they intend to. Finally, *activation gestures* cannot be assumed to only occur during interaction, they must be unique to each interface, and users either need to learn these obscure gestures or need a way of discovering which gestures activate each interface. This thesis proposes an alternative interaction technique — *rhythmic gestures* — which can avoid the issues of false-positive recognitions and ambiguity, while allowing users to direct their input and show intention to interact. They could be considered as a special kind of activation gesture, where rhythmic repetition is part of the gesture criteria.

Rhythmic gestures are hand movements which are repeated in time with a rhythmic 'beat'. For example, a rhythmic gesture may involve waving a hand from left-to-right, and back again, once every two seconds. These gestures would not be too cumbersome, as they would only require a few seconds (e.g. three or four repeated movements) prior to input. Rhythmic gestures have two components: (1) a hand movement; and (2) an interval. A variety of hand movements could be used for gestures, from simple ones like waving from side to side, to more complex finger movements, like waving an extended finger from side-to-side or extending and closing all fingers at once. Only simple hand movements are considered in this thesis (discussed later in Section 6.2.2), although rhythmic gestures have a large potential design space which extends beyond these. A rhythmic gesture interval is the time between successive hand movements. When users gesture in synchrony with the interval of a rhythmic gesture, they are *matching* its rhythm.

Users could use rhythmic gestures to direct their input towards gesture-sensing systems; each system would have its own rhythmic gesture which, when matched, would let that system know that subsequent actions should be treated as input. At the same time, other systems would know to ignore sensed movements, as their own rhythmic gestures were not performed.

Users would need to know which rhythmic gesture to perform to address the system they want to interact with. Gesture-sensing systems could communicate their gestures using simple visualisations, which would be shown on screens or interactive light displays; this means limited-display devices could also use rhythmic gestures. These visualisations have to reveal which movement pattern to use and at what speed it should be performed, so that users can gesture in time with the rhythm. Interactive light displays could use the location of light to show movement patterns, creating animations which users must mimic with their hands. Figure 6.1 shows an example of how an interactive light display may reveal gesture movements. The speed at which these animations repeat reveals the gesture interval; by moving their hands in time with the animation, users would match the rhythmic gesture. Audio and tactile displays could also be used to reveal rhythmic gestures. However, they would require spatial output capabilities which may not be available; that is, users may not be wearing headphones or a sophisticated enough tactile display. Since only light can easily communicate position and movement, it was treated specially in this case. Only interactive light will be investigated here, as it does not require users to have any additional accessories for feedback.

Gesture-sensing systems would have to coordinate their choice of rhythmic gestures with each other, ensuring that gestures are unique. Coordination between systems would also be required for other existing techniques, so this is not necessarily a disadvantage of using rhythmic gestures. For example, if activation gestures are used then coordination is required to make sure gestures are unique; for pointing- and gaze-based techniques, coordination is required so that systems can determine which one users are most likely to be addressing. This thesis does not investigate communication and coordination between gesture-sensing

**Figure 6.1:** An interactive light display could communicate gesture movements using light animations, showing how and how quickly to move the hand. This example shows how moving the hand from side to side could be communicated.

systems for rhythmic gestures; instead, it focuses on the usability of this interaction.

Using rhythmic gestures to address a gesture system is similar to using activation gestures, although it overcomes limitations with that technique. Rhythmic gestures are repeated at a known interval, so systems are able to distinguish between interactive and non-interactive hand movements, whereas activation gestures may still occur in non-interactive contexts; for example, O'Hara *et al.* [101] found that their carefully-chosen activation gesture was still performed accidentally. Rhythmic gestures allow the same gesture movements to be used by many systems, with interval being unique to each one; this means users need to learn and use less different gestures. Finally, gesture movements and intervals can be communicated using simple interactive light displays (or screens, if available), helping users discover which gestures to use; activation gestures are often complex hand or body poses which would be difficult to communicate without good visual output capabilities.

Research is needed to understand how well users can perform rhythmic gestures, so that gesture systems can choose movements and intervals which they can use with ease. Possible gesture movements were identified earlier in this section, although suitable gesture intervals must also be chosen. The following section reviews literature on two topics: (1) the use of rhythm in user interface design; and (2) our ability to perform movements in time with a rhythm. Knowledge from these areas will inform the selection of appropriate gesture intervals, leading to the design of a set of rhythmic gestures which will be evaluated in Experiment 5, discussed later in this chapter.

### 6.2.1 Related Work

#### Rhythm in User Interface Design

Research has investigated the use of rhythm in user interface design as rhythmic patterns in input can offer control when other input capabilities are limited. For example, *Motion-Pointing* [33] allowed users to select targets without pointing at them, so they could input when a pointing device was unavailable. Users selected targets with a mouse by imitating continuous elliptical movement patterns which were unique to each target. Their pointer-less pointing technique built on an earlier one by Williamson *et al.* [147], where users imitated erratic movements rather than elliptical ones. Evaluation of *Motion-Pointing* found that users were able to match rhythmic patterns without any visual feedback, relying instead on proprioception to modulate their movements. Users took around 1–1.5 seconds to match rhythmic movements, which they then sustained for around 1–2 seconds.

The use of continuous elliptical movements for input is described by Malacria *et al.* [82] with their *CycloStar* concept. They discussed how continuous elliptical movements can be used to modulate up to seven independent variables at once. Fundamental to the *Cyclostar* concept is that humans can easily perform and modulate circular movements in time with a rhythm. This allows users to control aspects of an interface through the properties of their circular movements, rather than the location of their movements. A benefit of this is that users could provide rich input when a limited amount of space is available for input.

Repeated elliptical movements are mechanically efficient because energy is conserved between motions; as users end one movement and begin the next, stored potential energy is turned into kinetic energy [39]. Elliptical movements, in this case, also include moving in a straight line, because a straight line is a fully eccentric ellipse. Users can also match rhythmic motions without any visual feedback about their own movements, relying on proprioception instead. Beek and Lewbel [10] demonstrate this with a look at juggling. They discuss how jugglers do not need visual feedback about their own hand movements; instead, a short glimpse of a ball's trajectory is sufficient for them to control their rhythmic hand motions. Fekete *et al.* [33] reported a similar finding with *Motion-Pointing*, where users could imitate rhythmic elliptical movement without any visual feedback about their own hand movements.

Bennett *et al.* [14] described *harmonic interactions* for tangible and mobile devices. These interactions are repetitive movements of a handheld device, performed in time with some stimuli. To see if harmonic interactions would work as an interaction technique, they asked users to resonate with virtual pendulums on a mobile phone. As users tilted the phone (towards or away from their body), pendulums on the display reacted to their movements. If users moved the phone in harmony with a particular pendulum, its amplitude would increase whilst the others' movements were dampened. This technique uses rhythm in a similar fashion to rhythmic gestures; users make a selection by synchronising with a rhythm. They found that users were able to isolate a target pendulum with greater accuracy when there were fewer pendulums to harmonise with and that users were able to control lower frequency pendulums with greater accuracy. They noted that users experienced difficulty and frustration at higher frequencies, although this is unsurprising as their maximum frequency (4 Hz, repeating every 250ms) exceeds the synchronisation threshold of a visual stimuli (460ms) [113]; this will be discussed more in the next section.

These interaction techniques all used rhythm to enhance interaction possibilities when input was limited. Repeating movements in time with a rhythm allowed users to select targets without a pointing device and when elliptical motions were used, users could manipulate several properties at the same time, by varying how they moved. Research in this area has highlighted the potential benefits of using rhythm in user interface design, although there is little advice given on *how* to use rhythm in interaction design. This review now turns to psychophysics research which investigates how humans perceive rhythm and move in time with it, better known as *sensorimotor synchronisation*. Knowledge from this area can provide insight into how to use rhythm in interaction and can provide a starting point for designing rhythmic gestures.

### Sensorimotor Synchronisation

Sensorimotor synchronisation is the process of perceiving rhythm and moving the body in time with it; a simple example of this is tapping a foot along with the beat in a piece of music. Research from this area can provide insight into the design of rhythmic gestures, as much is known about the limits of moving in time with rhythm and how to make synchronising with a rhythm easier. This review begins by looking at synchronising movements with visual and audio stimuli, as this can inform the design of effective feedback about gestures. It then considers the speed at which movements can be accurately synchronised, to inform the selection of rhythmic gesture intervals. Finally, it discusses cognitive processes for time-keeping, which can further inform the design of feedback about gestures. This review ends with a set of design recommendations based on knowledge from this literature.

#### *Synchronisation with Visual and Audio Cues*

Audio cues are more effective than visual cues for synchronising discrete movements, like finger taps [58, 115, 57], although it is unclear which modality performs best for synchronising continuous movements, like rhythmic gesture movements. Armstrong *et al.* [3] found that visual cues were more effective than audio cues for synchronising continuous movements, although earlier work found that discrete audio outperformed visual [138]. These contradictory results mean it is unclear which modality would be best for communicating gesture rhythm. However, visual cues can easily represent movement patterns, while communicating this information using sound would require complex spatial audio designs.

Research has found that hand movements are synchronised more accurately when a rhythm is presented as a continuous visualisation (for example, a light which moves between two points or a light which fades in and out; Figure 6.2), rather than presented as a discrete one (for example, a light which flashes briefly on each beat). While this is also true when visualisations do not have spatial cues about movement patterns [138], synchronisation is best when spatial and temporal information are combined [58, 4, 57, 3]; this suggests that the moving circle in Figure 6.2 would be more effective than the fading one, when the hand movement is from side to side. Continuous visualisations can also lead to more effective synchronisation for discrete movements, like finger tapping [57]. These findings suggest that rhythmic gestures should be revealed using continuous visualisations; these visualisations should reveal the movement pattern as this shows users which gesture to use and also leads to better synchronisation performance, by combining spatial and temporal information.

#### *Effective Synchronisation Speeds*

Repp [113] reviewed the sensorimotor synchronisation literature on timing limits for effective synchronisation. He notes that the lower interval limit for both audio and visual synchronisation is around 1,800ms; at slower intervals, movement becomes more of a response

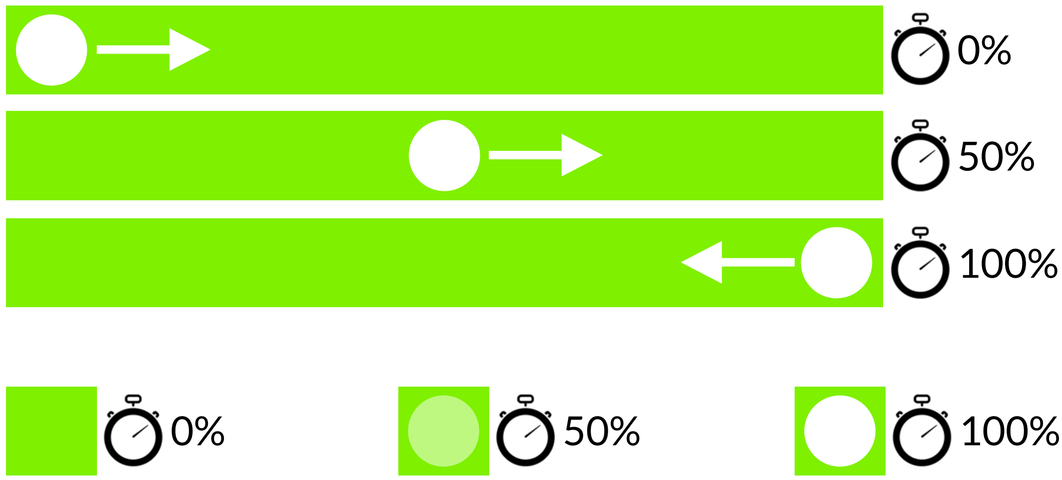



**Figure 6.2:** Two examples of continuous visualisations for rhythmic hand movements. Spatial and temporal cues (top) communicate movement pattern as well as timing information, whereas temporal cues (bottom) only show timing.

to stimulus than anticipated movement in time with a rhythm. Upper interval limits for synchronisation vary by modality: for audio cues, the threshold is around 120ms, whereas for visual cues, the threshold is around 460ms. Informed by these findings, rhythmic gestures should use intervals between 460ms and 1,800ms.

van der Wel *et al.* [137] found that as the interval of a rhythm increased, people avoided making increasingly slower movements. Instead, they moved at speeds closer to their preferred movement speed (PMS) and then paused until the end of the interval. These changes in movement kinematics started to occur when intervals increased past 800ms; the mean preferred interval in their study was 650ms (sd 175ms). This finding has implications for rhythmic gesture design: sensing algorithms should allow pauses between gestures, as users may move their hands closer to their PMS if an interval is too long. However, if they are allowed to pause for too long between movements then it will become difficult to determine if they are matching a rhythmic gesture interval, or not.

Armstrong *et al.* [4] found that synchronisation with continuous hand movements is worse at 80% PMS than at 120% PMS. In a later study [3], they reported similar findings: synchronisation was better at 120% PMS than at 80% and 100% PMS. These findings suggest that gesture intervals should be chosen which are similar to, or faster than, a user's PMS, as this is where movements are most effectively synchronised. This also avoids ambiguity caused by moving too quickly when slower intervals are used.

Rhythm interval also affects how people perform hand movements. Repp [114] looked at how repeated hand movements varied when participants were allowed to move without any constraints, finding that they made smaller movements at faster intervals and larger movements at slower intervals. This suggests that rhythmic gesture sensing should be flexible in terms of movement size, as users are likely to perform smaller gesture motions for faster rhythms.

### *Internal Timekeeping*

Separate cognitive processes are thought to control the synchronisation of discrete and continuous body movements. Discrete movements, like tapping a foot, are timed to coincide with predictable events, like the next beat in a rhythm. Continuous movements, like waving from side to side, are synchronised using an emergent timing process, which involves kinematic information about movement [129, 130, 80]. This means that movements start out of time with a rhythm and eventually become synchronous, with the body using its knowledge about how it is moving to correct errors in timing.

Studenka *et al.* [129] found that discrete perceptual events — short tactile cues, in this case — improved error correction when synchronising hand movements. Discrete movements had slower error correction when tactile feedback was removed. Continuous movements, which use an emergent timing process, described above, had faster error correction when tactile feedback was added. From this, the authors suggest that event-based timing depends on discrete perceptual events rather than being determined by knowledge of movement. In a later study [130], they found that discrete tactile feedback may lead to discrete, event-based timing for continuous movements; again, because of faster asynchrony correction. These findings suggest that discrete 'events' during continuous hand movements, feedback or otherwise, could improve synchronisation with a rhythm.

Some continuous movements may have 'events' in the form of movement reversal at physical limits [128]; for example, when waving from side to side in front of the body, reaching the comfortable limits of range of motion may be physical cues which help in synchronisation. Rhythmic gestures which feature these physical cues (like waving a hand from side to side, or continuously raising and lowering a hand, for example) may be easier to synchronise than those which do not (like moving the hand in a circular path, for example). Gestures which lack such physical cues may benefit from additional discrete feedback at the end of each complete movement [129, 130], as these can aid in correcting movement timing errors. Extra feedback should be given at the end of each rhythmic gesture movement as this may help users match the rhythms.

## Summary of Related Work

Research has found many benefits of using rhythm in human-computer interaction, including mechanical efficiency [39], effective feedback from proprioception [33], several degrees of freedom [82] and the ability for input when the space available for interaction is constrained [147, 33, 82]. However, that work gave little insight into how to use rhythm in interaction design effectively. The psychophysics literature on sensorimotor synchronisation, instead, provided many insights about moving in time with a rhythm. Initial design guidelines for rhythmic gestures were identified during the discussion of this literature, providing a start point for designing and implementing rhythmic gestures:

- Use continuous visual cues to reveal rhythmic gesture movements and intervals;
- Select gesture intervals between 460ms and 1,800ms;
- Allow short pauses between gesture movements if intervals are longer than preferred movement speed;
- Choose intervals which are close to, or faster than, preferred movement speed;
- Allow variation in gesture movement sizes;
- Give discrete feedback at the end of each gesture movement.

### 6.2.2 Rhythmic Gesture and Feedback Designs

This section now describes the rhythmic gestures chosen for this experiment. It begins by discussing the gesture design space and then describes feedback for the gestures.

#### Gestures

Five hand motions were chosen for rhythmic gestures: Side-to-Side (SS), Up-and-Down (UD), Forwards-and-Backwards (FB), Clockwise (C) and Anticlockwise (AC). These are illustrated in Figures 6.3–6.6; annotations show how gesture intervals relate to the movement patterns, using a 500ms interval as an example. These gestures were chosen because they are continuous elliptical movements (a straight line — found in SS, UD and FB — is just a fully eccentric ellipse). More complex gesture shapes, like moving the hand along a square or triangular path, require a sequence of discrete movements rather than one continuous movement. Although discrete movements could be used for rhythmic gestures, they lack the mechanical efficiency of continuous ones [39] so would require more effort.

Four intervals were chosen for the initial implementation of rhythmic gestures: 500ms, 700ms, 900ms and 1100ms. These intervals were selected because they fall within the range for effective synchronisation with a visual rhythm (460–1,800ms) but are not so long that interaction is unnecessarily time consuming. Rhythmic gestures use repetition as a means of showing intent to interact, avoiding the problem of normal hand movements being treated as input. Hand movements were therefore considered as intentional input after three repetitions. For circular gestures, this was three complete circles; for the others, this was three individual movements. If users performed three sequential repetitions of a gesture pattern in time with the rhythm interval then their gesture was considered to be successfully synchronised with the rhythm. Depending on the gesture interval, this means users must gesture for

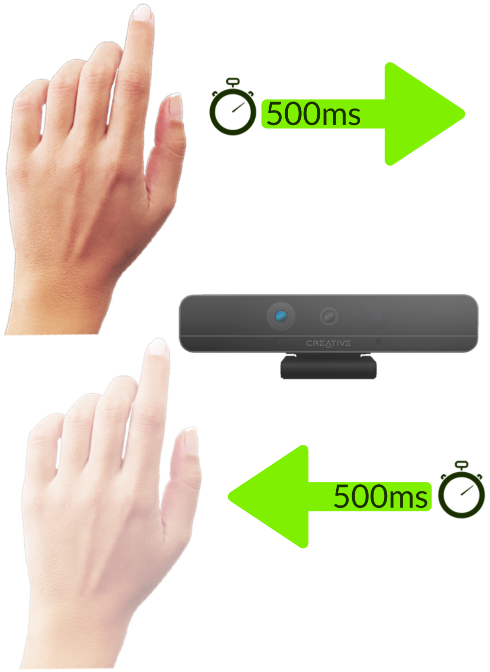



**Figure 6.3:** *Side-to-Side* (SS): continuous hand movement from left-to-right and back again. Here, a gesture interval is the time between starting and stopping a movement.

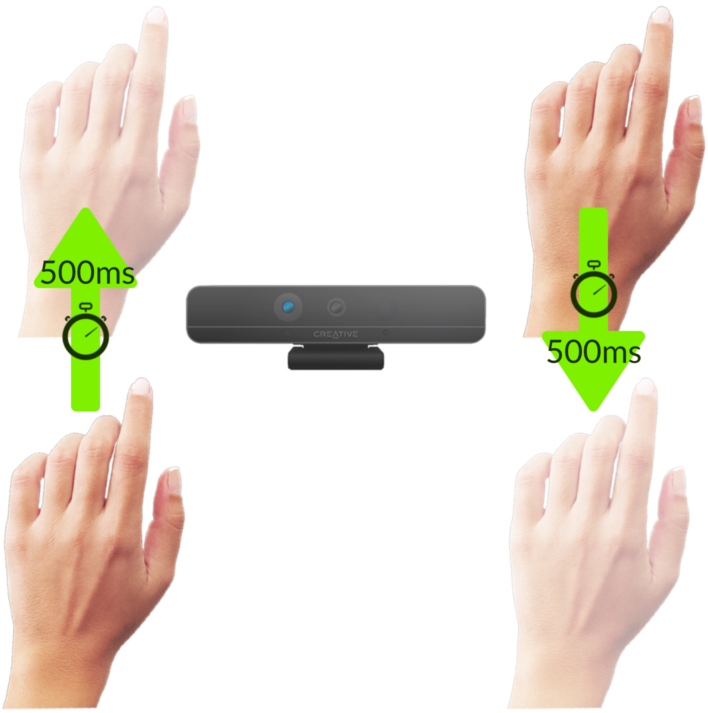



**Figure 6.4:** *Up-and-Down* (UD): raising then lowering one hand continuously. Here, a gesture interval is the time between starting and stopping a movement.

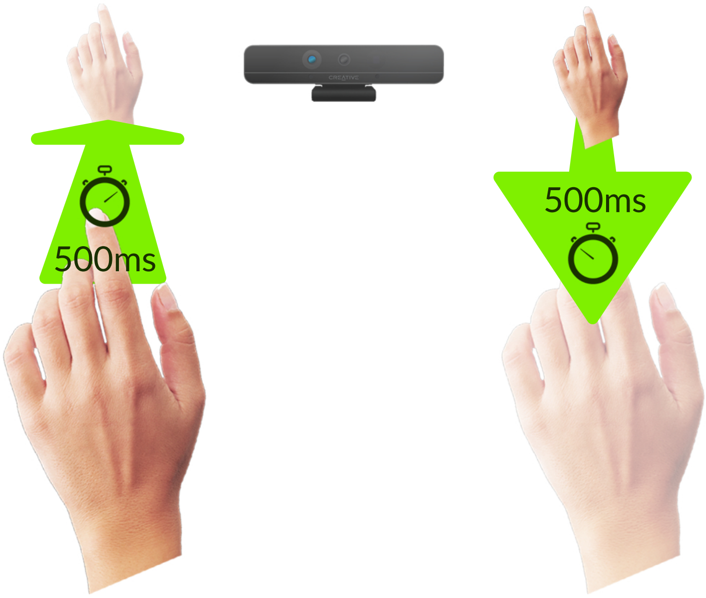



**Figure 6.5:** *Forwards-and-Backwards* (FB): continuous hand movement towards the sensor and away again. Here, a gesture interval is the time between starting and stopping a movement.

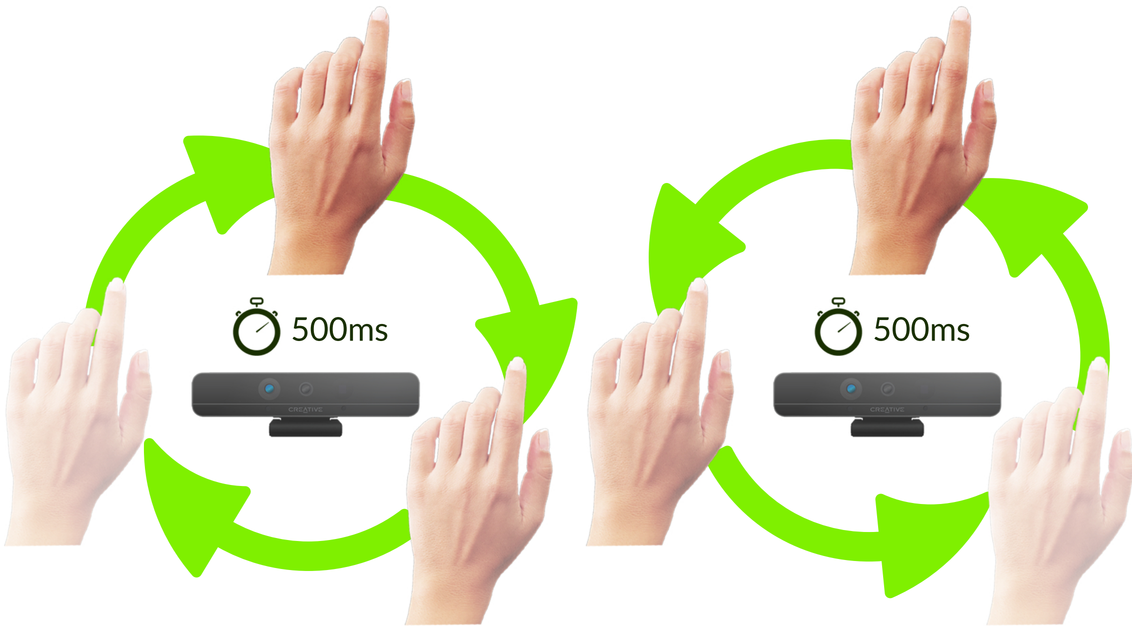



**Figure 6.6:** *Clockwise* (C, left) and *Anticlockwise* (AC, right): continuously moving the hand in a circle. Here, a gesture interval is one complete circular movement.

at least 1.5s–3.3s when addressing a gesture system. While this could be shorter (two repetitions), waiting for three repeated movements in time gives greater confidence that sensed movements are actually gestures.

#### Feedback

Based on the recommendation that gestures be revealed using continuous visualisations, animations were created for each of the five gesture movements. These animations were designed for interactive light displays which illuminate the space around devices, like those used in Experiments 3 and 4. Figures 6.7–6.11 illustrate these interactive light animations for the wall-mounted circular LED display used in Experiment 3. Animations used white light to show movement patterns to illuminate surrounding surfaces (as in Figure 6.1), although the illustrations use the colour blue. Once users match a rhythmic gesture, the animation is shown with green light instead (as in Figure 6.12).

Users were given audio and tactile feedback about their gesture movements, as recommended following the literature review earlier in this chapter. Feedback was given after each complete movement; for example, when their hands reached their leftmost or rightmost point when waving from side-to-side, or after completing a circular path. When users are gesturing in time with the rhythm, feedback will coincide with the "beat" at the end of each animation. This feedback was a 200ms tone (at 523 Hz, C5 on a piano) and a 200ms vibration (at 150 Hz, the resonant frequency of the actuator used). Discrete cues were used as these can lead to more accurate timing when moving in time with a rhythm.

### 6.2.3 Implementation and Apparatus

Rhythmic gesture animations were implemented for the Gesture Thermostat, the wall-mounted interactive light display used in Experiment 3. An Xbox Kinect sensor was used for tracking hand movements, as this would be able to sense gestures over the wide area used when interacting with the wall-mounted device prototype; both are shown in Figure 6.13. A C#

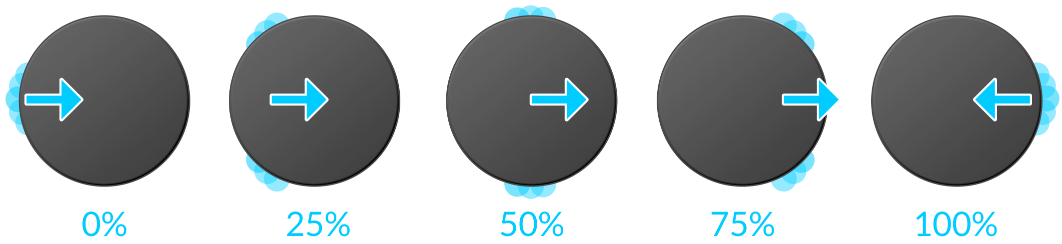



**Figure 6.7:** *Side-to-Side* (SS): an area of white light moves from left-to-right and then back again, in reverse; this animation uses blue light for illustration only.

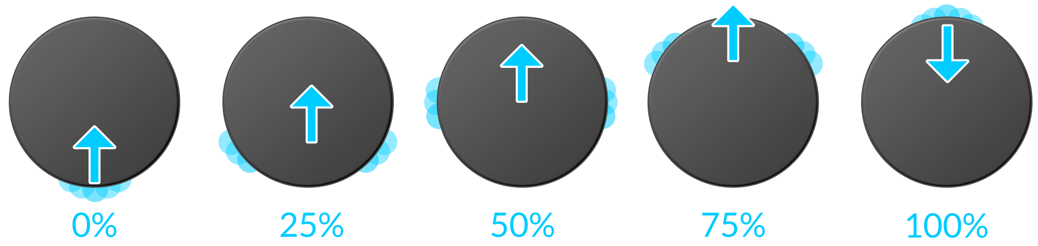



**Figure 6.8:** *Up-and-Down* (UD): an area of white light moves from bottom-to-top and then back again, in reverse; this animation uses blue light for illustration only.

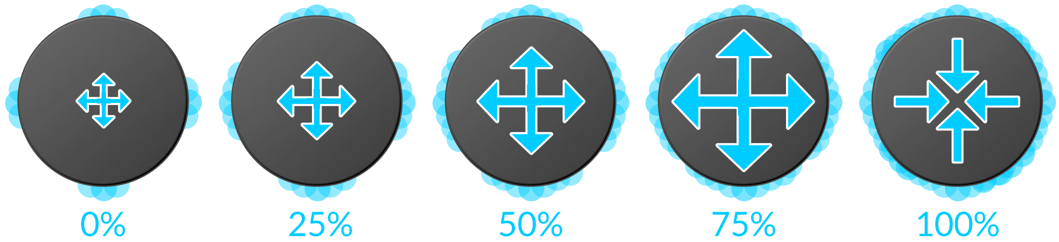



**Figure 6.9:** *Forwards-and-Backwards* (FB): areas of white light at the top, bottom, left and right of the device expand in size, until all lights are at full brightness. This then reverses. This animation uses blue light for illustration only.

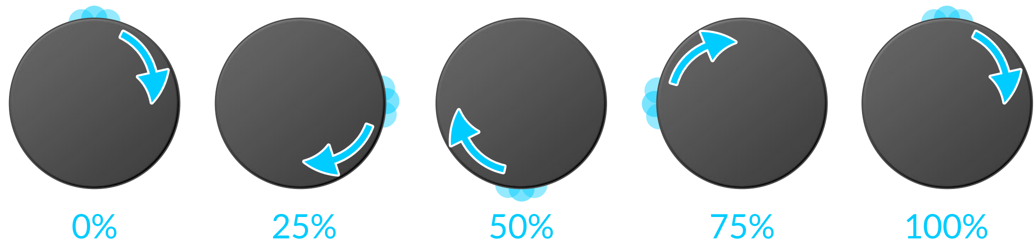



**Figure 6.10:** *Clockwise* (C): an area of white light moves clockwise around the device; this animation uses blue light for illustration only.

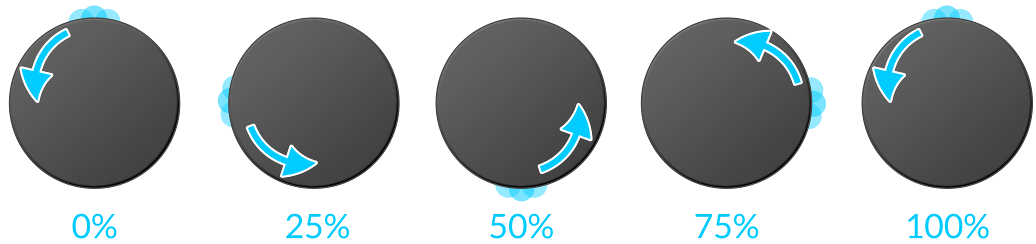



**Figure 6.11:** *Anticlockwise* (AC): an area of white light moves anticlockwise around the device; this animation uses blue light for illustration only.

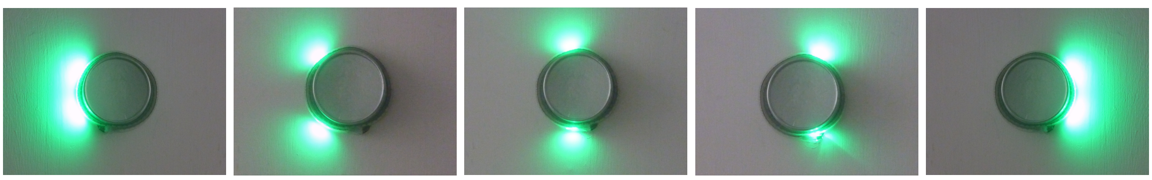


**Figure 6.12:** Once users match a rhythmic gesture, its animation is shown in green light. Here, parts of the *Side-to-Side* animation are shown (0%, 25%, 50%, 75% and 100%).

application running on a laptop computer received data from the Kinect sensor and sent commands to the Arduino device controlling the interactive light display, giving it instructions which controlled the gesture animations. Audio and tactile feedback were synthesised using Pure Data, which also communicated with the application on the laptop computer. Sound was presented using the laptop speakers and vibration was delivered using the wrist-based wearable device used in the previous experiments. Tactile feedback was delivered to the wrist as this location has been successful and acceptable to users in the other experiments.

A simple gesture-sensing algorithm was developed to allow the study of rhythmic gestures in Experiment 5. This research is interested in the usability of rhythmic gestures and their usefulness as a means of addressing gesture-sensing systems, so this algorithm was intended to be functional for the experiment but may not be robust enough for real use. Further research is needed to develop ways of detecting rhythmic gestures; however, this is outside of the scope of this thesis. Work by Lantz *et al.* [75] demonstrates that detecting rhythmic movements from sensor data can be done reliably, so accurate rhythmic gesture-sensing algorithms are possible. Appendix D provides an overview of the rhythmic gesture-sensing algorithm used here.

## 6.3 Experiment 5

### 6.3.1 Research Aims

Rhythmic gestures could be used as a way of addressing gesture-sensing systems, as they give users a way to direct their input towards a particular system and they give systems a way of detecting users' intention to interact. However, nothing is known about how effectively users can perform rhythmic gestures. Experiment 5 investigates rhythmic gesture performance, evaluating the rhythmic gestures identified in the previous section of this chapter. This experiment will compare gesture movements and intervals, to understand how performance is affected by these factors. As well as studying the use of rhythmic gestures, this experiment also investigates what effects feedback has on rhythmic gesture performance. Although interactive light animations are necessary to show users how to gesture, it is unknown if audio or tactile feedback will improve interaction. Research discussed in the literature review earlier in this chapter suggests that gestures might benefit from additional feedback, so this experiment aims to find out if this is the case.

Experiment 5 focuses on the use of rhythmic gestures as a means of directing input towards a gesture-sensing system, ignoring the issues discussed in the previous chapter about guiding users to gesture where their actions can be easily sensed. This is so that rhythmic gestures can be studied in detail on their own, first, before using them as part of more complex interaction

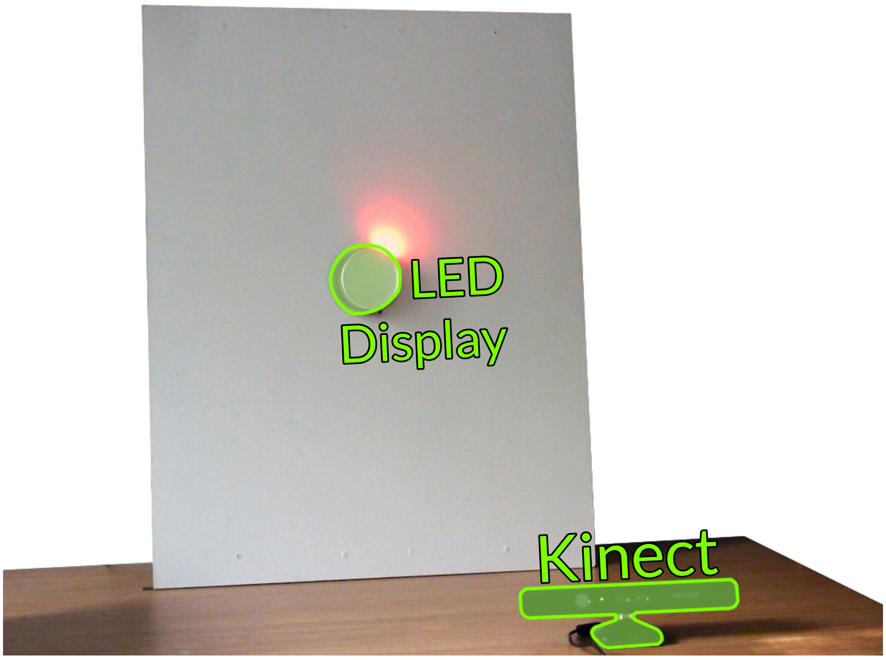



**Figure 6.13:** The wall-mounted interactive light display and Kinect sensor used in the implementation for Experiment 5.

techniques. Research in the next chapter of this thesis considers these issues together.

The main aims of this experiment are to: (1) investigate how gesture movement and interval affect the performance of rhythmic gestures; (2) study the effects of feedback on rhythmic gesture use. This research contributes an answer to the following research question:

> **RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

### 6.3.2 Experiment Design

During this experiment, participants performed all combinations of rhythmic gesture movements (**Gesture:** C, AC, SS, UD, FB) and intervals (**Interval:** 500ms, 700ms, 900ms, 1100ms). They were also given four types of **Feedback**, listed in Table 6.1, using the designs described in Section 6.2.2; interactive light feedback was always given, as this was necessary to show rhythmic gestures. This experiment therefore used a within-subjects design, with **Gesture**, **Interval** and **Feedback** as factors. Participants completed a block of tasks, described in the following section, for each type of **Feedback**. Within each block, they performed all **Gesture** and **Interval** combinations (as shown by Table 6.1). Block order was randomised using a Latin square.

| | Name | Feedback Type | Factors within condition |
|---|---|---|---|
| 1 | None | No additional feedback | Gesture: C, AC, SS, UD, FB<br>Interval: 500ms, 700ms, 900ms, 1100ms |
| 2 | Audio | Audio feedback only | Gesture: C, AC, SS, UD, FB<br>Interval: 500ms, 700ms, 900ms, 1100ms |
| 3 | Tactile | Tactile feedback only | Gesture: C, AC, SS, UD, FB<br>Interval: 500ms, 700ms, 900ms, 1100ms |
| 4 | Both | Audio and tactile feedback | Gesture: C, AC, SS, UD, FB<br>Interval: 500ms, 700ms, 900ms, 1100ms |

**Table 6.1:** Conditions in Experiment 5.

### Task Design

Each task in this experiment required participants to match a rhythmic gesture. There was one task within each block for each of the **Gesture** and **Interval** combinations; this meant participants completed twenty gestures per type of **Feedback**. Tasks had a twelve second limit from when participants started to gesture; this meant they would not spend too long trying to match difficult gestures, limiting overall fatigue from the experiment. Once participants successfully matched a rhythmic gesture, they were asked to continue for as long as possible; from this point, tasks were limited to eight gesture cycles, to limit fatigue and to keep the duration of the experiment within a reasonable time. Participants were asked to match a gesture for as long as possible as extended gesturing would allow more insight into the difficulty of using rhythmic gestures. This also means that findings from this experiment could inform the use of rhythmic gestures in other contexts; for example, as in-air gesture alternatives to the *CycloStar* [82] or *Motion-Pointing* [33] interactions, discussed earlier in this chapter.

### Measures

Table 6.2 presents a summary of the dependent measures in this experiment. These are described in more detail in the following sections. Participants were also asked to complete a survey, whose questions are not listed in Table 6.2 but are described later.

#### *Task Performance*

For each task, the time taken to match the gesture rhythm was measured (*Time-Match*). Timing started when participants first raised their hand at the start of a task and ended when they matched the gesture rhythm (after three synchronised movements). When times were recorded, the time required to match the gesture (three movements multipled by the gesture interval) was subtracted. The number of completed gesture cycles (*Cycles*) was also

| Name | Description |
|---|---|
| *Task performance measures* | |
| Time-Match | Time to match a rhythmic gesture. |
| Cycles | Count of completed rhythmic gesture movements. |
| Success | True if participant completed task, otherwise false. |
| *Task difficulty and workload ratings* | |
| Difficulty-Match | Difficulty rating from 1–10 (after every task). |
| Difficulty-Duration | Difficulty rating from 1–10 (after every task). |
| Workload | NASA-TLX workload from 0–100 (after every task block). |
| *Survey responses* | |
| Feedback | Preference rankings for feedback type (post-experiment). |
| Gesture | Preference rankings for gesture type (post-experiment). |
| Acceptability | Social acceptability score (from post-experiment survey). |

**Table 6.2:** Overview of dependent measures in Experiment 5.

recorded. This measurement started once the gesture had been matched and does not include the number of cycles required to match the rhythm. The *Success* of each task was also measured; participants successfully completed a task if they matched its rhythm within the given time limit.

### *Difficulty Ratings and Task Workload*

After each task, participants were asked to make two difficulty judgements: (1) the difficulty of matching the gesture rhythm (*Difficulty-Match*); and (2) the difficulty of continuing to match the rhythm for the duration of the task (*Difficulty-Duration*). They were not asked to provide the second rating if they did not successfully match the rhythmic gesture. Each of these ratings was on a ten-point scale, from 1 (easiest) to 10 (most difficult). These ratings were given verbally and were recorded electronically by the experimenter.

For each completed NASA-TLX survey, an overall *Workload* measure was calculated. This calculation was the same as in Experiment 1 and is described in Section 3.3.2. *Workload* ranged from 0 to 100, where higher values mean a greater task workload.

### *Post-Experiment Surveys*

At the end of the experiment, participants were asked to give preference rankings for types of feedback (*Feedback*) and for gestures (*Gesture*), from favourite to least favourite. They were asked to give preference rankings as this would show which gesture movements and feedback types were most liked by participants. Rankings were not sought for gesture intervals, as they were not expected to be able to identify a gesture interval during use. Participants were also given a survey, which asked them to rate their agreement with three statements about each gesture (survey shown in Appendix E). They responded using a five-point scale, from

'Strongly Disagree' (1) to 'Strongly Agree' (5). The three statements were:

1. I found the [...] movement easy to perform.
2. I found the animation for the [...] gesture easy to follow.
3. I found it easy to keep matching the [...] gesture.

These statements were chosen to provide insight into how difficult the rhythmic gesture movements were (Q1), how understandable their animations were (Q2), and how effortless (or fatiguing) it was to keep performing them (Q3).

Finally, participants were asked to complete a survey about the social acceptability of each gesture. This survey (shown in Appendix E) asked participants in which social situations they would find it acceptable to use each of the five gestures used in this experiment. Social acceptability was investigated in this study to see how willing users would be to use rhythmic gestures outside of a usability lab setting. Six social situations were asked about, listed below. Participants responded to each social situation with "Yes" (if the gesture was acceptable) or "No". This survey design was based on Rico's study [116] into the social acceptability of gestures, although this question format was chosen after discussion with Rico, who recommended it as an improvement on her original one. For each social situation, an overall *Acceptability* score was calculated as the proportion of "Yes" responses. Rico used the same calculation in her study for comparing gesture acceptability. The six social situations were:

- At home, alone
- At home, with family
- At work, alone
- At work, with colleagues
- In public, with friends
- In public, with strangers

### Procedure

Participants were given a short tutorial at the start of the experiment, which demonstrated the rhythmic gestures and gave them the chance to try performing them. They were shown each of the five gesture movements and the interactive light animations with which they had to synchronise their hand movements. They were also introduced to each of the types of feedback they would receive during the experiment. Finally, they were given twenty practice tasks where they performed one of each gesture movement with each type of feedback.

During the experiment, participants were seated 3m from the wall-mounted interactive light display and Kinect sensor. The experimenter sat next to participants, with a laptop computer for data entry. Participants were asked to gesture using their dominant hand. They wore the wrist-based tactile display on their other wrist, so that the cable from that device did not get in the way during gesture movements. Tasks started when participants raised their hand to at least shoulder height, so that task timing was consistent. Tasks ended when time/gesture cycle limits, described earlier, expired or when a gesture performance was no longer in time with the gesture rhythm. Participants were asked to lower their arms between tasks, minimising fatigue from keeping their hands raised for extended periods of time. After each block of tasks, participants were given a NASA-TLX survey to complete. At the end of the experiment, they were also asked to give preference rankings and complete a short survey; both are described in the following section.

### Hypotheses

- **H1:** *Time-Match* will be higher for circular gestures than for non-circular ones;
- **H2:** *Difficulty-Match* will be higher for circular gestures than for non-circular ones;
- **H3:** *Time-Match* will be lower when feedback is given than when not given;
- **H4:** *Difficulty-Match* will be lower when feedback is given than when not given;
- **H5:** *Difficulty-Match* and *Difficulty-Duration* will be lower for 700ms and 900ms intervals than for 500ms and 1100ms ones.

**H1** and **H2** predict that circular gestures (C and AC) will be harder to match than non-circular gestures (SS, UD and FB). This is because circular movements lack the proprioceptive feedback that the other gestures have. Research discussed earlier suggests that proprioceptive feedback from reaching comfortable joint limits may improve timekeeping with a rhythm for back-and-forth movements.

**H3** and **H4** expect better and easier rhythmic gesture performance when feedback is given. This is because feedback has been shown to improve sensorimotor synchronisation, as discussed earlier in this chapter.

**H5** anticipates higher difficulty ratings for gestures with 500ms or 1100ms intervals. Higher ratings are expected for 500ms gestures as this interval is close to the threshold for matching a continuous visual rhythm. Ratings are also expected to be higher for 1100ms gestures as this interval is likely to be slower than participants' preferred movement speed. Research discussed earlier in this chapter suggests that matching a rhythm is easier when the interval is close to, or faster than, the preferred movement speed.

### Participants

Sixteen people took part in this study. Of these sixteen participants, four were female and all were right-handed; their mean age was 28.9 years (sd 4.5 years). They were recruited using university email lists. Each experiment lasted for one hour and participants were paid £6.

### 6.3.3 Results

#### Success Rates

Participants successfully matched 1193 of 1280 rhythmic gestures (93.2%); Figure 6.14 shows *Success* for each **Gesture** and **Interval** combination. Logistic regression was used to analyse the effect of each experiment factor on *Success*, using a mixed-effects model with participant as a random factor. This approach was chosen as logistic regression can be used to model binary outcomes (gesture rhythm was matched, or not) as a function of predictor variables (the three experiment factors).

A repeated-measures ANOVA using the logistic regression model found that:

- **Feedback** was not a significant predictor of *Success*: $\chi^2(3) = 3.76$, $p = 0.29$
- **Gesture** was a significant predictor of *Success*: $\chi^2(4) = 65.71$, $p < 0.001$
- **Interval** was a significant predictor of *Success*: $\chi^2(3) = 55.48$, $p < 0.001$
- No interactions between factors were significant: all $p \geq 0.06$

*Post hoc* comparisons for **Gesture**, using Wilcoxon's tests, found that *Success* was higher for SS and UD than all other gestures: all $z \geq 3.50$, $p \leq 0.004$. No other comparisons were significant.

*Post hoc* comparisons for **Interval** found that *Success* was lower for 500ms than all other intervals (all $z \geq 2.70$, $p \leq 0.002$) and was also lower for 700ms than for 900ms ($z = 2.70$, $p = 0.03$). No other comparisons were significant.

#### Time-Match

Incomplete trials were excluded from analysis of *Time-Match*, as this would be strongly skewed by the task time limit being met. Measurements for *Time-Match* are also affected by the **Interval**, as gestures with shorter intervals can be matched in less time than gestures with longer intervals. All *Time-Match* values were normalised to an interval of 500ms, allowing a fair comparison of the time needed to match gestures at each interval.

Mean normalised *Time-Match* was 2204ms (sd 1548ms); Figure 6.15 shows mean times for each **Gesture** and **Interval** combination. Times were not normally-distributed (Shapiro-Wilk: $W = 0.73$, $p < 0.001$) so the Aligned-Rank Transform [153] was used prior to analysis; this transformation allows non-parametric data to be analysed using factorial parametric tests (such as ANOVA). Results from a repeated-measures ANOVA are shown in Table 6.3. **Gesture** and **Interval** both had a significant effect on the time taken to match a gesture rhythm, but **Feedback** did not. All interactions between the factors were both also significant.

*Post hoc* t-test comparisons for **Gesture** found the following significant differences: both circular gestures took significantly longer to match than all other gestures (all $t \geq 5.6$,

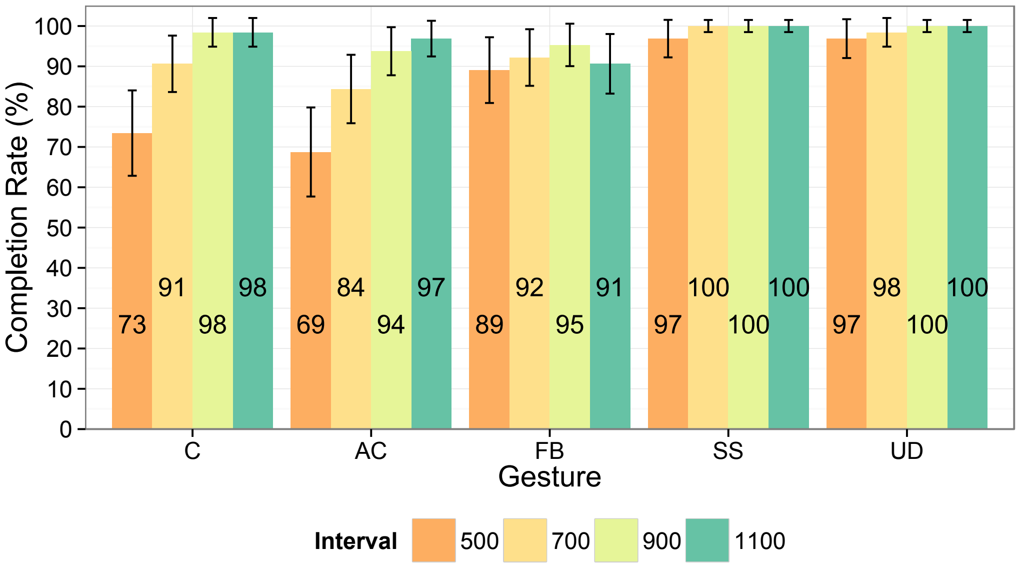



**Figure 6.14:** *Success* rates for each rhythmic gesture. Error bars show 95% CIs.

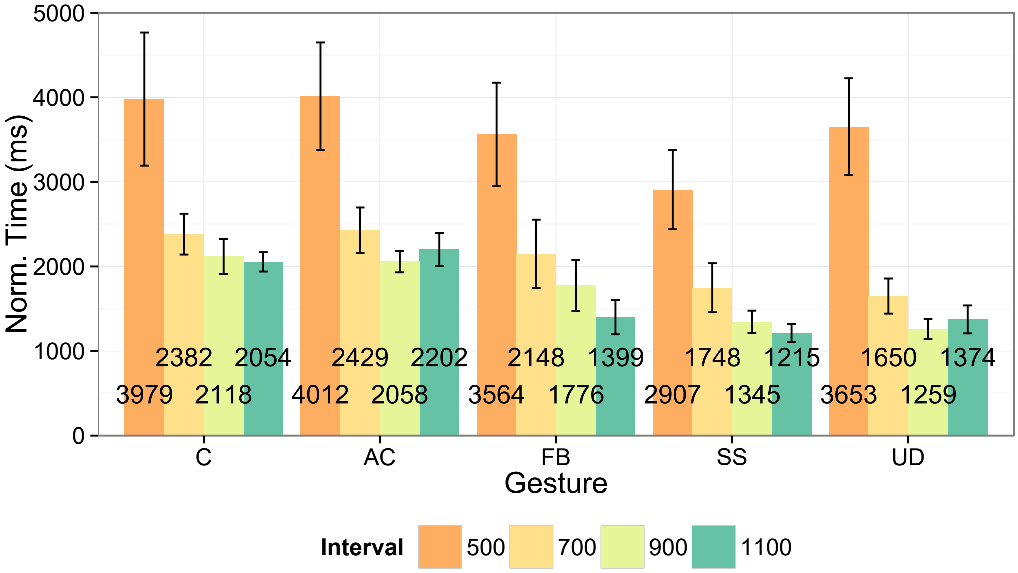



**Figure 6.15:** Mean *Time-Match*, with all times normalised to a 500ms interval. Error bars show 95% CIs.

| Factors | ANOVA Results | | | | |
|---|---|---|---|---|---|
| Feedback | F(3, 1099) | = | 0.98 | $p = 0.40$ | |
| Gesture | F(4, 1099) | = | 59.58 | $p < 0.001$ | ★ |
| Interval | F(3, 1099) | = | 153.78 | $p < 0.001$ | ★ |
| Feedback x Gesture | F(12, 1099) | = | 2.41 | $p = 0.004$ | ★ |
| Feedback x Interval | F(9, 1099) | = | 2.08 | $p = 0.03$ | ★ |
| Gesture x Interval | F(12, 1099) | = | 2.32 | $p = 0.006$ | ★ |
| Feedback x Gesture x Interval | F(36, 1099) | = | 1.67 | $p = 0.008$ | ★ |

**Table 6.3:** Repeated-measures ANOVA results for *Time-Match*. Significant effects are highlighted with '★'.

$p < 0.001$); FB took significantly longer than SS ($t = 5.58$, $p < 0.001$) and UD ($t = 2.74$, $p = 0.049$); and UD took significantly longer than SS ($t = 2.9$, $p = 0.03$).

*Post hoc* comparisons for **Interval** found the following significant differences: gestures with a 500ms interval took significantly more time to match than all others (all $t \geq 13.13$, $p < 0.001$); and gestures with a 700ms interval took significantly longer to match than those with 900ms and 1100ms intervals (both $t \geq 5.0$, $p < 0.001$).

*Post hoc* comparisons for the interaction between **Feedback** and **Interval** found the following significant differences: when no extra feedback was given, *Time-Match* was higher for 500ms intervals than for 700ms; and when audio feedback was given, *Time-Match* was higher for 700ms intervals than for 900ms.

*Post hoc* comparisons for the interaction between **Gesture** and **Interval** found that both circular gestures needed significantly more time to match at 500ms than at 1100ms (both $t \geq 3.63$, $p \leq 0.04$) and that SS needed significantly more time to match with a 500ms interval than with a 900ms or 1100ms interval (both $t \geq 3.9$, $p \leq 0.01$).

*Post hoc* comparisons for the three-factor interaction found no significant differences.

### Cycles

Mean number of *Cycles* was 6.8 (sd 2.2); Figure 6.16 shows the mean count for each **Gesture** and **Interval** combination. Cycles were not normally-distributed (Shapiro-Wilk: $W = 0.58$, $p < 0.001$) so were transformed using the Aligned-Rank Transform [153] prior to analysis. Results from a repeated-measures ANOVA are shown in Table 6.4. **Gesture** and **Interval** both had a significant effect on time taken to match a gesture rhythm. The interactions between **Feedback** and **Interval**, between **Gesture** and **Interval**, and between all three factors were also significant.

*Post hoc* t-test comparisons for **Gesture** found that both circular gestures were performed for significantly less gesture cycles than all other gestures (all $t \leq -3.04$, $p \leq 0.02$). FB was

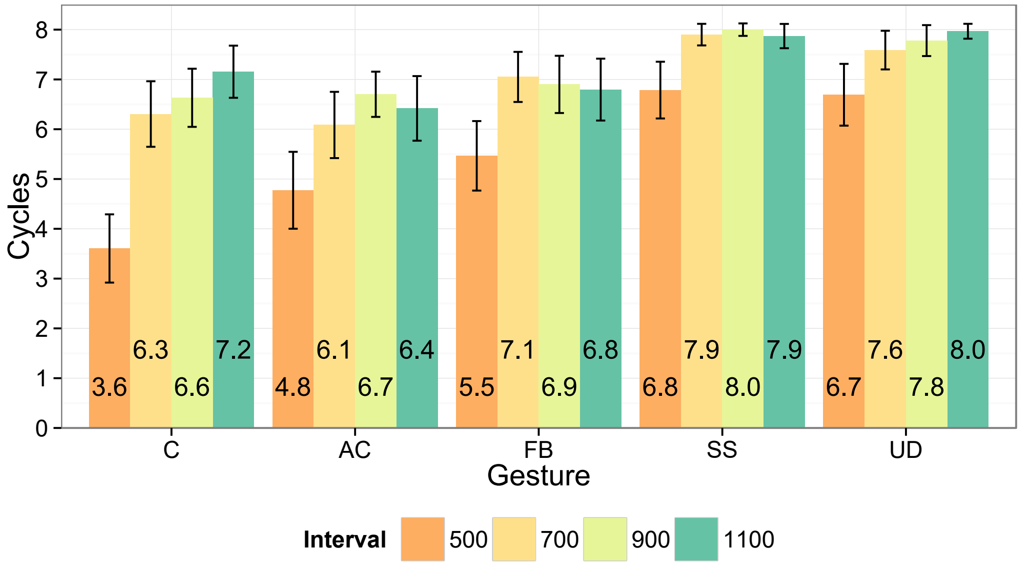



**Figure 6.16:** Mean *Cycles*; maximum allowed was 8. Error bars show 95% CIs.

| **Factors** | **ANOVA Results** | | | | |
|---|---|---|---|---|---|
| Feedback | F(3, 1099) | = | 1.58 | p = 0.19 | |
| Gesture | F(4, 1099) | = | 65.62 | $p < 0.001$ | ★ |
| Interval | F(3, 1099) | = | 72.54 | $p < 0.001$ | ★ |
| Feedback x Gesture | F(12, 1099) | = | 1.00 | p = 0.45 | |
| Feedback x Interval | F(9, 1099) | = | 2.10 | p = 0.03 | ★ |
| Gesture x Interval | F(12, 1099) | = | 13.17 | $p < 0.001$ | ★ |
| Feedback x Gesture x Interval | F(36, 1099) | = | 1.54 | p = 0.02 | ★ |

**Table 6.4:** Repeated-measures ANOVA results for *Cycles*. Significant effects are highlighted with '★'.

performed for significantly less cycles than SS ($t = -8.27$, $p < 0.001$) and UD ($t = -5.72$, $p < 0.001$). No other pairwise comparisons were significant.

*Post hoc* comparisons for **Interval** found that gestures with 500ms intervals were performed for significantly less cycles than those with all other intervals (all $t \leq -10.29$, $p < 0.001$), and that gestures with 700ms intervals were performed for less cycles than those with 1100ms intervals ($t = -3.56$, $p = 0.002$).

*Post hoc* comparisons for **Feedback x Interval** found only one significant difference: when audio feedback was used, gestures had less cycles with 500ms intervals than they did with 900ms intervals.

*Post hoc* comparisons for **Gesture x Interval** found a large number of significant differences. Those relevant to the experiment hypotheses are discussed here; the rest are listed in Section F.1 of Appendix F. There were no significant differences within-gestures between 700ms and 900ms intervals, between 700ms and 1100ms intervals, and between 900ms and 1100ms intervals. Most gestures had significantly less cycles at 500ms than at 900ms or 1100ms. The only gestures which had significantly less cycles at 500ms than at 700ms were C, SS and UD (all $t \geq 4.16$, $p \leq 0.005$).

*Post hoc* comparisons for the three-factor interaction found a large number of significant differences. Those relevant to the experiment hypotheses are discussed here; the rest are listed in Section F.1 of Appendix F. When tactile feedback was given, SS had less cycles at 500ms than at 700ms, and when audio and tactile feedback were given together, FB at 1100ms had less cycles than UD at 1100ms.

### Difficulty-Match

Mean *Difficulty-Match* was 3.41 (sd 2.07); Figure 6.17 shows mean ratings for each **Gesture** and **Interval** combination. Difficulty ratings were transformed using the Aligned-Rank Transform [153] prior to analysis; this means that parametric tests can be used to analyse the data, allowing for factorial analysis which is not possible with non-parametric tests [153]. Results from a repeated-measures ANOVA are shown in Table 6.5. **Feedback**, **Gesture** and **Interval** all had significant effects on *Difficulty-Match*, as did the interactions between **Feedback** and **Interval**, and between **Gesture** and **Interval**.

*Post hoc* t-test comparisons for **Feedback** found that difficulty ratings were lower for Audio and Both than for None (both $t \geq 3.88$, $p \leq 0.004$). Ratings were also significantly lower for Audio than Tactile ($t = 3.33$, $p = 0.005$). No other comparisons were significant.

*Post hoc* comparisons for **Gesture** found that difficulty ratings differed significantly between all gestures (all $t \geq 5.18$, $p < 0.001$), except between C and AC ($t = 1.79$, $p = 0.38$) and FB

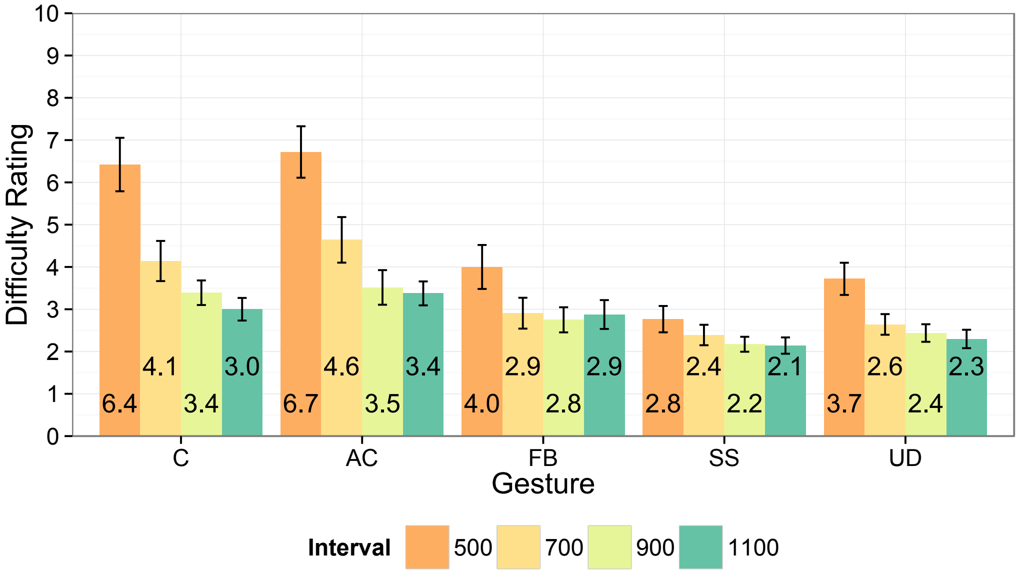



**Figure 6.17:** Mean *Difficulty-Match* ratings. Error bars show 95% CIs.

| **Factors** | **ANOVA Results** | | | | |
|---|---|---|---|---|---|
| Feedback | F(3, 1185) | = | 10.40 | p < 0.001 | ★ |
| Gesture | F(4, 1185) | = | 146.15 | p < 0.001 | ★ |
| Interval | F(3, 1185) | = | 138.02 | p < 0.001 | ★ |
| Feedback x Gesture | F(12, 1185) | = | 1.73 | p = 0.06 | |
| Feedback x Interval | F(9, 1185) | = | 1.94 | p = 0.04 | ★ |
| Gesture x Interval | F(12, 1185) | = | 14.75 | p < 0.001 | ★ |
| Feedback x Gesture x Interval | F(36, 1185) | = | 1.03 | p = 0.42 | |

**Table 6.5:** Repeated-measures ANOVA results for *Difficulty-Match*. Significant effects are highlighted with '★'.

and UD ($t = 2.57$, $p = 0.08$). Ratings for SS were lower than all other gestures, while FB and UD were rated easier than both circular gestures.

*Post hoc* comparisons for **Interval** found that all differences were significant (all $t \geq 4.55$, $p < 0.001$) except for the difference between 900ms and 1100ms ($t = 1.72$, $p = 0.31$). For the significant differences, difficulty ratings were higher for the lower intervals.

*Post hoc* comparisons for the interaction between **Feedback** and **Interval** found no significant differences.

*Post hoc* comparisons for the interaction between **Gesture** and **Interval** found a large number of significant differences. Those relevant to the experiment hypotheses are mentioned here; the others are listed in Section F.2 of Appendix F. There were no significant differences within-gestures between 700ms and 900ms, between 700ms and 1100ms, or between 900ms and 1100ms (all $p > 0.05$). The only gestures with significant differences between 500ms and 700ms were SS and AC (both $t \geq 3.62$, $p \leq 0.04$).

### Difficulty-Duration

Mean *Difficulty-Duration* was 3.10 (sd 1.81); Figure 6.18 shows mean ratings for each **Gesture** and **Interval** combination. Difficulty ratings were transformed using the Aligned-Rank Transform, as before. Results from a repeated-measures ANOVA are shown in Table 6.6. **Feedback**, **Gesture** and **Interval** all had significant effects on *Difficulty-Match*, as did the interactions between **Feedback** and **Gesture**, and between **Gesture** and **Interval**.

*Post hoc* t-test comparisons for **Feedback** found that ratings were significantly higher for None than all other types of feedback: all $t \geq 3.28$, $p \leq 0.006$. No others were significant.

*Post hoc* comparisons for **Gesture** found that circular gestures received significantly higher ratings than all other gestures: all $t \geq 14.22$, $p < 0.001$. SS was rated easier than UD and FB: both $t \geq 4.93$, $p < 0.001$. No other comparisons were significant.

| **Factors** | **ANOVA Results** | | | | |
|---|---|---|---|---|---|
| Feedback | F(3, 1099) | = | 10.67 | $p < 0.001$ | ★ |
| Gesture | F(4, 1099) | = | 184.32 | $p < 0.001$ | ★ |
| Interval | F(3, 1099) | = | 93.40 | $p < 0.001$ | ★ |
| Feedback x Gesture | F(12, 1099) | = | 2.16 | $p = 0.01$ | ★ |
| Feedback x Interval | F(9, 1099) | = | 0.67 | $p = 0.74$ | |
| Gesture x Interval | F(12, 1099) | = | 9.25 | $p < 0.001$ | ★ |
| Feedback x Gesture x Interval | F(36, 1099) | = | 1.18 | $p = 0.22$ | |

**Table 6.6:** Repeated-measures ANOVA results for *Difficulty-Duration*. Significant effects are highlighted with '★'. Note that the degrees of freedom for these analyses is different from those in Table 6.5; this is because ratings here were not made for incomplete tasks.

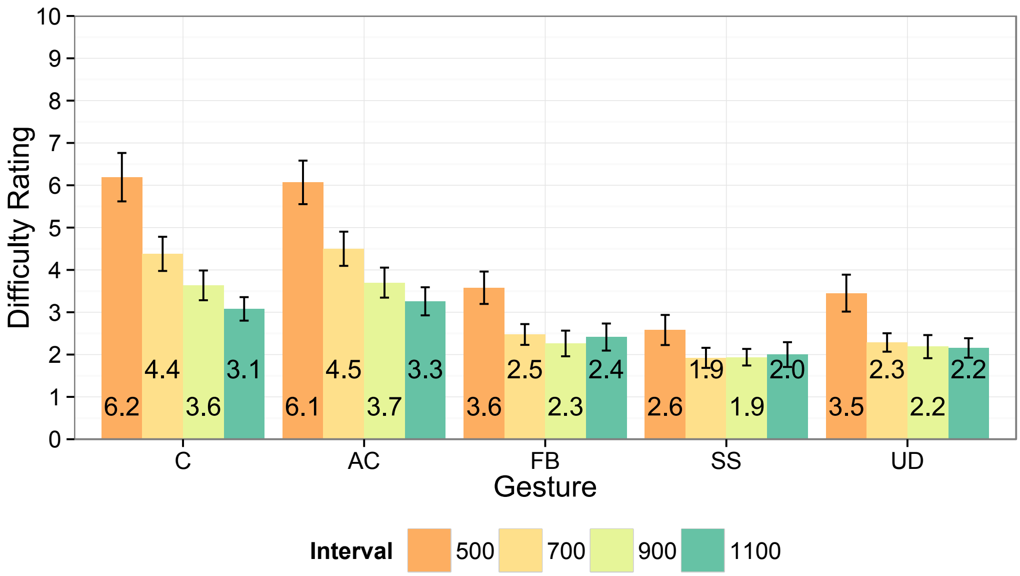



**Figure 6.18:** Mean *Difficulty-Duration* ratings. Error bars show 95% CIs.

*Post hoc* comparisons for **Interval** found that all differences were significant (all $t \geq 3.91$, $p < 0.001$) except for the difference between 900ms and 1100ms ($t = 1.93$, $p = 0.22$). Difficulty ratings were higher for shorter intervals.

*Post hoc* comparisons for the interaction between **Feedback** and **Gesture** found no significant differences.

*Post hoc* comparisons for the interaction between **Gesture** and **Interval** found that there were no significant differences within-gestures between 500ms and 700ms, 700ms and 900ms, 700ms and 1100ms, and 900ms and 1100ms: all $t \leq 3.31$, $p \geq 0.1$. Difficulty ratings for UD were not significantly different at different intervals: all $t \leq 2.12$, $p \geq 0.85$. In general, gestures with 500ms intervals were rated more difficult than those with 900ms or 1100ms intervals.

### Workload

Mean *Workload* was 41.20 (sd 11.74); Figure 6.19 shows mean *Workload* for each **Feedback**. Workload estimations were not normally-distributed (Shapiro-Wilk: $W = 0.96$, $p = 0.02$), so were transformed using the Aligned-Rank Transform [153] prior to analysis. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Workload*: $F(3, 45) = 12.08$, $p < 0.001$. *Post hoc* t-test comparisons found that *Workload* was significantly higher for None than all other types of feedback: all $t \geq 2.7$, $p \leq 0.04$.

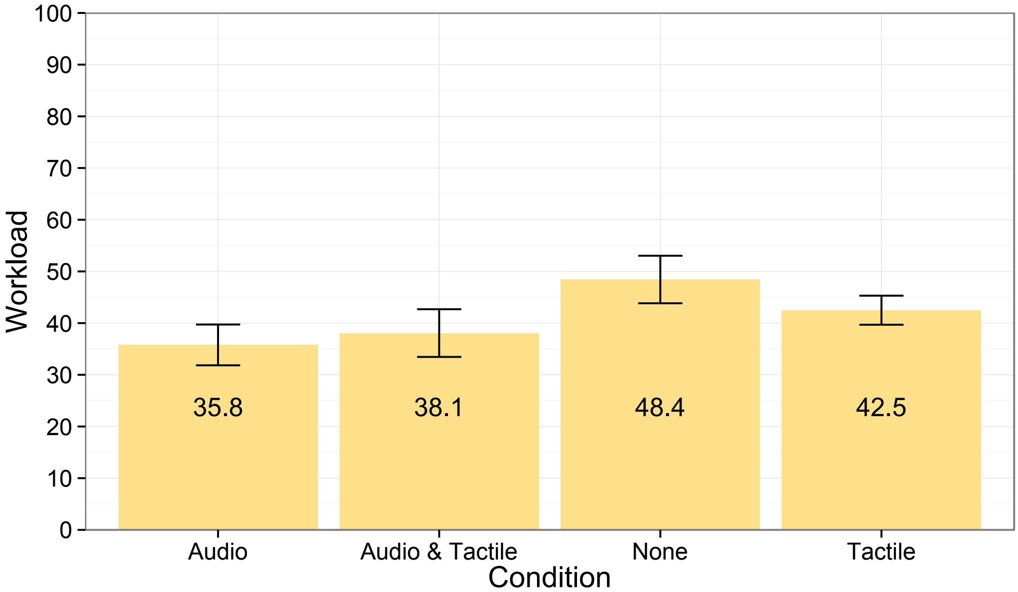



**Figure 6.19:** Mean *Workload* ratings. Error bars show 95% CIs.

### Preference

Table 6.7 shows the median ranks for feedback designs and gesture movements, where a rank of '1' was the most preferred option. Friedman's rank sum tests were used to see if these rankings were significant or not. Friedman's test found that **Feedback** rankings were significantly different: $\chi^2(3) = 19.88$, $p < 0.001$.

*Post hoc* comparisons using Bonferroni-corrected Wilcoxon's signed-rank tests found that None was ranked significantly lower than Audio (W = 122, p = 0.03) and Tactile (W = 123, p = 0.02). No other comparisons were significant. While Both had a higher median ranking than Tactile, the difference in rankings between Both and None was not significant (W = 112, p = 0.12). This is because participants showed less consensus in their ranking of Both (median absolute deviation of 0 for Tactile, versus 1.48 for Both); two participants ranked Both as their least favourite type of feedback, causing this disparity.

Friedman's test found that **Gesture** rankings were significant: $\chi^2(4) = 30.7$, $p < 0.001$. *Post hoc* comparisons using the same approach as before found that SS was ranked significantly higher than all other gestures (all $p \leq 0.01$). No other comparisons were significant.

| Feedback Rank | | | | Gesture Rank | | | | |
|---|---|---|---|---|---|---|---|---|
| None | Audio | Tactile | Both | C | AC | FB | SS | UD |
| 4 | 2 | 3 | 2 | 4 | 5 | 3 | 1 | 3 |

**Table 6.7:** Median ranks for feedback designs and gesture movements.

### Survey Ratings

This section discusses responses to the post-experiment survey questions about each rhythmic gesture movement. Median responses are shown in Table 6.8. Friedman's rank sum tests were used to compare ratings between gestures. When significant differences were found, *post hoc* Wilcoxon's signed-rank tests were used, with Bonferroni-corrected p-values.

#### *Q1: "I found the [...] movement easy to perform"*

Friedman's test found that responses were significantly different: $\chi^2(4) = 26.45$, $p < 0.001$. *Post hoc* comparisons found that SS was rated easier to perform than C, AC and FB: all $p \leq 0.04$. No other comparisons were significant.

#### *Q2: "I found the animation for the [...] gesture easy to follow"*

Friedman's test found that responses were significantly different: $\chi^2(4) = 26.51$, $p < 0.001$. *Post hoc* comparisons found that ratings were higher for SS than for C and AC: both $p \leq 0.035$. No other comparisons were significant.

#### *Q3: "I found it easy to keep matching the [...] gesture"*

Friedman's test found that responses were significantly different: $\chi^2(4) = 34.87$, $p < 0.001$. *Post hoc* comparisons found that SS was rated easier to keep matching than C and AC: both $p \leq 0.006$. No other comparisons were significant.

### Social Acceptability

Mean *Acceptability* was 59.8% (sd 14.1%); Table 6.9 shows *Acceptability* scores, summarised for each of the six social situations and the five rhythmic gestures. Cochran's Q test, with *post hoc* Bonferroni-corrected McNemar's tests, were used to compare acceptability responses for social situations; this approach was used because Rico *et al.* [116] also used it in their analysis of social acceptability data.

Cochran's Q test found that situation had a significant effect on *Acceptability*: $\chi^2(5) = 258.9$, $p < 0.001$. *Post hoc* comparisons found significant differences between all situations (all

| **Question** | **Gesture** | | | | |
|---|---|---|---|---|---|
| | C | AC | FB | SS | UD |
| *Q1: Easy to perform* | 3 | 3 | 4 | 5 | 4.5 |
| *Q2: Animation easy to follow* | 3 | 3 | 4 | 5 | 4 |
| *Q3: Easy to keep matching rhythm* | 3 | 3 | 4 | 5 | 4 |

**Table 6.8:** Median survey responses to each question for each gesture. Responses show agreement on a five-point scale, from "Strongly Disagree" (1) to "Strongly Agree" (5).

| Social Situation | Acceptability *mean* | *sd* | Gesture | Acceptability *mean* | *sd* |
|---|---|---|---|---|---|
| At home, alone | 100.0% | 0.0% | C | 47.9% | 16.0% |
| At home, with family | 83.8% | 18.2% | AC | 46.9% | 15.2% |
| At work, alone | 93.8% | 9.6% | FB | 66.7% | 10.5% |
| At work, with colleagues | 53.8% | 30.7% | SS | 82.3% | 17.7% |
| In public, with friends | 20.0% | 16.3% | UD | 55.2% | 14.6% |
| In public, with Strangers | 7.5% | 10.0% | | | |

**Table 6.9:** Mean *Acceptability* scores for each social situation and rhythmic gesture.

$p \leq 0.002$), except between *home, alone* and *work, alone* ($p = 0.06$), and between *home, with family* and *work, alone* ($p = 0.10$).

Wilcoxon's signed-rank test was used for pairwise comparison of *Acceptability* scores for all gestures. Bonferroni's correction was used for p-values, as multiple comparisons were being made. The following significant differences were found:

- SS was more acceptable than all other gestures: all $Z \geq 3.08$, $p \leq 0.001$
- UD was more acceptable than AC: $Z = 2.57$, $p = 0.01$
- FB was more acceptable than C and AC: both $Z \geq 2.91$, $p \leq 0.001$

### 6.3.4 Discussion

#### Rhythmic Gesture Designs

Five rhythmic gesture movements were investigated in this experiment. Of these five movements (shown in Figures 6.3 to 6.6), circular ones were found to be the most difficult to use. Participants took longer to match circular gesture rhythms (meaning **H1** can be accepted) and gave higher difficulty ratings for these gestures (meaning **H2** can also be accepted). Circular gestures were more difficult than non-circular ones because they required more complex hand movements; users had to trace a particular path rather than move their hands back and forth in one direction. These gestures were also more difficult to sense so the recognition algorithm may have been too strict. Survey responses to Q2 also suggest that the animations for these gestures were the most difficult to follow. Despite the increased difficulty, participants still performed these movements well, especially with slower intervals (900ms and 1100ms). Gesture-sensing systems should still use circular movements, as this gives more choice when choosing gestures.

Side-to-Side was the best performed gesture movement (in terms of gesture times and success rate) and also had the lowest difficulty ratings. Participants also rated it the most socially acceptable and preferred it to the others. During the experiment, participants were observed to use a smaller range-of-motion for Side-to-Side gestures than for Up-and-Down and Forwards-and-Backwards ones. Smaller movements may contribute to the better gesture performance, as participants could exert more control over their gestures. This may also explain the higher social acceptability of the Side-to-Side gesture; smaller hand movements would be more discreet, allowing users to gesture without attracting the attention of people nearby. Ahlström *et al.* [1] found that people were less willing to perform gestures in certain areas in front of them, so Side-to-Side may have been considered more socially acceptable because hand movements took place in a limited space directly in front of the body. For the other gestures, users moved their hands further away from this area. When selecting gestures, priority should be given to Side-to-Side; Up-and-Down was also performed well and was well liked by participants.

If smaller movements lead to more successful rhythmic gestures, then an interesting area for future research would be to investigate rhythmic gestures which feature minimal hand movement. For example, users could perform rhythmic gestures without actually moving their hands in front of their body, e.g. alternating between making a fist and extending all fingers, repeatedly pinching thumb and forefinger together, extending and bending a single finger repeatedly. Some of these findings suggest that two-handed gestures may be less successful and acceptable than one-handed gestures, as users would have to synchronise more complicated movements (two hands instead of one) in a potentially more obtrusive way (attracting more attention than had a single hand been used).

Gestures generally increased in difficulty as gesture interval decreased, shown through performance times and difficulty ratings. However, analysis of the interaction between gesture movement and interval found that significant differences were mostly found for larger interval increases (500ms–900ms, for example), rather than for stepwise increases (700ms–900ms, for example). Based on these results, **H5** cannot be accepted. In many cases, there were no differences found between 700ms, 900ms and 1100ms intervals. These findings are positive and suggest that interval can be used effectively as a design parameter for rhythmic gestures. When gesture-sensing systems coordinate their choice of rhythmic gestures, they may benefit more from using fewer gesture movements and more gesture intervals, rather than selecting from more difficult gesture movements.

### Gesture Feedback

Feedback about gestures had no effect on performance (in terms of time to match, gesture duration and success rate) but did make it easier to gesture, shown by lower difficulty ratings and lower task workload. These results mean **H3** cannot be accepted and **H4** can. Although participants benefitted from all types of feedback given, tactile feedback was less effective than audio feedback and was also less preferred. Similar findings were observed in Experiment 4; participants there performed better with sound than vibration, even though they used the same designs. Audio feedback may have been easier to perceive in these experiments, making it more useful during interaction. However, participants still found tactile feedback helpful and this modality may be preferred in some situations; for example, when users are near other people or in noisy environments.

## 6.4 Limitations

### 6.4.1 Gesture Matching Accuracy

This experiment investigated rhythmic gestures, to understand their usability and to see if they could be used for addressing gesture-sensing systems. Four gesture intervals were used, allowing an initial evaluation of how interval affects gesture performance. However, this experiment did not investigate *how well* users matched those rhythms, as that would require a more robust rhythmic gesture-sensing algorithm. More work is needed to understand how accurately users can perform rhythmic gestures, which will further inform their design and use. For example, knowing how accurate (or inaccurate) users are means designers can choose appropriate increases in rhythm interval and allows developers to make informed decisions about acceptable tolerances for rhythm matching. To facilitate this further study, research is needed to develop an accurate and robust way of sensing rhythmic gestures, combining the general problem of gesture recognition with the problem of detecting how accurately users are matching a rhythm. Work by Lantz *et al.* [75] demonstrates one feasible approach; they developed a technique for detecting rhythmic patterns in accelerometer data.

### 6.4.2 Device Form Factor

In this experiment, participants used rhythmic gestures to interact with a wall-mounted gesture-sensing system. As users faced the system when gesturing, their hand movements could be easily sensed by the input sensor. Some of the gesture movements may be less appropriate when interacting with other device setups, however. For example, the circular gestures used here may be difficult to sense when users are interacting over a device whose sensors face upwards; a mobile phone on a coffee table or a Leap Motion, for example. In such cases, horizontal circular movements may be more appropriate. It is not known how well users can perform rhythmic gestures when not directly facing the system; for example, if they are not parallel to the sensor and interactive light display when interacting.

The wall-mounted interactive light display used here had a circular layout, ideal for showing the circular gesture animations. However, other device shapes (rectangles, for example) may be less ideal for showing these animations. More research is needed to understand how best to pair rhythmic gestures with different types of interactive light display. However, the non-circular gestures (SS, UD, FB) represent the simplest movements possible in all three axes and should therefore be ideal for use with all devices which can sense in three dimensions.

## 6.5 Conclusions

This chapter described and investigated rhythmic gestures, a novel interaction technique which could be used for addressing gesture-sensing systems. Rhythmic gestures allow users to specify which system they wish to interact with, through the gestures they perform, while at the same time showing their intention to interact. These properties mean gesture-sensing systems can confidently determine if they are being addressed, or not, helping to avoid the *Midas Touch* problem [70]. An interaction technique for addressing gesture-sensing systems must work with a variety of devices, as such systems are now being introduced to more limited form factors. Limited-display devices can use rhythmic gestures effectively, as interactive light is used to communicate movement patterns and their rhythms. Minimal feedback about gestures is required; although participants in Experiment 5 benefit from audio and tactile feedback, they still performed gestures well without it. The following section presents recommendations for using rhythmic gestures, informed by findings from Experiment 5.

### 6.5.1 Design Recommendations for Rhythmic Gestures

#### DR1: Prioritise use of the Side-to-Side and Up-and-Down gestures

These gestures were performed best by participants in Experiment 5 and were also rated the easiest to perform. Side-to-Side was rated as the most socially acceptable gesture and should be used if users are likely to be gesturing around others. Gesture-sensing mobile devices, for example, should use this gesture when possible, as users are likely to use these devices in a variety of social situations.

#### DR2: Give users feedback about their gestures

Although feedback did not affect gesture performance, it did make it easier to use rhythmic gestures. Difficulty ratings and task workload ratings both show that participants benefitted from feedback about their movements. Gesture-sensing systems should use whatever output capabilities are available to present this feedback; audio feedback was especially effective in Experiment 5 and should be given if possible and if appropriate.

### DR3: Use a minimum interval of 700ms; 900ms for circular gestures

Gesture performance was generally worse with 500ms intervals than with longer intervals; this is unsurprising as 500ms is close to the threshold interval (460ms [113]) for being able to move in time with a visual rhythm. Participants successfully performed rhythmic gestures with 700ms intervals, however, suggesting that this would be an appropriate minimum interval for rhythmic gestures. Circular gestures were more difficult than the other gestures so users would benefit from a higher interval. Difficulty ratings suggest that 900ms intervals would be an appropriate starting point for circular gestures.

### DR4: Reserve faster intervals (500ms) for intentionally difficult interactions

Rhythmic gestures with 500ms intervals were rated as most difficult to perform; however, participants still had satisfactory performance with some of these gestures. These could be used for gesture systems which users are less likely to want to interact with. They could also be used in other interaction contexts; for example, for interactions which are intended to be difficult, requiring more effort and engagement from users. Such difficult gestures could be used for actions which cannot be easily reversed or which have significant consequences, to create *uncomfortable user experiences* [12] through physical effort, or for strenuous *inconvenient interactions* [112], for example.

## 6.5.2 Research Question 4

The research in this chapter contributes an answer to the following question:

> **RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

Experiment 5 found that rhythmic gestures could be an effective way of directing input when addressing gesture systems, as users were able to perform these gestures well and with ease. Gesture-sensing systems can choose rhythmic gestures from a varied design space, combining intervals and movement patterns to create unique gestures. These gestures can be revealed using minimal output capabilities, making them suitable for use with limited-display devices; interactive light displays can show users how to move and additional feedback, while beneficial, is not required for good gesture performance.

An interaction technique for addressing gesture-sensing systems must meet two criteria: (1) it must help users find where to perform gestures, so that their actions can be sensed; and (2) it must allow users to direct their input, so that their gestures do not unintentionally affect other systems. Research in the previous chapter investigated a solution to the first criterion, while research in this chapter provided a solution to the second. The following chapter investigates a combination of these solutions, developing and evaluating a single interaction technique which allows users to address gesture systems effectively.

### 6.5.3 Contributions

The research in this chapter makes the following contributions:

- It describes *rhythmic gestures*, a novel gesture interaction technique which can be used for directing input;
- It presents an experiment investigating rhythmic gestures, with design recommendations informed by the experiment findings.

# Chapter 7

# An Interaction Technique for Addressing Gesture-Sensing Systems

## 7.1 Introduction

So far, this thesis has treated the problem of addressing a gesture-sensing system as two separate parts: (1) knowing where to gesture; and (2) knowing how to direct input. Chapter 5 and Chapter 6 investigated interaction techniques which overcome each of these challenges, respectively. These techniques — *sensor strength feedback* and *rhythmic gestures* — were found to be effective on their own. However, when users start interacting with a gesture system, using these separate interactions in sequence could be needlessly time-consuming. Users will not always have to locate where to perform gestures and this may be something which they can do while they are directing their input using rhythmic gestures. Research in this chapter, therefore, considers how these interactions can be combined. It investigates an answer to the final research question of this thesis (**RQ5**), while also providing a more complete answer to **RQ3**:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?
>
> **RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

Chapter 5 focused on **RQ3** and presented sensor strength feedback, a technique which can guide users when they address gesture systems; however, research in that chapter only investigated use of this technique with smartphones, where users gestured from a short distance away. Research is needed to understand how effective this technique is for guiding users when they gesture from greater distances, using larger hand movements than were used in that experiment. This chapter describes research which uses sensor strength feedback with a wall-mounted gesture-sensing device, allowing a full answer to be given for **RQ3**.

### 7.1.1 Chapter Structure

Section 7.2 describes how sensor strength feedback and rhythmic gestures can be combined; it investigates how their feedback designs can be presented together and discusses some challenges which must be overcome when using these interactions in unison. Section 7.3 presents Experiment 6, which studies the usability of the resulting interaction technique. Section 7.4 discusses limitations with the experimental approach used here. Finally, Section 7.5 provides closing discussion on the research in this chapter; it also provides answers to each of the research questions identified previously.

## 7.2 Interaction Design

### 7.2.1 Introduction

Sensor strength feedback, described in Section 5.2, used interactive light, audio and tactile feedback to guide users to a location where they can be more easily sensed by gesture sensors. Rhythmic gestures, described in Section 6.2, used interactive light to reveal gesture movements and gave users audio and tactile feedback about their hand movements. Both of these techniques rely on their multimodal feedback, as this means they can be used effectively by limited-display devices. Combining the interactions, then, requires combining their feedback in a way which makes sense to users and remains usable. This section begins by considering how these types of feedback can be combined. It then discusses design problems which must be overcome for these interaction techniques to be used together. Finally, this section describes an implementation of the interaction techniques described in this chapter; this implementation will be used in Experiment 6.

### 7.2.2 Combining Feedback Designs

Sensor strength feedback and rhythmic gesture feedback use different properties of light, sound and vibration, shown in Table 7.1. Since different properties of interactive light are used, both types of feedback can be presented together without one obscuring the other.

When the interactive light designs are combined, the visibility of rhythmic gesture animations changes with sensor strength. When users are gesturing in a suitable position, gesture animations will be brighter and easier to see; when their movements are less noticeable to sensors, the animations will be less noticeable to them. Once a rhythmic gesture has been matched, the use of green hue remains unaffected by sensor strength feedback.

Audio and tactile feedback, for both gesture interactions, use temporal properties in their design: sensor strength feedback uses a Geiger counter metaphor, where the interval between 50ms signals tells users how well they can be sensed; and rhythmic gestures give feedback in response to movements, where the onset of 300ms signals tells users about their gestures. When these feedback designs are presented together, the feedback needs to be sufficiently different so that users can identify which cues are being presented. In their original designs, each feedback used 370 Hz audio signals (G#4 on piano) and 150 Hz tactile signals. If both types of feedback were given at once, one may *mask* the other; for example, sensor strength feedback may be obscured by the longer-duration rhythmic gesture tones. An alternative design is therefore required for one feedback type, so that users can reliably identify each type of signal. This is preferable to using different modalities to present different information; for example, using sound for sensor strength and vibration for rhythmic gestures. Instead, presenting the same information in both modalities gives users choice and gives systems greater flexibility; for example, if no vibrotactile display is available but speakers are, then information can be presented as audio feedback. The following two sections discuss refined designs for audio and tactile feedback, respectively.

### Audio Feedback

Audio feedback could be made discriminable by selecting a different frequency for one of the designs. A reliable rule of thumb for selecting sufficiently different frequencies is to select from different *critical bands* [159, 32]; these are bands of the frequency spectrum which are perceptually different. As 370 Hz is within the 350±50 Hz critical band, an alternative frequency could be chosen from the nearby 250±50 Hz, 450±50 Hz or 570±60 Hz critical bands. The rhythmic gesture audio feedback was changed to use a 523 Hz tone (C5 on piano); these tones should be easy to perceive over sensor strength feedback as this frequency is from a higher critical band. At frequencies below 1 kHz, sound intensity decreases

| Output | Sensor Strength | Rhythmic Gestures |
|---|---|---|
| *Light* | Brightness | Location and hue |
| *Sound* | Gap duration | Stimulus onset |
| *Vibration* | Gap duration | Stimulus onset |

**Table 7.1:** Output properties used by sensor strength and rhythmic gesture feedback.

as frequency decreases [48, 132], so the infrequent rhythmic gesture tones should also be perceived as marginally louder than the more frequent sensor strength tones.

Experiment 4 found that audio feedback was the best performing type of sensor strength feedback and was also most preferred by users. However, it was also considered annoying by many participants and this could be a barrier to its acceptable use. Therefore, an alternative audio design was considered for sensor strength feedback; this aimed to be playful and less monotonous, while still allowing accurate guidance. Instead of using a fixed frequency (of 370 Hz, as before), audio feedback for sensor strength will randomly sample a frequency from the C-Major chord (263 Hz, 330 Hz, 392 Hz; C4, E4 and G4, respectively). These frequencies fall within the 250±50 Hz and 350±50 Hz critical bands so remain perceptually different from rhythmic gesture feedback. This change affects the aesthetics of the audio feedback but does not alter its temporal structure, which is how information is encoded.

#### Tactile Feedback

Tactile feedback designs could be made discriminable by altering one of the frequencies used, like with audio feedback. However, the tactile modality has a more limited bandwidth which means perception of simultaneous tactile signals is difficult. Bensmaia *et al.* [15] found that tactile perception was poor when two waveforms, of different frequency, were presented at the same time. Hardware limitations also affect the use of different frequencies for feedback; vibrotactile actuators have limited frequency ranges with a narrow band where vibration is at its strongest [122].

Vibration intensity will be used to help participants identify feedback, instead of changing vibration frequency. Each type of information (sensor strength and feedback about rhythmic gestures) will be presented at 150 Hz[1] with reduced amplitude; when both types of feedback are presented together, the signal increases in intensity. Figure 7.1 shows how amplitude, but not frequency, is affected by presenting both signals at once. This design means audio and tactile feedback share the same temporal structure and can be presented together, or independently, with the same information encoding.

### 7.2.3 Multimodal Feedback Variations

The previous section described how sensor strength feedback and rhythmic gesture feedback can be combined, using three types of output. Gesture systems with multimodal output capabilities may wish to use all of their output modalities at once, presenting all information using interactive light, sound and vibration; this could be done using the designs discussed

[1] The optimal resonant frequency of the vibrotactile actuator used; this was the same actuator from the previous experiments.

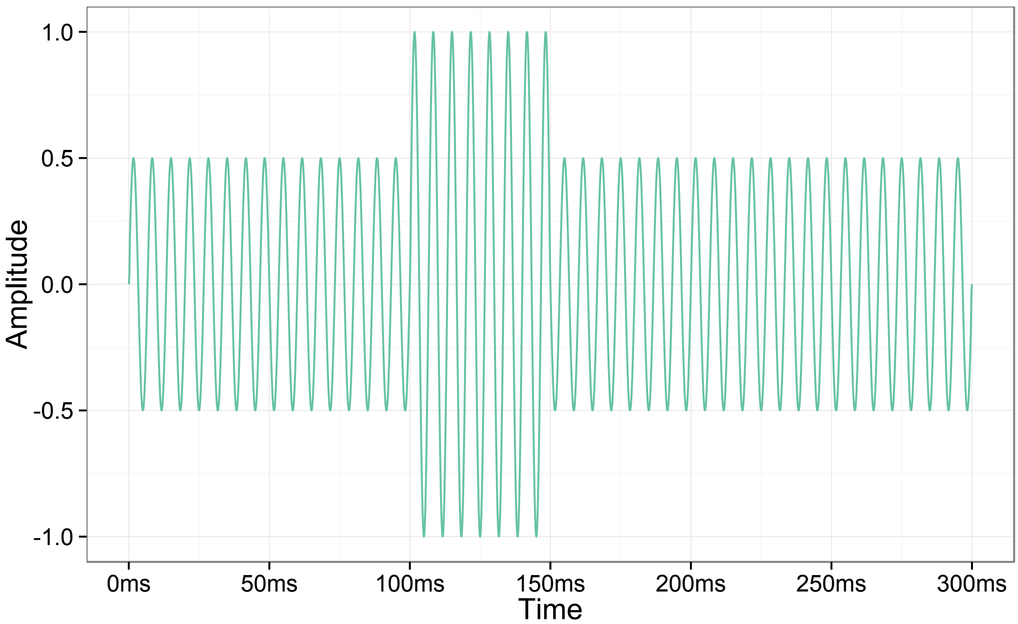



**Figure 7.1:** An example of both types of tactile feedback being presented at the same time. Here, sensor strength feedback (50ms long) is presented 100ms after a rhythmic gesture signal begins; frequency remains unaffected while amplitude is momentarily increased.

before. Alternatively, they may wish to use different modalities to present different types of information. For example, interactive light could be used to show rhythmic gesture animations and feedback, while audio and tactile outputs could be used to give sensor strength feedback. It is unknown how multimodal presentation affects the effectiveness of each type feedback, if at all.

To investigate this further, two types of combined feedback were created: *All* and *Split*. The All design presents sensor strength and rhythmic gesture feedback together using all three output types; in contrast, the Split design presents sensor strength feedback using audio and tactile output and rhythmic gesture animations and feedback using interactive light. There is not a third design where interactive light is used for sensor strength only, as this output type is needed to show rhythmic gesture animations. As discussed in the previous chapter, audio and tactile displays are less appropriate for revealing rhythmic gestures.

Sensor strength feedback helps users find a good location to perform gestures; however, this information may no longer be necessary once users have started performing gestures. To investigate if users find sensor strength feedback helpful in such situations, variations of All and Split were developed. *All-Short* and *Split-Short* present feedback like All and Split, respectively; however, once users begin gesturing, sensor strength feedback stops. All-Short presents rhythmic gesture feedback using all output modalities; Split-Short, on the other hand, only uses audio and tactile outputs for sensor strength feedback. Once sensor strength feedback stops for Split-Short, audio and tactile capabilities are no longer being used. A fifth combined feedback design — *Split-Swap* — therefore uses audio and tactile outputs for rhythmic gesture feedback, once sensor strength feedback stops. Table 7.2 provides a summary of the five feedback designs described here. These designs will be evaluated in Experiment 6.

### 7.2.4 Interaction Design Problems

This chapter has, so far, discussed ways in which sensor strength feedback and rhythmic gesture feedback can be combined. However, there are usability issues which need to be addressed when these interaction techniques are combined: sensor strength feedback changes as users move their hands, therefore it changes constantly during a rhythmic gesture performance; and some participants in Experiment 4 reported that tactile sensor strength feedback felt continuous, rather than discrete. Both of these problems, and their solutions, are discussed in the following sections.

#### Sensor Strength Changes During Gestures

As discussed in Section 5.2, sensor strength is calculated as a function of hand position relative to an input sensor. In vision-based approaches, this may be a function of the distance between hand position and a point where the sensor can capture most detail; in other approaches, this may be a function of sensor magnitude or how noisy a signal is. Regardless of the approach used in its calculation, sensor strength may vary during a rhythmic gesture movement, as illustrated by Figure 7.2. Constantly changing feedback would be ambiguous, as it does not tell users about their input position during a rhythmic gesture as a whole. To address this issue, sensor strength could be calculated using a filtered hand position, instead of using the actual hand position in real-time. This would mean sensor strength feedback tells users about where they are performing rhythmic gestures, rather than where their hand is during a gesture movement.

An appropriate filter must be selected so that sensor strength can be given about gestures

| **Design** | **Sensor Strength** | | **Rhythmic Gestures** | |
|---|---|---|---|---|
| | *Light* | *Audio & Tactile* | *Light* | *Audio & Tactile* |
| *All* | Yes | Yes | Yes | Yes |
| *All-Short* | Until gesturing | Until gesturing | Yes | Yes |
| *Split* | No | Yes | Yes | No |
| *Split-Short* | No | Until gesturing | Yes | No |
| *Split-Swap* | No | Until gesturing | Yes | After gesturing |

**Table 7.2:** A summary of the five multimodal feedback variations.

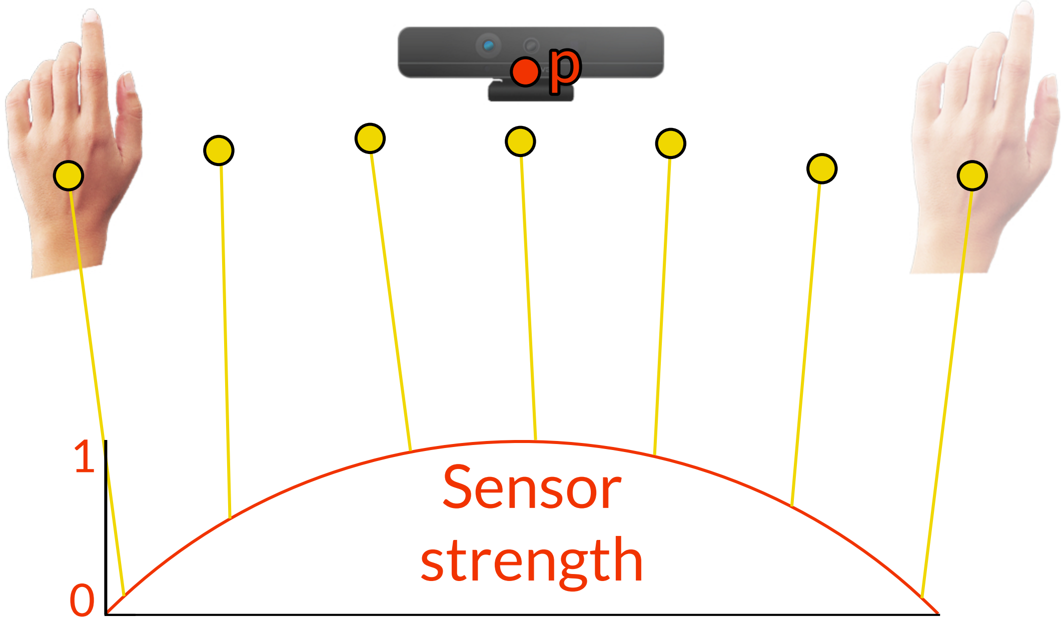



**Figure 7.2:** Sensor strength varies significantly during gestures when hand position is used in its calculation. Here, the yellow dots show hand position during a rhythmic gesture movement from left-to-right; the point **p** is where sensor strength is at its highest. Sensor strength increases as hands approach the midpoint of the gesture (closest to **p**), decreasing again as it moves past.

as a whole. For the implementation used in this thesis, which uses a depth camera to sense hand position, an exponentially-weighted moving average (EWMA) filter was selected. A moving average filter calculates an average value over time; an EWMA filter has the same effect, although more recent observations have a greater effect on the averaged value than older ones. An EWMA filter is therefore ideal for use during gesture input as it provides an averaged hand position during an entire gesture movement, but it also responds quickly when the hand starts to move away from its recent gesture path. When this filter is applied to hand position during a rhythmic gesture, sensor strength will be given about the mean position during the movement, illustrated by Figure 7.3. This position will not change much during a rhythmic gesture, as movements are repeated over time; as a result, sensor strength feedback is given about the overall gesture location.

### Tactile Feedback Feels Continuous

Tactile feedback did not perform as well in Experiment 4 as audio feedback did, despite both using the same temporal design to encode information. A small number of participants said tactile feedback often felt “continuous” when sensor strength was high, suggesting that the minimum gap between successive stimuli (70ms) was too short. To address this issue for Experiment 6, the minimum gap was increased to 100ms. This interval was chosen as it increases the gap between stimuli, but is not a significant enough increase that users would

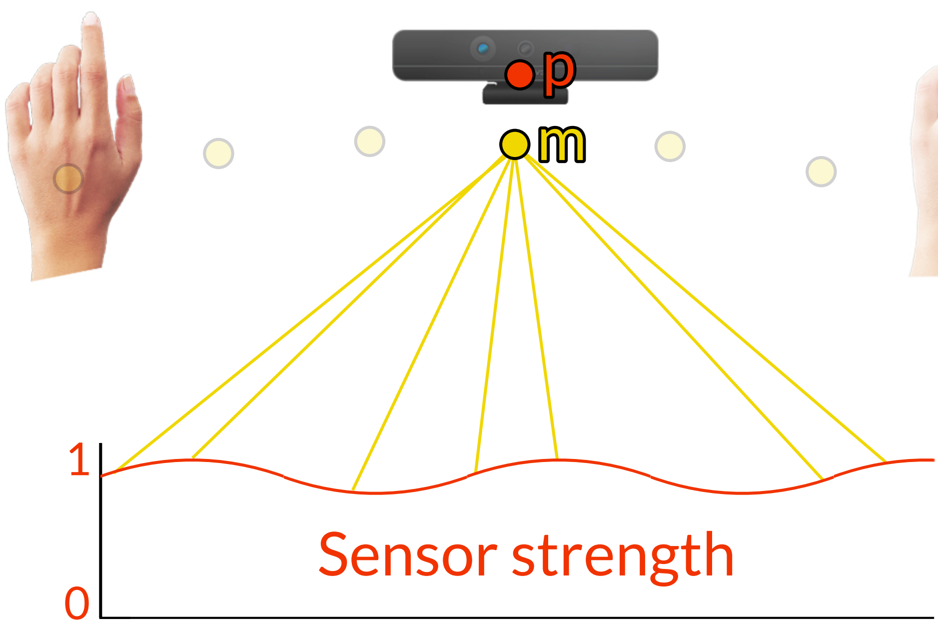



**Figure 7.3:** Sensor strength is more stable when hand position is averaged during a rhythmic gesture. Here, **m** shows the mean hand position during the movement. Since rhythmic gestures consist of repeated cyclical movements, this does not change much. Sensor strength varies less than in Figure 7.2 because the mean position remains near **p**.

find the Geiger counter feedback too time-consuming. Audio and tactile feedback, therefore, used minimum and maximum gap durations of 100ms and 400ms, respectively.

### 7.2.5 Implementation and Apparatus

The wall-mounted gesture system used in Experiments 3 and 5 was used here; this comprises a circular LED display, a Kinect depth sensor and a laptop computer. Interactive light feedback is given using the LED display, which is controlled by an Arduino device. Audio and tactile feedback are synthesised using Pure Data and are delivered through laptop speakers and a wearable tactile display, respectively. The wearable tactile display is the same device used in the previous experiments. Software from the rhythmic gestures experiment (Experiment 5) was adapted to provide sensor strength feedback.

## 7.3 Experiment 6

### 7.3.1 Research Aims

Sensor strength feedback and rhythmic gestures have both been investigated independently in earlier experiments; however, these interaction techniques should be used together when users are addressing gesture-sensing systems. Research is therefore needed to understand

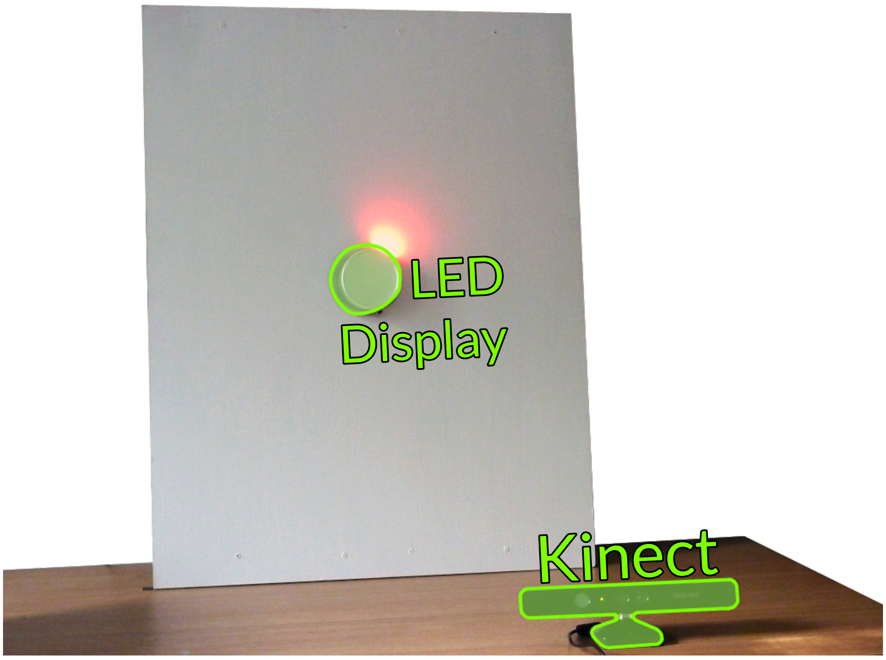



**Figure 7.4:** The wall-mounted interactive light display and Kinect sensor used in the implementation for Experiment 6.

how well these techniques work together, if at all. Experiment 6 investigated the combined use of these interaction techniques, using the five feedback designs discussed earlier in this chapter. By studying the effectiveness of these designs, this experiment gave insight into how best to present different types of information when users are addressing a gesture-sensing system; it also investigated answers to questions raised in Section 7.2.2, which motivated each of the feedback designs. This experiment used the wall-mounted gesture system used in Experiment 5; therefore, it is also the first evaluation of sensor strength feedback when users gesture from across the room, rather than immediately over a device, as in Experiment 4.

The main aims of this experiment were to: (1) investigate if sensor strength feedback and rhythmic gestures are still usable and effective when used together; (2) study effective feedback designs, to see when each type of information is useful; and (3) investigate if sensor strength feedback works for larger movements than used in Experiment 4. This research contributes answers to the following research questions:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?
>
> **RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

### 7.3.2 Experiment Design

This experiment investigated simultaneous use of sensor strength feedback and rhythmic gestures, using the feedback designs described earlier in this chapter. Participants had to do two things for each task: (1) find a target point with their hand (similar to Experiment 4); and (2) perform a rhythmic gesture as close to that target point as possible. Rhythmic gesture performance was similar to in Experiment 5: participants had to match the rhythmic gesture shown by the system and were to continue gesturing afterwards.

A within-subjects design was used, with two factors (**Feedback:** All, All-Short, Split, Split-Short, Split-Swap; and **Gesture:** Side-to-Side, Up-and-Down). The following 'Task Design' section discusses choice of **Gesture** types. Participants completed a block of tasks for each type of **Feedback**; block order was randomised using a Latin square design. Table 7.3 lists the blocks, showing the combinations of **Feedback** and **Gesture**. There were sixteen tasks per block, eight using each **Gesture**; this number was chosen after pilot testing, so that the experiment would last for around one hour.

#### Task Design

Tasks from Experiments 4 and 5 were combined for this experiment: participants were required to match rhythmic gestures, like in Experiment 5; however, they also had to perform them as close to a target point as possible, similar to Experiment 4. Target points were positioned randomly at the beginning of each task; they were placed relative to hand position, which was held in front of the body between tasks, within a 300x150x50mm volume, as illustrated by Figure 7.5. This volume was chosen because it meant target points and rhythmic gestures around them would be within comfortable reach. Points were always placed at the corners of this volume, meaning participants always had to search the space around them to find where to gesture. The z-axis of the target point volume was only 50mm; this was because participants would be seated during the experiment and larger variation in that axis may require reaching too far forward or moving backwards to find targets.

| | Name | Feedback Type | Gestures |
|---|---|---|---|
| 1 | All | All | Side-to-Side (SS), Up-and-Down (UD) |
| 2 | All-Short | All-Short | Side-to-Side (SS), Up-and-Down (UD) |
| 3 | Split | Split | Side-to-Side (SS), Up-and-Down (UD) |
| 4 | Split-Short | Split-Short | Side-to-Side (SS), Up-and-Down (UD) |
| 5 | Split-Swap | Split-Swap | Side-to-Side (SS), Up-and-Down (UD) |

**Table 7.3:** Conditions in Experiment 6. There was one condition for each **Feedback** level, with participants experiencing both levels of **Gesture** within each condition.

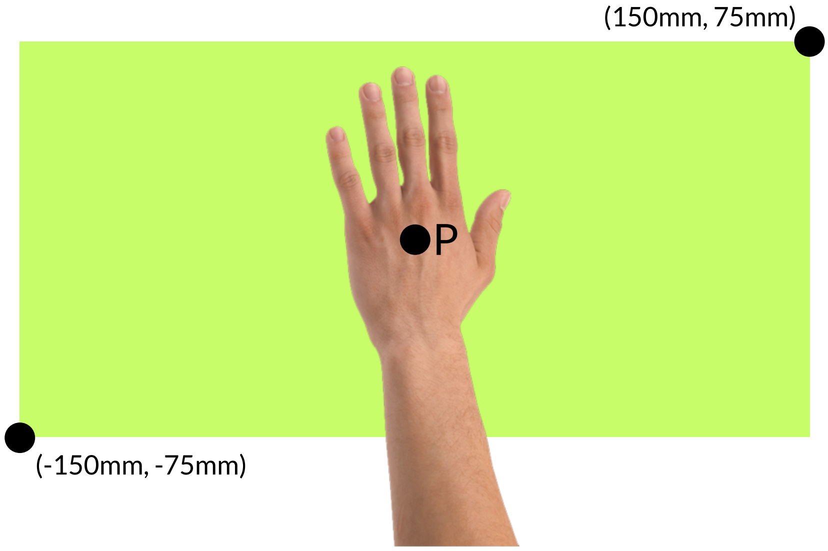



**Figure 7.5:** Target points were positioned at one of the eight corners of a 300x150x50mm volume, centred around a participant's hand position, **P**. This image illustrates the xy plane of this volume; target points were positioned 25mm in front of, or behind, this plane.

Participants were given a maximum of twelve seconds from when they began gesturing to match the rhythm. Once they matched the rhythm, they were asked to continue performing the gesture for as long as possible, for a maximum of eight gesture movement cycles. These limits were the same as in Experiment 5.

Two of the five rhythmic gestures from Experiment 5 were selected for this experiment: Side-to-Side and Up-and-Down. Only two gesture movements were chosen as this reduced the complexity of the experiment for participants; two gestures were used, rather than one, to see if findings were dependent on gesture. Side-to-Side and Up-and-Down were selected based on design recommendation **DR1** from Chapter 6 (see Section 6.5.1). Gesture interval was fixed at 700ms, the minimum recommended by **DR3**; only one interval was used so that gesture interval did not become an experimental factor, increasing the complexity of the experiment design.

### Measures

Table 7.4 presents a summary of the dependent measures in this experiment. These are described in more detail in the following sections. Participants were also asked to complete a survey, whose questions are not listed in Table 7.4 but are described later.

#### *Task Performance*

Two times were measured for each task: (1) the time spent finding where to provide input before starting a rhythmic gesture (*Time-Locate*); and (2) the time taken to match a gesture rhythm (*Time-Match*). Timing for *Time-Locate* started when participants first raised their

| Name | Description |
|---|---|
| *Task performance measures* | |
| Time-Locate | Time before starting a rhythmic gesture. |
| Time-Match | Time to match a rhythmic gesture. |
| Cycles | Count of completed rhythmic gesture movements. |
| Success | True if participant completed task, otherwise false. |
| Distance | Distance between mean hand position and task target point. |
| *Task difficulty and workload ratings* | |
| Difficulty-Locate | Difficulty rating from 1–10 (after every task). |
| Difficulty-Match | Difficulty rating from 1–10 (after every task). |
| Difficulty-Duration | Difficulty rating from 1–10 (after every task). |
| Workload | NASA-TLX workload from 0–100 (after every task block). |
| *Survey responses* | |
| Feedback | Preference rankings for feedback type (post-experiment). |
| Modality | Preference rankings for feedback modalities (post-experiment). |
| Gesture | Preference rankings for gesture type (post-experiment). |
| Acceptability | Social acceptability score (from post-experiment survey). |

**Table 7.4:** Overview of dependent measures in Experiment 6.

hand at the start of a task and ended after they completed three gesture cycles; this was the same point at which sensor strength feedback ceased in the All-Short, Split-Short and Split-Swap conditions. This measurement was taken as it gives insight into how sensor strength feedback affects users' actions before they begin a rhythmic gesture. *Time-Match* and *Time-Locate* were measured as in Experiment 5; however, timing for *Time-Match* started when *Time-Locate* ended, rather than at the start of the task as before. This was because users also had to locate a target point, something that was not required in Experiment 5.

Task *Success* was also measured, as before, showing if participants successfully matched a rhythmic gesture within the given task time limit. The number of completed gesture *Cycles* was also measured, as in Experiment 5, counting how many times they performed the gesture in synchrony after matching it successfully.

For each task, the distance between the target point and the mean hand position during a rhythmic gesture (*Distance*) was calculated. This gives an indication of how accurately sensor strength feedback guided users. The mean hand position was used, since the hand moved continuously during a rhythmic gesture performance. *Distance* was calculated as it was in Experiment 4, as the Euclidean distance in three dimensions: $\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. This was calculated in mm, using the real-world coordinates provided by the Kinect depth sensor; this will allow comparison between Experiments 4 and 6.

### *Difficulty Ratings and Task Workload*

After each task, participants were asked to make three difficulty judgements: (1) the difficulty of finding where to start the gesture (*Difficulty-Locate*); (2) the difficulty of matching the gesture rhythm (*Difficulty-Match*); and (3) the difficulty of continuing to match the rhythm for the duration of the task (*Difficulty-Duration*). They were not asked to provide *Difficulty-Duration* if they did not successfully match the rhythmic gesture. Ratings were on a ten-point scale, from 1 (easiest) to 10 (most difficult). These ratings were given verbally and were recorded electronically by the experimenter. Participants were asked about the difficulty of finding where to gesture as this could provide insight about the usability of each type of feedback. They were also asked about the difficulty of performing rhythmic gestures, as they were in Experiment 5, so that results from this experiment could be compared with earlier findings; comparing these difficulty ratings could show how, if at all, sensor strength feedback affects the ease of using rhythmic gestures.

For each completed NASA-TLX survey, an overall *Workload* measure was calculated. This calculation was the same as in Experiment 1 and is described in Section 3.3.2. *Workload* ranged from 0 to 100, where higher values mean a greater task workload.

### *Post-Experiment Surveys*

At the end of the experiment session, participants were asked to give preference rankings for feedback design (*Feedback*), sensor strength feedback type (*Modality*: Light, Audio & Tactile) and gesture (*Gesture*), from favourite to least favourite. Participants were given descriptions of each feedback design to help them when providing rankings. They were also asked to complete a short survey (shown in Appendix G), which asked them to rate their agreement with three statements about sensor strength and rhythmic gesture feedback. In particular, these questions focused on sensor strength feedback *after* they had found where to gesture and started providing input. Participants responded to these using a five-point scale, from 'Strongly Disagree' (1) to 'Strongly Agree' (5). The three statements, with alternating descriptions depending on which type of information was being asked about, were:

1. Feedback about [where to gesture, after I started gesturing / my hand movements] was useful.
2. Feedback about [where to gesture, after I started gesturing / my hand movements] was distracting.
3. Feedback about [where to gesture, after I started gesturing / my hand movements] let me know my hands were being tracked.

These statements were chosen as they could provide insight about how useful each type of information was (Q1), how distracting or disruptive they were (Q2), and how they supported awareness of system attention (Q3).

Participants were also asked to complete a shortened version of the survey from Experiment 5 (described in Section 6.3.2). This asked them to rate their agreement with statements about the usability of the two rhythmic gestures and it also investigated how socially acceptable they are. An *Acceptability* score was calculated from these responses as it was in Experiment 5. Participants were asked about the usability and acceptability of gestures to see if their responses support findings from the past study.

### Procedure

Participants were given a short tutorial at the start of the experiment; this tutorial introduced them to sensor strength feedback and the two rhythmic gestures, and gave them a chance to try each interaction technique separately. Once participants had familiarised themselves with these interactions and their feedback, they were given the chance to try both of them together. This part of the tutorial let participants experience each of the five experiment conditions.

Experiment setup was identical to Experiment 5: participants were seated 3m from the wall-mounted interactive light display and Kinect sensor, in a usability lab. The experimenter sat beside them, using a laptop computer for data entry. Participants gestured with their dominant hand and wore the tactile display on their other wrist; this was so that the cable to the tactile display did not constrain their hand movements.

Tasks started when participants said they were ready and ended after tasks were completed, or after the time limits were reached. Between tasks, participants were asked to hold their dominant hand in front of their body, at shoulder height. This was so that task target points could be generated in front of the body within comfortable reach. After each block of tasks, participants were asked to complete a NASA-TLX survey. At the end of the experiment, they were also asked to give preference rankings and complete a survey; both are described in the following section.

### Hypotheses

- **H1:** *Distance* will be lower for All and All-Short than for the other three designs;
- **H2:** *Time-Locate* will be higher for All and All-Short than for the other three designs;
- **H3:** *Time-Match* will be higher for All and Split than their -Short/-Swap versions;
- **H4:** *Difficulty-Match* will be higher for All and Split than their -Short/-Swap versions.

**H1** expects participants to be more accurate when using All or All-Short feedback. This is because sensor strength affects rhythmic gesture animations for these designs; participants are expected to focus on being more accurate so that gesture animations are more visible. This is expected because if they are not accurate in their movements, the rhythmic gesture animations will be difficult to see (due to their lower brightness). An effect of this is that participants are also expected to spend longer finding where to gesture before starting a rhythmic gesture, hence **H2** predicting higher *Time-Locate* for these conditions.

All and Split continue giving sensor strength feedback even after users have started performing a rhythmic gesture. This additional feedback is expected to place more demands on participants, as there is a greater amount of feedback for them to process. As such, **H3** and **H4** predict that matching a rhythmic gesture will take longer and will be rated more difficult when sensor strength is given throughout an entire interaction.

#### Participants

Twenty people participated in this study (five were female). Their mean age was 26.7 years (sd 3.6 years). Four participants also took part in Experiment 4 and six did Experiment 5. Participants were recruited using university email lists. Each experiment lasted for one hour and participants were paid £6.

### 7.3.3 Results

#### Success Rates

Participants successfully matched 1598 of 1600 rhythmic gestures (99.88%). Both unsuccessful trials were with the *Side-to-Side* gesture. No further analysis was performed because of insufficient data (for failing to match a gesture).

#### Time-Locate

Mean *Time-Locate* was 3686ms (sd 1577ms), which includes at least three gesture movements (2100ms). Figure 7.6 shows mean times for type of **Feedback**. Times were not normally-distributed (Shapiro-Wilk: $W = 0.98$, $p = 0.008$) so the Aligned-Rank Transform [153] was used prior to analysis. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Time-Locate*: $F(4, 171) = 40.69$, $p < 0.001$. **Gesture** did not have a significant effect ($F(1, 171) = 0.87$, $p = 0.35$), nor did the interaction between these factors ($F(4, 171) = 0.22$, $p = 0.93$). *Post hoc* comparisons for **Feedback** found that times for All and All-Short were significantly lower than all of the split designs: all $t(171) \leq -7.38$, $p < 0.001$. No other comparisons were significantly different.

#### Time-Match

Mean *Time-Match* was 2106ms (sd 984ms); Figure 7.7 shows mean times for each type of **Feedback**. Unlike *Time-Locate*, data had a normal distribution (Shapiro-Wilk: $W = 0.99$,


**Figure 7.6:** Mean *Time-Locate*. Error bars show 95% CIs.

p = 0.09) so no transformation was applied. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Time-Match*: F(4, 171) = 2.86, p = 0.02. **Gesture** did not have a significant effect (F(1, 171) = 0.03, p = 0.86), nor did the interaction between these factors (F(4, 171) = 1.17, p = 0.33). *Post hoc* comparisons for **Feedback** found one significant difference: the time taken to match a gesture was longer for All than for Split-Swap (t(171) = 2.98, p = 0.03).

### Cycles

Mean *Cycles* was 7.7 (sd 1.1); Figure 7.8 shows mean number of gesture cycles with each type of **Feedback**. These data were not normally-distributed (Shapiro-Wilk: W = 0.7, $p < 0.001$) so were transformed using the Aligned-Rank Transform [153]. A repeated-measures ANOVA found that **Feedback** had no significant effect on *Cycles*: F(4, 171) = 2.34, p = 0.06. **Gesture** did not have a significant effect either (F(1, 171) = 0.03, p = 0.85), nor did the interaction between **Feedback** and **Gesture** (F(4, 171) = 1.7, p = 0.15).

### Accuracy

Mean *Distance* was 80mm (sd 34mm); Figure 7.9 shows mean distances for all combinations of **Feedback** and **Gesture**. *Distance* was not normally-distributed so was transformed using the Aligned-Rank Transform [153] prior to analysis. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Distance*: F(4, 171) = 9.17, $p < 0.001$. **Gesture** also

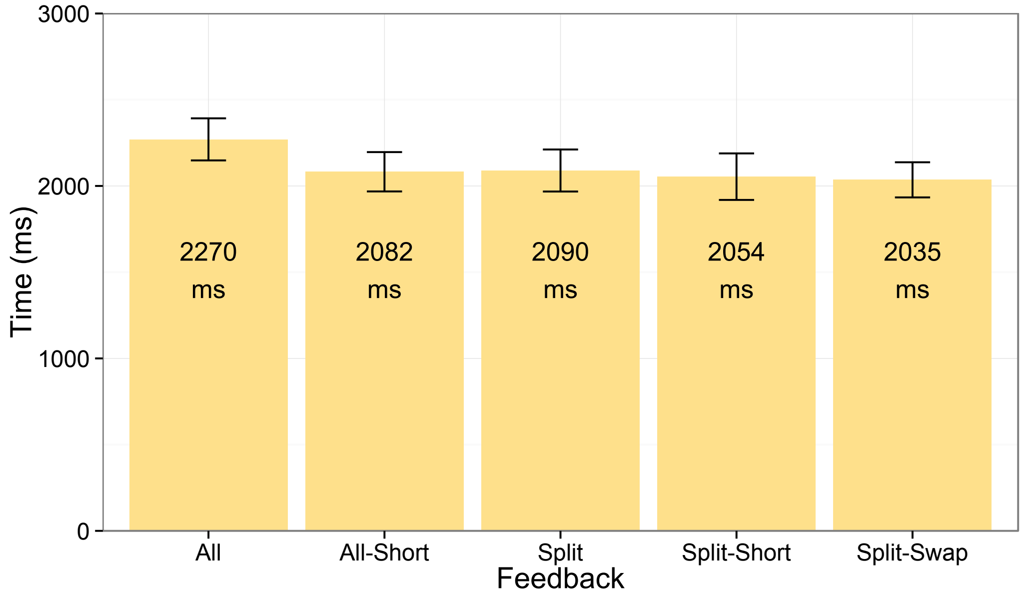



**Figure 7.7:** Mean *Time-Match*. Error bars show 95% CIs.

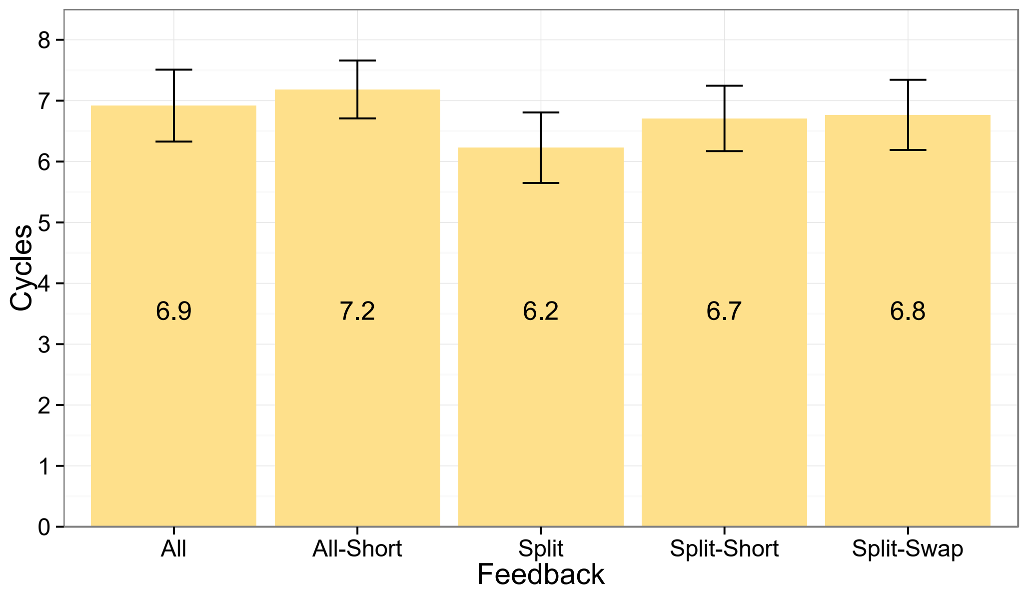



**Figure 7.8:** Mean number of gesture *Cycles*. Error bars show 95% CIs.

had a significant effect: F(1, 171) = 30.32, $p < 0.001$. The interaction between **Feedback** and **Gesture** did not have a significant effect on *Distance*: F(4, 171) = 0.41, p = 0.80.

*Post hoc* comparisons for **Feedback** found that participants were more accurate with All than all three split designs (all t(171) $\leq$ -3.93, $p \leq 0.001$); they were also more accurate with All-Short than Split-Short and Split-Swap (both t(171) $\leq$ -3.15, $p \leq 0.02$). No other differences were significant. *Post hoc* comparisons for **Gesture** found that accuracy was greater with Up-and-Down than with Side-to-Side: t(171) = 5.51, $p < 0.001$.

### Difficulty-Locate

Mean *Difficulty-Locate* was 3.70 (sd 0.95); Figure 7.10 shows mean ratings for each type of **Feedback**. All difficulty rating data was transformed using the Aligned-Rank Transform, as in Experiment 5, allowing factorial analysis using a repeated-measures ANOVA. The ANOVA found that **Feedback** had a significant effect on *Difficulty-Locate*: F(4, 171) = 34.85, $p < 0.001$. **Gesture** had no significant effect on difficulty ratings (F(1, 171) = 0.0005, p = 0.98), nor did the interaction between **Feedback** and **Gesture** (F(4, 171) = 0.47, p = 0.76). *Post hoc* comparisons for **Feedback** found that All and All-Short had significantly lower difficulty ratings than all other types of feedback: all t(171) $\leq$ -5.99, $p < 0.001$. No other differences were significant.

### Difficulty-Match

Mean *Difficulty-Match* was 3.29 (sd 0.78); Figure 7.11 shows mean ratings for each type of **Feedback**. A repeated-measures ANOVA on transformed difficulty ratings found that **Feedback** had a significant effect on *Difficulty-Match*: F(4, 171) = 56.56, $p < 0.001$. **Gesture** had no significant effect on difficulty ratings (F(1, 171) = 1.10, p = 0.30), nor did the interaction between **Feedback** and **Gesture** (F(4, 171) = 0.09, p = 0.98). *Post hoc* comparisons for **Feedback** found that difficulty ratings were lower for All and All-Short than for all other types of feedback (all t(171) $\leq$ -3.91, $p < 0.001$) and that ratings were lower for Split-Swap than for Split and Split-Short (both t(171) $\leq$ -5.95, $p < 0.001$). No other comparisons were significant.

### Difficulty-Duration

Mean *Difficulty-Duration* was 3.21 (sd 0.87); Figure 7.12 shows mean ratings for each type of **Feedback**. A repeated-measures ANOVA on transformed difficulty ratings found that **Feedback** had a significant effect on *Difficulty-Duration*: F(4, 171) = 45.61, $p < 0.001$. **Gesture** also had a significant effect: F(1, 171) = 4.65, p = 0.03. The interaction between

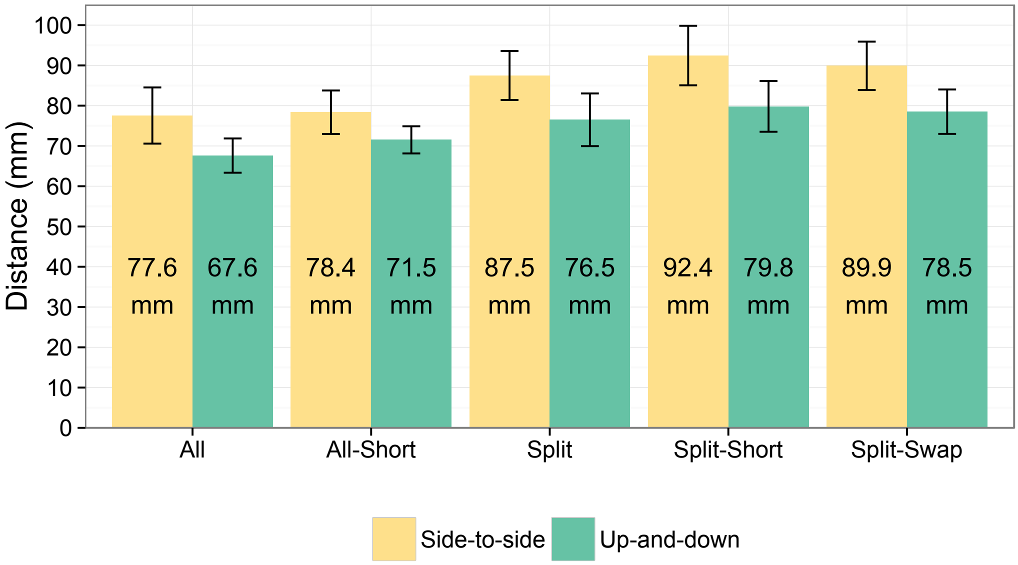



**Figure 7.9:** Mean *Distance* for each **Feedback** and **Gesture**. Error bars show 95% CIs.

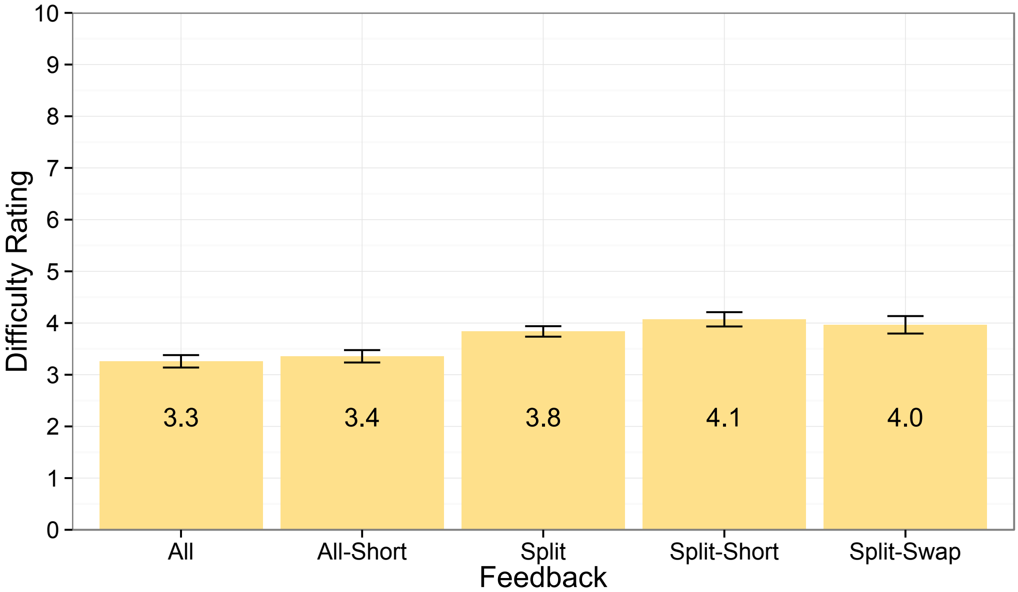



**Figure 7.10:** Mean *Difficulty-Locate*. Error bars show 95% CIs.

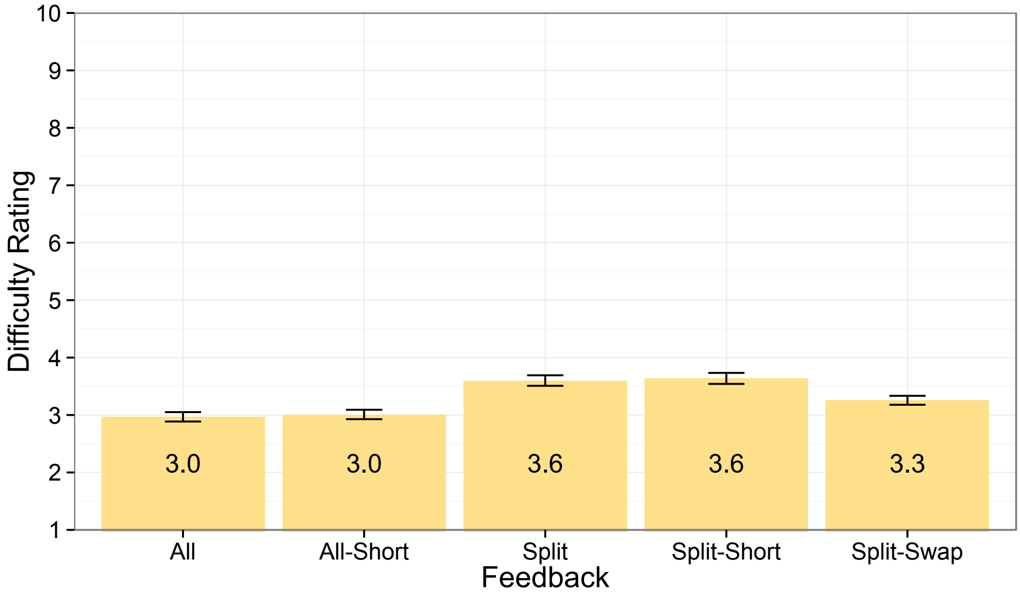



**Figure 7.11:** Mean *Difficulty-Match*. Error bars show 95% CIs.

**Feedback** and **Gesture** did not have a significant effect: F(4, 171) = 0.19, p = 0.94. *Post hoc* comparisons for **Feedback** found that difficulty ratings were higher for Split and Split-Short than all other types of feedback: all $t(171) \leq -6.55$, $p < 0.001$. No other comparisons were significant. *Post hoc* comparisons for **Gesture** found that difficulty ratings for Side-to-Side were significantly lower: t(171) = -2.16, p = 0.03.

### Workload

Mean *Workload* was 32.75 (sd 9.14); Figure 7.13 shows mean *Workload* for each type of **Feedback**. A repeated-measures ANOVA found that **Feedback** had a significant effect on *Workload*: F(4, 76) = 11.45, $p < 0.001$. *Post hoc* comparisons found that All had significantly lower task workload than Split, Split-Short and Split-Swap (all $t(76) \leq -3.59$, $p \leq 0.005$); All-Short also had significantly lower workload than Split-Short (t(76) = -4.76, $p < 0.001$). No other differences were significant.

### Preference

Table 7.5 shows median preference ranks for feedback designs, sensor strength feedback modality and gesture movements; a rank of '1' was the most preferred option. The Friedman's rank sum test was used to see if these rankings were significant or not.

A Friedman's test found that preference rankings for *Feedback* were significant: $\chi^2(4) = 52.28$, $p < 0.001$. *Post hoc* comparisons using Wilcoxon's signed-rank test found that All and All-

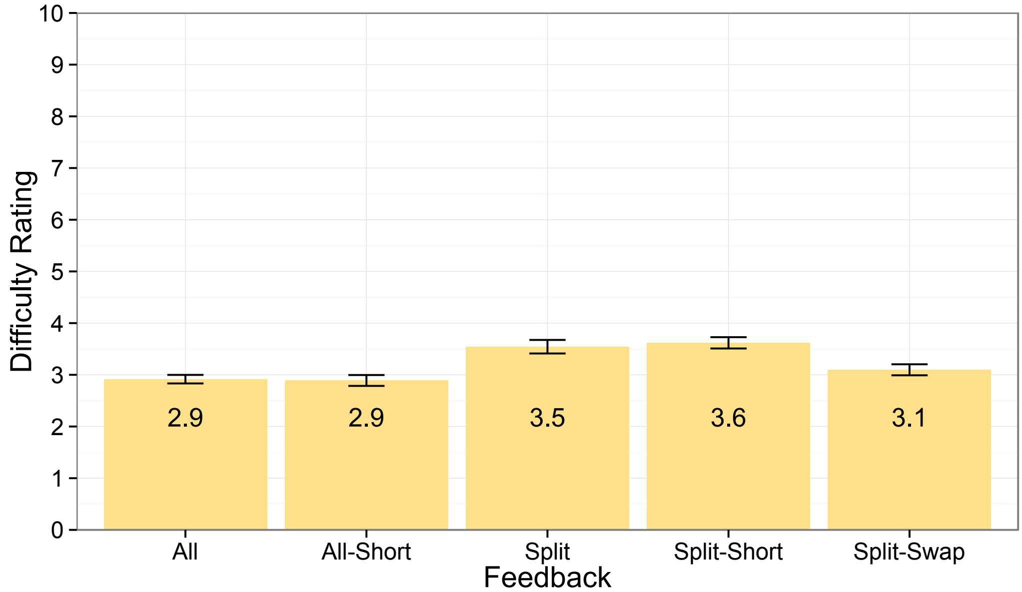



**Figure 7.12:** Mean *Difficulty-Duration*. Error bars show 95% CIs.

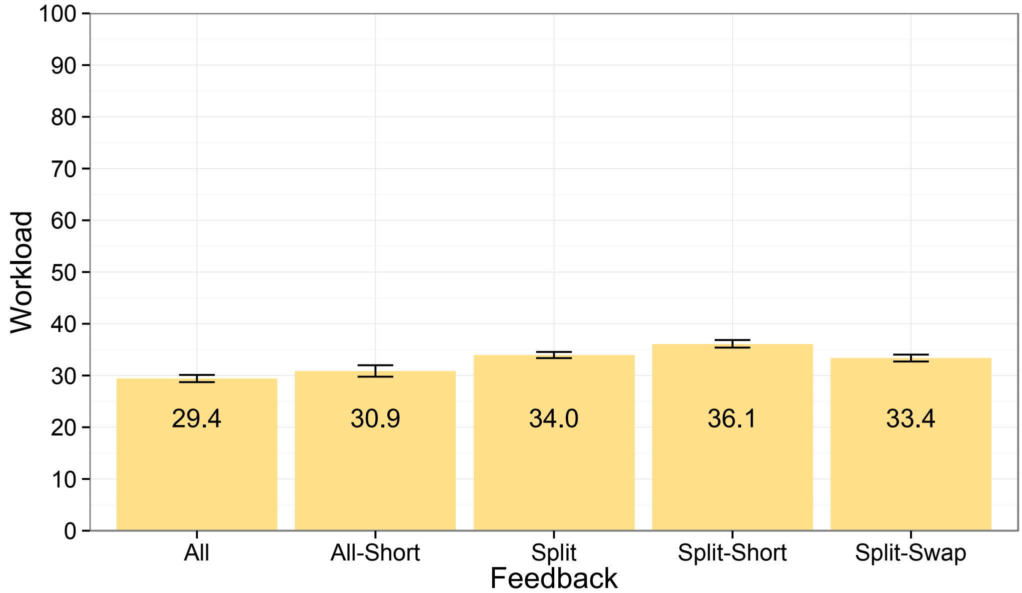



**Figure 7.13:** Mean *Workload*. Error bars show 95% CIs.

| **Feedback Design** | | | | | **Modality** | | **Gesture** | |
|---|---|---|---|---|---|---|---|---|
| All | All-Short | Split | Split-Short | Split-Swap | Light | Non-Visual | SS | UD |
| 2 | 1 | 4 | 5 | 3 | 1 | 2 | 1 | 2 |

**Table 7.5:** Median ranks for *Feedback*, *Modality* and *Gesture*.

Short were ranked significantly higher than all other types of feedback (all $p \leq 0.02$) and Split-Swap was ranked significantly higher than Split-Short ($p < 0.001$).

A Friedman's test found that preference rankings for *Modality* were not significant: $\chi^2(1) = 0.8$, p = 0.37. Preference rankings for *Gesture* were significant: $\chi^2(1) = 16.2$, $p < 0.001$; Side-to-Side was most preferred, with only one participant preferring the Up-and-Down gesture.

### Survey Ratings

Table 7.6 shows median responses to the statements about sensor strength and rhythmic gesture feedback. Friedman's rank sum tests were used to compare ratings between feedback types. Rhythmic gesture feedback was rated more useful than sensor strength feedback, once users had began perform gestures: $\chi^2(1) = 16$, $p < 0.001$. Sensor strength feedback was rated more distracting than rhythmic gesture feedback: $\chi^2(1) = 14$, $p < 0.001$. Rhythmic gesture feedback also provided greater awareness of tracking: $\chi^2(1) = 12.25$, $p < 0.001$.

Table 7.7 shows median responses to the questions about each rhythmic gesture used in this experiment. Friedman's rank sum tests were used to compare ratings between gestures, as this approach was also used in Experiment 5. No significant differences were found for any question: all $\chi^2 \leq 4.0$, $p \geq 0.05$.

### Social Acceptability

Mean *Acceptability* was 69.6% (sd 14.7%); Table 7.8 shows *Acceptability* scores, summarised for each of the six social situations and the two rhythmic gestures used in this experiment. Cochran's Q test, with *post hoc* McNemar's tests, were used to analyse *Acceptability*, as in Experiment 5. Cochran's Q test found that situation had a significant effect on *Acceptability*: $\chi^2(5) = 141.78$, $p < 0.001$. *Post hoc* comparisons found that *home, alone*, *home, family* and *work, alone* had significantly higher acceptability scores than the other three social situations (all $p \leq 0.002$), but not each other (all p = 1.0). *Work, others* also had higher greater acceptability scores than both *public* situations (both $p < 0.001$). No other

| **Question** | **Type of Feedback** | |
|---|---|---|
| | Sensor Strength | Rhythmic Gestures |
| *Q1: Feedback was useful* | 4 | 5 |
| *Q2: Feedback was distracting* | 2 | 1 |
| *Q3: Feedback made aware of tracking* | 3.5 | 5 |

**Table 7.6:** Median survey responses to each question for each feedback. Responses show agreement on a five-point scale, from "Strongly Disagree" (1) to "Strongly Agree" (5).

| Question | Gesture | |
|---|---|---|
| | Side-to-Side | Up-and-Down |
| *Q4: Easy to perform* | 5 | 5 |
| *Q5: Animation easy to follow* | 5 | 5 |
| *Q6: Easy to keep matching rhythm* | 5 | 5 |

**Table 7.7:** Median survey responses to each question for each gesture. Responses show agreement on a five-point scale, from "Strongly Disagree" (1) to "Strongly Agree" (5).

pairwise comparisons were significant. Wilcoxon's signed-rank test was used to compare *Acceptability* between gestures. No significant difference was found (Z -1.70, p = 0.09).

### 7.3.4 Comparison of Experiments 4, 5 and 6

#### Finding Where to Gesture in Experiments 4 and 6

In Experiment 4, each experiment task required participants to locate a target point using sensor strength feedback. In Experiment 6, locating a target point was just one part of the experiment task design; participants also had to perform a rhythmic gesture. It was therefore expected that they would spend less time locating target points in Experiment 6 than they did in Experiment 4, as they also had another activity to complete for each task. The Mann-Whitney U test was used to compare times from these experiments; this test was chosen because it is appropriate for comparing non-normal data (both Shapiro-Wilk's $W \leq 0.95$, $p < 0.001$) from independent samples. As expected, time spent finding where to gesture was significantly lower in Experiment 6: $Z = 18.59$, $p < 0.001$. The mean times for Experiments 4 and 6 were 7174ms and 3686ms, respectively.

| Social Situation | Acceptability | | Gesture | Acceptability | |
|---|---|---|---|---|---|
| | *mean* | *sd* | | *mean* | *sd* |
| At home, alone | 100.0% | 0.0% | SS | 73.3% | 16.6% |
| At home, with family | 100.0% | 0.0% | UD | 65.8% | 16.6% |
| At work, alone | 100.0% | 0.0% | | | |
| At work, with colleagues | 75.0% | 25.6% | | | |
| In public, with friends | 25.0% | 38.0% | | | |
| In public, with Strangers | 17.5% | 24.5% | | | |

**Table 7.8:** Mean *Acceptability* scores for each social situation and rhythmic gesture.

### Distance from Targets in Experiments 4 and 6

The *Distance* from target points in Experiments 4 and 6 were compared, using the Mann-Whitney U test. This test was used as *Distance* was not normally-distributed (both Shapiro-Wilk W $\leq$ 0.96, p $<$ 0.001). The mean *Distance* for Experiments 4 and 6 was 51.4mm and 81.5mm, respectively. The difference between these is significant: Z = -9.5, p $<$ 0.001.

### Rhythmic Gesture Cycles in Experiments 5 and 6

Participants used the Side-to-Side and Up-and-Down gestures with 700ms intervals in Experiment 6; these rhythmic gestures were also used in some trials in Experiment 5. The number of *Cycles* (synchronised gesture movements) for these rhythmic gestures was compared between Experiments 5 and 6, to see if the addition of sensor strength feedback in Experiment 6 was detrimental to gesture performance. The time taken to match a rhythmic gesture (*Match*) could not be compared due to differences in how these values were calculated between the experiments; in Experiment 6, this timing was different as participants also had to find where to gesture first.

The Mann-Whitney U test was used to compare *Cycles* between these experiments; this test was used because *Cycles* was not normally-distributed in either experiment: both Shapiro-Wilk's W $\leq$ 0.58, p $<$ 0.001. The mean number of *Cycles* for these gestures in Experiments 5 and 6 were 6.8 and 7.7, respectively. The Mann-Whitney test found no significant difference in *Cycles*: Z = 1.08, p = 0.28.

### Rhythmic Gesture Difficulty Ratings in Experiments 5 and 6

Difficulty ratings from Experiment 6 (*Difficulty-Match* and *Difficulty-Duration*) were compared to difficulty ratings from Experiment 5, for the gestures mentioned in the previous section comparing performance times. Mann-Whitney's U test was used to compare difficulty ratings, as it was also used for performance data. The mean *Difficulty-Match* ratings were 2.47 and 3.29, respectively. The mean *Difficulty-Duration* ratings were 2.10 and 3.21, respectively. *Difficulty-Match* was significantly higher in Experiment 6 than in Experiment 5: Z = -10.21, p $<$ 0.001. *Difficulty-Duration* was also significantly higher in Experiment 6: Z = -11.82, p $<$ 0.001.

### 7.3.5 Discussion

#### Finding Where to Gesture

Five feedback designs were compared in this experiment. Of these five, All and All-Short performed the best. Participants gestured closer to the target area when given these types of feedback and they also spent less time finding where to start gesturing. These findings support accepting **H1**; however, **H2** cannot be accepted. It was expected that participants would spend longer finding where to gesture with All and All-Short, as the visibility of gesture animations was affected by sensor strength. This was not the case and participants were even faster with these designs. They also gave All and All-Short lower difficulty ratings for finding where to gesture. This may have been because visual feedback changes more quickly than the audio and tactile feedback, allowing users to rapidly locate the right area for input.

Similar findings were observed in Experiment 4: interactive light led to faster task times when searching for a target point. However, there was a tradeoff between accuracy and time in Experiment 4; interactive light feedback led to lower accuracy but faster interaction. There was no tradeoff between speed and accuracy here, as All and All-Short had the greatest accuracy and the fastest times. The reason for this difference may have been because target points were easier to find in Experiment 6. Target points were always positioned within reach of users' hands (for practical reasons), limiting the amount of searching necessary. Had a much wider area been used, requiring users to reposition their body in the room, the results may not necessarily have been the same. However, many participants did report that interactive light helped them see immediately if they were in a good gesturing position and this would likely be the case when moving in a wider area. Based on these findings, interactive light should be used for sensor strength feedback.

Participants in Experiment 6 gestured less accurately (with respect to target points) than those in Experiment 4. However, this was expected. In Experiment 6, participants were interacting from a much greater distance (3m across the room, rather than 10cm over a device) and they also performed gestures, whereas they only had to locate a target with their index finger in Experiment 4. Although the difference was significant, it was less than 30mm. Participants in Experiment 6 still achieved a high degree of accuracy, suggesting that sensor strength feedback is an effective way of guiding users over longer distances as well as shorter ones. It was also expected that participants would spend less time finding where to gesture in Experiment 6, since that was only part of the experiment task. Analysis supports this, finding that participants spent almost half as long as they did in Experiment 4.

### Effects of Feedback on Interaction

Although feedback design had an effect on how long users spent finding where to gesture, it did not have much impact on gesture times; there were significant effects on time but *post hoc* tests found few significant comparisons. Participants gave lower difficulty ratings for All, All-Short and Split-Swap than they did for Split and Split-Short; they also preferred those three feedback designs the most. These findings suggest that they found it easier to perform gestures when given audio and tactile feedback about their movements, as those designs were the only ones to give non-visual rhythmic gesture feedback. Survey responses to Q1 support this, with participants strongly agreeing that rhythmic gesture feedback was useful. Based on these findings, audio and tactile feedback about gestures should be given when possible.

The finding that feedback had little impact on rhythmic gesture times but improved user experience is consistent with other findings presented in this thesis. For example, Experiments 1 and 2 found that tactile feedback had no impact on gesture selection times, but participants reported that it improved their awareness of how the systems were responding to their gestures. It would seem that in these cases, the main benefits of giving feedback is that it keeps users informed about system attention and gives them confidence that gesture systems are responding to them as they expect. Improving user experience is more important than gesture performance times, which are only a few seconds long here, so it is still worth giving feedback about gestures.

Participants were expected to find it more difficult to match rhythmic gestures when given sensor strength feedback, as that information was thought to be a distraction and of little use once they started gesturing. However, findings from this experiment did not support this idea; there were no differences in time or difficulty ratings between All and All-Short, and between Split and Split-Short. The only significant difference was in difficulty ratings between Split and Split-Swap. Based on these findings, **H3** and **H4** cannot be accepted.

A comparison between Experiments 5 and 6 found that the number of completed gesture cycles was not significantly different, although gesture difficulty was higher in Experiment 6. This could suggest that sensor strength feedback made it more difficult to perform rhythmic gestures; however, this contradicts the findings discussed in the previous paragraph. Instead, difficulty ratings may have been higher in Experiment 6 because participants were also focusing on where they were gesturing. Regardless, the mean difficulty ratings were less than 3.3 out of 10, suggesting that participants did not find interaction difficult. Based on these findings, sensor strength feedback could be given throughout an entire interaction without significant detriment to usability. Users may find this feedback helpful when they stop addressing a gesture interface and start interacting with it, as it could help them continue gesturing where their actions can be sensed.

### Comparing Rhythmic Gestures

Two rhythmic gestures were used in this experiment: Side-to-Side and Up-and-Down, with 700ms intervals. Gesture type had no impact on performance times, although participants were more accurate (with respect to target points) when using the Up-and-Down gesture than they were when using Side-to-Side. From observation of participants during the experiment, this may have been because their Up-and-Down movements were more controlled than the Side-to-Side ones. When performing Side-to-Side gestures, participants' hand movements typically followed an arc from one side to the other, whereas the Up-and-Down gestures followed a straighter path. This difference in movement patterns may have caused the greater inaccuracy for Side-to-Side, as gestures could drift over time. However, the difference between the gestures was small (mean distance from target was 74.79mm and 85.15mm for Up-and-Down and Side-to-Side, respectively); this would be unlikely to cause problems during interaction.

## 7.4 Limitations

### 7.4.1 Device Form Factor

A wall-mounted gesture system was used in Experiments 5 and 6, meaning the use of rhythmic gestures with other device form factors — like mobile phones — has yet to be studied. Rhythmic gestures are not device-specific and their use of interactive light displays means that many limited-display devices can use them effectively. None of the findings from this research suggests that rhythmic gestures would not be as usable with other device types. More work is needed to investigate the usability of rhythmic gestures with mobile phones, as well as appropriate choice of gesture movements (as discussed in Section 6.4.2); however, this is out of the scope of this thesis research.

## 7.5 Conclusions

This chapter described research which investigated the combination of sensor strength feedback with rhythmic gestures, creating an approach which allows users to address gesture-sensing systems. Experiment 6 found that these interaction techniques remain usable when combined; participants gestured with a high level of accuracy and they performed rhythmic gestures well. Comparisons to earlier experiments show that usability was not significantly affected by the simultaneous use of these interactions. The use of interactive light, audio and tactile feedback means that the approach described here can be used effectively by limited-display devices. The following section provides design recommendations for how such devices can make the most of these interactions. Finally, this chapter concludes by discussing answers to **RQ3** and **RQ5**.

### 7.5.1 Design Recommendations for Address Interactions

In addition to the four design recommendations for using rhythmic gestures, presented in Section 6.5.1, this chapter contributes recommendations relating to the combined use of sensor strength feedback and rhythmic gestures.

#### DR5: Use interactive light for sensor strength feedback

When interactive light was used to show sensor strength (the All and All-Short feedback designs), participants found it easier to find where to gesture; this is reflected by performance data (time to match, gesture duration and success rates), difficulty ratings and preference rankings. Interactive light changed quickly in response to hand movements and the direct effect on the rhythmic gesture animations meant that users were aware of how well they were being sensed. Gesture-sensing systems should use interactive light to show sensor strength as well as gesture animations, even if audio and tactile sensor strength feedback are also being given.

#### DR6: Give rhythmic gesture feedback from the start of an interaction

Lower difficulty ratings were given for feedback designs where rhythmic gesture feedback was given using the audio and tactile modalities, especially when feedback was given from the beginning of a gesture. This recommendation extends **DR2** ("*Give users feedback about their gestures*", Section 6.5.1) with the further recommendation that rhythmic gesture feedback be given from the start of an interaction, rather than from when users' intent to gesture is established. The Split-Swap feedback design considered giving rhythmic gesture feedback later in an interaction; while participants preferred this to receiving no feedback about their gestures, they preferred it less than when they were given audio and tactile gesture feedback from the start (with All and All-Short).

#### DR7: Give short sensor strength feedback with the audio and tactile modalities

Experiment 6 compared constant sensor strength feedback to 'short' versions, which stopped once users started performing a rhythmic gesture. There were no significant differences found between full and shortened versions of this feedback. As such, it is recommended that minimal audio and tactile sensor strength feedback be used. Unnecessary audio feedback may be obtrusive and annoying to others and, as found in Experiment 1, tactile feedback can also be obtrusive when too much is given. Interactive light feedback is the exception to this recommendation; as discussed with **DR5**, participants benefit from that and it should not change much during interaction if users continue to gesture in a good position.

### 7.5.2 Research Question 3

Research in Chapter 5 began investigating an answer to **RQ3**:

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

Experiment 4 found that sensor strength feedback was an effective way of guiding users, with a mean accuracy of 51mm (sd 15mm). However, the task design in that experiment was not representative of how sensor strength feedback would be used when addressing a gesture system. It also only evaluated it in a context where users gestured directly over a mobile phone. More research was needed to evaluate its use in a realistic scenario and in a different context, where users gesture from a greater distance.

Findings from this chapter allow a more complete answer to be given to **RQ3**. When used alongside rhythmic gestures, performed from up to 3m from an input device, sensor strength feedback achieved a mean accuracy of 80mm (sd 34mm). Participants in Experiment 6 were found to be less accurate than those in Experiment 4; however, this was with the additional demands of performing gestures. They also took significantly less time to find where to gesture. These findings suggest that limited-display devices can guide users to within 80mm of a target point, with greater accuracy possible when they interact from shorter distances and take more time.

### 7.5.3 Research Question 5

This chapter investigated an answer to the final research question in this thesis:

> **RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

Based on findings from Experiment 6, yes, limited-display devices can help users find where to gesture while also indicating which system they intend to interact with. By combining sensor strength feedback with rhythmic gestures, gesture systems can guide users and show them how to direct their input. Experiment 6 found this combination of interactions effective, with users gesturing close to target points while also matching and performing gestures with ease. These interaction techniques only used simple interactive light, audio and tactile displays, meaning that they are suitable for limited-display devices. Gesture-sensing systems with richer output capabilities could also use these techniques to help users address them, meaning these interactions are suitable for use across a wide range of device types.

### 7.5.4 Contributions

The research in this chapter makes the following contributions:

- It describes and evaluates a new interaction technique for addressing gesture interfaces, by combining sensor strength feedback with rhythmic gestures;
- It investigates sensor strength feedback in a different usage context, demonstrating its effectiveness when interacting from a greater distance than in Experiment 3;
- It presents design recommendations based on the experiment findings; these will help others use these interaction techniques effectively.

# Chapter 8

# Conclusions

## 8.1 Introduction

This thesis made the following statement in its Introduction:

> This thesis asserts that multimodal feedback can help users address and interact with in-air gesture systems, even when their devices have limited or no screens. Tactile, audio and interactive light displays offer gesture systems a variety of ways of presenting information during gestural interactions, overcoming the need for visual feedback on screens. This thesis presents novel interaction techniques which use these modalities to help users address in-air interfaces, showing them where and how to gesture.

In the chapters which followed, research was presented which supports this statement, investigating answers to the thesis research questions. Chapters 3 and 4 studied tactile feedback and interactive light feedback for gesture systems, respectively, demonstrating their potential effectiveness and contributing a better understanding of how these novel output modalities can be used by limited-display devices for gesture feedback. They also informed the later use of these modalities for helping users address gesture systems. Chapters 5 to 7 then investigated interaction techniques which allow users to address gesture-sensing systems. The experiments in these chapters found users could use these interaction techniques successfully. This chapter now summarises this research and revisits each of the research questions, discussing how they were addressed and summarising their answers. It also summarises the main contributions of this research and discusses areas for future work.

## 8.2 Research Questions

### 8.2.1 Research Question 1

**RQ1:** How can in-air gesture interfaces present tactile feedback to users?

Tactile displays offer gesture systems, especially those with limited output capabilities, an alternative way of giving users feedback about their gestures. However, little was known about how to use this output modality for gesture feedback. This thesis contributes a first detailed investigation of tactile feedback for gesture systems, as well as a collection of interaction techniques using this modality. Chapter 3 described two experiments which studied how such feedback may be designed and presented. Experiment 1 compared an ultrasound haptic display with wearable tactile displays at different body locations. It found divided preferences for feedback location and there was little effect of location on task performance. Experiment 2 focused on tactile feedback design and ways of encoding information. It found that users benefit from dynamically changing feedback, as changes in the vibration, expected or otherwise, gave them cues about what was happening. Findings from both experiments show that users benefit from tactile feedback about their gestures, as it provides assurance that they are interacting correctly and it reduces their reliance on visual information. These findings informed the use of tactile feedback in interaction techniques for addressing gesture systems, investigated later in the thesis.

### 8.2.2 Research Question 2

**RQ2:** Can interactive light be used to present gesture feedback to users?

Interactive light displays are a novel way of presenting visual information from devices with limited display capabilities. However, it was not known if — and how — they could be used effectively by gesture systems. This thesis contributes an initial evaluation of interactive light feedback about gestures, as well as a set of interaction techniques which use interactive light for feedback and for helping users address gesture systems. Chapter 4 described the Gesture Thermostat, an example of a limited-display device which has no screen but has other output capabilities. Interactive light was used to give feedback about gesture sensing, as well as for displaying information about the thermostat. Experiment 3 found that interactive light feedback from the thermostat was useful when addressing and interacting with it. It also found that interactive light displays might be enhanced by audio and tactile feedback. Interactive light was also successfully used as part of multimodal feedback in Experiments 4–6, demonstrating its potential as a feedback modality.

### 8.2.3 Research Question 3

> **RQ3:** To what extent can limited-display devices guide users when finding where to perform gestures?

To address a gesture-sensing system, users need to know where to gesture and how to direct their input towards only that system. Chapter 5 presented sensor strength feedback, an interaction technique which helps users find where to perform gestures. Experiment 4 evaluated the use of this technique with a gesture-sensing mobile phone, finding that it could guide users with 51mm accuracy in a localisation task. Chapter 7 combined sensor strength feedback with rhythmic gestures, another interaction technique, telling users where to gesture and how to gesture at the same time. Experiment 6 evaluated this new technique with a gesture system using a wall-mounted interactive light display, finding that users performed gestures within 80mm of target points. These experiments show that sensor strength feedback can be used effectively alongside other gesture interaction techniques, guiding users with 51mm accuracy, when gesturing close by, or 80mm accuracy, when gesturing from across the room. Multimodal sensor strength feedback was especially effective, with audio leading to greater precision and interactive light leading to faster interaction times.

### 8.2.4 Research Question 4

> **RQ4:** How can users direct their gestures towards a gesture-sensing system with limited display capabilities?

Users need to direct their input appropriately when addressing a gesture system, so that their gestures do not unintentionally affect other systems. Chapter 6 described rhythmic gestures, an interaction technique which can be used to direct input. Rhythmic gestures are hand movements repeated in time with a rhythmic beat. Experiment 5 studied how well users can perform such gestures, as well as the effects of feedback on gesture performance. It found that users can match rhythmic gestures with ease, with certain designs yielding very high performance. They can also match them well without feedback about their hand movements, although this extra feedback made gesturing easier. Chapter 6 contributes design recommendations for rhythmic gestures, informed by the experiment findings.

### 8.2.5 Research Question 5

> **RQ5:** Can limited-display devices help users find where to gesture while also directing their input towards a gesture-sensing system?

Chapter 7 combined several aspects of the research in this thesis. It described an interaction technique which combined sensor strength with rhythmic gestures, using interactive light, audio and tactile displays for feedback. This interaction technique aimed to help users address gesture systems by showing them where and how to gesture at the same time. Experiment 6 evaluated its usability, also investigating how to best use multimodal display capabilities for feedback. The experiment findings showed that this technique was successful; users were able to perform rhythmic gestures well (99.9% completion) while also locating target points with good accuracy (80mm). This was done with just interactive light, audio and tactile feedback, demonstrating the success of the interaction with limited-display devices. Chapter 7 also contributed design recommendations for using sensor strength feedback and rhythmic gestures together.

## 8.3 Contributions

This thesis makes novel contributions which inform the design of gesture-sensing systems. Its main contributions are: (1) a study of tactile feedback for in-air gestures; (2) investigation of interactive light feedback about gestures; and (3) new interaction techniques for addressing gesture systems. This section summarises these contributions.

### Tactile Feedback for Gesture Systems

This thesis contributes an experimental comparison of an ultrasound haptic display with wearable tactile displays for tactile feedback. It also presents an initial investigation of body locations for gesture feedback from wearable devices, comparing tactile feedback on the wrist to feedback on fingers. The experiment findings show that tactile feedback can improve the user experience of in-air gesture interaction, while also helping users address gesture systems.

### Interactive Light Feedback for Gesture Systems

This thesis contributes an initial investigation of the usability of interactive light feedback about gestures. It shows that interactive light displays can be used successfully for displaying visual feedback and shows that they can be used effectively in interaction techniques for addressing gesture systems.

## Interaction Techniques for Addressing Gesture Systems

This thesis focused on the problem of addressing gesture systems, a "fundamental" [11] part of interaction. It contributes two new interaction techniques — *sensor strength feedback* and *rhythmic gestures* — which can be used together, or independently, when addressing a gesture system. These were investigated in three experiments, which found the interaction techniques successful and led to a set of design recommendations, informed by their findings. Overall, the findings of this thesis show that the interaction techniques, supported by novel multimodal feedback, can be used to help users address gesture systems. Although the techniques focus on addressing gesture systems, the feedback investigated in this thesis could continue to be given throughout the remainder of the interaction, e.g. to continue informing users of how well they can be sensed by the system.

### 8.3.1 Design Recommendations

This thesis made seven design recommendations for using sensor strength feedback and rhythmic gestures, informed by the experiment findings. These design recommendations are summarised here; for further discussion about them, see Section 6.5.1 (for **DR1–4**) and Section 7.5.1 (for **DR5–7**).

#### DR1: Prioritise use of the Side-to-Side and Up-and-Down gestures

The *Side-to-Side* and *Up-and-Down* movements were the easiest rhythmic gestures to use and also had the best performance, so should be used for rhythmic gestures when possible. *Side-to-Side* was also rated the most socially acceptable movement, so users may feel most comfortable using it when addressing gesture systems around others.

#### DR2: Give users feedback about their gestures

Rhythmic gesture feedback from audio and tactile displays had no effect on gesture performance but did make interaction easier. As such, users should be given this non-visual feedback about their rhythmic gestures when possible.

#### DR3: Use a minimum interval of 700ms; 900ms for circular gestures

Rhythmic gestures were more difficult with 500ms intervals than with longer intervals, so such short intervals should not normally be used (however, see **DR4**). Rhythmic gestures should use a minimum interval of 700ms (or 900ms for the more difficult circular movements) as users performed gestures well with these intervals.

### DR4: Reserve faster intervals (500ms) for intentionally difficult interactions

Despite the increased difficulty of rhythmic gestures with 500ms intervals, users still achieved satisfactory performance with some of them. These more difficult gestures could be useful when more challenging interaction is desirable; for example, for actions which cannot be easily undone.

### DR5: Use interactive light for sensor strength feedback

Users found it easier to find where to address a gesture system when interactive light was used for sensor strength feedback. Interactive light feedback responded quickly to hand movements and affected the appearance of the rhythmic gesture animations, allowing users to easily see the effects of gesturing where they are.

### DR6: Give rhythmic gesture feedback from the start of an interaction

One of the issues investigated in Experiment 6 was how to present sensor strength and rhythmic gesture feedback together in the audio and tactile modalities. The findings suggest that rhythmic gesture feedback should be given from the beginning of an interaction, as users gave lower difficulty ratings than when rhythmic gesture feedback was given after sensor strength feedback stopped.

### DR7: Give short sensor strength feedback with the audio and tactile modalities

'Short' sensor strength feedback stops once users have stopped finding where to gesture and have started using rhythmic gestures to address their chosen system. The audio and tactile modalities should use short feedback, so that feedback is minimal and unobtrusive.

## 8.4 Revised Gesture Usability Scenarios

In Section 2.3.1 of the Literature Review, two scenarios were used to illustrate gesture usability problems. In these scenarios, Katie and Bobby encountered difficulty and uncertainty arising from a lack of feedback and from being unable to properly address the systems they wanted to interact with. What follows are updated versions of these scenarios, showing how the contributions of this thesis could lead to better outcomes for Katie and Bobby. By addressing the gesture systems, which also benefit from novel output modalities, they are able to accomplish their interaction goals without difficulty.

**Scenario 1:** Katie is in her living room watching television. Her mobile phone is on the coffee table a few feet in front of her. Her phone starts ringing; after glancing at its screen and seeing an unrecognised number, she raises her hand to gesture, with the intention of dismissing the call. As she raises her hand, the table surface around her phone faintly lights up, showing a rhythmic gesture animation (as in Figure 8.1). She moves her hand until the animation becomes brighter. Her television also begins showing a gesture animation on its screen. She gestures in time with the animation shown around her mobile phone; a bracelet on her wrist vibrates in time with her movements. After a couple of seconds, both animations stop and the space around her phone remains illuminated. Katie is now addressing the phone. With a flick of her wrist, she dismisses the incoming call and the phone stops ringing.

**Scenario 2:** Bobby is getting ready to leave the house. As he walks past the thermostat in the hallway, he raises his hand; the thermostat lights up, showing a gesture animation. Bobby matches the rhythmic gesture and then waves, as though saying goodbye, to turn off the heating. His smartwatch vibrates and the thermostat plays a sound to confirm his gesture was recognised; it also shows feedback using its interactive light display (as in Figure 8.2). He does this once more, this time gesturing at the hallway lights. They turn off and his smartwatch vibrates again.

## 8.5 Limitations and Future Work

Limitations of the experiment designs were discussed at the end of each experiment chapter (Chapters 3 to 7). Discussed here are more general limitations of the thesis research, along with areas for future work which address them and further investigate the topics presented in this work.

### 8.5.1 Audio and Tactile Feedback

The interaction techniques presented in this thesis used abstract audio and vibrotactile feedback. These feedback types were chosen as they are commonly used and are widely supported by consumer products. For example, many wearable devices like smart-watches and activity trackers have vibrotactile displays for delivering notifications and loudspeakers are common in many devices for presenting sounds. By studying these feedback types, this thesis makes contributions which are relevant to hardware and capabilities which are available

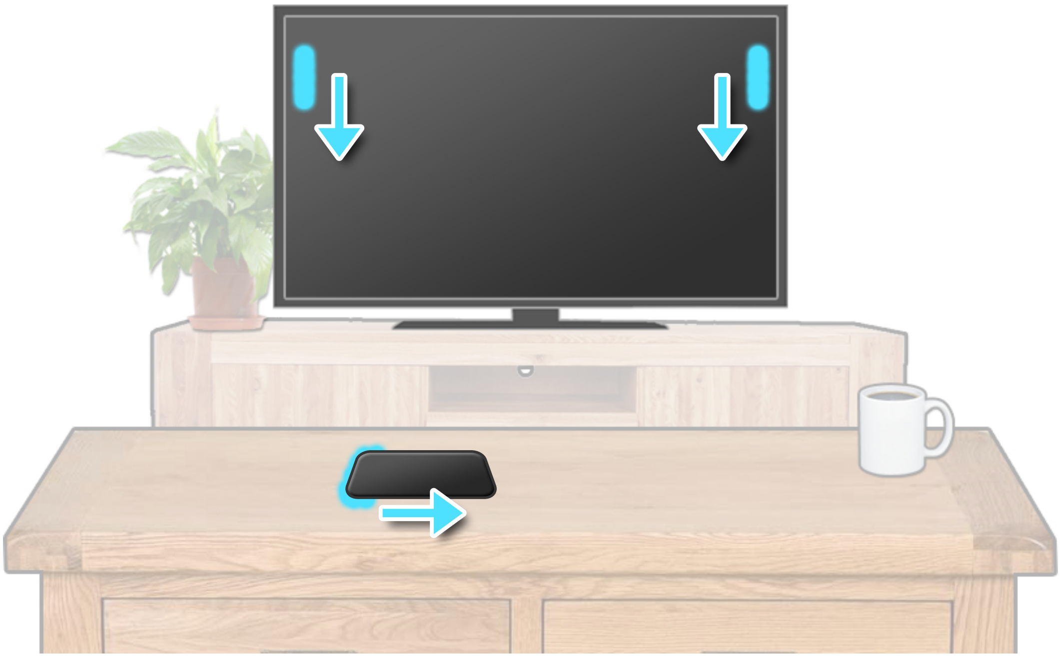


**Figure 8.1:** In the revised version of **Scenario 1**, Katie's phone and television both display rhythmic gesture animations; her phone uses its interactive light display, while her television shows visual cues on its screen. Blue light has been used here for illustration only; rhythmic gesture animations would be shown with white light, as in Experiments 5 and 6. Interactive light could also be used behind the television; however, its display is not as limited and its large size means on-screen feedback can be seen easily.

**Figure 8.2:** In the revised version of **Scenario 2**, Bobby knows his gestures have been recognised because the thermostat displays interactive light feedback and plays a sound, while his smartwatch delivers tactile feedback.

now. However, the human-computer interaction field has demonstrated a variety of technologies which use other capabilities of the audio and tactile modalities. Speech output and skin-drag displays [61], for example, provide other ways of giving audio and tactile feedback, respectively. Future research could build on the work presented here by investigating how other technologies could also be used for gesture feedback.

#### The Efficacy of Tactile Feedback

Tactile feedback had little impact on task performance in Experiments 1 and 2. This may have been because the interactions were too short or too easy for there to have been a noticeable effect. Users were positive about tactile feedback in these experiments, reporting that it gave improved awareness of how the system was responding to their actions. This improved the user experience of the interactions but had no effect on the task performance. In the later experiments, tactile feedback was generally less effective (in terms of improving task performance) and was less preferred than audio feedback. This may have been because the audio feedback was more easily noticed than the tactile feedback or because users are more used to receiving audio feedback about their input.

Although tactile feedback had underwhelming efficacy in these experiments, there may be situations where it is more appropriate than other output modalities. For example, it could be used in situations where audio feedback would be socially unacceptable (e.g. when around others) or difficult to notice (e.g. when in noisy public spaces). It could also be more useful in collaborative situations where many people are interacting with the one system, as the feedback would be noticeable only to the person providing the input.

### 8.5.2 Interactive Light Displays

The interactive light displays used in this thesis were created using LEDs positioned around device edges. This meant the space around these devices could be illuminated for displaying information. This approach was used because LEDs are cheap, have low power requirements, and can be unobtrusively integrated with devices without changing their form factor significantly. However, there are other technologies which could also be used for interactive light displays and this could be an area for future work. Steerable light sources, like the *Beamatron* [150], or projection displays, like *IllumiRoom* [64], could be used to project interactive light feedback around devices as users gesture towards them. Interactive household lighting, such as *Philips Hue*[1] lamps and light bulbs, could also be investigated as interactive light displays (as in Figure 8.3). Investigating these alternatives was outside the scope of

[1]Philips Hue: `www.meethue.com` Accessed 14/07/15

this thesis; instead, it focused on interactive light *around* devices, as such displays have been demonstrated in other work [110, 93] and have been used in consumer products[2].

Another area for future work would be to investigate the use of future headset displays — like Microsoft's *HoloLens*[3] augmented reality headset — for displaying feedback while users address and interact with gesture systems. Such devices could be used to render visual feedback around limited-display devices, like interactive light displays, except the feedback would only be visible to the wearer. This would allow more complex feedback than is possible with an LED display and the amount of feedback given could be personalised depending on how much the user needs. However, such technologies may not always be available or worn by users; the interactive light displays used in this research do not require users to have any additional technologies in order to experience feedback from their devices.

### 8.5.3 Sensor Strength Feedback

Sensor strength feedback helped users find where to perform gestures by giving feedback encoding a single dimension of information: an estimation of how well they could be sensed. This resulted in simple feedback which was found to be effective; however, future work could look at other ways of helping users find where to gesture. Spatial cues (feedback showing to "move left", for example) could be used instead to guide movements, similar to the feedback given by *LightGuide* [124], where projected arrows were used to guide movement. Such alternative feedback may be more effective by giving more explicit guidance, although it may require more complex display technology. It also may not work well with other interaction techniques, like rhythmic gestures, which already use spatial elements in their feedback.

Sensor strength feedback is something which users may become less reliant on over time, as they gain experience with a particular system and learn where they should perform gestures in order to be sensed reliably. The feedback supports novice and experienced users, alike, however, as experienced users could simply ignore it as they immediately begin providing input without 'searching' for the best input space. In such circumstances, experienced users may prefer to disable or reduce the amount of feedback given. Future work could investigate how the interaction techniques presented in this thesis could be adapted as users gain experience, providing users with support but without being unacceptable or irritating. Research could also investigate ways of doing this automatically over time; for example, if a system sees that users are becoming more competent and accurate in their input, it could gradually start to reduce the amount of feedback given, supporting users in their transition from novice users to expert users.

---

[2] Sony Illumination Bar Developer API: `developer.sonymobile.com/knowledge-base/experimental-apis/illumination-bar-api` Accessed 11/05/15

[3] Microsoft HoloLens: `www.microsoft.com/microsoft-hololens` Accessed 04/08/15

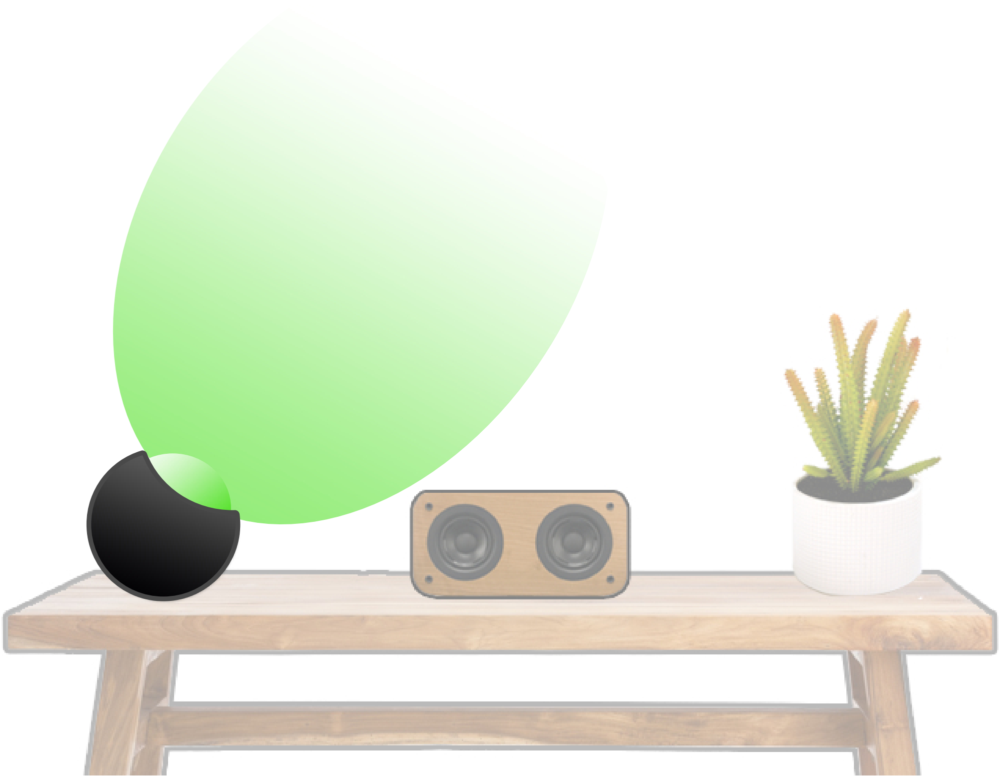


**Figure 8.3:** Interactive lighting, like the lamp shown here, could be used to give feedback about gesture interaction with limited display devices, such as the music system.

### 8.5.4 Rhythmic Gestures

Five rhythmic gesture movements were evaluated in this thesis; however, there is a much larger design space from which other gesture movements can be chosen. The five gestures used here were selected because they could be easily and reliably sensed by the chosen sensor technology. Alternative types of movement may be more appropriate for other sensor technologies; for example, wearable sensors like Thalmic Lab's *Myo*[4] could detect users repeatedly pinching their thumb and index finger together. Future work could build on this thesis research and further investigate the rhythmic gesture design space.

Rhythmic gestures could also be used in other interaction scenarios, not just for addressing gesture systems. Future research could investigate other uses for this interaction technique. For example, continuous rhythmic gestures could be used in similar contexts to *CycloStar* [82], allowing users to provide continuous input.

Interactive light displays were used in this thesis for revealing rhythmic gestures, as these offer limited-display devices a way of presenting visual information. However, devices with other output capabilities could also show gestures in other ways — on screen or using multi-dimensional tactile displays, for example. Further research could build on the work here by investigating new ways of revealing rhythmic gestures.

[4]Myo Gesture Control Armband: `www.myo.com` Accessed 03/08/15

### 8.5.5 Other Interaction Modalities

This thesis focused on usability problems with in-air gesture systems, although other sensing systems — like those using speech or tangible objects for input — will have many of the same issues. These sensing system problems, discussed by Bellotti *et al.* [11], could benefit from the research in this thesis. For example, limited-display devices with speech input could use interactive light displays to show system attention while users speak commands, letting them see that they are addressing the correct device. Gesture input, using the interactions investigated in this thesis, could also be used by speech systems; for example, to allow users to show which device they are addressing with their spoken commands. Speech systems could also use wearable devices to deliver tactile feedback about interaction, giving Tactons when spoken commands have been accepted or when they are not understood. Tangible user interfaces could, for example, give sensor strength feedback to help guide users as they move interactive objects. There are many possibilities for future work to investigate how the novel feedback modalities and interaction techniques investigated here could be applied to other types of sensing system.

### 8.5.6 Address in Multi-User Situations

This thesis focused on single-user interactions with gesture systems, although there are situations where multiple users may be interacting with a gesture system at the same time. For example, O'Hara *et al.* [101] describe a gesture system for use in the operating theatre and *StrikeAPose* [141] was deployed in public spaces; both systems had many potential users at the same time. Such situations make the address problem more complicated, as users need to know who is in control, if *they* are in control, and what they must to do take (or relinquish) control of the system. An interesting area for future work would be to extend the interaction techniques of this thesis to multi-user situations; for example, using rhythmic gestures to take control of a collaborative system. If there are many users within the sensor space then the one who performs a rhythmic gesture could be given control of the system; they could then relinquish control explicitly (perhaps through a certain gesture, or by repeating the rhythmic gesture) or lose it to other users (who, perhaps, perform the rhythmic gesture themselves). Appropriate feedback would also have to be given to indicate who is in control of the system, which would be more challenging when multiple users interact with a limited display device.

## 8.6 Conclusions

Users must be able to address gesture-sensing systems in order to interact with them and this thesis contributes interaction and feedback techniques which allow them to do this. This thesis investigated tactile displays and interactive light displays, two novel ways in which limited-display devices can give feedback to users. These displays, along with the more widely researched and better understood audio displays, can be used to give feedback during gesture interaction. This thesis also studied two novel interaction techniques — sensor strength feedback and rhythmic gestures — as well as their combined use, helping users address gesture systems by showing them where to gesture, so they can be reliably sensed, and how to direct their input, so their gestures do not affect other systems unintentionally. The results show that these techniques are successful. Their use in future gesture systems will allow those systems to provide expressive feedback and will help users address those systems properly.

# Appendix A

# Experiment 3 Task Screenshots

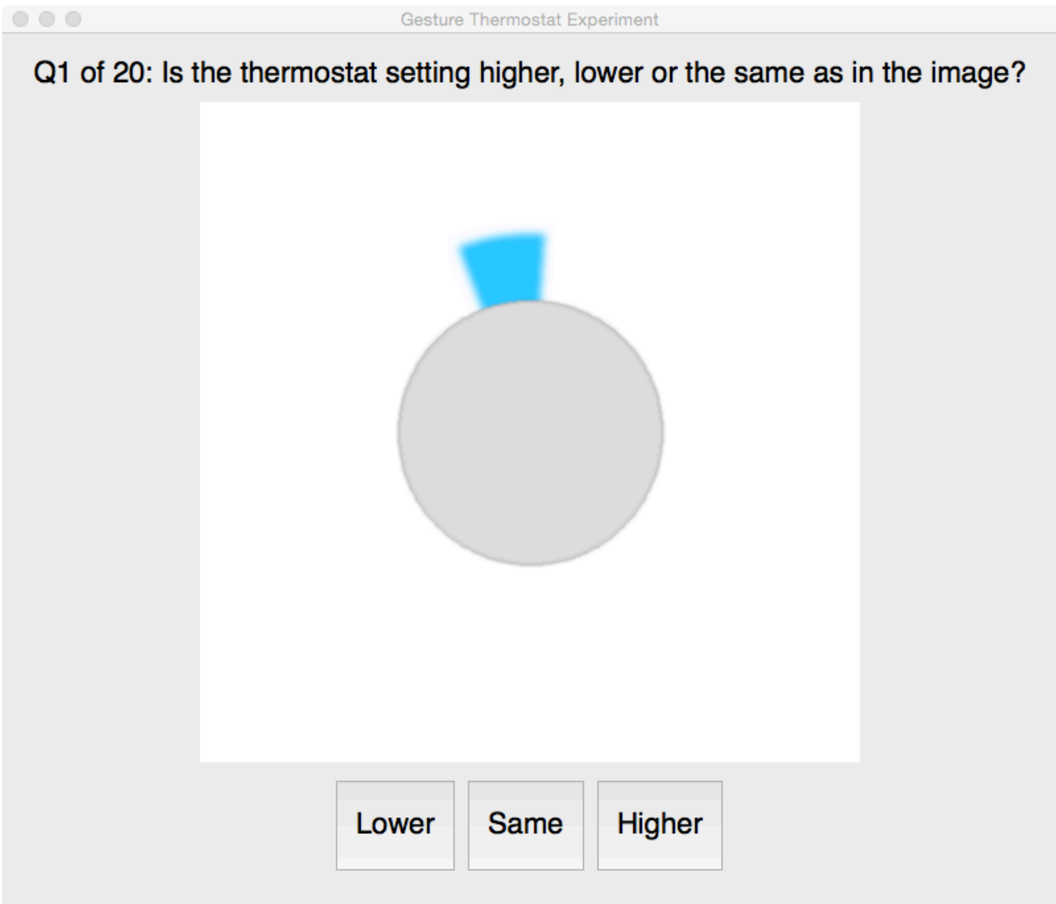



**Figure A.1:** Instructions and user interface for the *Query* tasks in Experiment 3.

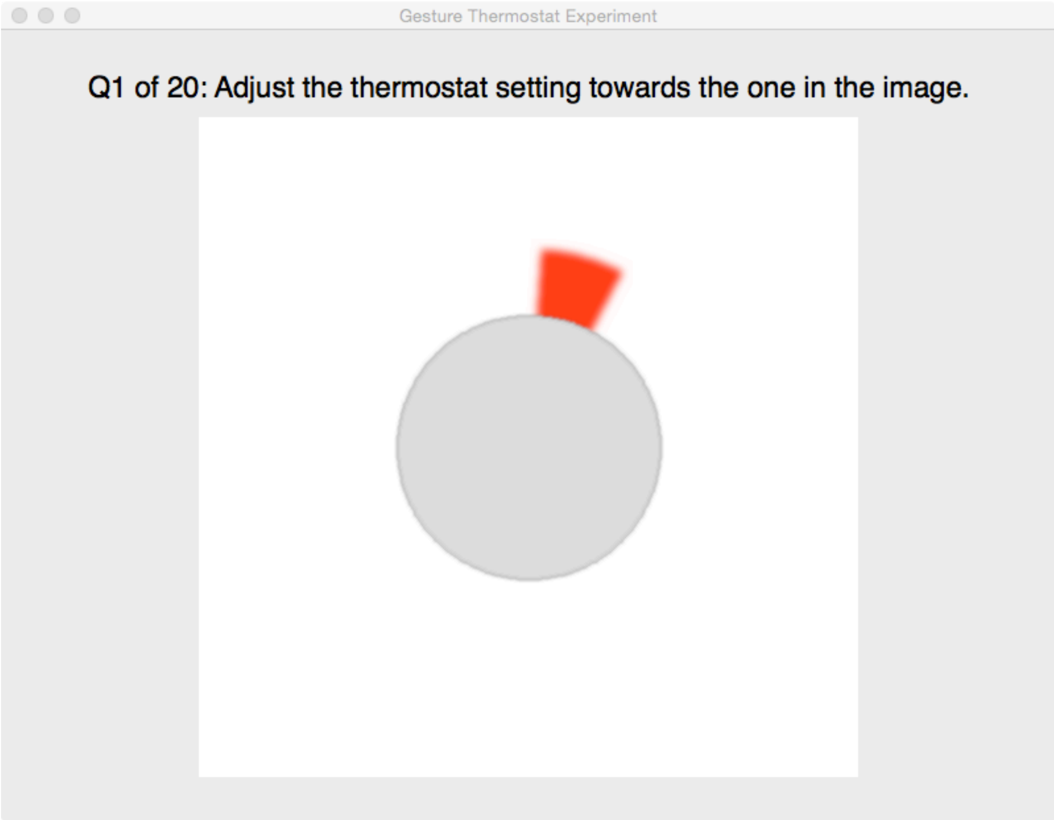



**Figure A.2:** Instructions and user interface for the *Quick Change* tasks in Experiment 3.

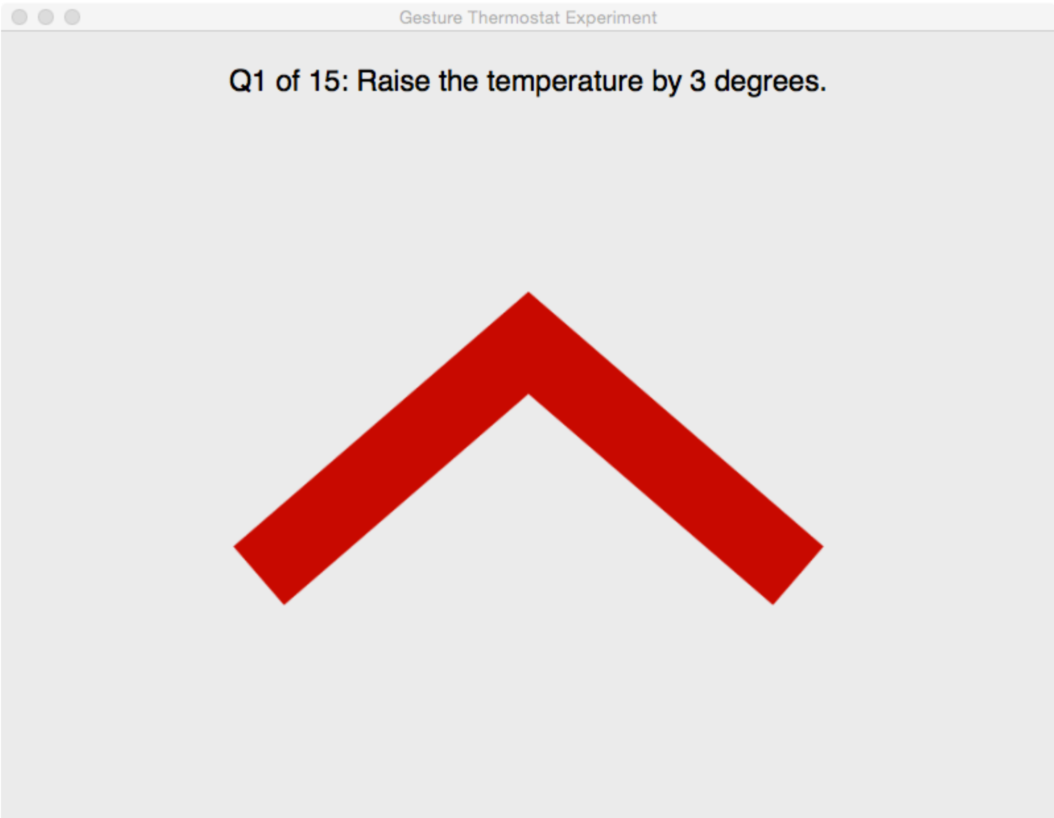
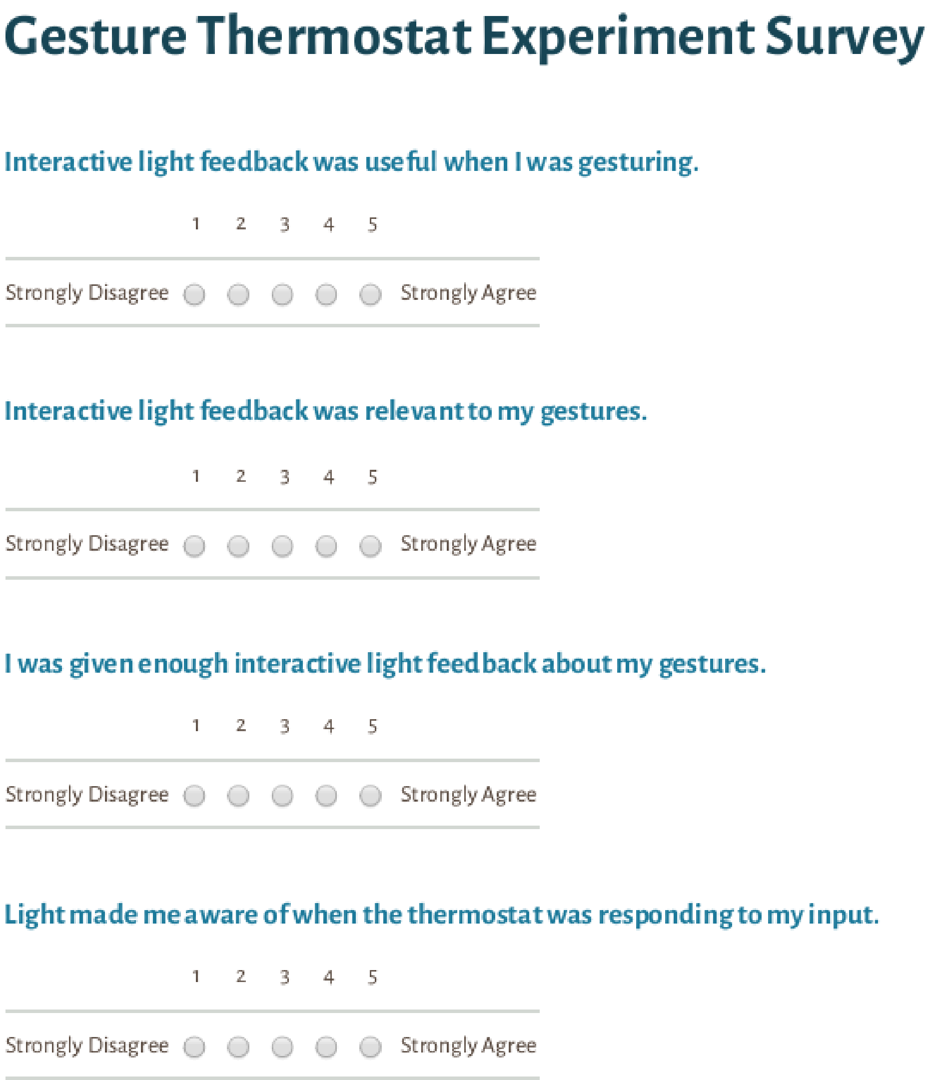



**Figure A.3:** Instructions and user interface for the *Precise Change* tasks in Experiment 3.

# Appendix B

# Experiment 3 Surveys

Figures B.1 and B.2 show the first survey used in Experiment 3 (S1Q1–S1Q8). Figures B.3 to B.5 show the second survey used in Experiment 3 (S2Q1–S2Q10). Both surveys were administered electronically using a laptop.



# Gesture Thermostat Experiment Survey

**Interactive light feedback was useful when I was gesturing.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Interactive light feedback was relevant to my gestures.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

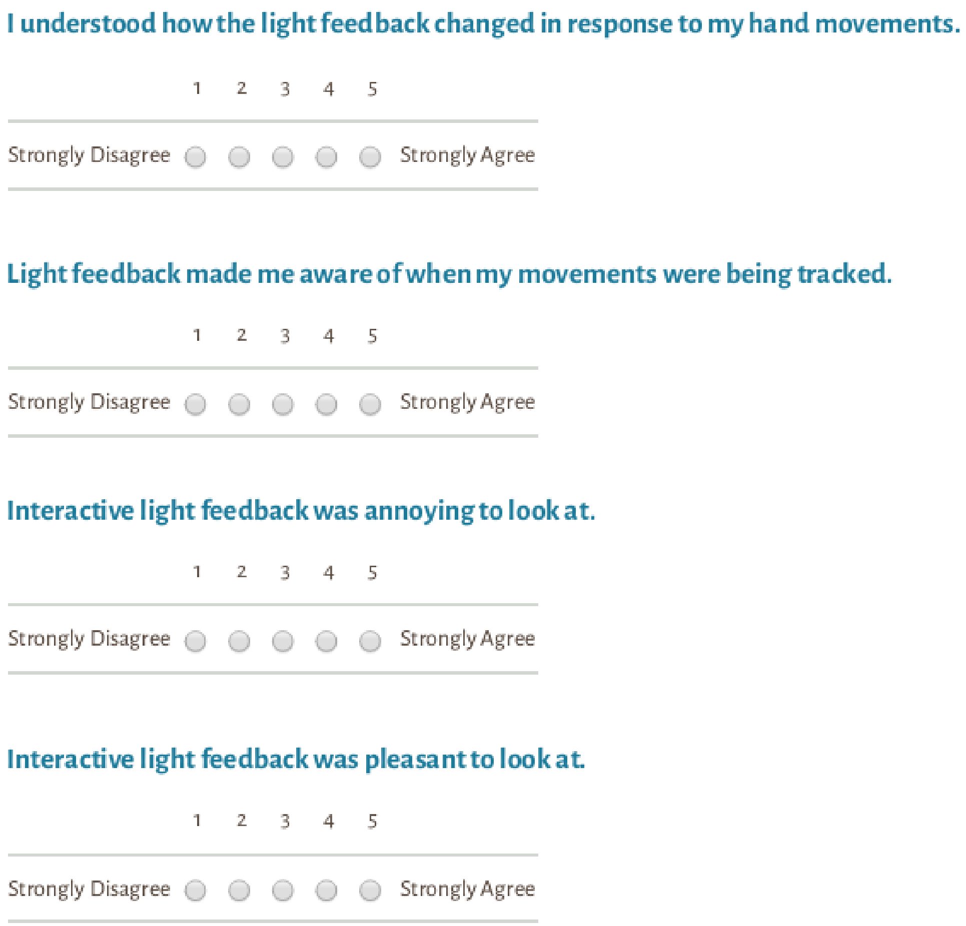
**I was given enough interactive light feedback about my gestures.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Light made me aware of when the thermostat was responding to my input.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure B.1:** First survey used in Experiment 3 (p1 of 2).



I understood how the light feedback changed in response to my hand movements.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Light feedback made me aware of when my movements were being tracked.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Interactive light feedback was annoying to look at.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Interactive light feedback was pleasant to look at.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure B.2:** First survey used in Experiment 3 (p2 of 2).

# Gesture Thermostat Experiment Survey 2

### Audio feedback was useful when I was gesturing.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### Tactile feedback was useful when I was gesturing.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### Audio feedback was relevant to my gestures.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

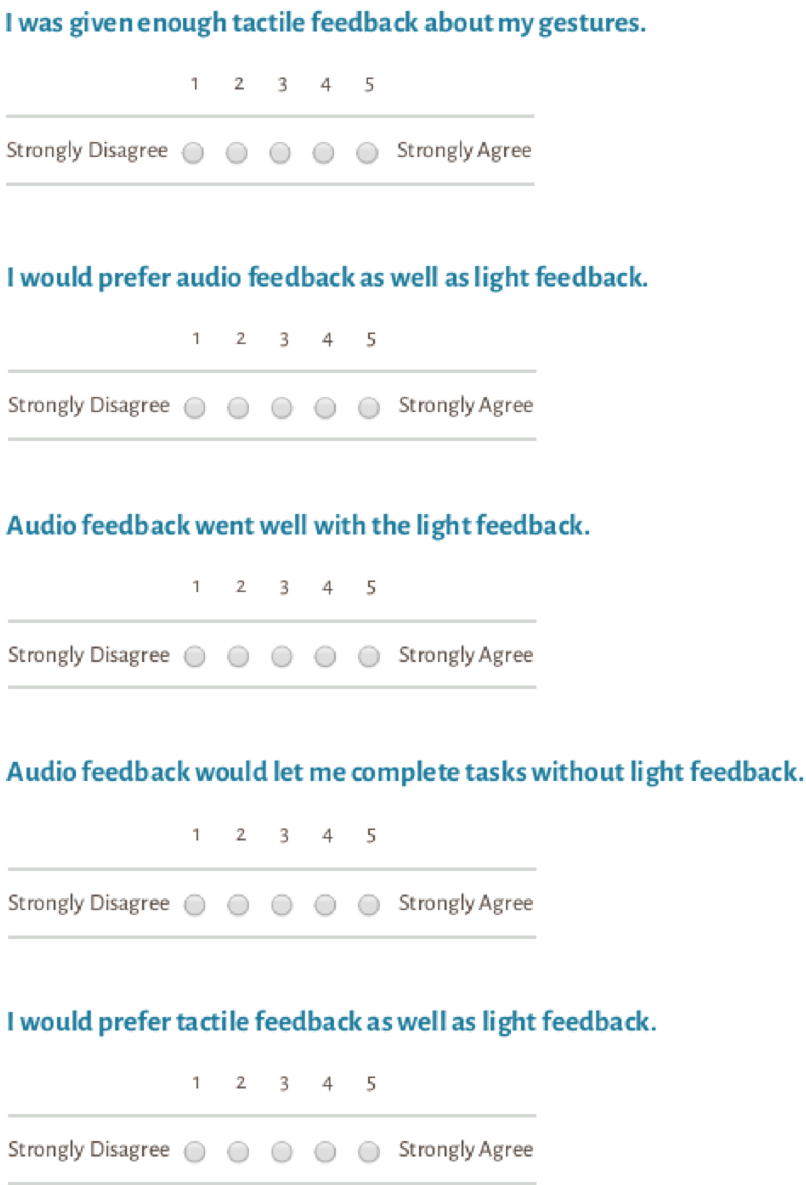
### Tactile feedback was relevant to my gestures.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### I was given enough audio feedback about my gestures.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure B.3:** Second survey used in Experiment 3 (p1 of 3).



I was given enough tactile feedback about my gestures.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

I would prefer audio feedback as well as light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Audio feedback went well with the light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

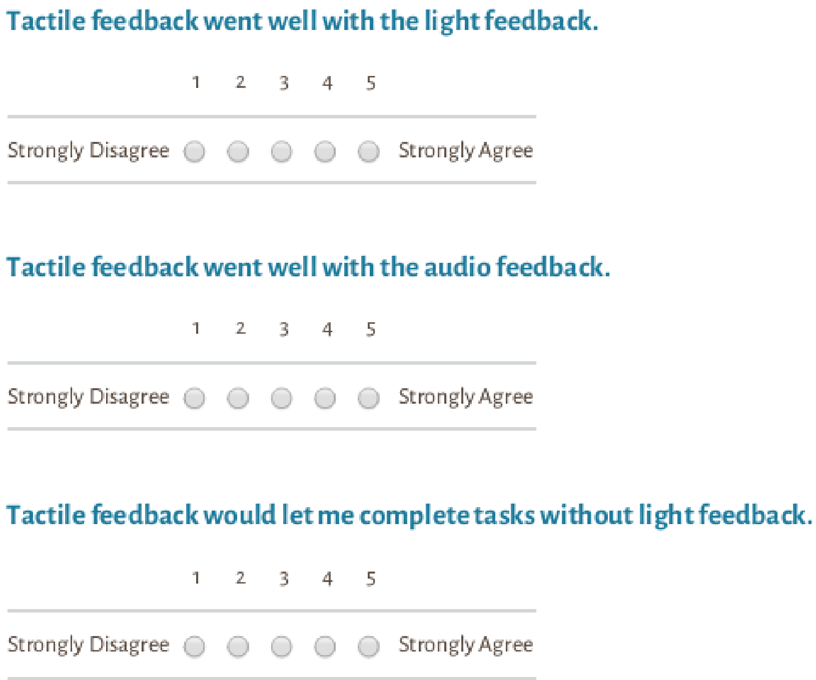
Audio feedback would let me complete tasks without light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

I would prefer tactile feedback as well as light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure B.4:** Second survey used in Experiment 3 (p2 of 3).



Tactile feedback went well with the light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Tactile feedback went well with the audio feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

Tactile feedback would let me complete tasks without light feedback.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure B.5:** Second survey used in Experiment 3 (p3 of 3).

# Appendix C

# Experiment 4 Survey

Figures C.1 to C.3 show the survey used in Experiment 4.

# Gesture Feedback Survey

### 1 a) Changes in the light feedback were easy to notice.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

### 1 b) Changes in the audio feedback were easy to notice.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

### 1 c) Changes in the tactile feedback were easy to notice.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

### 2 a) I understood how light feedback changed with my hand movements.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

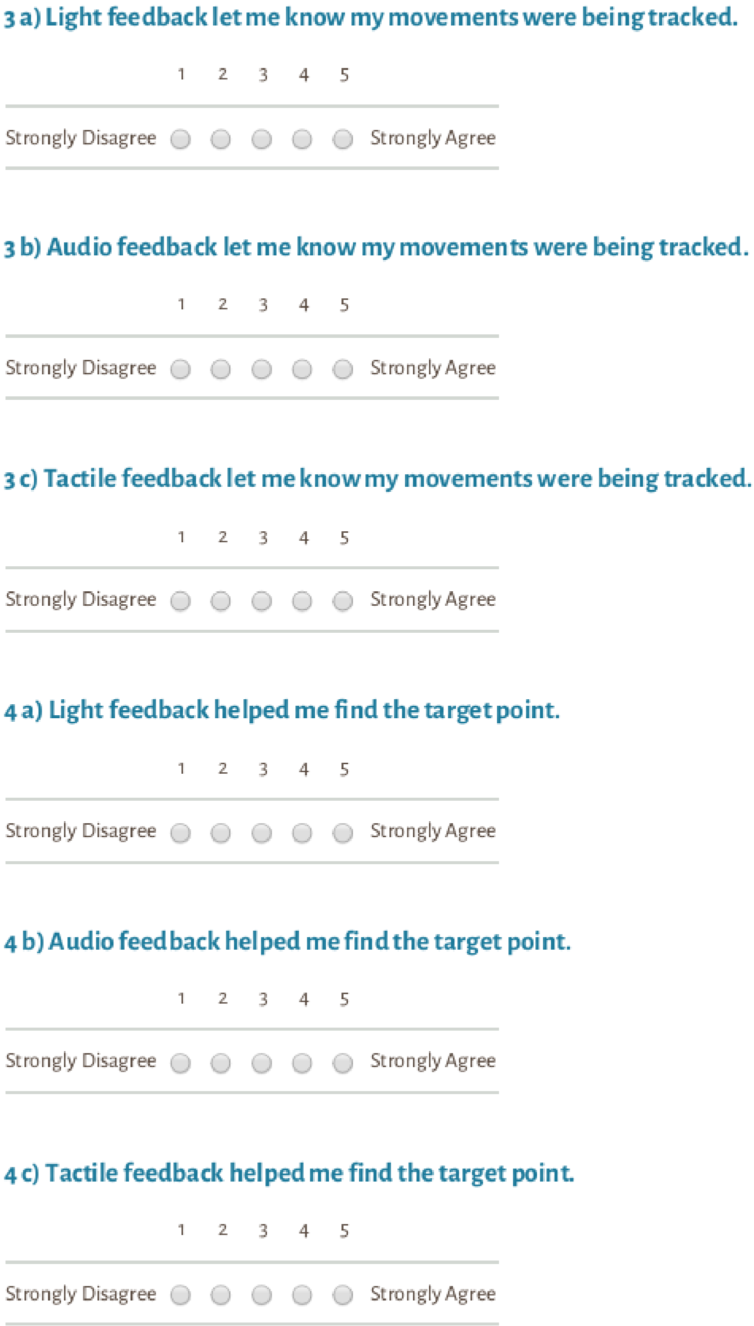
### 2 b) I understood how audio feedback changed with my hand movements.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

### 2 c) I understood how tactile feedback changed with my hand movements.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

**Figure C.1:** Survey used in Experiment 4 (p1 of 3).



3 a) Light feedback let me know my movements were being tracked.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 b) Audio feedback let me know my movements were being tracked.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 c) Tactile feedback let me know my movements were being tracked.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

4 a) Light feedback helped me find the target point.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

4 b) Audio feedback helped me find the target point.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

4 c) Tactile feedback helped me find the target point.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure C.2:** Survey used in Experiment 4 (p2 of 3).


**Figure C.3:** Survey used in Experiment 4 (p3 of 3).

# Appendix D

# Rhythmic Gesture Sensing Summary

As discussed in Section 6.2.3, a simple rhythmic gesture-sensing algorithm was created for use in Experiment 5. This section provides a summary of the sensing approach.

Rhythmic gestures consist of a gesture movement and a gesture interval. Since the gesture interval is known, a gesture-sensing algorithm does not always have to be running; instead, it can just check for gestures at the end of each gesture interval. This approach is used here; after each gesture interval, the algorithm checks recent hand movements for gestures. Hand movement over the most recent interval is compared to the gesture movement template. For example, for the Side-to-Side gesture, hand movement should be a straight left-to-right or right-to-left path, with no change in direction. If recent hand movement does not match the expected gesture movement, then no rhythmic gesture has been performed. If hand movements do match the gesture template then one *cycle* of a rhythmic gesture has been performed. Further analysis is then needed to determine if that cycle is part of a complete rhythmic gesture performance.

When a gesture cycle has been detected, it is compared to previous gesture intervals. If there is a sufficiently long sequence of valid gesture cycles, a rhythmic gesture is being matched correctly. As discussed in Section 6.2.3, at least three gesture cycles are required for a rhythm to be matched. Different gesture movements have different criteria for what makes a valid sequence of cycles: for circular gestures, successive cycles are complete circles; for the other gestures, successive cycles are movements in alternating directions.

The approach described here looks for sequences of valid gestures after known intervals; while this is able to detect rhythmic gesture performances with a good degree of accuracy, it is not able to detect gesture performances which are significantly out of time with the gesture interval. It is also unable to detect *how well* users are matching the rhythm. These limitations are discussed in Chapter 6.

# Appendix E

# Experiment 5 Surveys

Figures E.1 to E.3 show the rhythmic gesture survey used in Experiment 5. Figures E.4 to E.6 show the social acceptability survey used in this experiment.

# Rythmic Gestures Survey

### 1 a) I found the Side-to-Side movement easy to perform.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### 1 b) I found the Up-and-Down movement easy to perform.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### 1 c) I found the Forwards-and-Backwards movement easy to perform.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

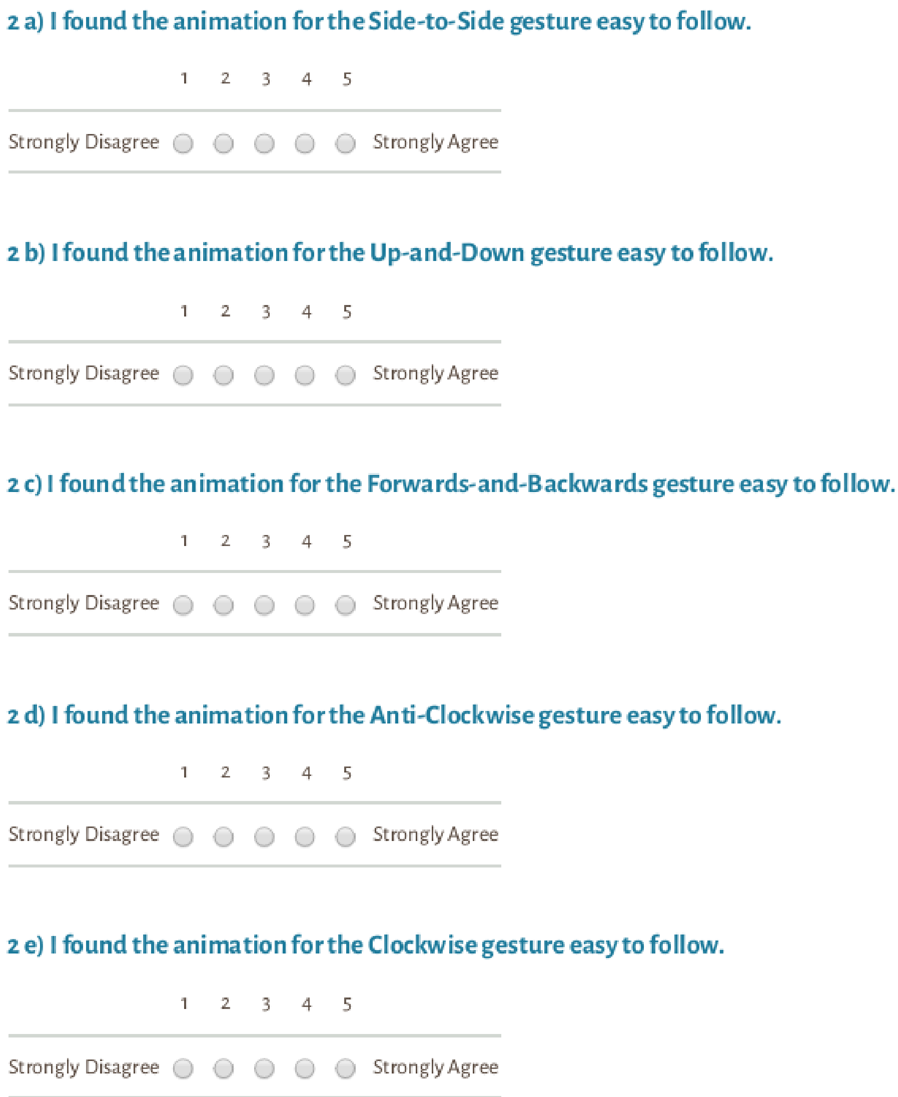
### 1 d) I found the Anti-Clockwise movement easy to perform.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

### 1 e) I found the Clockwise movement easy to perform.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure E.1:** Rhythmic gestures survey used in Experiment 5 (p1 of 3).



2 a) I found the animation for the Side-to-Side gesture easy to follow.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

2 b) I found the animation for the Up-and-Down gesture easy to follow.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

2 c) I found the animation for the Forwards-and-Backwards gesture easy to follow.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

2 d) I found the animation for the Anti-Clockwise gesture easy to follow.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

2 e) I found the animation for the Clockwise gesture easy to follow.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure E.2:** Rhythmic gestures survey used in Experiment 5 (p2 of 3).

3 a) I found it easy to keep matching the Side-to-Side gesture.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 b) I found it easy to keep matching the Up-and-Down gesture.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 c) I found it easy to keep matching the Forwards-and-Backwards gesture.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 d) I found it easy to keep matching the Anti-Clockwise gesture.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

3 e) I found it easy to keep matching the Clockwise gesture.

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure E.3:** Rhythmic gestures survey used in Experiment 5 (p3 of 3).

# Rhythmic Gestures Survey

## Please indicate where and with whom you would feel comfortable using the side-to-side gesture.

Select the "Yes" box if you would feel comfortable using this gesture in these situations.

| | Yes | No |
|---|---|---|
| Alone at home. | ○ | ○ |
| Around family at home. | ○ | ○ |
| Alone at work. | ○ | ○ |
| Around others at work. | ○ | ○ |
| Around friends in public. | ○ | ○ |
| Around strangers in public. | ○ | ○ |

## Please indicate where and with whom you would feel comfortable using the up-and-down gesture.

Select the "Yes" box if you would feel comfortable using this gesture in these situations.

| | Yes | No |
|---|---|---|
| Alone at home. | ○ | ○ |
| Around family at home. | ○ | ○ |
| Alone at work. | ○ | ○ |
| Around others at work. | ○ | ○ |
| Around friends in public. | ○ | ○ |
| Around strangers in public. | ○ | ○ |

**Figure E.4:** Social acceptability survey used in Experiment 5 (p1 of 3).

**Please indicate where and with whom you would feel comfortable using the back-and-forth gesture.**

Select the "Yes" box if you would feel comfortable using this gesture in these situations.

| | Yes | No |
|---|---|---|
| Alone at home. | ○ | ○ |
| Around family at home. | ○ | ○ |
| Alone at work. | ○ | ○ |
| Around others at work. | ○ | ○ |
| Around friends in public. | ○ | ○ |
| Around strangers in public. | ○ | ○ |

**Please indicate where and with whom you would feel comfortable using the clockwise-circle gesture.**

Select the "Yes" box if you would feel comfortable using this gesture in these situations.

| | Yes | No |
|---|---|---|
| Alone at home. | ○ | ○ |
| Around family at home. | ○ | ○ |
| Alone at work. | ○ | ○ |
| Around others at work. | ○ | ○ |
| Around friends in public. | ○ | ○ |
| Around strangers in public. | ○ | ○ |

**Figure E.5:** Social acceptability survey used in Experiment 5 (p2 of 3).

**Please indicate where and with whom you would feel comfortable using the anti-clockwise-circle gesture.**

Select the "Yes" box if you would feel comfortable using this gesture in these situations.

| | Yes | No |
|---|---|---|
| Alone at home. | | |
| Around family at home. | | |
| Alone at work. | | |
| Around others at work. | | |
| Around friends in public. | | |
| Around strangers in public. | | |

**Figure E.6:** Social acceptability survey used in Experiment 5 (p3 of 3).

# Appendix F

# Extended Results from Experiment 5

## F.1 Cycles

In Section 6.3.3 (Experiment 5 Results), it was found that the interaction effect between **Gesture** and **Interval** had a significant effect on the number of rhythmic gesture cycles performed. However, there were too many significant *post hoc* comparisons to list them all in that section, so only those relevant to the experiment hypotheses were presented. Table F.1 lists all significant differences found by *post hoc* t-tests for this interaction.

The interaction between all three experiment factors (**Feedback**, **Gesture** and **Interval**) was also found to be significant. There were too many significant *post hoc* comparisons to list them all in that section. Table F.2 lists all significant differences found by *post hoc* t-tests for the three-factor interaction.

## F.2 Difficulty-Match

The interaction between **Gesture** and **Interval** also had a significant effect on difficulty ratings for matching a rhythmic gesture. However, there were too many significant *post hoc* comparisons to list them all. Table F.3 lists all significant differences found by *post hoc* t-tests.

| Comparison | Result | Comparison | Result |
|---|---|---|---|
| C,500 - FB,500 | t = -4.5, p = 0.0013 | C,700 - SS,900 | t = 4.92, p < 0.001 |
| C,500 - SS,500 | t = -6.51, p < 0.001 | C,700 - UD,900 | t = 4.07, p = 0.0078 |
| C,500 - UD,500 | t = -7, p < 0.001 | C,700 - SS,1100 | t = 5.76, p < 0.001 |
| C,500 - C,700 | t = -5.08, p < 0.001 | C,700 - UD,1100 | t = 3.59, p = 0.044 |
| C,500 - AC,700 | t = -3.82, p = 0.02 | AC,700 - SS,1100 | t = 4.36, p = 0.0023 |
| C,500 - FB,700 | t = -4.98, p < 0.001 | FB,700 - SS,900 | t = 4.82, p < 0.001 |
| C,500 - C,900 | t = -5.44, p < 0.001 | FB,700 - UD,900 | t = 3.97, p = 0.012 |
| C,500 - AC,900 | t = -4.58, p < 0.001 | FB,700 - SS,1100 | t = 5.66, p < 0.001 |
| C,500 - FB,900 | t = -4, p = 0.01 | SS,700 - C,1100 | t = -4.66, p < 0.001 |
| C,500 - C,1100 | t = -6.97, p < 0.001 | UD,700 - C,1100 | t = -5.11, p < 0.001 |
| C,500 - AC,1100 | t = -5.11, p < 0.001 | C,900 - SS,900 | t = 5.32, p < 0.001 |
| C,500 - FB,1100 | t = -3.77, p = 0.024 | C,900 - UD,900 | t = 4.46, p = 0.0016 |
| AC,500 - UD,500 | t = -3.78, p = 0.023 | C,900 - SS,1100 | t = 6.18, p < 0.001 |
| AC,500 - C,1100 | t = -3.74, p = 0.027 | C,900 - UD,1100 | t = 3.97, p = 0.011 |
| FB,500 - SS,900 | t = 4.3, p = 0.0031 | AC,900 - SS,900 | t = 4.38, p = 0.0022 |
| FB,500 - SS,1100 | t = 5.13, p < 0.001 | AC,900 - SS,1100 | t = 5.22, p < 0.001 |
| SS,500 - SS,700 | t = 4.16, p = 0.0054 | FB,900 - SS,900 | t = 3.75, p = 0.026 |
| SS,500 - UD,700 | t = 4.61, p < 0.001 | FB,900 - SS,1100 | t = 4.6, p < 0.001 |
| SS,500 - SS,900 | t = 6.49, p < 0.001 | SS,900 - C,1100 | t = -7, p < 0.001 |
| SS,500 - UD,900 | t = 5.63, p < 0.001 | SS,900 - AC,1100 | t = -4.97, p < 0.001 |
| SS,500 - SS,1100 | t = 7.34, p < 0.001 | UD,900 - C,1100 | t = -6.13, p < 0.001 |
| SS,500 - UD,1100 | t = 5.14, p < 0.001 | UD,900 - AC,1100 | t = -4.1, p = 0.0069 |
| UD,500 - SS,700 | t = 4.71, p < 0.001 | C,1100 - SS,1100 | t = 7.86, p < 0.001 |
| UD,500 - UD,700 | t = 5.15, p < 0.001 | C,1100 - UD,1100 | t = 5.64, p < 0.001 |
| UD,500 - SS,900 | t = 7.03, p < 0.001 | AC,1100 - SS,1100 | t = 5.82, p < 0.001 |
| UD,500 - UD,900 | t = 6.17, p < 0.001 | AC,1100 - UD,1100 | t = 3.62, p = 0.04 |
| UD,500 - SS,1100 | t = 7.89, p < 0.001 | FB,1100 - SS,1100 | t = 4.34, p = 0.0026 |
| UD,500 - UD,1100 | t = 5.69, p < 0.001 | | |

**Table F.1:** Significant comparisons for gesture cycles for **Gesture** and **Interval**.

| Comparison | Result |
|---|---|
| Audio,FB,500 - Tactile,SS,700 | t = 4.74, p = 0.0059 |
| Audio,FB,500 - None,SS,900 | t = 4.46, p = 0.019 |
| Audio,FB,500 - Both,SS,900 | t = 4.32, p = 0.034 |
| Audio,FB,500 - Both,UD,1100 | t = 4.6, p = 0.011 |
| Tactile,SS,500 - Tactile,SS,700 | t = 4.55, p = 0.014 |
| Tactile,SS,500 - None,SS,900 | t = 4.25, p = 0.044 |
| Tactile,SS,500 - Both,UD,1100 | t = 4.4, p = 0.025 |
| Tactile,UD,500 - Tactile,SS,700 | t = 5.17, p < 0.001 |
| Tactile,UD,500 - None,SS,900 | t = 4.87, p = 0.0033 |
| Tactile,UD,500 - Both,SS,900 | t = 4.72, p = 0.0065 |
| Tactile,UD,500 - None,UD,900 | t = 4.56, p = 0.013 |
| Tactile,UD,500 - Tactile,SS,1100 | t = 4.41, p = 0.024 |
| Tactile,UD,500 - Both,SS,1100 | t = 4.56, p = 0.013 |
| Tactile,UD,500 - Audio,UD,1100 | t = 4.23, p = 0.048 |
| Tactile,UD,500 - Both,UD,1100 | t = 5.02, p = 0.0017 |
| Tactile,C,700 - Tactile,SS,700 | t = 5.16, p < 0.001 |
| Tactile,C,700 - None,SS,900 | t = 4.87, p = 0.0033 |
| Tactile,C,700 - Both,SS,900 | t = 4.72, p = 0.0065 |
| Tactile,C,700 - None,UD,900 | t = 4.57, p = 0.013 |
| Tactile,C,700 - Tactile,SS,1100 | t = 4.43, p = 0.023 |
| Tactile,C,700 - Both,SS,1100 | t = 4.57, p = 0.013 |
| Tactile,C,700 - Audio,UD,1100 | t = 4.25, p = 0.045 |
| Tactile,C,700 - Both,UD,1100 | t = 5.01, p = 0.0017 |
| Audio,AC,700 - Tactile,SS,700 | t = 4.53, p = 0.015 |

| Comparison | Result |
|---|---|
| Audio,AC,700 - None,SS,900 | t = 4.25, p = 0.045 |
| Audio,AC,700 - Both,UD,1100 | t = 4.38, p = 0.027 |
| Tactile,SS,700 - None,C,900 | t = -4.22, p = 0.049 |
| Tactile,SS,700 - Audio,C,900 | t = -4.55, p = 0.014 |
| Tactile,SS,700 - None,FB,900 | t = -4.41, p = 0.024 |
| Tactile,SS,700 - None,AC,1100 | t = -5.39, p < 0.001 |
| Tactile,SS,700 - Tactile,AC,1100 | t = -4.81, p = 0.0044 |
| Tactile,SS,700 - None,FB,1100 | t = -4.24, p = 0.046 |
| Tactile,SS,700 - Both,FB,1100 | t = -4.54, p = 0.014 |
| Audio,C,900 - None,SS,900 | t = 4.25, p = 0.044 |
| Audio,C,900 - Both,UD,1100 | t = 4.4, p = 0.025 |
| None,FB,900 - Both,UD,1100 | t = 4.26, p = 0.043 |
| None,SS,900 - None,AC,1100 | t = -5.09, p = 0.0011 |
| None,SS,900 - Tactile,AC,1100 | t = -4.52, p = 0.016 |
| None,SS,900 - Both,FB,1100 | t = -4.25, p = 0.044 |
| Both,SS,900 - None,AC,1100 | t = -4.94, p = 0.0023 |
| Both,SS,900 - Tactile,AC,1100 | t = -4.37, p = 0.029 |
| None,UD,900 - None,AC,1100 | t = -4.78, p = 0.005 |
| None,AC,1100 - Tactile,SS,1100 | t = 4.64, p = 0.0094 |
| None,AC,1100 - Both,SS,1100 | t = 4.78, p = 0.005 |
| None,AC,1100 - Audio,UD,1100 | t = 4.45, p = 0.02 |
| None,AC,1100 - Both,UD,1100 | t = 5.24, p < 0.001 |
| Tactile,AC,1100 - Both,UD,1100 | t = 4.66, p = 0.0086 |
| Both,FB,1100 - Both,UD,1100 | t = 4.39, p = 0.026 |

**Table F.2:** Significant comparisons for gesture cycles for **Feedback x Gesture x Interval**.

| Comparison | Result | Comparison | Result |
|---|---|---|---|
| C,500 - FB,500 | t = 5.13, p < 0.001 | UD,500 - FB,1100 | t = -4.12, p = 0.0064 |
| C,500 - SS,500 | t = 6.68, p < 0.001 | UD,500 - SS,1100 | t = -5.93, p < 0.001 |
| C,500 - UD,500 | t = 4.14, p = 0.0058 | UD,500 - UD,1100 | t = -4.11, p = 0.0066 |
| C,500 - AC,900 | t = 4.4, p = 0.002 | C,700 - SS,900 | t = -4.2, p = 0.0046 |
| C,500 - C,1100 | t = 4.83, p < 0.001 | C,700 - SS,1100 | t = -5.05, p < 0.001 |
| C,500 - AC,1100 | t = 4.42, p = 0.0018 | AC,700 - SS,900 | t = -4.39, p = 0.002 |
| AC,500 - FB,500 | t = 5.29, p < 0.001 | AC,700 - UD,900 | t = -3.72, p = 0.028 |
| AC,500 - SS,500 | t = 6.84, p < 0.001 | AC,700 - SS,1100 | t = -5.25, p < 0.001 |
| AC,500 - UD,500 | t = 4.3, p = 0.003 | FB,700 - SS,900 | t = -3.62, p = 0.04 |
| AC,500 - AC,700 | t = 3.62, p = 0.04 | FB,700 - SS,1100 | t = -4.48, p = 0.0014 |
| AC,500 - AC,900 | t = 4.56, p < 0.001 | SS,700 - AC,900 | t = 3.62, p = 0.04 |
| AC,500 - C,1100 | t = 4.99, p < 0.001 | SS,700 - C,1100 | t = 4.05, p = 0.0083 |
| AC,500 - AC,1100 | t = 4.58, p < 0.001 | SS,700 - AC,1100 | t = 3.65, p = 0.036 |
| FB,500 - SS,700 | t = -4.36, p = 0.0024 | C,900 - SS,900 | t = -4.16, p = 0.0054 |
| FB,500 - UD,700 | t = -3.56, p = 0.049 | C,900 - SS,1100 | t = -5.02, p < 0.001 |
| FB,500 - FB,900 | t = -4.59, p < 0.001 | AC,900 - FB,900 | t = -3.85, p = 0.018 |
| FB,500 - SS,900 | t = -6.07, p < 0.001 | AC,900 - SS,900 | t = -5.34, p < 0.001 |
| FB,500 - UD,900 | t = -5.4, p < 0.001 | AC,900 - UD,900 | t = -4.66, p < 0.001 |
| FB,500 - FB,1100 | t = -5.11, p < 0.001 | AC,900 - FB,1100 | t = -4.37, p = 0.0022 |
| FB,500 - SS,1100 | t = -6.93, p < 0.001 | AC,900 - SS,1100 | t = -6.19, p < 0.001 |
| FB,500 - UD,1100 | t = -5.1, p < 0.001 | AC,900 - UD,1100 | t = -4.37, p = 0.0023 |
| SS,500 - FB,700 | t = -4, p = 0.01 | FB,900 - C,1100 | t = 4.28, p = 0.0033 |
| SS,500 - SS,700 | t = -5.91, p < 0.001 | FB,900 - AC,1100 | t = 3.88, p = 0.016 |
| SS,500 - UD,700 | t = -5.11, p < 0.001 | SS,900 - C,1100 | t = 5.76, p < 0.001 |
| SS,500 - FB,900 | t = -6.14, p < 0.001 | SS,900 - AC,1100 | t = 5.36, p < 0.001 |
| SS,500 - SS,900 | t = -7.62, p < 0.001 | UD,900 - C,1100 | t = 5.09, p < 0.001 |
| SS,500 - UD,900 | t = -6.95, p < 0.001 | UD,900 - AC,1100 | t = 4.69, p < 0.001 |
| SS,500 - FB,1100 | t = -6.66, p < 0.001 | C,1100 - FB,1100 | t = -4.8, p < 0.001 |
| SS,500 - SS,1100 | t = -8.48, p < 0.001 | C,1100 - SS,1100 | t = -6.62, p < 0.001 |
| SS,500 - UD,1100 | t = -6.65, p < 0.001 | C,1100 - UD,1100 | t = -4.8, p < 0.001 |
| UD,500 - FB,900 | t = -3.6, p = 0.043 | AC,1100 - FB,1100 | t = -4.4, p = 0.002 |
| UD,500 - SS,900 | t = -5.08, p < 0.001 | AC,1100 - SS,1100 | t = -6.22, p < 0.001 |
| UD,500 - UD,900 | t = -4.4, p = 0.0019 | AC,1100 - UD,1100 | t = -4.39, p = 0.002 |

**Table F.3:** Significant comparisons for difficulty for **Gesture** and **Interval**.

# Appendix G

# Experiment 6 Survey

Figure G.1 shows part of the survey used in Experiment 6. See Section 7.3.2 for a full description of the rest of the survey design, which incorporated some of the questions from Experiment 5.

# Gestures Survey

**1 a) Feedback about where to gesture, after I started gesturing, was useful.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**1 b) Feedback about my hand movements was useful.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**2 a) Feedback about where to gesture, after I started gesturing, was distracting.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**2 b) Feedback about my hand movements was distracting.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**3 a) Feedback about where to gesture, after I started gesturing, let me know my hands were being tracked.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**3 b) Feedback about my hand movements let me know my hands were being tracked.**

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Strongly Disagree | ○ | ○ | ○ | ○ | ○ | Strongly Agree |

**Figure G.1:** Survey used in Experiment 6.

# Bibliography

[1] AHLSTRÖM, D., HASAN, K., AND IRANI, P. Are You Comfortable Doing That?: Acceptance Studies of Around-Device Gestures in and for Public Settings. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services - Mobile HCI '14* (2014), ACM Press, pp. 193–202.

[2] ALEXANDER, J., MARSHALL, M. T., AND SUBRAMANIAN, S. Adding Haptic Feedback to Mobile TV. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (2011), ACM Press, pp. 1975–1980.

[3] ARMSTRONG, A., AND ISSARTEL, J. Sensorimotor synchronization with audio-visual stimuli: limited multisensory integration. *Experimental brain research 232*, 11 (2014), 3453–63.

[4] ARMSTRONG, A., ISSARTEL, J., VARLET, M., AND MARIN, L. The supplementation of spatial information improves coordination. *Neuroscience Letters 548* (2013), 212–216.

[5] ASAO, T., HAYASHI, H., HAYASHI, M., KOTANI, K., AND HORII, K. A Study on Fundamental Information Transmission Characteristics of an Air-Jet Driven Tactile Display. *Lecture Notes in Computer Science 5614* (2009), 48–57.

[6] ASHBROOK, D., BAUDISCH, P., AND WHITE, S. Nenya: Subtle and Eyes-Free Mobile Input with a Magnetically-Tracked Finger Ring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11* (2011), ACM Press, pp. 2043–2046.

[7] AUMI, M. T. I., GUPTA, S., GOEL, M., LARSON, E., AND PATEL, S. DopLink: Using the Doppler Effect for Multi-Device Interaction. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '13* (2013), ACM Press, pp. 583–586.

[8] BAILLY, G., WALTER, R., MÜLLER, J., NING, T., AND LECOLINET, E. Comparing Free Hand Menu Techniques for Distant Displays Using Linear, Marking and Finger-

Count Menus. In *Proceedings of INTERACT '11* (2011), Springer Berlin Heidelberg, pp. 248–262.

[9] BAUDEL, T., AND BEAUDOUIN-LAFON, M. Charade: Remote Control of Objects Using Free-Hand Gestures. *Communications of the ACM 36*, 7 (1993), 28–35.

[10] BEEK, P. J., AND LEWBEL, A. The Science of Juggling. *Scientific American 273*, 5 (1995), 92–97.

[11] BELLOTTI, V., BACK, M., EDWARDS, W. K., GRINTER, R. E., HENDERSON, A., AND LOPES, C. Making Sense of Sensing Systems: Five Questions for Designers and Researchers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '02* (2002), ACM Press, pp. 415–422.

[12] BENFORD, S., GREENHALGH, C., GIANNACHI, G., WALKER, B., MARSHALL, J., AND RODDEN, T. Uncomfortable User Experience. *Communications of the ACM 56*, 9 (2013), 66–73.

[13] BENFORD, S., SCHNÄDELBACH, H., KOLEVA, B., ANASTASI, R., GREENHALGH, C., RODDEN, T., GREEN, J., GHALI, A., PRIDMORE, T., GAVER, W. W., BOUCHER, A., WALKER, B., PENNINGTON, S., SCHMIDT, A., GELLERSEN, H., AND STEED, A. Expected, Sensed, and Desired: A Framework for Designing Sensing-Based Interaction. *ACM Transactions on Computer-Human Interaction 12*, 1 (2005), 3–30.

[14] BENNETT, P., CATER, K., AND FRASER, M. Resonant Bits: Harmonic Interaction with Virtual Pendulums. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction - TEI '15* (2015), ACM Press, pp. 49–52.

[15] BENSMAÏA, S. J., AND HOLLINS, M. Complex tactile waveform discrimination. *The Journal of the Acoustical Society of America 108*, 3 (2000), 1236–1245.

[16] BOLT, R. A. "Put-That-There": Voice and Gesture at the Graphics Interface. *ACM SIGGRAPH Computer Graphics 14*, 3 (1980), 262–270.

[17] BREWSTER, S., AND BROWN, L. M. Tactons: Structured Tactile Messages for Non-Visual Information Display. In *Proceedings of the 5th Australasian User Interface Conference - AUIC '04* (2004), Australian Computer Society, Inc., pp. 15–23.

[18] BREWSTER, S., CHOHAN, F., AND BROWN, L. M. Tactile feedback for mobile interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07* (2007), ACM Press, pp. 159–162.

[19] BROWN, L. *Tactons: Structured Vibrotactile Messages for Non-Visual Information Display*. Phd thesis, University of Glasgow, 2007.

[20] BROWN, L. M., BREWSTER, S., AND PURCHASE, H. C. A First Investigation into the Effectiveness of Tactons. In *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems - WHC '05* (2005), Ieee, pp. 167–176.

[21] BROWN, L. M., BREWSTER, S., AND PURCHASE, H. C. Multidimensional tactons for non-visual information presentation in mobile devices. In *Proceedings of the 8th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '06* (New York, New York, USA, 2006), ACM Press, pp. 231–238.

[22] BUDDE, M., BERNING, M., BAUMGÄRTNER, C., KINN, F., KOPF, T., OCHS, S., REICHE, F., RIEDEL, T., AND BEIGL, M. Point & Control - Interaction in Smart Environments: You Only Click Twice. In *Adjunct Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing - UbiComp '13 Adjunct* (2013), ACM Press, pp. 303–306.

[23] BUTLER, A., IZADI, S., AND HODGES, S. SideSight: Multi-touch Interaction Around Small Devices. In *Proceedings of the 21st Symposium on User Interface Software and Technology - UIST '08* (2008), ACM Press, pp. 201–204.

[24] CARTER, T., SEAH, S. A., LONG, B., DRINKWATER, B., AND SUBRAMANIAN, S. UltraHaptics: Multi-Point Mid-Air Haptic Feedback for Touch Surfaces. In *Proceedings of the 26th Symposium on User Interface Software and Technology - UIST '13* (2013), ACM Press, pp. 505–514.

[25] CHALMERS, M., MACCOLL, I., AND BELL, M. Seamful Design: Showing the Seams in Wearable Computing. In *Proceedings of Eurowearable '03* (2003), IEE, pp. 11–16.

[26] CHARBONNEAU, E., HUGHES, C. E., AND LAVIOLA JR, J. J. Vibraudio Pose: An Investigation of Non-Visual Feedback Roles for Body Controlled Video Games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games - Sandbox 2010* (2010), ACM Press, pp. 79–84.

[27] CHEN, X., SCHWARZ, J., HARRISON, C., MANKOFF, J., AND HUDSON, S. E. Air+Touch: Interweaving Touch & In-Air Gestures. In *Proceedings of the 27th Symposium on User Interface Software and Technology - UIST '14* (2014), ACM Press, pp. 519–525.

[28] DELAMARE, W., COUTRIX, C., AND NIGAY, L. Pointing in the Physical World for Light Source Selection. In *Proceedings of the Designing Interactive Lighting Workshop* (2012).

[29] DELAMARE, W., COUTRIX, C., AND NIGAY, L. Mobile Pointing Task in the Physical World: Balancing Focus and Performance while Disambiguating. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices & Services - Mobile HCI '13* (2013), ACM Press, pp. 89–98.

[30] DESLOGE, J. G., REED, C. M., BRAIDA, L. D., PEREZ, Z. D., DELHORNE, L. A., AND VILLABONA, T. J. Auditory and tactile gap discrimination by observers with normal and impaired hearing. *The Journal of the Acoustical Society of America 135*, 2 (2014), 838–850.

[31] DJAJADININGRAT, T., GEURTS, L., BONT, J. D., AND CHAO, P.-Y. Grace: A gesture-controlled wake-up light. In *Proceedings of the 7th International Workshop on the Design & Semantics of Form & Movement - DeSForM '12* (2012), pp. 130–136.

[32] FASTL, H., AND ZWICKER, E. Critical Bands and Excitation. In *Psychoacoustics: Facts and Models*. Springer, 2007, ch. 6, pp. 149–173.

[33] FEKETE, J.-D., ELMQVIST, N., AND GUIARD, Y. Motion-Pointing: Target Selection using Elliptical Motions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '09* (2009), ACM Press, pp. 289–298.

[34] FLEER, D., AND LEICHSENRING, C. MISO: A Context-Sensitive Multimodal Interface for Smart Objects Based on Hand Gestures and Finger Snaps. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology - UIST '12* (2012), ACM Press, pp. 93–94.

[35] FREEMAN, E., BREWSTER, S., AND LANTZ, V. Towards Usable and Acceptable Above-Device Interactions. In *Mobile HCI '14 Posters* (2014), ACM Press, pp. 459–464.

[36] GARZOTTO, F., AND VALORIANI, M. "Don't touch the oven": Motion-based Touchless Interaction with Household Appliances. In *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '12* (2012), ACM Press, pp. 721–724.

[37] GOLOD, I., HEIDRICH, F., MÖLLERING, C., AND ZIEFLE, M. Design Principles of Hand Gesture Interfaces for Microinteractions. In *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces - DPPI '13* (2013), ACM Press, pp. 11–20.

[38] GROSSE-PUPPENDAHL, T., BECK, S., WILBERS, D., ZEISS, S., VON WILMSDORFF, J., AND KUIJPER, A. Ambient Gesture-Recognizing Surfaces with Visual Feedback. In *Proceedings of the 2nd International Conference on Distributed, Ambient, and Pervasive Interactions - DAPI '14* (2014), Springer, pp. 97–108.

[39] GUIARD, Y. On Fitts's and Hooke's laws: Simple harmonic movement in upper-limb cyclical aiming. *Acta Psychologica 82*, 1-3 (1993), 139–159.

[40] GUPTA, S., MORRIS, D., PATEL, S., AND TAN, D. AirWave: Non-Contact Haptic Feedback Using Air Vortex Rings. In *Proceedings of UbiComp '13* (2013), ACM Press, pp. 419–428.

[41] GUSTAFSON, S., HOLZ, C., AND BAUDISCH, P. Imaginary Phone: Learning Imaginary Interfaces by Transferring Spatial Memory from a Familiar Device. In *Proceedings of the 24th Symposium on User Interface Software and Technology - UIST '11* (2011), ACM Press, pp. 283–292.

[42] HARRISON, C., HORSTMAN, J., HSIEH, G., AND HUDSON, S. Unlocking the Expressivity of Point Lights. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '12* (2012), ACM Press, pp. 1683–1692.

[43] HARRISON, C., AND HUDSON, S. E. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. In *Proceedings of the 22nd Symposium on User Interface Software and Technology - UIST '09* (2009), ACM Press, pp. 121–124.

[44] HART, S. G. Nasa-Task Load Index (NASA-TLX): 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting - HFES '06* (2006), 904–908.

[45] HART, S. G., AND STAVELAND, L. E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. North Holland Press, Amsterdam, 1988, pp. 139–183.

[46] HASAN, K., AHLSTRÖM, D., AND IRANI, P. AD-Binning: Leveraging Around Device Space for Storing, Browsing and Retrieving Mobile Device Content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 899–908.

[47] HASEGAWA, K., AND SHINODA, H. Aerial Display of Vibrotactile Sensation with High Spatial-Temporal Resolution using Large-Aperture Airborne Ultrasound Phased Array. In *Proceedings of World Haptics 2013* (2013), IEEE, pp. 31–36.

[48] HELLMAN, R. P., AND ZWISLOCKI, J. J. Loudness Determination at Low Sound Frequencies. *The Journal of the Acoustical Society of America 43*, 1 (1968), 60–64.

[49] HINCAPIÉ-RAMOS, J. D., GUO, X., MOGHADASIAN, P., AND IRANI, P. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '14* (2014), ACM Press, pp. 1063–1072.

[50] HINCKLEY, K., PAUSCH, R., GOBLE, J. C., AND KASSELL, N. F. A Survey of Design Issues in Spatial Input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology - UIST '94* (1994), ACM Press, pp. 213–222.

[51] HOGGAN, E., AND BREWSTER, S. Designing Audio and Tactile Crossmodal Icons for Mobile Devices. In *Proceedings of the 9th International Conference on Multimodal iIterfaces - ICMI '07* (2007), ACM Press, pp. 162–169.

[52] HOLLAND, S., MORSE, D. R., AND GEDENRYD, H. AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal and Ubiquitous Computing 6* (2002), 253–259.

[53] HOSHI, T. Development of aerial-input and aerial-tactile-feedback system. In *Proceedings of the 2011 IEEE World Haptics Conference* (2011), IEEE, pp. 569–573.

[54] HOSHI, T. Compact Ultrasound Device for Noncontact Interaction. In *Proceedings of Advances in Computer Entertainment - ACE '12* (2012), Springer, pp. 502–505.

[55] HOSHI, T., NISHIYAMA, Y., AND TORIGOE, I. Observations of airflow arising from airborne Ultrasound Tactile Display. In *Proceedings of the SICE Annual Conference - SICE '10* (2010), IEEE, pp. 384–385.

[56] HOSHI, T., TAKAHASHI, M., IWAMOTO, T., AND SHINODA, H. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Transactions on Haptics 3*, 3 (2010), 155–165.

[57] HOVE, M. J., FAIRHURST, M. T., KOTZ, S. A., AND KELLER, P. E. Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *NeuroImage 67* (2013), 313–321.

[58] HOVE, M. J., SPIVEY, M. J., AND KRUMHANSL, C. L. Compatibility of Motion Facilitates Visuomotor Synchronization. *Journal of experimental psychology. Human perception and performance 36*, 6 (2010), 1525–1534.

[59] HUMES, L. E., BUSEY, T. A., CRAIG, J. C., AND KEWLEY-PORT, D. The effects of age on sensory thresholds and temporal gap detection in hearing, vision, and touch. *Attention, Perception, & Psychophysics 71*, 4 (2009), 860–871.

[60] HWANG, S., BIANCHI, A., AHN, M., AND WOHN, K.-Y. MagPen: Magnetically Driven Pen Interaction On and Around Conventional Smartphones. In *Proceedings of the 15th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '13* (2013), ACM Press, pp. 412–415.

[61] ION, A., WANG, E. J., AND BAUDISCH, P. Skin Drag Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '15* (2015), ACM Press, pp. 2501–2504.

[62] IWAMOTO, T., TATEZONO, M., AND SHINODA, H. Non-contact method for producing tactile sensation using airborne ultrasound. In *Proceedings of EuroHaptics 2008* (2008), Springer, pp. 504–513.

[63] JONES, B., SODHI, R., FORYSTH, D., BAILEY, B., AND MACIOCCI, G. Around Device Interaction for Multiscale Navigation. In *Proceedings of the 14th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '12* (2012), ACM Press, pp. 83–92.

[64] JONES, B. R., BENKO, H., OFEK, E., AND WILSON, A. D. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 869–878.

[65] KAJASTILA, R., AND LOKKI, T. Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human Computer Studies 71*, 5 (2013), 627–640.

[66] KAMAL, A., LI, Y., AND LANK, E. Teaching Motion Gestures via Recognizer Feedback. In *Proceedings of the 19th International Conference on Intelligent User Interfaces - IUI '14* (2014), ACM Press, pp. 73–82.

[67] KARAM, M., AND SCHRAEFEL, M. C. A Taxonomy of Gestures in Human Computer Interactions. Tech. rep., University of Southamptom, Electronics and Computer Science, 2005.

[68] KELLOGG, B., TALLA, V., AND GOLLAKOTA, S. Bringing Gesture Recognition To All Devices. In *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation - NSDI 14* (2014), USENIX.

[69] KETABDAR, H., ROSHANDEL, M., AND YÜKSEL, K. A. Towards Using Embedded Magnetic Field Sensor for Around Mobile Device 3D Interaction. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '10* (2010), ACM Press, pp. 153–156.

[70] KJELDSEN, R., AND HARTMAN, J. Design Issues for Vision-Based Computer Interaction Systems. In *Proceedings of the Workshop on Perceptive User Interfaces - PUI '01* (2001), ACM Press, pp. 1–8.

[71] KOSKELA, T., AND VÄÄNÄNEN-VAINIO-MATTILA, K. Evolution towards smart home environments: Empirical evaluation of three user interfaces. *Personal and Ubiquitous Computing 8* (2004), 234–240.

[72] KRATZ, S., AND BALLAGAS, R. Unravelling Seams: Improving Mobile Gesture Recognition with Visual Feedback Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '09* (2009), ACM Press, pp. 937–940.

[73] KRATZ, S., AND ROHS, M. HoverFlow: Expanding the Design Space of Around-Device Interaction. In *Proceedings of the 11th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '09* (2009), ACM Press, p. Article 4.

[74] KRATZ, S., ROHS, M., GUSE, D., MÜLLER, J., BAILLY, G., AND NISCHT, M. PalmSpace: Continuous Around-Device Gestures vs. Multitouch for 3D Rotation Tasks on Mobile Devices. In *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '12* (2012), ACM Press, pp. 181–188.

[75] LANTZ, V., AND MURRAY-SMITH, R. Rhythmic Interaction with a Mobile Device. In *Proceedings of the third Nordic Conference on Human Computer Interaction - NordiCHI '04* (2004), no. 2, ACM Press, pp. 97–100.

[76] LEE, J., HAN, J., AND LEE, G. Investigating the Information Transfer Efficiency of a 3x3 Watch-back Tactile Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '15* (2015), ACM Press, pp. 1229–1232.

[77] LEE, S. C., LI, B., AND STARNER, T. AirTouch: Synchronizing In-air Hand Gesture and On-body Tactile Feedback to Augment Mobile Gesture Interaction. In *Proceedings of the 15th International Symposium on Wearable Computers ISWC '11* (2011), IEEE, pp. 3–10.

[78] LÖCKEN, A., HESSELMANN, T., PIELOT, M., HENZE, N., AND BOLL, S. User-centred process for the definition of free-hand gestures applied to controlling music playback. *Multimedia Systems 18*, 1 (2011), 15–31.

[79] LONG, B., SEAH, S. A., CARTER, T., AND SUBRAMANIAN, S. Rendering Volumetric Haptic Shapes in Mid-Air using Ultrasound. *ACM Transactions on Graphics 33*, 6 (2014), Article 181.

[80] LORÅS, H., STENSDOTTER, A. K., ÖHBERG, F., AND SIGMUNDSSON, H. Individual differences in timing of discrete and continuous movements: a dimensional approach. *Psychological Research 78*, 2 (2014), 289–299.

[81] MAGLIO, P. P., MATLOCK, T., CAMPBELL, C. S., ZHAI, S., AND SMITH, B. A. Gaze and Speech in Attentive User Interfaces. In *Proceedings of the International Conference on Multimodal Interaction in LNCS 1948 - ICMI 2000* (2000), Springer, pp. 1–7.

[82] MALACRIA, S., LECOLINET, E., AND GUIARD, Y. Clutch-Free Panning and Integrated Pan-Zoom Control on Touch-Sensitive Surfaces: The CycloStar Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '10* (2010), ACM Press, pp. 2615–2624.

[83] MATSCHEKO, M., FERSCHA, A., RIENER, A., AND LEHNER, M. Tactor Placement in Wrist Worn Wearables. In *Proceedings of the International Symposium on Wearable Computers - ISWC '10* (2010), IEEE, pp. 1–8.

[84] MCADAM, C., AND BREWSTER, S. Distal Tactile Feedback for Text Entry on Tabletop Computers. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers - BCS-HCI '09* (2009), British Computer Society, pp. 504–511.

[85] MCADAM, C., AND BREWSTER, S. Mobile phones as a tactile display for tabletop typing. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '11* (2011), ACM Press, pp. 276–277.

[86] MCGOOKIN, D., BREWSTER, S., AND PRIEGO, P. Audio Bubbles: Employing Non-speech Audio to Support Tourist Wayfinding. In *Proceedings of Haptic and Audio Interaction Design - HAID '09* (2009), Springer-Verlag, pp. 41–50.

[87] MCNEILL, D. *Gesture and Thought*, 1st ed. University of Chicago Press, 2005.

[88] MISTRY, P., MAES, P., AND CHANG, L. WUW - Wear Ur World: A Wearable Gestural Interface. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (2009), ACM Press, pp. 4111–4116.

[89] MONNAI, Y., HASEGAWA, K., FUJIWARA, M., YOSHINO, K., INOUE, S., AND SHINODA, H. HaptoMime: Mid-Air Haptic Interaction with a Floating Virtual Screen. In *Proceedings of the 27th Symposium on User Interface Software and Technology - UIST '14* (2014), ACM Press, pp. 663–667.

[90] MORRISON, C., SMYTH, N., CORISH, R., O'HARA, K., AND SELLEN, A. Collaborating with Computer Vision Systems: An Exploration of Audio Feedback. In *Proceedings of the 2014 Conference on Designing Interactive Systems - DIS '14* (2014), ACM Press, pp. 229–238.

[91] MORRISON-SMITH, S., AND RUIZ, J. Using Audio Cues to Support Motion Gesture Interaction on Mobile Devices. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (2014), ACM Press, pp. 1621–1626.

[92] MÜLLER, H., KAZAKOVA, A., PIELOT, M., HEUTEN, W., AND BOLL, S. Ambient Timer - Unobtrusively Reminding Users of Upcoming Tasks with Ambient Light. In *Proceedings of INTERACT '13* (2013), Springer Berlin Heidelberg, pp. 211–228.

[93] MÜLLER, H., LÖCKEN, A., HEUTEN, W., AND BOLL, S. Sparkle: An Ambient Light Display for Dynamic Off-Screen Points of Interest. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction - NordiCHI '14* (2014), ACM Press, pp. 51–60.

[94] NESSELRATH, R., LU, C., SCHULZ, C. H., FREY, J., AND ALEXANDERSSON, J. A Gesture Based System for Context-Sensitive Interaction with Smart Homes. In *Proceedings of Ambient Assisted Living '11* (2011), pp. 209–219.

[95] NIELSEN, M., STÖRRING, M., MOESLUND, T. B., AND GRANUM, E. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In *5th International Gesture Workshop* (2004), Springer Berlin Heidelberg, pp. 409–420.

[96] NORMAN, D. A. *The Psychology of Everyday Things*. Basic Books, 1988.

[97] NORMAN, D. A. Natural User Interfaces Are Not Natural. *Interactions 17*, 3 (2010), 6–10.

[98] OAKLEY, I., KIM, Y., LEE, J., AND RYU, J. Determining the Feasibility of Forearm Mounted Vibrotactile Displays. In *Proceedings of the 14th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (2006), IEEE, pp. 27–34.

[99] OFFERMANS, S. A. M., VAN ESSEN, H. A., AND EGGEN, J. H. User interaction with everyday lighting systems. *Personal and Ubiquitous Computing 18*, 8 (2014), 2035–2055.

[100] OH, U., KANE, S. K., AND FINDLATER, L. Follow That Sound: Using Sonification and Corrective Verbal Feedback to Teach Touchscreen Gestures. In *Proceedings of the 15th International Conference on Computers and Accessibility - ASSETS '13* (2013), ACM Press, p. Article 13.

[101] O'HARA, K., GONZALEZ, G., SELLEN, A., PENNEY, G., VARNAVAS, A., MENTIS, H., CRIMINISI, A., CORISH, R., ROUNCEFIELD, M., DASTUR, N., AND CARRELL, T. Touchless Interaction in Surgery. *Communications of the ACM 57*, 1 (2014), 70–77.

[102] PANEELS, S., ANASTASSOVA, M., STRACHAN, S., VAN, S. P., SIVACOUMARANE, S., AND BOLZMACHER, C. What's Around Me Multi-Actuator Haptic Feedback on the Wrist. In *Proceedings of World Haptics 2013* (2013), IEEE, pp. 407–412.

[103] PANGER, G. Kinect in the Kitchen: Testing Depth Camera Interactions in Practical Home Environments. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (2012), ACM Press, pp. 1985–1990.

[104] PARK, Y., KIM, J., AND LEE, K. Effects of Auditory Feedback on Menu Selection in Hand-Gesture Interfaces. *IEEE Multimedia 22*, 1 (2015), 32–40.

[105] PASQUERO, J., STOBBE, S. J., AND STONEHOUSE, N. A Haptic Wristwatch for Eyes-Free Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11* (2011), ACM Press, pp. 3257–3266.

[106] PIELOT, M., POPPINGA, B., HEUTEN, W., AND BOLL, S. A Tactile Compass for Eyes-Free Pedestrian Navigation. In *Proceedings of INTERACT '11* (2011), pp. 640–656.

[107] PITT, I. J., AND EDWARDS, A. D. Navigating the Interface by Sound for Blind Users. In *People and Computers VI: Proceedings of HCI '91* (1991), Cambridge University Press, pp. 373–383.

[108] POHL, H., AND MURRAY-SMITH, R. Focused and Casual Interactions: Allowing Users to Vary Their Level of Engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 2223–2232.

[109] POHL, H., AND ROHS, M. Around-Device Devices: My Coffee Mug is a Volume Dial. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services - Mobile HCI '14* (2014), ACM Press, pp. 81–90.

[110] QIN, Q., ROHS, M., AND KRATZ, S. Dynamic Ambient Lighting for Mobile Devices. In *Adjunct Proceedings of the 24th Symposium on User Interface Software and Technology - UIST '11 Adjunct* (2011), ACM Press, pp. 51–52.

[111] REEVES, S., BENFORD, S., O'MALLEY, C., AND FRASER, M. Designing the Spectator Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05* (2005), ACM Press, pp. 741–750.

[112] REKIMOTO, J., AND TSUIJITA, H. Inconvenient Interactions: An Alternative Design Approach to Enrich our Lives. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI '14* (2014), ACM Press, pp. 225–228.

[113] REPP, B. H. Rate Limits of Sensorimotor Synchronization. *Advances in Cognitive Psychology 2*, 2 (2006), 163–181.

[114] REPP, B. H. Comfortable synchronization of cyclic drawing movements with a metronome. *Human Movement Science 30*, 1 (2011), 18–39.

[115] REPP, B. H., AND SU, Y.-H. Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review 20*, 3 (2013), 403–52.

[116] RICO, J., AND BREWSTER, S. Usable Gestures for Mobile Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '10* (2010), ACM Press, pp. 887–896.

[117] ROGERS, Y., AND MULLER, H. A Framework for Designing Sensor-Based Interactions to Promote Exploration and Reflection in Play. *International Journal of Human-Computer Studies 64*, 1 (2006), 1–14.

[118] ROVELO, G., DEGRAEN, D., VANACKEN, D., LUYTEN, K., AND CONINX, K. Gestu-Wan - An Intelligible Mid-Air Gesture Guidance System for Walk-up-and-Use Displays. In *Proceedings of INTERACT '15 in LNCS 9297* (2015), pp. 368–386.

[119] RUKZIO, E., LEICHTENSTERN, K., AND SCHMIDT, A. Mobile Interaction with the Real World: An Evaluation and Comparison of Physical Mobile Interaction Techniques. In *Proceedings of the 2007 European Conference on Ambient Intelligence - AmI '07* (2007), Springer, pp. 1–18.

[120] SCHMIDT, D., MOLYNEAUX, D., AND CAO, X. PICOntrol: Using a Handheld Projector for Direct Control of Physical Devices through Visible Light. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology - UIST '12* (2012), ACM Press, pp. 379–388.

[121] SCHWARZ, J., MARAIS, C. C., LEYVAND, T., HUDSON, S. E., AND MANKOFF, J. Combining Body Pose, Gaze, and Gesture to Determine Intention to Interact in Vision-Based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '14* (2014), ACM Press, pp. 3443–3452.

[122] SHERRICK, C. E. A scale for rate of tactual vibration. *The Journal of the Acoustical Society of America 78*, 1 (1985), 78–83.

[123] SHINODA, H. Noncontact Haptic Interface Using Ultrasound. In *Proceedings of Haptic and Audio Interaction Design - HAID '11* (2011), Springer, pp. 120–127.

[124] SODHI, R., BENKO, H., AND WILSON, A. D. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '12* (2012), ACM Press, pp. 179–188.

[125] SODHI, R., POUPYREV, I., GLISSON, M., AND ISRAR, A. AIREAL: Interactive Tactile Experiences in Free Air. *ACM Transactions on Graphics 32*, 4 (2013), Article 134.

[126] SØRENSEN, T., ANDERSEN, O. D., AND MERRITT, T. "Tangible Lights": In-Air Gestural Control of Home Lighting. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction - TEI '15 Work-in-Progress* (2015), ACM Press, pp. 727–732.

[127] STARNER, T., AUXIER, J., ASHBROOK, D., AND GANDY, M. The Gesture Pendant: A Self-Illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. In *Proceedings of the Fourth International Symposium on Wearable Computers - ISWC '00* (2000), IEEE, pp. 87–94.

[128] STUDENKA, B. E., AND ZELAZNIK, H. N. Circle Drawing Does Not Exhibit Auditory-Motor Synchronization. *Journal of Motor Behavior 43*, 3 (2011), 185–191.

[129] STUDENKA, B. E., AND ZELAZNIK, H. N. Synchronization in repetitive smooth movement requires perceptible events. *Acta Psychologica 136*, 3 (2011), 432–441.

[130] STUDENKA, B. E., ZELAZNIK, H. N., AND BALASUBRAMANIAM, R. The distinction between tapping and circle drawing with and without tactile feedback: an examination of the sources of timing variance. *Quarterly Journal of Experimental Psychology 65*, 6 (2012), 1086–1100.

[131] SUZUKI, Y., AND KOBAYASHI, M. Air Jet Driven Force Feedback in Virtual Reality. *IEEE Computer Graphics and Applications 25*, 1 (2005), 44–47.

[132] SUZUKI, Y., AND TAKESHIMA, H. Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America 116*, 2 (2004), 918–933.

[133] SWINDELLS, C., INKPEN, K. M., DILL, J. C., AND TORY, M. That one there! Pointing to establish device identity. In *Proceedings of the 15th Annual Symposium on User Interface Software and Technology - UIST '02* (2002), ACM Press, pp. 151–160.

[134] TAHIROLU, K., SVEDSTRÖM, T., WIKSTRÖM, V., OVERSTALL, S., KILDAL, J., AND AHMANIEMI, T. SoundFLEX: Designing Audio to Guide Interactions with Shape-Retaining Deformable Interfaces. In *Proceedings of the 16th International Conference on Multimodal Interaction - ICMI '14* (2014), ACM Press, pp. 267–274.

[135] TAKAHASHI, M., AND SHINODA, H. Large aperture Airborne Ultrasound Tactile Display using distributed array units. In *Proceedings of the SICE Annual Conference - SICE '10* (2010), IEEE, pp. 359–362.

[136] VÄLKKYNEN, P., NIEMELÄ, M., AND TUOMISTO, T. Evaluating Touching and Pointing with a Mobile Terminal for Physical Browsing. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction - NordiCHI '06* (2006), ACM Press, pp. 28–37.

[137] VAN DER WEL, R. P. R. D., STERNAD, D., AND ROSENBAUM, D. A. Moving the Arm at Different Rates: Slow Movements are Avoided. *Journal of Motor Behavior 26*, 5 (jan 2009), 29–36.

[138] VARLET, M., MARIN, L., ISSARTEL, J., SCHMIDT, R. C., AND BARDY, B. G. Continuity of Visual and Auditory Rhythms Influences Sensorimotor Coordination. *PLoS ONE 7*, 9 (2012), e44082.

[139] VERMEULEN, J., LUYTEN, K., CONINX, K., MARQUARDT, N., AND BIRD, J. Proxemic Flow: Dynamic Peripheral Floor Visualizations for Revealing and Mediating Large Surface Interactions. In *Proceedings of INTERACT '15 in LNCS 9299* (2015), pp. 264–281.

[140] VO, D.-B., AND BREWSTER, S. Touching the Invisible: Localizing Ultrasonic Haptic Cues. In *Proceedings of World Haptics Conference 2015 - WHC '15* (2015), IEEE, pp. 368 – 373.

[141] WALTER, R., BAILLY, G., AND MÜLLER, J. StrikeAPose: Revealing Mid-Air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), ACM Press, pp. 841–850.

[142] WELCH, R. B., AND WARREN, D. H. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin 88*, 3 (1980), 638–667.

[143] WENSVEEN, S. A. G., DJAJADININGRAT, J. P., AND OVERBEEKE, C. J. Interaction Frogger: A Design Framework to Couple Action and Function Through Feedback and Feedforward. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques - DIS '04* (2004), ACM Press, pp. 177–184.

[144] WIGDOR, D., AND WIXON, D. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*, 1 ed. Morgan Kaufmann Publishers Inc., 2011.

[145] WIKSTRÖM, V., OVERSTALL, S., TAHIROLU, K., KILDAL, J., AND AHMANIEMI, T. MARSUI: Malleable Audio-Reactive Shape-Retaining User Interface. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (2013), ACM Press, pp. 3151–3154.

[146] WILLIAMSON, J., AND MURRAY-SMITH, R. Audio feedback for gesture recognition. Tech. rep., Technical Report TR-2002-127, Department of Computing Science, University of Glasgow, 2002.

[147] WILLIAMSON, J., AND MURRAY-SMITH, R. Pointing Without a Pointer. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (2004), ACM Press, pp. 1407–1410.

[148] WILLIAMSON, J., AND MURRAY-SMITH, R. Sonification of Probabilistic Feedback through Granular Synthesis. *IEEE Multimedia 12*, 2 (apr 2005), 45–52.

[149] WILLIAMSON, J. R., CROSSAN, A., AND BREWSTER, S. Multimodal Mobile Interactions: Usability Studies in Real World Settings. In *Proceedings of ICMI 2011* (2011), ACM Press, pp. 361–368.

[150] WILSON, A., BENKO, H., IZADI, S., AND HILLIGES, O. Steerable Augmented Reality with the Beamatron. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology - UIST '12* (2012), ACM Press, pp. 413–422.

[151] WILSON, G., CARTER, T., SUBRAMANIAN, S., AND BREWSTER, S. Perception of Ultrasonic Haptic Feedback on the Hand: Localisation and Apparent Motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '14* (2014), ACM Press, pp. 1133–1142.

[152] WOBBROCK, J. O., AUNG, H. H., ROTHROCK, B., AND MYERS, B. A. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (2005), ACM Press, pp. 1869–1872.

[153] WOBBROCK, J. O., FINDLATER, L., GERGLE, D., AND HIGGINS, J. J. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11* (2011), ACM Press, pp. 143–146.

[154] XU, C., AND LYONS, K. Shimmering Smartwatches: Exploring the Smartwatch Design Space. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction - TEI '15* (2015), ACM Press, pp. 69–76.

[155] YANG, X.-D., HASAN, K., BRUCE, N., AND IRANI, P. Surround-See: Enabling Peripheral Vision on Smartphones during Active Use. In *Proceedings of the 26th Symposium on User Interface Software and Technology - UIST '13* (2013), ACM Press, pp. 291–300.

[156] ZAMBORLIN, B., BEVILACQUA, F., GILLIES, M., AND D'INVERNO, M. Fluid Gesture Interaction Design: Applications of Continuous Recognition for the Design of Modern Gestural Interfaces. *ACM Transactions on Interactive Intelligent Systems 3*, 4 (2014), Article 22.

[157] ZHANG, C., GUO, A., ZHANG, D., SOUTHERN, C., ARRIAGA, R., AND ABOWD, G. BeyondTouch: Extending the Input Language with Built-in Sensors on Commodity Smartphones. In *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15* (2015), ACM Press, pp. 66–77.

[158] ZHAO, S., DRAGICEVIC, P., CHIGNELL, M., BALAKRISHNAN, R., AND BAUDISCH, P. earPod: Eyes-free Menu Selection using Touch Input and Reactive Audio Feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07* (2007), ACM Press, pp. 1395–1404.

[159] ZWICKER, E. Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). *The Journal of the Acoustical Society of America 33*, 2 (1961), 248.