Investigating Factors Affecting Haptic Information Transfer and a Comparative Analysis of Audio and Audio-Haptic Maps

Sri Gannavarapu

Department of Electrical & Computer Engineering
McGill University
Montréal, Québec, Canada

April 2024

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science.

2024/04/15

Abstract

Maps enable location and navigational inferences by providing individuals with a layout of an environment. However, comprehension of traditional maps relies heavily on visual cues, the perception of which is diminished in blind and low vision (BLV) individuals. Members of this community can enhance their perception of spatio-visual information through use of touch in haptic interactions, compensating for the diminished visual references. Motivated by such a use case, we aimed to better understand how people perceive haptic feedback in relation to maps through a) the perceived separation between two tactile stimuli as a possible representation of distance and b) user-based feedback of audio and audio-haptic stimuli and the overall improvement of map comprehension.

In the first part of this thesis, we explored the role of distance between actuators and its impact on the frequency discrimination ability in a general population. We used the adaptive staircase methodology to calculate the just-noticeable-difference (JND) measures for three actuator separations ("close" [5 – 10 cm], "mid" [30 – 40 cm], and "far" [70 – 80 cm]). The lower JND threshold for the "far" separation suggests that the ability to discriminate haptic frequency increases with the distance between actuators.

However, we note that the sample size ($n = 9$) of the study was too small to achieve statistical significance. Our findings from this study contribute to the broader understanding of the fundamental mechanisms of haptic perception. Future work geared towards characterizing other body regions with a high density of mechanoreceptors can provide comprehensive information on ideal targets for haptic devices.

Next, we characterized the efficacy of unimodal audio and two distinct audio-haptic (force feedback and tactile) renderings in conveying map information to BLV users using semi-structured interviews. We found that participants interacting with these renderings revealed six major themes that influenced the user's preference for each rendering approach: *initial and lasting impressions*, *multimodal congruence*, *cognitive load*, *perceived affordances*, *learnability*, and *influence of background*. Subsequently, participant descriptions were scored based on accuracy to evaluate how well each map rendering was understood. Analysis of these scores suggested an improvement in map comprehension when participants interacted with the tactile refreshable pin array over both the audio-only and force feedback interactions. In our analysis of the specific quotations, underlying themes, and task scores, we offer a discussion of the implications of the resulting themes on future renderings. Our findings highlight a need for more targeted training sessions to better educate users about the affordances of each device. Additionally, improving crossmodal congruence by focusing on sensory illusions, such as the vertical–horizontal illusion, needs to be taken into account to enhance the multimodal congruence of the audio-haptic experiences. Our observations and findings enable the generation of more interactive renderings through continued iteration of our current designs.

Résumé

Les cartes permettent de localiser un lieu et de s'y orienter en fournissant aux individus un plan de l'environnement. Cependant, la compréhension des cartes traditionnelles repose en grande partie sur des indices visuels, dont la perception est réduite chez les personnes BLV. Les membres de cette communauté peuvent améliorer leur perception des informations spatiales et visuelles en utilisant le toucher dans les interactions haptiques, ce qui compense la diminution des références visuelles. Motivés par un tel cas d'utilisation, nous avons cherché à mieux comprendre comment les gens perçoivent le retour d'information haptique en relation avec les cartes à travers a) la séparation perçue entre deux stimuli tactiles comme une représentation possible de la distance et b) le retour d'information basé sur l'utilisateur des stimuli audio et audio-haptiques et l'amélioration globale de la compréhension de la carte.

Dans la première partie de cette thèse, nous avons étudié le rôle de la distance entre les actionneurs et son impact sur la capacité de discrimination des fréquences dans une population générale. Nous avons utilisé la méthode de l'escalier adaptatif pour calculer les mesures de JND pour trois séparations d'actionneurs ("proche" [5 – 10 cm], "moyen" [30 – 40 cm], et "éloigné" [70 – 80 cm]). Le seuil inférieur de JND pour la séparation "éloigné" suggère que la capacité à discriminer la fréquence haptique augmente avec la distance entre les actionneurs. Cependant, nous notons que la taille de l'échantillon ($n = 9$) de l'étude était trop petite pour atteindre une signification statistique. Les résultats de cette étude contribuent à une meilleure compréhension des mécanismes fondamentaux de la perception haptique. Nos résultats contribuent à une meilleure compréhension des mécanismes fondamentaux de la perception haptique. De futurs travaux visant à caractériser d'autres régions du corps présentant une forte densité de mécanorécepteurs peuvent fournir des informations complètes sur les cibles idéales pour les

dispositifs haptiques.

Ensuite, nous avons caractérisé l'efficacité de l'audio unimodal et de deux rendus audio-haptiques distincts (retour de force et tactile) pour transmettre des informations cartographiques aux utilisateurs de BLV à l'aide d'entretiens semi-structurés. Nous avons constaté que les participants interagissant avec ces rendus ont révélé six thèmes majeurs qui ont influencé la préférence de l'utilisateur pour chaque approche de rendu : *impression initiale et durable*, *congruence multimodale*, *charge cognitive*, *perception des possibilités*, *apprenabilité*, et *influence de l'arrière-plan*. Par la suite, les descriptions des participants ont été notées sur la base de la précision afin d'évaluer la compréhension de chaque rendu de carte. L'analyse de ces scores suggère une amélioration de la compréhension de la carte lorsque les participants interagissent avec le réseau tactile d'épingles rafraîchissables par rapport aux interactions audio uniquement et au retour de force. Dans notre analyse des citations spécifiques, des thèmes sous-jacents et des scores des tâches, nous proposons une discussion sur les implications des thèmes résultants sur les rendus futurs. Nos résultats mettent en évidence la nécessité d'organiser des sessions de formation plus ciblées afin de mieux informer les utilisateurs sur les possibilités offertes par chaque dispositif. En outre, l'amélioration de la congruence intermodale en se concentrant sur les illusions sensorielles, telles que l'illusion verticale-horizontale, doit être prise en compte pour améliorer la congruence multimodale des expériences audio-haptiques. Nos observations et nos résultats permettent de générer des rendus plus interactifs grâce à l'itération continue de nos conceptions actuelles.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Jeremy Cooperstock, for the invaluable guidance, unwavering support, and mentorship throughout the course of my thesis. Their insightful feedback and encouragement have been instrumental in shaping the outcome of this research.

I extend my heartfelt thanks to my dedicated labmates, whose collaborative spirit and shared enthusiasm created an enriching research environment. Special appreciation is extended to Cyan Kuo and Linnea Kirby, for constantly lifting my spirits.

I also wish to acknowledge my sister for her exceptional contributions to this endeavour. Her countless hours of life coaching, meticulous editing assistance, and unwavering moral support have significantly helped me in getting me over the finish line.

Last but certainly not least, I extend my deepest appreciation to my mom for her enduring encouragement and steadfast support during some of the most challenging periods of my life. Her belief in my capabilities has been a constant source of motivation, and I am truly grateful for her calming presence throughout this academic journey.

Contents

List of Figures

List of Tables

Acronyms

ANOVA analysis of variance.

API application programming interface.

BLV blind and low vision.

df degrees of freedom.

DOF degrees of freedom.

EE end-effector.

FF force feedback.

HCI human-computer interaction.

IMAGE Internet Multimodal Access to Graphical Exploration.

IT information transfer.

JAWS Job Access with Speech.

JND just-noticeable-difference.

JPEG joint photographic experts group.

MIDI musical instrument digital interface.

ML machine learning.

OSM OpenStreetMap.

PID proportional derivative integral.

POI points of interest.

PSE point of subjective equality.

REB research ethics board.

TTS text-to-speech.

Chapter 1

Introduction

Vision, hearing, taste, touch, and smell are the five of the sensory modalities that influence the human experience and shape the perception of our environment, with touch being an essential, and sometimes overlooked part of our daily lives [2]. Haptics refers to the study of touch sensing and its application in design [3, 4, 5]. A common implementation of haptics is observed in wearables, such as smartwatches and VR accessories, which often make use of inexpensive vibration motors. These haptic actuators remain favoured in wearables due to their small size and low complexity.

Previous research efforts focused on investigating the factors that influence vibrotactile perception predominantly concentrated on the fingertips, primarily due to the high density of mechanoreceptors at this location [6, 7]. However, considering the prevalence of wearable devices placed on the wrist, it is imperative that we comprehensively understand the effectiveness of vibrotactile stimuli in this particular area, which has a comparatively lower density of mechanoreceptors. Moreover, owing to our consistent movement throughout the day, wearables located on the wrist encounter spatial variations as a consequence of postural changes arising from routine activities. This requires an investigation of the impact of postural changes on vibrotactile perception, with a specific focus on examining the spatial distance between the actuators located on the wrist. Based on this need, the first part of this thesis assesses the effects of changes in distance facilitated through postural changes on vibrotactile frequency discrimination. A limitation of vibrotactile stimuli is its lower information transfer (IT) bandwidth when used unimodally [8]. Characterizing the factors that influence haptic IT can inform more complex rendering designs. Specifically, we aimed to characterize the ability to discriminate vibration frequencies on the wrist by examining the impact of varying distances between actuators on haptic IT. Understanding this effect can be beneficial to designers of haptic interactions, as it would indicate whether information such as the position of actuators is important to consider.

Although maximizing the IT capabilities of the haptic channel can expand the range of possible renderings, designers often supplement haptic feedback with other sensory modalities, such as audio and visual, to provide a more comprehensive understanding of the data presented [9]. The IT limitations of the haptic channel are particularly relevant in the case of visual maps, which often contain a high density of information that can be challenging to convey solely through haptic feedback. The use of traditional haptic maps, such as raised line maps, is an effective method for improving comprehension of geographic data for BLV individuals [10]. However, existing methods for generating these maps are often laborious and expensive. Fortunately, recent advances in digital haptic technology have enabled designers to transfer dense visual renderings despite the limited IT bandwidth present in the haptic modality. Specific examples include the development of commercially available refreshable pin arrays, such as the Dot Pad, as well as low-cost force feedback devices, such as the Haply

2DIY [11].[1]

To address the need for improved maps accessibility for BLV individuals, this thesis primarily focuses on the design and development of unimodal and multimodal renderings of maps. Previous surveys of the BLV individuals [12] indicate that autonomy in navigation is highly valued by this population. A prevalent approach to enhancing autonomy involves the use of preplanning. This method encompasses the gathering of pertinent information about an environment before physically reaching a designated destination. Based on this literature, along with early user interviews, we narrowed our scope to transit maps and focused on two prominent features: route shapes and points of interest. These features aid in preplanning navigation, which can contribute to a sense of autonomy and comfort in self-navigation [13]. To achieve this, our design approach incorporated a combination of speech and nonspeech audio along with two haptic devices: a tactile refreshable pin array and a force feedback device. We evaluated the usability of these renderings through semi-structured interviews, which allowed us to gather qualitative feedback directly from users about their experience with the renderings. In this study, participants were provided with audio-only and audio-haptic renderings. The goal was to understand whether the audio-haptic renderings were able to communicate information more effectively than the audio-only renderings. In particular, we hypothesized that the use of multimodality would enhance the communication of map information. By engaging users in open-ended discussions, we gained user input on the efficacy of multimodal maps on perception and understanding of geographic information. Specifically, our objective was to answer the following research questions through the interviews.

- RQ1: "What are the differences in user experience and effectiveness between spatial audio renderings, force feedback (FF), and refreshable pin arrays for rendering maps?"

[1] https://pad.dotincorp.com/

- RQ2: “How can the use of multimodality enhance audio-only renderings?”

The results of this study including the user feedback on the usability of our rendering strategies were applied to the next iteration of multimodal renderings on the refreshable pin array.

1.1 Overview

Chapter 2 reviews the literature relevant to this work. *Chapter 3* presents the manuscript for a perception study that investigates how the distance between actuators affects perception. *Chapter 4* introduces the technologies and strategies used to design prototypes for accessible audio and audio-haptic maps, with a small sample of users participating in a qualitative user study to provide feedback on the designs. The insights from this study are used to propose future directions for prototyping more interactive audio-haptic maps. The conclusion in *Chapter 5* summarizes the thesis and outlines directions for future research, emphasizing the potential for further exploration of haptic technology in accessible maps and other applications.

1.2 Summary of Contributions

This thesis contributes to the growing body of haptics research in two ways. Firstly, by investigating the differences in user experience and effectiveness of three map rendering techniques—spatial audio, force feedback, and refreshable pin arrays—we can better understand how to combine audio with haptics to improve the fidelity of our proposed renderings. Second, by exploring the effect of distance changes on frequency discrimination, we advance our understanding of the fundamentals of haptic perception, informing an improved design of future haptic technologies.

Chapter 2

Related Work

This thesis focuses on advancing our understanding of haptic perception and exploring the application of haptics in accessibility technology. As such, this background section is divided into two main parts. The first section provides an overview of the mechanisms of haptic perception and relevant studies that investigate the role of the distance between haptic actuators. The second section addresses the background relevant to the development of accessible maps, including the mental models necessary for understanding our space, and provides an overview of current efforts to convey maps to BLV individuals.

2.1 Haptic Perception

Haptic perception is the result of a complex process that enables the translation of sensory stimuli into purposeful motor responses and spatial orientations. The process begins with mechanoreceptors that detect touch in the form of vibrations, pressure, temperature, and texture [14]. Four types of cutaneous mechanoreceptors specialize in different forms of touch

(Table 2.1); Meissner's corpuscle, Merkel cell, Pacinian corpuscle, and Ruffini endings [15].

Table 2.1: The four main mechanoreceptors and their specialized roles.

Mechanoreceptor	**Categorization**	**Specializations**
Meissner's corpuscle	Fast Adapting	light touch, grasp, vibrations between 40-60 Hz
Pacinian corpuscle	Fast Adapting	vibrations between 200-300 Hz
Merkel cell	Slow Adapting	fine skin deformations, texture, sustained pressure
Ruffini endings	Slow Adapting	skin stretch, tension, direction of motion, shape, position

2.1.1 On-body Reference Frames

Despite the high density of mechanoreceptors on the fingers [16], wearable devices are often geared towards the forearms and wrists as these locations are less obstructive to daily tasks [17]. However, these sites remain far less studied compared with more sensitive sites, such as the fingers. Often, haptic perception studies include multiple actuators and require the participants to identify where they felt the stimuli and characterize the sensation. Haptic localization refers to the ability to determine the position or location of an object by touch or through tactile feedback and can be categorized into two coordinate systems: spatiotopic (i.e., physical space coordinates) or somatotopic (i.e., along the body coordinates) [18, 19].

Lakatos & Shepard are one of the first groups to note the use of a spatiotopic reference frame by human subjects during a haptic perception study. They instructed and measured the participants' reaction time as they paid attention to a particular location on their body while the target stimulus is presented either at that location or at a different location on their body. Participants detect the target stimulus more efficiently when shifting their attention between nearby locations on their body compared to locations that are farther apart. The time taken to shift attention between two locations on the body depends on the distance between those locations, measured along a straight line. The findings of this study suggest that the localization of haptic stimuli may be conducted using a spatiotopic reference frame [19].

Contrastingly, other studies imply that people use a somatotopic frame of reference [6, 7]. Roder et al. [6] provide tactile stimuli either to the index or middle finger on both hands. They show that reaction times in response to the tactile stimuli are slowest when the cue and target stimuli are presented from the same or adjacent fingers on the same hand, suggesting the use of a somatotopic frame of reference instead of an external one. This result suggests a greater complexity in how reference frames are employed in our perception of haptic stimuli.

Consolidating both perspectives, Shore et al. [20] suggest that individuals switch between two proprioceptive frames of reference in a context-dependent manner. They requested participants to adopt different postures, varying the distance between stimuli on single and opposing hands. The just-noticeable-difference (JND), commonly used in sensory perception research, was used as a measure of tactile acuity to quantify the minimum amount of difference detected between two stimuli [21]. In this case, a smaller JND indicates greater sensitivity to tactile stimuli [22].

The researchers found that the JNDs decrease with distance, indicating greater tactile acuity with spatial separation when opposite sides of the body are stimulated and suggests that people may switch reference frames depending on the context of the stimuli. Specifically, when stimulating a smaller region localized to one side of the body, a somatotopic reference frame may be employed, whereas an external reference frame may be used when stimulating sites on different sides of the body.

To the best of our knowledge, all such studies (i.e., those studying the effect of distance on vibrotactile perception) focus on the fingers. We believe that if the same phenomenon were better characterized at the wrist, this information would be meaningful for the design of haptic devices in this region.

2.2 Maps for Blind and Partially Sighted People

Cognitive mapping, originally proposed by Tolman [23], refers to the process by which individuals create mental representations of geographic space, allowing them to connect with and understand their environment. Humans encode spatial information by building a cognitive map either through direct or indirect experience [24].

With direct experience, individuals explore their environment and transform its features and routes into mental maps. Siegel and White suggested that this process occurs in stages, beginning by noting major landmarks and connecting them to form small routes. Through greater examination, these routes are connected into minimaps, and subsequently combined into larger surveys [25]. Although useful, direct experiences are time-consuming and require multiple explorations of an area.

Cognitive maps are also created through indirect experiences, by interpreting verbal, written, tactile, or graphical materials [26]. Graphical maps are most commonly used for navigation and orientation among sighted individuals and transfer an understanding of geographical space more quickly than direct experience [27, 28]. However, graphical content remains inaccessible to BLV individuals. Consequently, this population tends to rely more heavily on direct experience [29]. In the sections below I discuss the existing literature on mental models for spatial encoding and how this relates to the use of maps in the audio and haptic modalities.

2.2.1 Spatial Encoding for Blind and Low Vision Individuals

Individuals who are congenitally blind rely heavily on their auditory and haptic perception for spatial encoding [30], as they lack visual input for spatial information. Members of this population recruit and reorganize their visual cortex to support auditory and haptic processing, even during nonvisual tasks [31], thereby benefiting from improved auditory and haptic acuity. The lack of exposure to visual media renders descriptions of visual scenes meaningless to congenitally blind individuals. Likewise, visual concepts such as colour and perspective cause confusion due to a lack of reference points for these concepts [32][33]. By contrast, individuals with acquired blindness often rely heavily on visual references [34]. Interestingly, this population also benefits from improved auditory and haptic acuity following the onset of blindness [35, 36]. As this thesis focuses on designs for a mixed population of individuals with acquired and congenital blindness, it is important to acknowledge the differences in the sensory acuity across our participants.

Visual information is received and processed primarily in the visual cortex and forms an integral checkpoint in the perception of one's environment [37]. Despite the strong link between the visual cortex and the development of spatial models [38, 39, 40], at present, there is no consensus on whether the absence of the visual modality impacts spatial memory. Sighted individuals rely heavily on the visual modality when generating mental models of space, however, visio-spatial memory may also originate from other sensory modalities (i.e., sound or touch) [41, 42]. This suggests that sight may not be necessary to form mental models. However, limitations exist in the mapping of spatial information for BLV individuals, particularly when memorizing spatial layouts [43, 44].

Spatial frames of reference are often used to ground mental models and encode information for later recall. Frames of reference may be either egocentric (from the observer's perspective) or allocentric (from an external perspective) [45, 46]. Increasingly, it is clear that the size of the space we are trying to understand dictates the reference frame we employ [47, 48, 49]. Shelton and McNamara show that an egocentric frame might apply when focusing on smaller layouts such as a room [48]. Participants in their study learn the layout of objects presented in two orthogonal views (anchored to one of the objects) and are queried about the relative orientation between objects (i.e., "Imagine you are at the book and facing the wood. Point to the skillet."). Response times to these questions suggest the participants encode the information from an egocentric reference frame, as they respond faster to questions asked from the anchor viewpoint. Mou et al. [50] further demonstrate that individuals frequently employ an intrinsic reference frame, defined by the spatial arrangement. This reference frame is independent of the observer's position and orientation and consists of a primary (north-south) axis and a secondary (east-west) axis. Mou et al. [50] suggest that the identification of these axes is not purely based on egocentric cues.

The use and understanding of reference frames also differ between the sighted and blind populations [51]. Iachini et al. [52] studied the use of reference frames for small and large areas and compared spatial memory for blind and sighted people. Congenitally blind participants exhibit lower accuracy in processing allocentric information pertaining to large spaces. This was tested by asking participants to learn the configuration of objects in both small and large spaces, and then evaluating their understanding of the locations from egocentric (e.g., "which object was closest to you") and allocentric (e.g., "which object was closest to a target object") frames. The congenitally blind population respond slower to spatial judgements for large spaces. By contrast, this population responds faster than the sighted population for egocentric representations in smaller spaces, suggesting a limitation in congenitally blind people for generating allocentric mental representations in larger spaces. As visual maps are often generated using an allocentric representation [53], this limitation is important to note. For communicating allocentric spatial information, the multisensory cueing approach of Conig et al. [54] enhances performance compared to a unimodal approach. Such techniques address the limitations of BLV individuals in understanding allocentric spatial information.

2.2.2 Audio Maps

With the increased use of handheld personal devices such as smartphones with built-in speakers, audio description is among the most accessible options for BLV individuals. In this section, I introduce how audio is currently being used to represent maps.

2.2.2.1 Verbal Descriptions

The most common audio representation of spatial information is a verbal description. These descriptions are often provided with an egocentric frame of reference, such as GPS instructions for real-time navigation. For example, the Drishti system uses real-time sensor data to identify and speak out obstacles and features close to the person [55]. More recent projects using spatialized audio labels enable localization of the sources [56]. Though speech can be powerful in providing context and targeted information about objects, it remains limited in its ability to communicate complex information in a succinct manner. Common voice-based applications, such as GPS or VoiceOver, suffer from long speech descriptions.

Additionally, Tran et al. [57] demonstrate a preference for nonspeech sounds over speech for navigation. Similarly, Bjork [58] reports participant preference for natural sounds, such as the song of a bird over speech.

Despite the enhanced speech-processing capacities observed in individuals with BLV [59], verbal descriptions still exhibit limitations in terms of information transfer bandwidth when compared to nonspeech-based renderings. Moreover, studies [60] indicate that high levels of noise significantly reduce the comprehension of speech-based information.

2.2.2.2 Sonfications and Soundscapes

Nonspeech sounds act as sensory substitution tools, by mapping a visual environment into an auditory soundscape. This mapping is achieved using various sonification techniques by carefully and deliberately mapping parameters such as location, size, colour, and texture to perceivable audio parameters [61, 62].

Another way to create an auditory map is to associate audio icons with objects and triggering the playback of these icons upon their activation. Localization is achieved using signal processing techniques that emulate sources arriving at specified angles, to mimic how sound propagates to an individual's ear. BLV people generally have a superior ability to discriminate frequencies and localize sounds in the azimuthal (horizontal) plane compared with that of sighted individuals [63, 64, 65]. Extending the capabilities of audio icons, sonification techniques produce continuous sounds by mapping parameters such as volume, timbre, and pitch, to visual attributes [66]. Specifically, for spatial information, pitch is commonly used as a metaphor to communicate vertical changes [67, 68]. Several studies, successfully show that sonfications perform well with BLV individuals to communicate spatial information. For example, Josselin et al. [69] designed a system that generated sonified geographic features of a topographic map by the proportion of their dominant colours.

2.2.3 Haptic Maps

Haptic effects are used in a variety of ways to render maps, ranging in complexity from a printed raised line drawing to a fully interactive multimodal system. As the bandwidth of the haptic channel is fairly limited, designers often use a joint audio-haptic approach to better communicate map details. In this section, I introduce the three main types of perception targeted in haptic maps: tactile, vibrotactile, and force-feedback.

2.2.3.1 Tactile

Tactile maps are designed to be understood through touch, using raised lines and textures to represent different surfaces and objects. These maps are often custom-made, following guidelines such as those published by the Braille Authority [70]. Creating tactile maps requires specialized equipment, such as microcapsule paper and fusers, making the process both difficult and expensive. Despite these barriers, tactile maps are useful tools for navigation and wayfinding [71, 72, 73].

Custom 3D-printed maps have emerged as a popular substitute for conventional tactile maps. However, these maps are frequently limited in size due to material variations and technological constraints, such as the use of plastics instead of paper and the size limitations imposed by 3D-printers, respectively. Consequently, this technology is primarily used for generating smaller-scale maps, with an emphasis on providing users with an insight into specific areas, such as street crossings.

Holloway et al. [74] use participatory design methods to develop 3D-printed representations of street crossings that are tested on seven participants. User interviews reveal that the models are well received due to their portability, engagement, and professional appearance.

Giraud et al. [75] compare the performance of 3D-printed maps to traditional raised line maps, and report that 3D-printed maps foster more accurate representations of spatial layouts aiding memorization. Given these benefits, there is great enthusiasm in the community about the use of 3D-printers and maker spaces to foster the development of tactile maps, with more recent projects developing audio-haptic renderings [76, 77, 78]. Despite these benefits, 3D-printed and traditional tactile maps still require great effort to model and are generated using specialized equipment.

Another promising direction is the development of digital versions of tactile maps with refreshable graphical Braille displays. These devices generally take the form factor of a tablet with a surface covered with cells of raised pins. By selectively raising and lowering these pins, elements such as object contours can be represented.

2.2.3.2 Vibrotactile

Vibrotactile stimuli use vibrations, often from small electrical motors, to communicate information through the strength and frequency of the vibration. These motors are commonly used in digital devices, such as phones and video game controllers. Previous work takes advantage of the vibration motors in smartphones to aid BLV individuals in navigating cities [79, 80]. Alternatively, other researchers have developed their own body-worn devices with these motors. For example, Wacker et al. [81] developed a vest equipped with an array of vibrotactile motors that intensify vibration to indicate the proximity of nearby objects to the wearer.

A key aspect of manipulating haptic technologies to improve perception relies on efficacious signal transduction, and this is often dependent on the concentration of mechanoreceptors. For example, hands have a large density of these receptors, and some designs take advantage of this fact to increase the information transfer of their devices. To improve communication for deaf-blind users, Ozioko et al. [82] designed a haptic glove based on finger Braille, a system that utilizes the ring, index, and middle fingers of both hands to represent the six dots of a Braille cell. Their system uses sensors and actuators placed over the tips of the wearer's fingers to communicate short vibrations that are encoded using Braille letters. However, given that hands are used throughout the day for various important tasks, others have opted to investigate the use of vibrations on other parts of the body, such as the wrist [83] or waist [84].

2.2.3.3 Force Feedback

Force feedback devices employ a combination of force sensors and encoders to measure the user's motion and exerted force. By capturing these inputs, the device can generate corresponding motion and force feedback. These devices simulate different surfaces, objects, textures, and physical properties such as weight and viscosity, using proprioception and kinesthetic sensations [85].

Haptic-enabled computer mice, also known as force feedback mice, are used in research projects to design interactions for various types of graphics, including maps. For example, Tornil et al. [86] employ the Logitech WingMan force feedback mouse to teach participants a map of the United States of America using force guidance and texture rendering. Their results demonstrate that haptic mice could significantly improve the accessibility of digital maps for BLV individuals. However, the authors also acknowledge the need for further refinement of the technology.

As many attempt to create high-fidelity renderings, Konig et al. [54] took a different approach where they intentionally made nonrealistic haptic renderings of buildings by simplifying the information to only display salient properties. For example, they created "2.5D" structures by stacking the outlines of 3D buildings, forgoing the height information and focusing on the layout.

Magnusson and Rasmuss-Grohn demonstrate the ability of force feedback devices to communicate dynamic elements in their maps by incorporating cars and bikes [87]. Although commercially available, force feedback devices are seldom targeted toward the BLV population. As such, all the projects discussed above are conducted in a research setting using prototypes.

Despite significant efforts to create accessible maps using haptic devices, the primary issue remains the prohibitive cost for end-users. Thousands of dollars can be spent on a high-quality force feedback device or a refreshable pin array. Furthermore, map representations discussed in this section are heavily customized and optimized for specific devices, presenting a challenge in adopting and iteratively refining the designs on other systems.

Despite the demonstrated effectiveness of both force feedback and refreshable pin arrays in conveying map information, their similarities and differences remain poorly understood. Specifically, as it pertains to combining either device with the audio modality.

Chapter 3

Effects of Spatiotopic Distances on Vibrotactile Perception

Preface

This chapter describes a perceptual experiment that was conducted to investigate how postural variations affect vibrotactile frequency discrimination. As discussed in Chapter 2, numerous studies indicate that body position impacts haptic perception through the use of different reference frames. However, to the best of our knowledge, no work has studied the impact of body position on vibrotactile stimuli delivered at the wrist.

In this chapter, we detail the design of the experiment, methodology, and discuss key findings along with their relevance and applications.

Author's Contribution

Sri Gannavarapu designed and conducted the study. Yongjae Yoo advised on the methodology. Prof. Cooperstock supervised the research.

3.1 Purpose of the Study

There is a growing interest in understanding how to convey information through haptic feedback on different regions of the body; particularly, in the context of wearables, in the region of the forearm. However, there remains a lack of consensus on the IT bandwidth of multiple haptic actuators placed at different sites on the body [88, 89]. Furthermore, the consequences of a change in distance between actuators at different sites on the body are not completely understood. Recognizing these effects enables us to account for the natural shifts in body posture that occur during gestures or when typing with hand movements. These commonplace actions have the potential to change the distance between two haptic actuators positioned on various body areas.

As discussed in Chapter 2, the way we process distance through haptics may be explained through one of two frames, somatotopic (through the body) and spatiotopic (external). Currently, there is no widely accepted consensus on the use of spatiotopic reference frames for the perception of haptic stimuli. Additionally, there is a similar lack of consensus when an individual exhibits a preference for somatotopic or spatiotopic reference frames in any given environment. To address these open questions, we investigated the impact of the distance between two actuators on the ability to discriminate vibration frequencies in a general population.

Specifically, we aimed to determine the following:

- **RQ1:** What is the baseline minimum change in frequency that causes a perceptible change in sensation on the wrist?
- **RQ2:** How does the threshold from RQ1 change with varying mediolateral distance between the actuators via postural changes?

3.2 Experimental Design

We measured the JND of vibrotactile frequency on the wrist, which represents the smallest change in frequency that is perceivable [21, 90].

To calculate this value, we used the adaptive staircase methodology [21] enabling us to use the participant's response to the current stimulus to determine the subsequent stimulus. Below, we detail the design of the staircase, the calculation of the frequency JND, and the design of a calibration protocol.

3.2.1 Designing the Staircase

The adaptive staircase methodology works by providing the participant with a comparison between two stimuli. When participants correctly identify a contrast between the presented stimuli, the next pair becomes more similar. In our specific experiment, participants were tasked with comparing two stimuli and asked to indicate whether the stimuli were perceived as "same" or "different". Whenever a participant's response differed from the previous one, the staircase direction was reversed, meaning that the contrast between subsequent frequencies was increased.

This approach, wherein one of the stimuli remains constant throughout every trial, leverages the reminder paradigm. This paradigm diminishes the importance of participants needing to be highly familiar with the stimuli, as the constant stimulus serves as a memory aid, as noted in Macmillan's study on detection tasks [21]. By using the reminder paradigm, we streamlined the familiarization process, resulting in a more efficient overall experiment. Additionally, it is worth mentioning that prolonged exposure to vibration often leads to adaptation, underscoring the benefits of minimizing the experiment's duration.

In a staircase procedure, the optimal termination point is typically reached when the participant attains what is known as the point of subjective equality (PSE), indicating a situation where their likelihood of guessing correctly reaches 50 percent. However, alternative termination criteria can also be employed. For our specific experiment, we adopted a termination approach based on either detecting six reversals or conducting a total of 40 trials (as illustrated in Fig. 3.1 for a single staircase). If six reversals were not recorded within this limit, the staircase procedure would be repeated. With the adaptive staircases, there is no threshold calculator built into the procedure. However, the estimation through the averaging of the reversal points, proposed by Wetherill et al. [91], is commonly used. The resultant JND, often expressed as a Weber fraction [22], serves as a fundamental metric to gauge the relative sensitivity of individuals to changes in stimuli. The Weber fraction indicates the proportional change required in a stimulus to achieve a noticeable difference, and it is calculated using Equation 3.1.

We determined the rules for the step size, duration, and starting and reference frequencies using pilot testing. A larger step size made it difficult for the procedure to converge on the PSE, while a smaller step size resulted in excessively long session durations. To overcome this, we adopted a dynamic step size rule, which proved to be more effective. This approach involved starting with a larger step size to quickly approach the reference stimulus intensity, and decreasing the step size after each reversal. Using a dynamic step size allowed us to conduct shorter sessions while still obtaining a precise range for the PSE.

$$K = \Delta I / I \tag{3.1}$$

where:

K is the Weber fraction

ΔI is the JND

I is reference intensity

To increase the rigour of the experiment, we chose an interleaving model, where we simultaneously ran two staircases in different directions. Each trial switched between the ascending and descending staircase, as seen in Fig. 3.3. The starting frequencies and the initial direction were randomized to ensure that the participant was not habituated to the beginning portion of the staircase in between sessions. We note the concerns of response bias in answers of "same" in a staircase procedure response. Macmillan et al. [21] propose a modification to this procedure, called the ABX discrimination task, in which a third variable (X) is used as an intermediary for comparing the two variables of interest (A and B). Nevertheless, the interleaved staircase method is generally well accepted as a means of reducing such bias through its approach of alternating the presentation of stimuli from different conditions.

3.2.2 Calibrating Perceptual Intensity

We decided to use Lofelt L5 actuators for this experiment, as they are typical of the hardware found in end-user products in virtual reality and wearables. Fig. 3.2 shows the frequency response curve for the Lofelt L5 actuator used in the experiment, with a resonant frequency around 70 Hz. Hence, a calibration procedure was needed to equalize the variation in amplitude through the frequencies. We conducted a test using 200 Hz as the reference frequency, which generated 1 G of acceleration, and calibrated the amplitudes of frequencies in the range of 40–400 Hz to match the perceived intensity of vibration. Fig. 3.2 shows the correction factors as a function of the frequency.

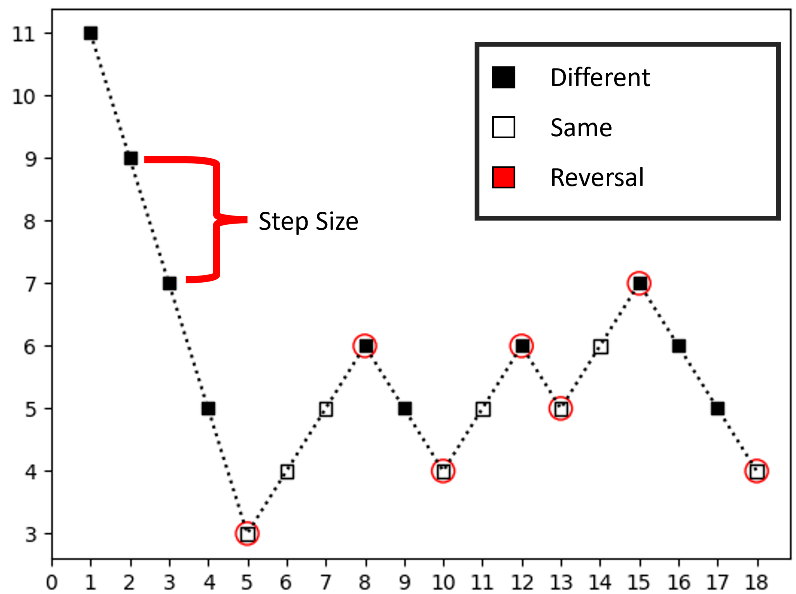


Fig. 3.1: A sample 1-up-1-down staircase with the participant responses denoted by white and black squares. Reversal points are highlighted with red circles. A same-different comparison task was presented in each trial.

As the function appears to be an inverse of the frequency response curve, this validates our use of these multiplication factors. We used these data as our initial values for the amplitude adjustment in the final experiment. To determine these multiplication factors, we recorded subjective intensity at 12 frequencies, as shown in Fig. 3.5. To account for frequencies in between the discrete points, we conducted a linear interpolation to obtain the intensity for the values using the closest lower and upper calibration values.

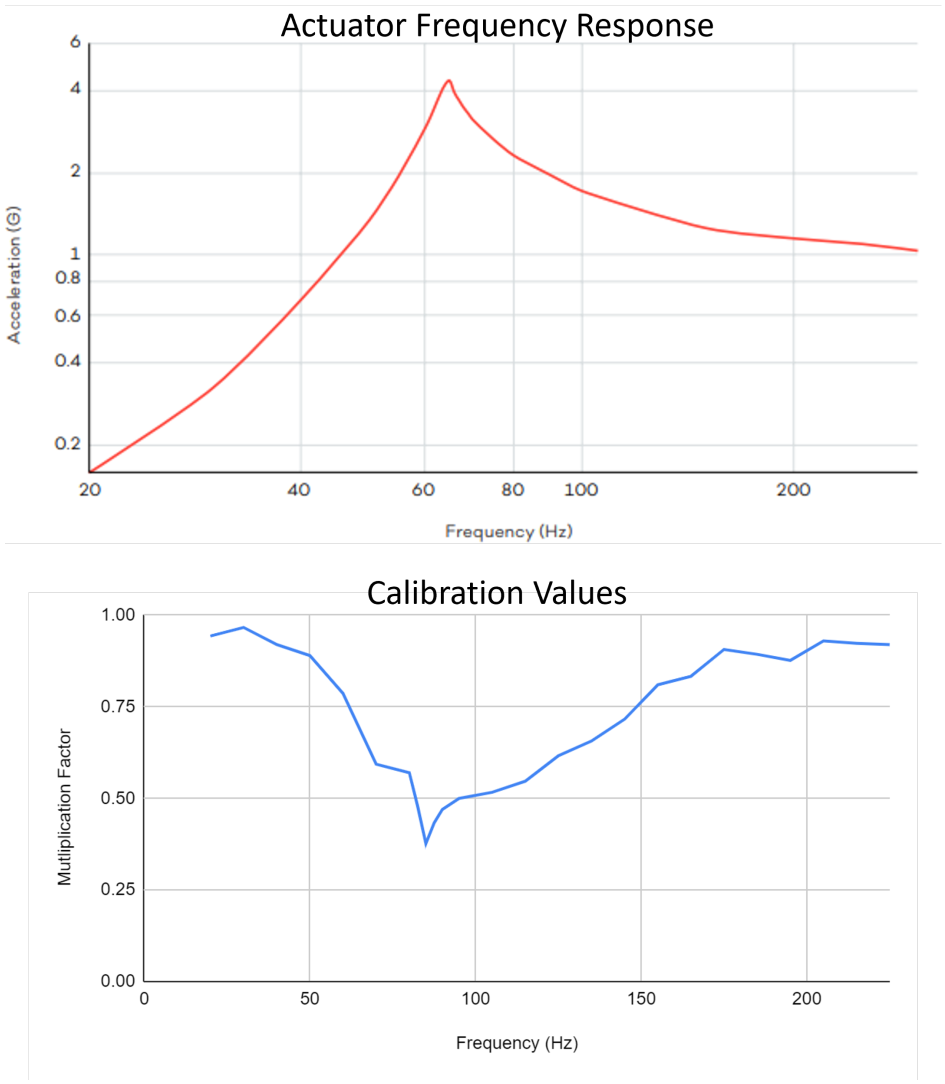


Fig. 3.2: The process of perceptual intensity calibration using the frequency response curve as a reference. The top graph displays the frequency response for the Lofelt L5 actuators, with a peak intensity of 70 Hz. The bottom graph shows the multiplication factors used to calibrate the perceived intensity of each frequency in reference to 200 Hz.

3.3 Methods

3.3.1 Participants

Ten participants (5 male, 5 female; 19–36 years in age; mean = 24.9; SD = 6.7 years) were recruited from the McGill University community. All participants were healthy and reported no deficits in tactile perception. Eight reported prior experience with participating in user studies; however, none had experience with either haptics or psychophysics studies, specifically. This experiment was approved by the McGill Research Ethics Board (file #20-10-041) and participants were compensated CAN$10/hour.

3.3.2 Apparatus

Participants were seated with both arms outstretched over the apron of a wooden table with their palms facing up. Vibrotactile stimuli were delivered through two 5 mm linear resonant actuators (Lofelt, L5), attached to the palmar side of the participant's wrists with Velcro bands. Each band was placed approximately 5 cm proximally from the wrist toward the forearm. The actuators were centred on the band, and participants were instructed to fasten each band such that the device was tightly coupled to the arm yet comfortable to wear. The actuators were driven by an amplifier board connected to a laptop computer running a Python script to control the frequency and amplitude of the stimuli as well as the presentation order for vibrations. Participants wore earplugs and noise-cancelling headphones playing pink noise to mask the sound of the vibrations throughout the experiment. We included a short test procedure to ensure that participants perceived vibrations at both ends of the detectable frequency range on the forearm.

3.3.3 Procedures

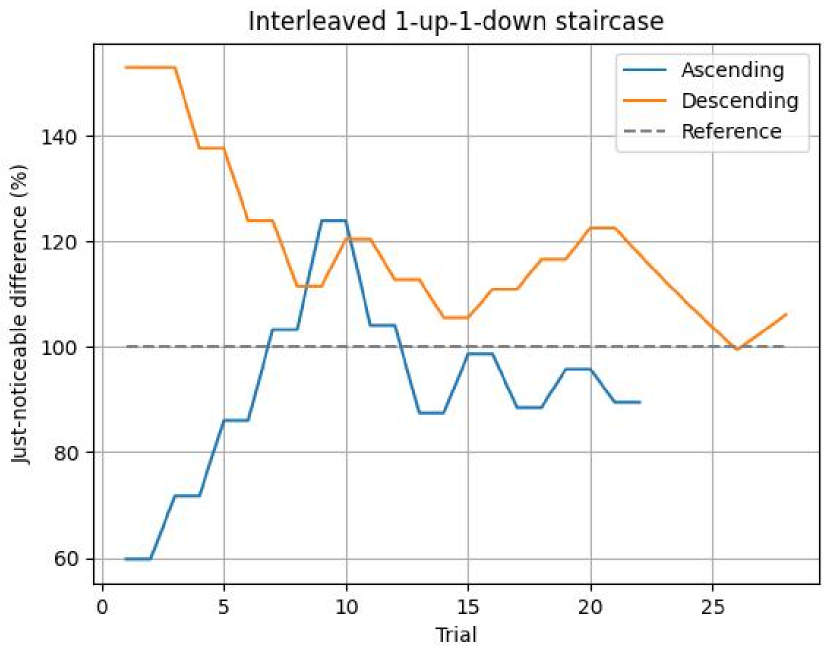


Fig. 3.3: A single interleaved staircase for one posture.

This experiment was conducted over two consecutive days, taking approximately 45 minutes for the first session and one hour for the second.

All stimuli were delivered as sinusoidal waves with an 800 ms duration, as per Berglund et al. [92]. We included an initial familiarization stage in which participants were presented with a simple interface giving them the option to select one of two stimuli (see Fig. 3.4). The participants chose to deliver these stimuli individually or in sequence. They were also provided with on-screen sliders to manipulate both frequency and amplitude according to their preference. The range of the frequency was limited to those sufficiently perceivable on the forearm (40–200 Hz) [93]. Participants were told to take as long as they needed to become

Fig. 3.4: The user interface for familiarization. The participants used the sliders to adjust the values for frequency and amplitude. The radio buttons were used to select the order of presentation for the vibrations.

comfortable with the vibrations, but typically only spent five minutes during the first session.

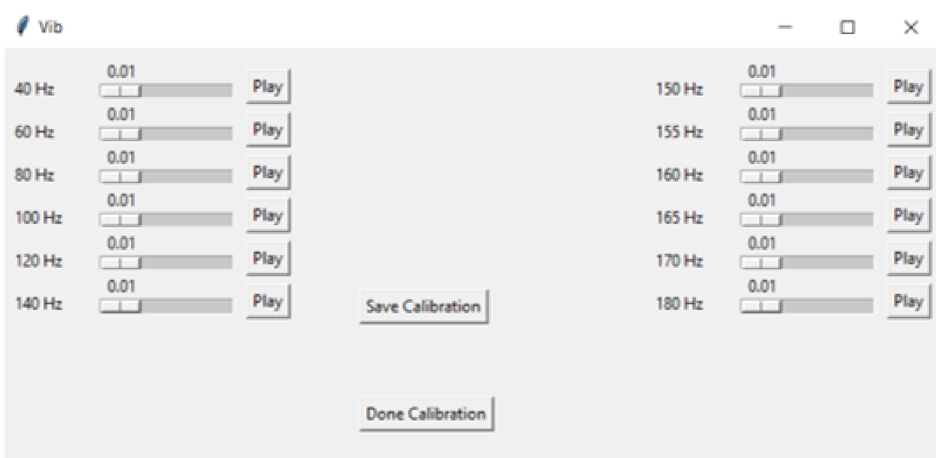


Fig. 3.5: The user interface for the perceptual intensity calibration. The participants used the slider to adjust the intensity of each frequency in comparison to a reference.

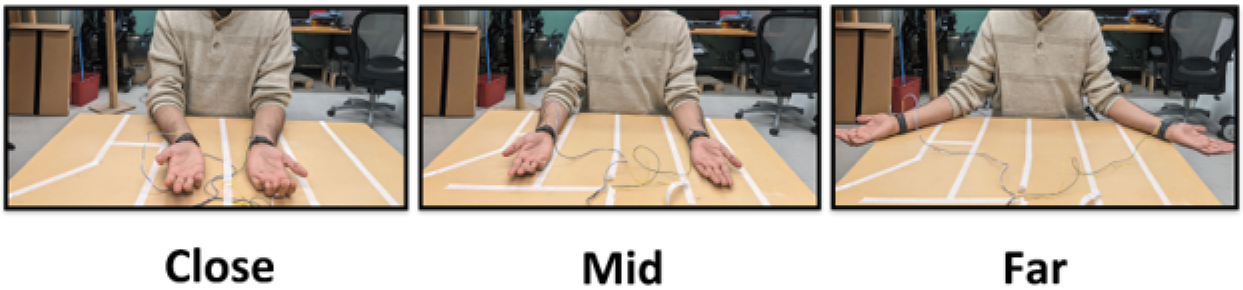


Fig. 3.6: A depiction of the three postures used in the experiment, with vibrotactile actuators placed above the wrist.

Next, the participants completed the calibration procedure described in Section 3.2.2 to minimize the effect of frequency on perceived amplitude. Following this calibration, the main experiment required participants to adopt and maintain one of the three postures, as shown in Fig. 3.6.

While maintaining each specified posture, participants were presented with consecutive pairs of vibrations, and asked to indicate whether the two were perceived as the same or different. Stimuli were presented according to an interleaved 1-up 1-down adaptive staircase, as specified in Section 3.2.1. Each staircase consisted of at least 20 trials and terminated once six reversals were recorded or 40 trials were completed. Following the completion of the staircase, participants were instructed to switch to the next posture and the experiment was repeated. The completion of one full staircase for each posture marked the end of the first day. On the second day, participants repeated the adaptive staircase procedure twice for each posture. Three staircases were conducted in total for each posture, and the ordering of the postures was counterbalanced. For each staircase, every reversal point was recorded and the last six reversals were averaged to determine the frequency JND. Participants were provided with an explanation of the purpose of the study following the completion of the experiment and post-experiment questionnaire.

3.4 Data Analysis

Outliers were identified using the Interquartile Range (IQR) Method in R [94]. This method calculates the range within which the middle 50% of the data lies and flags data points beyond a threshold, determined by multiplying the IQR by a factor of 1.5. Data points that fell outside this threshold were deemed outliers and were excluded from the subsequent analysis. Our analysis detected a single outlier across all three groups, as depicted in Fig. 3.7. Consequently, the data corresponding to this particular participant were excluded from the subsequent analysis.

We chose the LME4 library in R [95] to conduct a linear mixed-effects regression (LMER) analysis due to several advantages over traditional repeated measures analysis of variance (ANOVA). First, LMER allows for the modelling of random effects, such as participant-level variability, common in experiments where participants are exposed to multiple conditions. This accounts for potential heterogeneity among participants, leading to a more accurate estimation of fixed effects. Additionally, as our experiment involved measuring pseudo-replicates by repeating each condition on the same participant, LMER properly accounts for this non-independence of data and provides reliable estimates of fixed effects.

3.5 Results and Discussion

In a post-experiment questionnaire, no participant indicated difficulty understanding the experimental procedure or discomfort from the actuators during the experiments. They were also asked to report the conditions under which they felt they were able to best distinguish

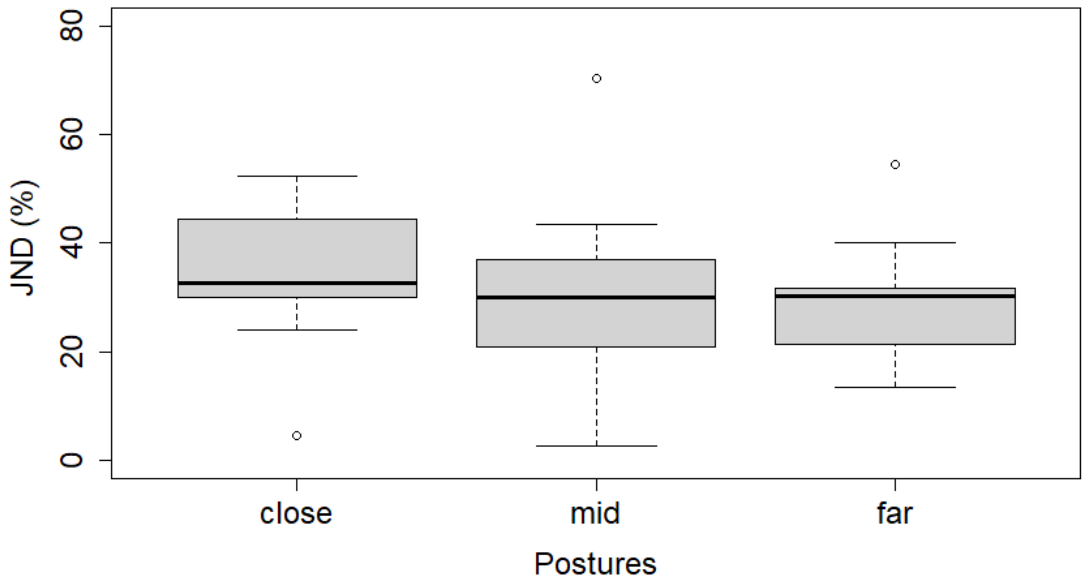


Fig. 3.7: Average JNDs for each posture with outliers denoted by circles.

between frequency pairs. Although there were mixed reports, most reported no differences. A couple of participants noted that the “far" condition appeared to be the most straightforward for discerning vibration frequencies, while one participant specifically pointed out that the“mid" condition was the most straightforward for them in this regard. No correlation was found between participant demographic details and their discrimination of frequencies during the experiment.

The JND percentages are presented in Table 3.1, reported as Weber fractions, for each participant and condition. In particular, the far condition exhibited the lowest average JND at 25.3%, indicating a higher sensitivity to frequency changes. This was closely followed by the mid condition with a JND of 29.0%, and then the close condition with a JND of 31.6%.

Table 3.1: Average JND for each participant and condition.

Participant	Far	Mid	Close
1	30.1	33.3	30.5
2	54.6	80.3	62.3
3	30.0	19.8	4.5
4	36.7	42.4	37.6
5	13.7	12.7	25.3
6	31.7	32.9	52.3
7	18.3	15.4	18.9
8	20.3	29.1	24.0
9	21.3	26.5	29.6
10	30.7	28.5	31.1

A lower JND value is indicative of a higher acuity in perceiving changes in frequency. The variances between the far and close ($F = 0.90$; $df = 18$; $p = 0.43$), the far and mid ($F = 0.52$; $df = 18$; $p = 0.66$), and the close and mid ($F = 0.60$; $df = 18$; $p = 0.82$) group pairs were not significant. Despite the lack of statistical significance across the conditions, we observed a trend, in line with Shore et al. [20], of a linear correlation between JND and distance (Fig. 3.7).

The location of the actuation (wrist) may also be partly responsible for the inconclusive results. As noted earlier, any effect of postural changes on vibrotactile haptic perception has only been observed on the fingers, which have a high density of mechanoreceptors [16]. Germann et al. [96] revealed that the fingertips contain more than 20 Pacinian corpuscles nodules, whereas only a few were observed in the wrist. Our results suggest that a location with a high density of mechanoreceptors, such as the fingers, may be necessary to observe the effect of postural changes on haptic perception. Additionally, increasing the distance of the most extreme condition and only comparing the two extreme conditions might also allow us to observe the effect.

In our opinion, accessing a range of postural changes with actuators on different sites, such as the forearms, wrists, and fingers would be worthwhile. This would allow us to examine the role of mechanoreceptor density as it pertains to the effect of postural changes. Of note, the wrist contains a lower density of mechanoreceptors compared with that of fingers. If found that postural changes require a higher density to elicit a significant effect on haptic perception, then wearable devices such as smartwatches may benefit from incorporating frequency characteristics without extensively considering the impact of a user's daily activity.

Chapter 4

Accessible Maps

Preface

Based on early user interviews, we discovered that digital maps presented a significant challenge for the BLV community. To overcome this accessibility issue, we designed renderings that use three communication mechanisms: audio, force feedback haptics, and tactile pin arrays. Each of these options has its unique features, cost points, and accessibility levels. Audio was expected to be the easiest solution to implement as headphones are widely available to provide audio playback. In contrast, both haptic options require more specialized equipment.

In this chapter, we outline our efforts toward rendering digital maps. In Section 4.1, we present an initial effort focused on unimodal rendering. In Section 4.2, these renderings were presented to BLV in a qualitative user study. In Section 4.3, we present an expanded solution, using lessons from the user study to incorporate multiple modalities.

Author's Contribution

The force-feedback renderings were designed and implemented by Sri Gannavarapu, while the initial renderings for the refreshable pin array were developed by Yongaje Yoo. Florian Grond was responsible for designing and implementing the audio renderings. The multimodal study and thematic analysis were conducted collaboratively by Cyan Kuo, Corentin Conan, Rayan Isran, and Sri Gannavarapu. Rayan Isran contributed to the analysis of transcripts and the generation of themes, while Corentin Conan was responsible for translating the study and contributing to the analysis of transcripts and the generation of themes. Cyan Kuo provided guidance and advice on the experimental design and analysis methods. In addition, the prototyping environment discussed in Section 4.3 was co-designed by Rayan Isran and Sri Gannavarapu. Prof. Cooperstock supervised the research.

The Internet Multimodal Access to Graphical Exploration (IMAGE) web extension allows users to request additional information when encountering a graphic, and provides them with multiple renderings on the audio and haptic modalities. The architecture follows a server-client model consisting of three main components [97]:

- Preprocessors—a set of machine learning (ML)-based services that extract data from a graphic.
- Handlers—a set of microservices that transform the raw data into a usable form.
- Renderers—client-side programs to create experiences that describe the graphic.

The IMAGE extension currently offers renderings for photos, charts and embedded Google maps, all of which rely on a combination of ML-based feature extraction and open-source application programming interface (API)s to provide users with text and audio renderings.

As previously discussed in Chapter 2, maps are particularly information-dense, in order to provide users with both a general overview of an area while allowing for a further detailed explanation of regions of interest [98]. Our BLV participants expressed excitement during early interviews and conversations when discussing the ability of our audio renderings to provide a general layout of a map. Although audio renderings are useful, audio-haptic renderings often provide more detail and reinforce the information conveyed [99]. Encouraged by the initial participant response, we developed renderings that offered a comprehensive overview and increased accessibility to more detailed information by users. This user-feedback was obtained through semi-structured interviews, discussed further in Section 4.2.

As IMAGE is intended for use on desktop and laptop computers, we chose to focus on applications that do not require the user's real-time location and orientation. Instead, we aimed to design renderings that convey information about the features of the map, enabling users to learn about an area or plan a route to navigate.

Two primary features of street maps are the streets themselves and the points of interest (POI) located along them. Although transit maps serve a different purpose from street maps, they often simplify the representation of these features to aid in wayfinding and navigation. Therefore, we chose to use transit maps as an exploratory space for the design of our renderings. By studying the principles and design techniques used in transit maps, we can gain insight into how to communicate information effectively to users.

At present, there is no easily accessible open-source data for transit map coordinates and information about POI, namely transit stations, that can be used for the design of a preprocessor. Current open-source offerings, such as OpenStreetMap (OSM), are often sparse and rely on user-input data, which might not be up-to-date or complete. This makes it challenging to extract and interpret the necessary data accurately. It also limits the ability to design effective and reliable renderings for transit maps. As a result, we have decided to use public images of transit maps to identify the intersections and line contours. This approach allowed us to extract the information we needed for the rendering design. We developed a *Processing* sketch to annotate lines and POIs on top of a joint photographic experts group (JPEG) image. An example of a marked-up map is shown in Fig. 4.1.

4.1 Design of Renderings

4.1.1 Sonification Approach

To create the audio renderings for our transit maps, we used a combination of nonspeech audio and text-to-speech (TTS) announcements. This approach allowed us to provide users with an

Fig. 4.1: A sample output from the markup tool. The normalized coordinates of the lines and points were used for generating audio and audio-haptic renderings. Each coloured line depicts a transit line with interpolated coordinates saved to a JSON file.

understanding of the features of the map, including key POIs such as intersections.

As previously stated in Chapter 2, BLV individuals have superior auditory discrimination in the azimuthal plane, and therefore we expected this population to respond well to audio content that was processed to move, virtually, along this plane. Pitch was used as a metaphor to communicate changes in the north-south direction. The rendering iterated through each transit line individually, sonifying the shape of the line using a combination of spatialized audio to convey movements from west to east, and changes in pitch represented movements from north to south. At each intersection of transit lines, the sonification stopped and announced the location with a ping sound preceding the speech announcement.

The nonspeech audio was developed using a technique known as granular synthesis. This method involves creating complex sounds by stitching together small samples of sounds, called grains [100]. We synthesized the audio using the Ambisonic Toolkit in SuperCollider 3[1]. The generated samples were output as binaural audio, using the technique and measurements proposed by Bernschuetz et al. [101]. The mapping was expressed in musical instrument digital interface (MIDI) notes, ranging from 70 (∼460 Hz) to 118 (∼7450 Hz). Subsequently, the MIDI values were converted to a filter's center frequency, which shaped a broadband noise signal during the sound synthesis process. In addition to nonspeech audio, we incorporated spatialized TTS announcements to provide users with more specific and detailed information about the transit map.

4.1.2 Force Feedback Design

We employed two haptic devices to convey map information: a proprioceptive force feedback (FF) device and a refreshable tactile pin array, in conjunction with audio cues.

The renderings on the proprioceptive (FF) device provided users with a sense of physical interaction and response, enabling them to perceive the forces of virtual objects, such as walls, as if they were real. In contrast, the tactile refreshable pin array offered a more familiar experience, akin to existing nondigital solutions, such as raised line diagrams.

To create map renderings for the proprioceptive device, we used the 2DIY (a planar robotic device from Haply Robotics, which has two degrees of freedom (DOF)) [11]. The device includes an Arduino microcontroller board, which interfaced with sensors and actuators to con-

[1] https://supercollider.github.io/

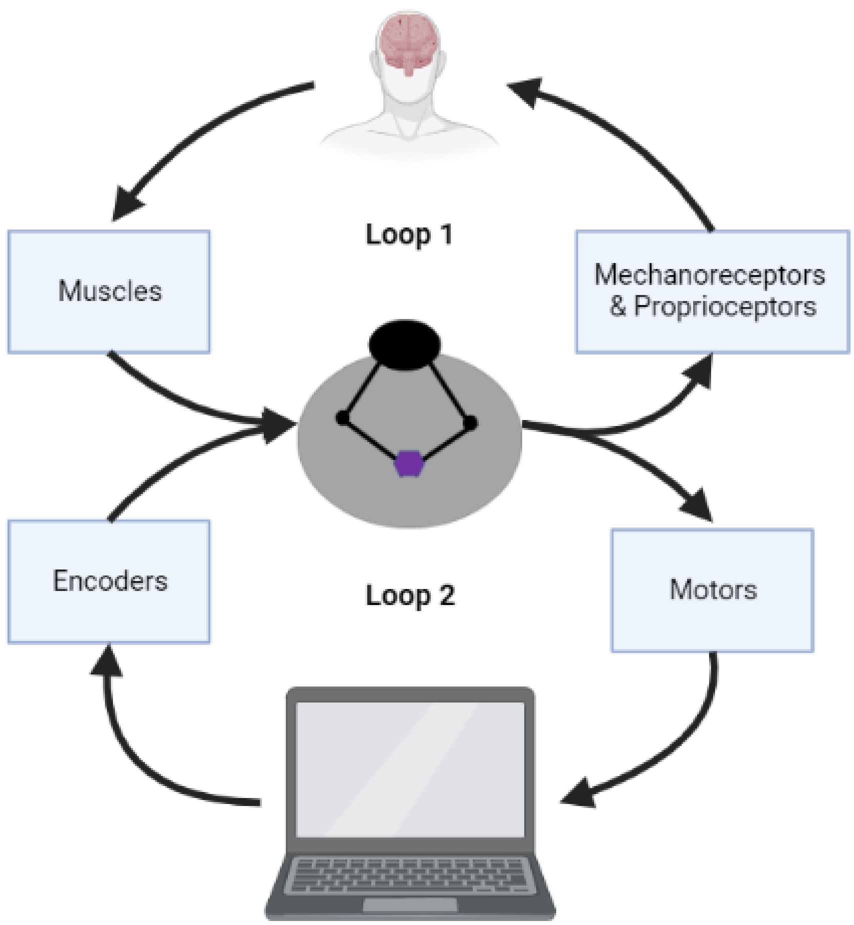


Fig. 4.2: The haptic loop. This figure is based on previously published work [1].

trol the movements.

We maintained a haptic loop to ensure effective control of the 2DIY. Briefly, a haptic loop is a fundamental concept in the design of FF haptic interactions and it is comprised of two interconnected feedback loops (see Fig. 4.2). The first loop involves the sensory system receiving information from the environment through mechanoreceptors and proprioceptors, allowing users to perceive and respond to external stimuli. The processed information is sent to the human motor system, which generates appropriate motor commands to produce a response, such as movement or change in grip force. The second loop describes how the haptic interface senses changes in the system based on user input, and responds using actuators, such as motors, to change the forces provided. This feedback loop occurs rapidly and continuously, enabling users to have a seamless and immersive experience while interacting with the environment [1].

Using the haptic loop, we generate two main types of interaction: one where the user drives the action, and the other where the user takes a more passive role. We employed this rendering approach to create two distinct modes on 2DIY: guidance and exploration.

4.1.2.1 Guided Tour

This operating mode of the 2DIY involved the user grasping the end-effector (EE) of the device, which was then guided by the 2DIY along the path of a transit line. We implemented a proportional derivative integral (PID) controller in *Processing* to regulate and maintain the path of the 2DIY device. This type of controller adjusts the output position of the EE by using feedback from the 2DIY's sensors to minimize the error between the desired and actual position of the EE. We dedicated a thread in our *Processing* sketch to maintain the haptic loop, ensuring a refresh rate of 1 kHz [102]. Each transit line was discretized into an array of coordinates, which were used as set points for the PID control.

Given the generally low fidelity of the 2DIY device, its control was prone to positional errors, which can be attributed to the implementation of a low-resolution mechanical actuation system and a constrained workspace.

To enhance the performance, we upsampled the coordinates between the set points. This technique reduces the force exerted at any given time by decreasing the distance between the points. As a result, the controller's response was observed to be more consistent.

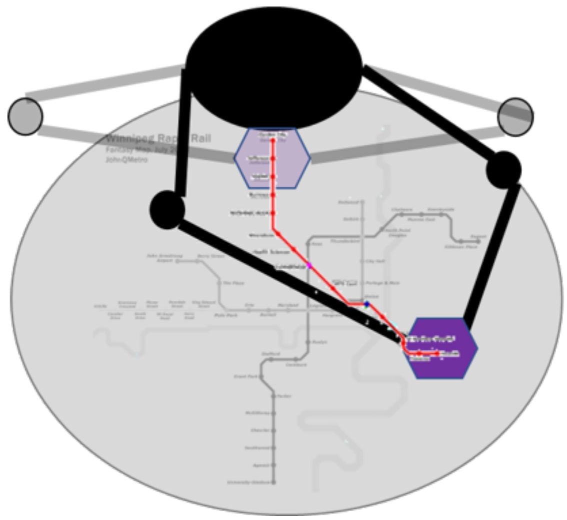


Fig. 4.3: The guidance mode of the 2DIY for a single line. The highlighted red line denotes the path taken during guidance. The placement of the device EE at the beginning and end of the guidance is shown.

4.1.2.2 Free Exploration

In the *Free Exploration* mode, users were able to freely move the 2DIY's EE and experience forces as they interacted with virtual objects. This rendering was generated using the Fisica physics library[2] in *Processing* to create a virtual environment that dynamically responds to user movements. In this virtual world, we used the transit map data, generated through our markup tool, to create force walls on either side of each transit line, as shown in Fig. 4.4. These walls acted as guide rails that restricted user exploration. We hypothesized that the presence of the walls provided perceptual and physical constraints, grounding user exploration and minimizing any distractions or disorientations. The intersections were further highlighted by introducing increased resistance at these specific locations. This effect was achieved through the use of force damping, which is a mechanism that opposes the motion of the EE by exerting a force

[2] http://www.ricardmarxer.com/fisica/

proportional to the EE's velocity. This damping effect was generated by introducing variations in viscosity within the Fisica physics engine. Increasing the viscosity value induced greater opposition to motion, correlated with the velocity magnitude. Alternative rendering strategies, such as magnetic forces, were initially explored for the free exploration mode. However, they were ultimately disregarded due to two main reasons. Firstly, the lower fidelity of the 2DIY made precise motor control challenging. Secondly, technical obstacles arose in synchronizing the magnetic effects with the audio announcements.

Table 4.1: A summary of the force feedback effects employed with the Haply 2DIY for map renderings.

Mode	Description of Haptic Effects
Free Exploration	Force walls on either side of the transit lines Force damping at intersections
Guided Tour	PID Control of the EE.

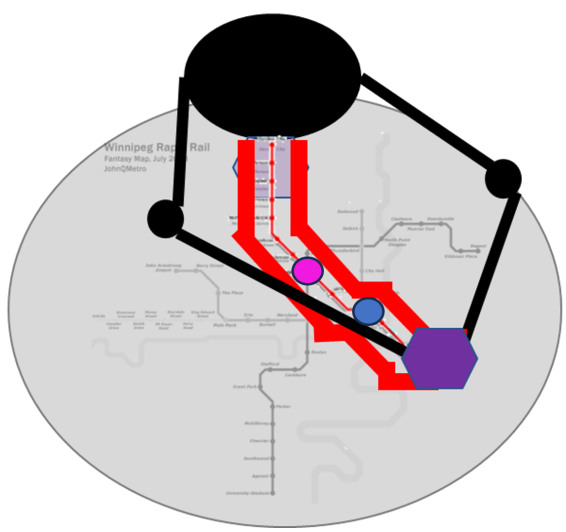


Fig. 4.4: The 2DIY with a map overlay on the surface. The thick red lines illustrate the force walls generated in the virtual world. The purple and blue dots indicate intersection points which produce a haptic damping effect upon interaction.

4.1.2.3 Multimodal Approach

Preliminary user testing indicated that the exploration mode on its own was disorienting and confusing. The single point of contact made it challenging for users to quickly find areas of interest. To address this issue, we proposed starting with a guided tour to provide users with a more general overview of the map's contents and to provide context for their exploration.

We initially planned to synchronize the guided tour with the spatialized audio samples to reinforce the information using multimodality. However, during development, we realized that our current control strategy did not precisely guarantee a set pace for the haptic guidance, making this synchronization technically infeasible. Therefore, we designed a serialized multimodal interaction that started with the user experiencing the audio, followed by a guided tour, and finally allowed them to explore areas of interest. To avoid lengthy rendering times and given noticeable technical issues with prolonged guidance, we restricted this process to one transit line at a time. The interaction flow is depicted in Fig. 4.5.

4.1.3 Refreshable Pin Array Design

We used the Dot Pad, a refreshable pin array manufactured by Dot Incorporated, as the tactile haptic device in our project.[3] This device has a resolution of 40 × 60 pins, which are raised or lowered via Bluetooth. We employed a web application hosted on a local server to develop and deploy renderings to the Dot Pad. This application translated the output of the markup tool into a bitmap that was subsequently transmitted to the Dot Pad.

To avoid overloading the user with information, we decomposed the transit map into layers that correspond to individual transit lines. This approach allowed users to navigate between

[3] https://pad.dotincorp.com/

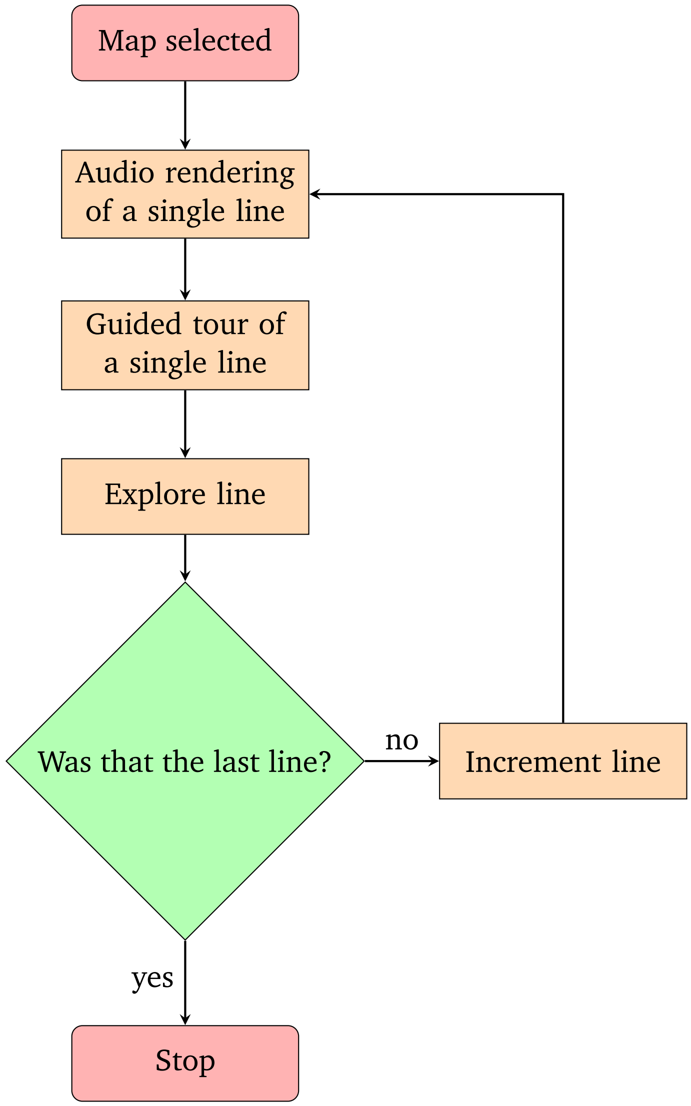


Fig. 4.5: Flowchart showing the serialized interaction for a single transit line.

different layers, or to view the entire map, depending on their needs. By providing this flexibility, users can better understand how the different elements within each layer relate to one another. For an example of this, please refer to Fig. 4.6.

Due to the Dot Pad's lack of position-tracking capabilities, dynamic interactions were not

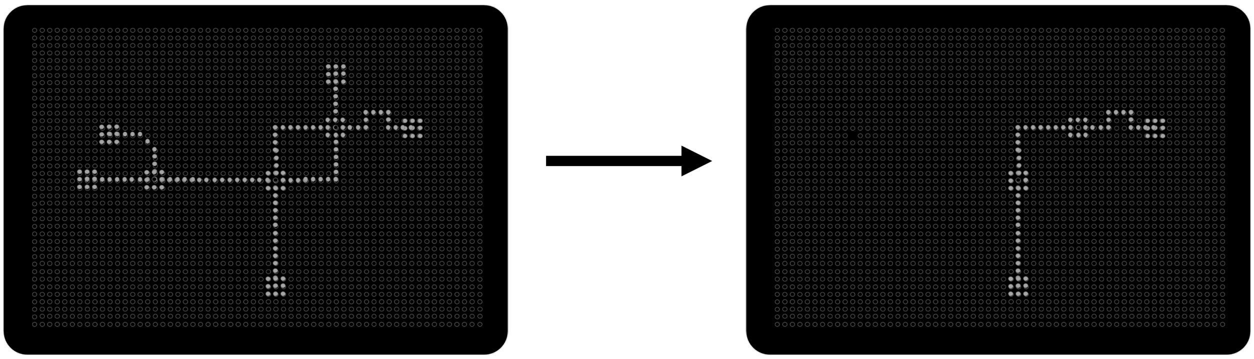


Fig. 4.6: The image demonstrates the layering technique used for transit maps on the Dot Pad. The left side of the image shows all the layers, whereas the right side displays a single layer in isolation. Note that the isolated layer is intentionally not centred, as this maintains its relative position in comparison to the other transit lines.

possible. To overcome this limitation, we employed a serialized approach to multimodality, similar to the one used in our 2DIY renderings. This involved presenting the audio rendering before allowing the user to explore the tactile version. It is worth noting that the Dot Pad does not refresh with weight on the pins, meaning that users needed to pause their exploration between layers.

4.2 Multimodal Study

We conducted a user study with BLV participants and provided them with our designed renderings. Our research aimed to answer the following questions:

> **RQ1:** "What are the differences between the way users perceive renderings of maps designed using audio, FF, and refreshable pin arrays?"
>
> **RQ2:** "How does the use of multimodality enhance audio-only renderings?"

We conducted semi-structured interviews consisting of a framework of questions related to user experience with no leading queries, to capture both qualitative feedback and rendering-specific metrics such as the detection of intersections.

The IMAGE project focuses on the design of multimodal renderings of photos, maps, and charts. Specifically, the study discussed in this thesis examined the effectiveness of the haptic feedback system for the comprehension of maps.

4.2.1 Participants

Nine participants were recruited through one of three channels: the Regroupement des aveugles et amblyopes du Montréal métropolitain (RAAMM), the Canadian Council of the Blind, and word of mouth. None of the participants reported any hearing impairments.

The participant pool displayed a spectrum of blindness. Four participants reported congenital blindness and five reported either partial or legal blindness. All participants were aged over 50, with a mean age of 64 years. One of the participants was unable to complete the experiment, and therefore their data and feedback were not used in the analysis. Since the RAAMM recruits were francophone and bilingual, we ran the experiment in English and French, based on the participant's preferred language. The McGill research ethics board (REB) approved this experiment under REB #21-10-031 and participants were compensated CAN$10/hour.

4.2.2 Setup

The experiment took place at McGill University (in the Shared Reality Lab) and the Institut Nazareth et Louis-Braille (INLB) in Longueuil. The participants were seated with the audio and haptic devices placed in front of them, as seen in Fig. 4.7a. A WASP 9907 action camera was used to record video of the user's hands, and the participant responses were recorded using a laptop. The audio samples were delivered using a pair of open-back headphones, as well, two haptic devices were used in this study: a 2DIY and a Dot Pad.

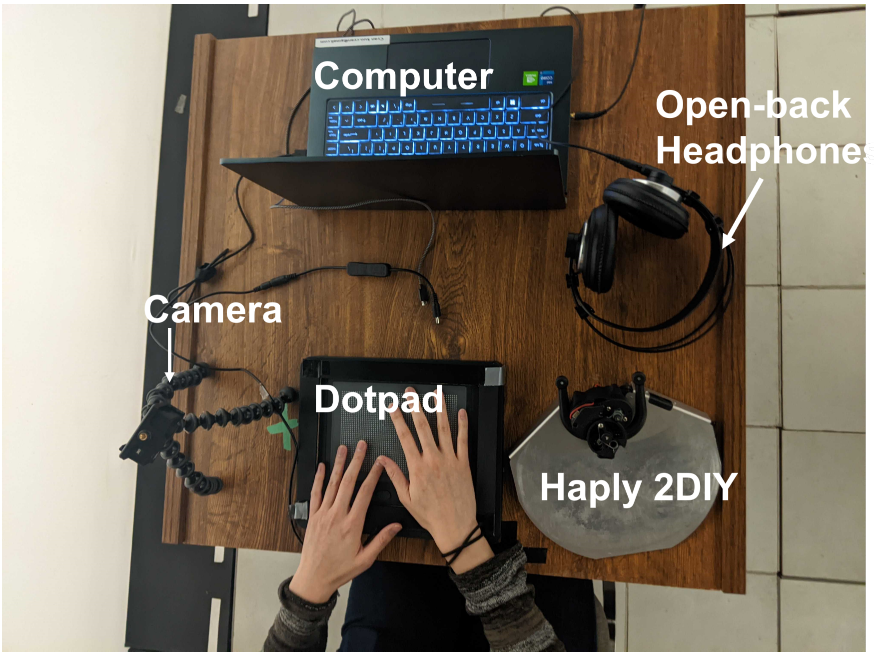
4.2.3 Procedure

Prior to the study, we introduced each participant to the project and an overview of any associated risks. The experimenter read aloud the ethics consent form, and participants were then given a pretest questionnaire to collect demographic data and determine their familiarity with the technology used in the experiment. Specifically, the questions targeted their experience with maps to allow researchers to better understand the current behaviours and challenges of the participants.

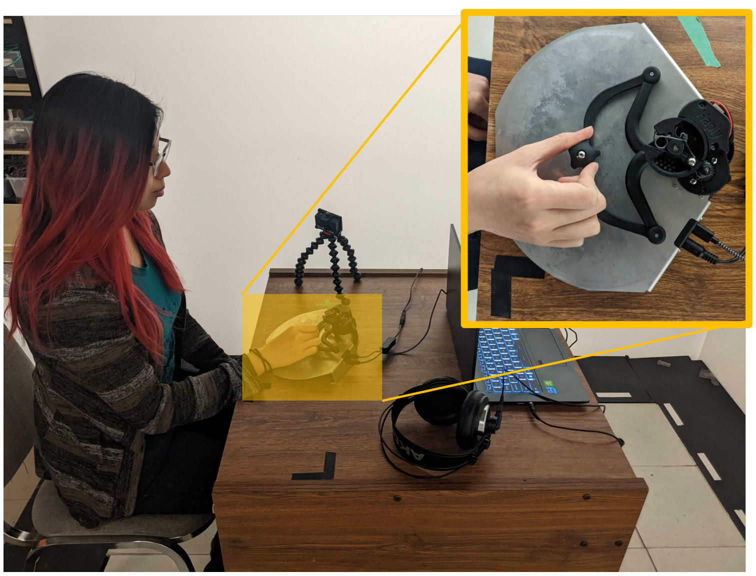
Participants were then assigned an order of device and stimuli presentation, using pseudo-randomization to counteract learning effects. The experiment was divided into two blocks: FF maps, tactile maps.

For each block, the audio rendering was played first followed by the haptic compliment. Within a block, the participants interacted with two to three renderings and were given breaks between block transitions. Each block lasted about 20 minutes, and the entire study was completed in less than two hours. To control for the general familiarity of location, we chose transit maps from geographically distant areas as the stimuli—Baltimore and Winnipeg. The Baltimore metro system is relatively simple, consisting of only two lines, whereas, the Winnipeg system is more complex, featuring four lines. This deliberate selection allowed us to capture maps of varying levels of complexity.

The participants always experienced an audio rendering to start; they were then asked to describe what they heard. The interviewer (researcher) was instructed to ask the following questions after each rendering:



(a) The experimental setup showing all the materials used.



(b) A participant grasping the Haply 2DIY by pinching the end-effector using their thumb and index finger.

Fig. 4.7: The experimental setup showing the 2DIY device and its grip, along with all the necessary materials used in the experiment. The grip of the 2DIY device is demonstrated in detail to provide a clear understanding of how the participants held and interacted with the device.

- Did you notice any trends or patterns?
- Can you describe the shape of the trend?
- Did you hear or identify any intersection points?
- Can you recall how many intersections you heard about/found?

The path to these questions depended on how the participants answered the initial broad questions. For example, if a participant had any confusion with the trend being conveyed in the rendering, the interviewer focused on this source of confusion and tried to elicit a description of what they were experiencing. In instances where participants provided responses with insufficient detail, the interviewer proactively sought additional information by asking for elaboration or clarification to ensure a comprehensive understanding of their responses.

At the start of a new haptic block, an introductory example was provided for the participants to get used to the respective devices. For the 2DIY, we presented an example to train the participants on the proper way to grasp the EE and understand how much force to exert on it when being guided (Fig. 4.7b). In this example, the participants were asked to place their hands on the 2DIY EE and a *Processing* sketch was responsible for moving the EE in a circular motion. For the Dot Pad, participants were given the opportunity to explore the device and familiarize themselves with its tactile interface. To communicate the difference between raised and unraised dots, a small sample of dots was raised for participants to feel. Participants were instructed to keep their hands on the side of the device when refreshing the Dot Pad to avoid interfering with the raising mechanism.

Audio stimuli were presented across all blocks, and the same questions and interactions were used to introduce and evaluate the audio-haptic devices. In the case of the 2DIY device, the guidance mode was experienced first, where the user was guided through the contour of a single line. The participants were allowed to replay this rendering as many times as they liked. Next, the exploration mode was loaded, and the user could freely move along the line. Between each mode (audio, guidance, and exploration), the participants were asked about their experience, as described above.

At the end of the final block, participants were questioned about their preferences between the devices and renderings. Finally, the participants were given the opportunity to share their general thoughts and re-experience any specific renderings.

4.2.4 Data Analysis

Audio and video recordings were collected from each semi-structured interview. The camera only recorded the hands of the participants to provide context about the discussion of their interactions. Transcripts were generated using Microsoft Word and manually reviewed by the research team to identify the speakers.

We conducted a thematic analysis to gain a deeper understanding of the individual interactions with each rendering style. To ensure a rigorous analysis, we followed a reflexive approach informed by guidelines proposed by Braun et al. [103].

To mitigate interpretation bias, three team members were involved in the analysis. Each member independently reviewed and coded the transcripts using *Taguette* [104] (an open-source qualitative analysis software) to associate tags with the emerging categories. Next, we discussed the independently identified categories as a group to further refine our coding scheme to ensure accuracy, comprehensiveness, and fidelity to the data.

Following the finalization of our transcript organization, we grouped similar codes together and provided each group with a descriptive label. This enabled us to identify major themes as well as sub-themes, providing a higher-level understanding of the patterns observed in our data.

Throughout the analysis, we engaged in ongoing reflection and discussion to refine our codes and categories. We were mindful of the iterative nature of our analysis and ensured that each phase informed the next. Finally, we presented our findings to the group and discussed the themes identified and the evidence supporting them. This allowed us to arrive at a shared understanding of the patterns in the data and to draw conclusions about the phenomena under investigation. This process is summarized in Fig. 4.8.

As a secondary analysis, the participant's performance in each example was evaluated by assigning binary scores to the following tasks: intersection identification and line shape description. To obtain a point for intersection identification, participants had to correctly identify the number of intersections in a line. To obtain a point for the verbal description of the shape, the participant had to properly identify the starting location, ending location and general trends. The scoring for the shape description was divided into two categories: north-south and east-west. This distinction was motivated by previous research [64] that demonstrates that discrimination of changes along the azimuth is superior in 3D sound. We aimed to determine if this effect manifested within our current renderings.

4.2.5 Results

4.2.5.1 Current Behaviours

Maps provide extensive visual information that BLV individuals often struggle to perceive as completely as individuals with an intact ability to receive visual cues. The findings discussed in the section were obtained through guided discussions during the semi-structured interviews.

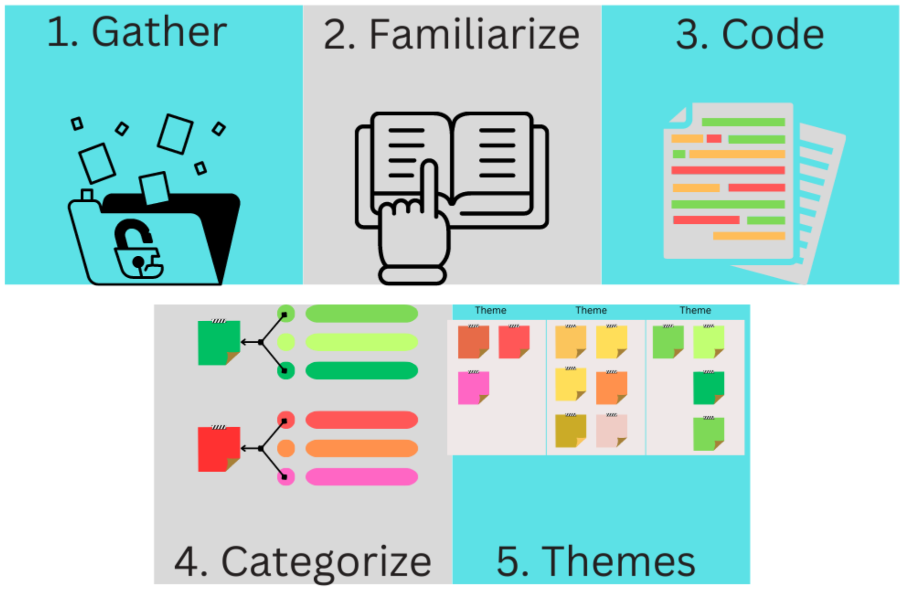


Fig. 4.8: Phases of the thematic analysis process, including data preparation, initial coding, and theme development. The process involved iterative cycles of coding and theme refinement to identify meaningful patterns and insights from the data.

Accessibility issues often prevented participants from using maps in their daily lives, with many indicating that online maps, such as embedded Google Maps, when used on a computer, are incompatible with screen readers. As a result, most of the participants relied primarily on their GPS software, such as Google Maps or Apple Maps, on their mobile devices. Our results showed that there was a diversity in the tools that were used within our participant pool. Some participants reported using alternative methods, such as tactile maps and sonifications, with an individual regularly using 3D-printed tactile cards to navigate the local metro system. However, the current offerings for accessible desktop applications for maps remain limited.

4.2.5.2 Accuracy Scores and User Preferences

To track how well the participants understood the shape of each transit line and the location of the intersections in that line, we scored their verbal responses. To obtain a point for intersection identification, participants had to correctly identify the number of intersections on a line. To obtain a point for the verbal description of the shape, the participant had to correctly identify the starting location, the ending location, and general trends. These scores were recorded as binary values based on criteria detailed in Section 4.2.4. Averages for each condition are reported in Table 4.3. All participants were able to clearly describe the shape of the transit line using the Dot Pad, and the majority were able to identify all intersection points.

Users with the Dot Pad had markedly better performance than users with both the audio-only and 2DIY conditions. However, we noticed that the 2DIY guidance was generally better than the audio-only mode (Table 4.3). This was especially true in the north-south direction. The 2DIY exploration had the poorest performance of all conditions (50% for both north-south and east-west) for trend understanding.Based on the user input, we conjecture that the poor performance can be attributed to both the lower fidelity of the device and the increased complexity of the rendering design. More detailed discussion on the specific user feedback regarding these factors is provided in the following sections.

On average, participants in the audio-only condition recalled all intersections 57.1% of the time, while demonstrating markedly better performance in both haptic conditions.

They identified all intersections with the 2DIY exploration and Dot Pad 93.0% and 93.7% of the time, respectively.

To compare the performance of three pairs of devices (Audio-2DIY, Audio-Dot Pad, and 2DIY-Dot Pad) in terms of their scores, we conducted pairwise t-tests. Alpha (α), known as

Table 4.2: Demographic characteristics of participants, map usage summary, and preferred mode of use.

ID	Visual Acuity	Age (Years)	Current Maps Usage	Preferred Condition	Preferred FF mode
1	Congenitally Blind	70	Does not use digital maps, uses GPS	Dot Pad	Guidance
2	Partially Sighted	69	Uses digital maps by zooming in to be able to see with contrast	Dot Pad	Guidance
3	Visually Impaired	70	Trained on how to invoke GPS. set destination and that's it	Dot Pad	Guidance
4	Noncongenitally Blind	51	Uses GPS for point-point navigation and relies on learning through direct experience	Dot Pad	Guidance
5	Congenitally Blind	65	Has knowledge of their city and doesn't need to rely as much on real-time GPS	Dot Pad	Exploration
6	Noncongenitally Blind	61	Aids the driver with navigation using GPS	Dot Pad	Exploration
7	Congenitally Blind	58	Uses 3D mock-ups and GPS for navigation	Dot Pad	Guidance
8	Noncongenitally Blind	64	Mostly relies on iPhone GPS but has used tactile cards for metro before	Dot Pad	Guidance
9	Congenitally Blind	65	N/A (did not complete the study)	N/A	N/A

Table 4.3: Average accuracy percentages by mode. These values represent the average based on binary scores.

	Audio-only	**2DIY (guidance)**	**2DIY (exploration)**	**Dot Pad**
Intersection	57.1	N/A	93.0	93.7
East-West	83.4	85.7	50.0	100
North-South	78.5	83.7	50.0	100

the significance level, represents the likelihood of erroneously rejecting the null hypothesis when it is true. The p-value gauges the probability of obtaining a more extreme result than the one observed in the experiment. When the p-value exceeds alpha, the null hypothesis is accepted and the finding is not significant. For the study in this thesis, $\alpha = 0.05$ was used. As a result, any p-value < 0.05 was considered statistically significant. The results of the tests are presented in Table 4.4 and Table 4.5.

We found that none of the multimodal conditions reached statistical significance compared to the audio-only condition for trend identification. Similarly, neither of the haptic conditions showed statistical significance for trend identification. However, for intersection identification, both of the multimodal conditions showed significance compared to the audio-only condition. Moreover, the intersection detection scores exhibited a statistically significant increase on the Dot Pad compared to the 2DIY.

Table 4.4: degrees of freedom (df), t-stat, and p-values for intersection identification between pairs of conditions, conducted at $\alpha = 0.05$. Statistically significant values are indicated by *.

Multimodal Category	df	t-stat	p-value
Dot Pad vs Audio	13	1.4	0.0036*
2DIY(exploration) vs Audio	13	1.7	0.0040*
2DIY (exploration) vs Dot Pad	13	1.7	0.0040*

After the completion of each block, participants were asked to indicate their preference among the options of audio, Dot Pad, and FF. All eight participants indicated a preference for the Dot Pad renderings. Additionally, participants were asked to express their preference between FF exploration and guidance. Six participants (out of eight) indicated a preference for the FF guidance mode, while two participants preferred the exploration mode. These results, along with additional details, are summarized in Table 4.2 which presents the preferences expressed by the participants.

Table 4.5: df, t-stat, and p-values for trend comparisons between pairs of conditions, conducted at $\alpha = 0.05$, which refers to the significance level in hypothesis testing. Statistically significant values are indicated by *.

Multimodal Category	Direction	df	t-stat	p-value
Dot Pad vs Audio	East-west	13	1.5	0.54
Dot Pad vs Audio	North-south	13	1.8	0.60
2DIY (guidance) vs Audio	East-west	13	2.2	0.25
2DIY (guidance) vs Audio	North-south	13	1.6	0.99
2DIY (exploration) vs Audio	East-west	13	1.2	0.33
2DIY (exploration) vs Audio	North-south	13	-1.6	0.31
2DIY (guidance) vs Dot Pad	East-west	13	1.8	0.84
2DIY (guidance) vs Dot Pad	North-south	13	1.5	0.87
2DIY (exploration) vs Dot Pad	East-west	13	4.2	0.020*
2DIY (exploration) vs Dot Pad	North-south	13	3.6	0.010*

4.2.5.3 Thematic Analysis

Our thematic analysis revealed six themes that together provided insights into how well our rendering strategies were able to communicate spatial information. These themes are elaborated below, with a discussion of the connecting threads shown in Fig. 4.9.

When discussing these themes, user-driven interactions (audio and haptic guidance) are referred to as *active user modes*, and non-user-driven (haptic exploration on the 2DIY and Dot Pad) as *passive user modes*.

4.2.5.3.1 Theme 1: Initial and lasting impressions

The users' initial impressions of each condition can significantly impact their comfort and willingness to engage, both during and after experiencing the rendering. During our observations of participants' interactions, we identified one or two key factors for each device that had a

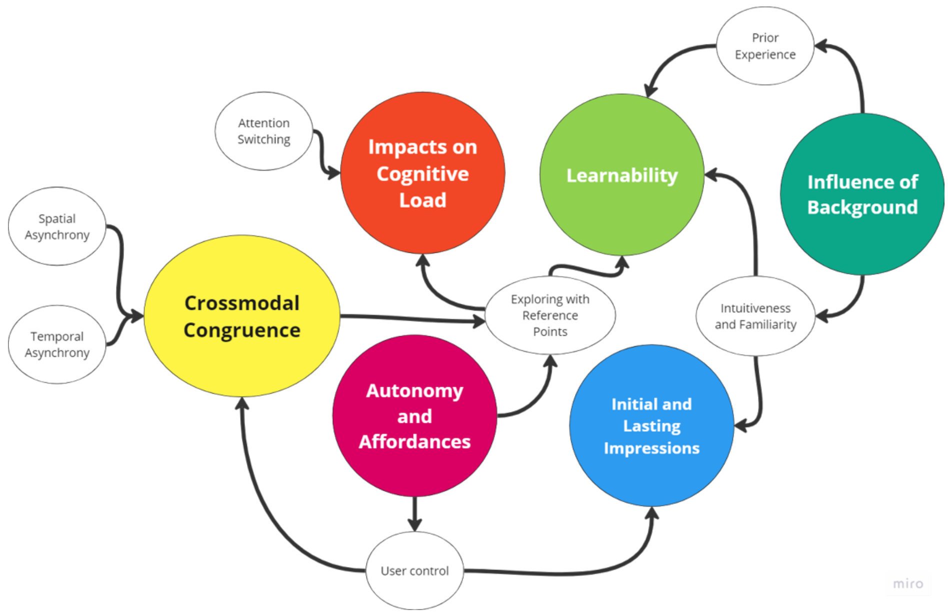


Fig. 4.9: Results of the thematic analysis, with emerging themes represented by circles. Each major theme is marked with a different colour to distinguish them. The smaller ovals depict the connecting ideas that link the themes together. The arrows indicate the direction of cause or outcome. The connecting threads highlight the inter-relatedness of the themes and the direction of their respective arrowheads illustrate the granularity of the theme.

notable impact on their initial impressions.

Regarding audio renderings, we found that the sound tonality had the greatest impact on users. Instead of a pure tone, we used a broadband sound that elicited mixed emotions. Some participants noted that the audio quality made it difficult to follow sonification trends. Specifically, the complexity of the synthesized audio samples appeared to induce confusion in some participants, as they exhibited an expectation for a more simplified pure-tone presentation.

For instance, P2 described their first experience with sonification as follows:

> P2: "Well, of course, the sound really wasn't much, didn't really do anything for me, it's just, but as you know, it's his words. I understood the sound behind the

> sound afterwards meant nothing."

In the case of P2, they understood the meaning conveyed by the TTS, however, this participant found it difficult to decipher the sonification, which they attributed to the tonality of the sound being used.

Other participants tended to associate real-life sounds with the audio provided, which acted as distractors. For instance, P2, P3, and P8 referred to some of the sonifications as "wind sounds".

> P2: "...it sounds like winds at the beginning ..."

> P3: "...I have no idea. It's like a sound of wind that starts on my left ear ..."

> P8: "Ça ressemblait je dirais un peu à du vent ou une soufflerie." [translated to: "I would say it resembled wind or a wind tunnel."]

P6, who self reported as being experienced with sonification, had very positive things to say about the tonality, remarking that it was pleasant to listen to for extended periods of time. This participant self-identified the trade-offs of using pleasant sounds instead of a pure tone when designing sonfications.

> P6: "And those sounds you're using are good sounds, they are relaxing, they are almost cozy, I would think.... With that kind of sound you're using, it is neater to the ear, but it would prevent me from evaluating such a thing, but then do I need to?"

Regarding the FF renderings, the perceptible impact of motion smoothness emerged as the most significant factor influencing both their initial and lasting impressions. Unpleasant interactions with the FF renderings were often due to rapid changes in directions or oscillations experienced at the EE. Improving the fidelity of the device could potentially improve stability by mitigating factors like friction, joint backlash, or enhancing onboard computational capabilities [105]. However, without testing identical force feedback renderings on a higher fidelity device, it remains uncertain how the rendering design directly influences the discomfort reported by the user. Even variations in smoothness between different examples of the same mode had a large impact on the participants' overall impression of the device and mode.For example, P6 initially experienced a guided mode example where there were more sharp turns in the path. Subsequent to encountering a more refined motion profile, their observations were as follows:

> P6: "The other one was a hormonized teenager, this one is a little more grown-up. [this example] didn't surprise me with jerks of all sorts…it engages more slowly, so it doesn't startle you. So jerkiness is very important. I felt it was smooth."

Regarding the Dot Pad renderings, we found that the participants' initial and lasting impressions were influenced primarily by two factors: familiarity with the device's affordances and the static nature of the renderings.

Participants tended to be more comfortable exploring on the Dot Pad, whereas with the 2DIY they were more cautious and tended to prematurely end their exploration. As most participants had some familiarity with Braille and Braille displays, we hypothesized that they understood the affordances of the Dot Pad more easily, giving them greater confidence in exploring on the device. Given that the Dot Pad does not have constantly moving parts, it provides a more relaxed exploration experience.

Theme Summary: This theme revealed that a couple of factors were responsible for forming initial and lasting impressions with interactions in the three conditions. For audio, the main factor was tonality, which is subjective. If the participants found the tone to be pleasant, they could focus more easily on interpreting the data conveyed through the sonification. In addition, associations with real-life sounds were a source of distraction for some users. For the 2DIY, the smoothness of the EE's motion was the greatest contributor to the participants' impressions. Participants often reported unpleasant interactions with this device, likely due to rough motion profiles as a consequence of the low resolution of the actuators. Finally, in the case of the Dot Pad, users reported greater confidence in their explorations on the device, which in turn left them with pleasant initial and lasting impressions. This confidence can be attributed to two factors: familiarity with the device's affordances and the static nature of the device. Examining and focusing on these factors shows the potential to enhance rendering effectiveness. In addition, mitigating the identified negative aspects may contribute to a more favourable user adoption rate for such renderings.

4.2.5.3.2 Theme 2: Crossmodal congruence

This theme was related to perceived mismatches between audio and haptic experiences. Comparing the two passive modes (audio and haptic guidance) we observed that some users found it challenging to reconcile due to perceived differences in the duration of the experiences. This was most evident when P2 asked to experience both the audio rendering and the haptic guidance together.

> P2: "I heard the beginning, but that doesn't help. . . That was sort of hard to interpret. Like the audio and this doesn't mesh."

While the audio consistently maintained pacing, haptic guidance demonstrated variations influenced by factors such as grip and load on the EE. Despite participants being primed with the audio, they consistently noted discrepancies. Additionally, it was observed that this mismatch served as a source of distraction for the participants.

Several participants also noted disparities in the magnitudes communicated between the audio and haptic renderings. Most accounts indicated that the audio renderings led them to underestimate the extent of changes along the y-axis (the north-south direction). Whereas, participants reported that the haptic feedback gave them the impression of reduced movement in the north-south direction compared to the actual extent. The participant anecdotes below exemplify this observation.

> P6: "I was trying to equate my speed of movement on your map with the audio. And to me the audio lingered on, going north more than my actual speed would have me go north on this thing [pointing to Dot Pad]"

> P4: "I could almost, got my finger as it's doing it now [tracing tactile while listening to audio]; I can kind of get an idea of what it's trying to represent now. If I weren't able to feel it, I think I would be able to now have a better idea of the sound was trying to convey . . . it made it clearer. A clear understanding of what was going across 'cause it had so many ups and downs. So it made it clearer on the trends. "

> P2: "I mean it changed maybe a little bit, but here [pointing to the 2DIY] definitely is more of an up or down, whereas in the audio you don't get that big huge difference up or down."

However, one participant found instead that during haptic guidance the movement in the vertical direction seemed subdued when compared to the audio.

The discrepancies in the magnitudes communicated between the two modalities may also have implications for the design of future guidance systems.

Theme Summary: This theme illustrated that there may be a significant discrepancy in the information conveyed by different modalities, which can affect the overall user experience. The notable instances of this discrepancy were evident in the pacing differences between audio and haptic guidance, along with spatial asynchrony between the haptic and audio modalities. Considering these sources of incongruence is important when combining audio and haptic modalities.

4.2.5.3.3 Theme 3: Cognitive load

Our observations indicated that the renderings had a varying impact on the cognitive load imposed on the participants. These impacts appeared to be the result of increased reliance on short-term memory, as well as distractions in the renderings that forced users to switch their attention from their primary task.

Exploration and short-term memory

Allowing users to explore a haptic feedback system may reduce their dependence on short-term memory, as they actively engage with the system and retrace areas of interest at their own pace. This was observed in both the Dot Pad and the 2DIY interactions.

By giving users control over the pace and order in which information was being conveyed, they were less likely to feel overwhelmed by the information density. It is worth noting that, we observed a significant difference between the exploration experiences on the 2DIY and Dot Pad. The Dot Pad allows multi-touch interactions, whereas the 2DIY is constrained to a single point of contact. This limitation makes it difficult for the user to ground their exploration using a reference point on the 2DIY, unlike the Dot Pad.

> P5: "When you're dealing with these types of things, with the guided step with the fingers [2DIY] and you're only able to get a sense of where you are in that particular point, you don't necessarily have a good reference of where you were through it 'cause you're having to keep a mental representation of the whole thing. Whereas, the tactile thing is always in front of you at all times and it's a great reference the whole time."

> P4: "Oh, like I said. I can go back anytime I want. I can explore the whole, get out a whole ballpark image of what's around. Whereas the [2DIY] I have to use my memory to remember where I came from."

Interrupts increase cognitive load

Interruptions during an interaction increased the cognitive load on the participants by forcing them to switch their attention. This presented differently between the audio renderings and the 2DIY. With the audio renderings, the announcement of intersections would interrupt the specialization. The disruption of the participant's concentration made it more difficult for them to remember the trends in the sonification preceding the announcement. In the 2DIY exploration mode, the use of force walls appeared to disrupt exploration and seemed to cause disorientation.

> P2: "I can't tell the difference. I can't tell the difference because when she [referring to speech audio] interrupts to talk."

This was further exacerbated by the fact that users did not ground their exploration by using a reference point.

> P4: "Well, because the walls that I met took away my memory of where I came from. I know it's instant, but meeting the wall made me forget if I was higher or lower."

The last source of interruptions that we identified was with force guidance, where the way users handled the EE either distracted from or interrupted the trace. For example, if the user placed more weight on the EE or gripped it more tightly this would increase the lag during guidance. This lag can be significant enough for users to be distracted, such as P7.

> P7: "I got hacked by the lag like the thing was doing this in my hands was after 1/4 of a second delay or whatever and I was destroyed. Part of my concentration. . . . I think I'm obstructing its movement unintentionally because I don't know like my wrist is not smooth enough for it. When you don't know what you're going to see, I unintentionally resist."

Theme Summary: This theme highlighted the effect of interruptions on the understanding of contours represented through our renderings. Interruptions, whether they are active or passive, can have a severe negative impact on the learning process. This is because shape learning requires significant cognitive resources, and interruptions can disrupt the mental model being developed. Multiple iterations through the renderings may be necessary to create a complete and accurate mental model, and interruptions can hinder this process. As well, a heavy reliance on short-term memory was found to also hinder the participant's learning abilities. To mitigate this issue, the incorporation of memory aids, such as tactile reference points, presents a promising avenue for resolution. Taken together, both of these factors should be considered when iterating on our designs.

4.2.5.3.4 Theme 4: Perceived affordances and user autonomy

Users expressed ideas about their expectations and a general understanding of the devices they were asked to interact with while participating in the study. They also shared their personal input on what they felt would improve their experience. We discuss these ideas in more detail below.

For the 2DIY, users' understanding of the device's affordances increased as they interacted with more renderings. Contrastingly, for the Dot Pad and audio they generally understood the behaviour and limitations of both devices with a single example.

A common missing affordance noticed by multiple participants was the lack of control over the speed of the audio renderings.

> P2: "Since it's very fast - At what point it's up or down? It's very hard to tell."

When asked more about the playback control for these interactions, P1 said the following:

> P1: "And if [the speed] is user adjustable, I'd say you know. Then you could. You go several times over it and decide what level you're getting more information. If it's a fixed value, then you don't have that control."

Similarly, participants encountered a comparable challenge concerning the lack of control over the pacing of the 2DIY guidance mode. Furthermore, those who raised the issue of lack of control over playback speed associated negative emotions, such as frustration or confusion, with this overlooked aspect.

Expanding on this theme, it became evident that autonomy, the ability to exercise control over one's actions, held significant value. More specifically, the absence of control over the pace of interaction appears to originate from a deficiency in user autonomy throughout the interaction.

> P6:“ you have to be a slave to the thing. I much prefer the Braille display if there’s a question to be asked.”

> P4: “made you feel as if you could explore rather than being forced where to go. Maybe I just don’t like being pushed around.”

Regarding the Dot Pad, the ability to use multiple points of contact and control the speed of the interaction resulted in a unanimously positive experience. For example, P8 said the following about the affordances they noticed with the Dot Pad:

> P8: “Je dirais probablement la première forme avec les points. Je trouve que ce serait plus rapide probablement avec la première. Parce que c’est toujours le même. Et tu peux le toucher rapidement, tu peux le toucher plus longtemps, tu peux le prendre monter. Tu peux t’attarder à certains endroits et moins à d’autres. Comme quand c’est des lignes droites et quand ça change de direction . . . , je pense que les points c’est plus simple. Et on peut faire avec les deux mains aussi.” [translated to: “I would probably say the first device with the dots . . . Because you have it right away, if you can do it once, twice. And you can touch it fast, you can touch it longer, you can take it up. You can linger in some places and less in others. Like when it’s straight lines and when it changes direction. . . , I think the points are easier. And you can do it with both hands too.”

Theme Summary: The ability to use multiple points of contact during exploration on the Dot Pad significantly improved participants’ confidence and understanding of the global shapes in the rendering. This affordance not only provided a quick overview of the data but also allowed users to query more detailed shapes and spatial information. On the contrary, both the 2DIY and the audio modes failed to offer this degree of exploratory freedom. Users instead expressed a desire for increased control over the passive user modes.

4.2.5.3.5 Theme 5: Learnability

The learning curves for each condition varied significantly between the participants. Of the three modalities, audio renderings posed the greatest challenge to comprehension, probably attributed to the unfamiliarity with the sounds used. The pace and density of information in the audio renderings appeared to be overwhelming. Users extracted information more meaningfully when they paid attention to a single dimension.

> P1: " I would think you know over time there'd be more of a learning curve as you got used to it. I think you know you become more acutely aware of what all that was meaning. So you know after a trial period, I think you would become more efficient with it."

Despite the fact that the 2DIY was equally unfamiliar to numerous participants, they found it easier to interpret the haptic renderings. This was evident from the elevated shape understanding scores observed for both the haptic devices, in contrast to those of the audio-only condition.

Theme summary: The disparity in learning curves between modalities highlighted the importance of targeted training. It also demonstrated the potential for using a multimodal rendering to teach unimodal renderings, notably, using an audio-haptic tutorial to train users on the audio-only renderings.

4.2.5.3.6 Theme 6: Influence of background

Interviews revealed that the prior experiences related to spatialized audio and usage of Braille in users may greatly impact their attitudes toward each rendering style. Some of these factors include the level of blindness, the age at which they acquired blindness, and their comfort with technology. Furthermore, the user's background can also affect how much training they require to understand the renderings. Although this theme shares similarities with *Theme 5: Learnability*, it remains distinct as it focuses on how easily a user can understand and use a rendering.

In particular, participants' prior experience with sonifications or any form of musical training can profoundly influence their ability to comprehend the audio renderings.

> P4: "And it's a sound language. I'm sure artists or musicians would probably do really well with this. Because their ears are fine-tuned in a certain way, I wonder. But learning the acoustic aspect of it, there's a learning curve involved with that, whereas less so with the tactile device."

Both P6 and P8 had significant experience with sonifications (rated themselves at level 5 in the pre-test questionnaire). They also demonstrated greater confidence in their interpretations of the audio renderings. Furthermore, they paid attention to the more subtle qualities of the rendering such as the granularity of the spatialization. For instance, P6 was the only one to notice the spatialization of the speech in the renderings.

> P6: "What I find particularly neat is that the description follows the panning of the sound, and that's just neat"

P8 was able to discern fairly small changes in the azimuthal plane and constantly clarified the details of the rendering strategy.

> P8: "What would be the difference earwise between going straight up north and going up north towards east…If it goes straight up north or if it goes up north at ten or fifteen degrees…is it that important?"

Similarly, we noticed that Braille literacy significantly informed how well the participants understood the tactile renderings. P1 self-reported to be highly Braille literate and was able to easily identify the difference between intersection points (hollow square) and terminus points (filled-in square) without being trained on the rendering scheme. Whereas P4, who was not Braille literate, was initially confused with the intersection points.

> P4: "Although I heard the sound coming up and down and pitching, it feels as if it's in this little section here [points to 1st and 2nd intersection area]. It's gone up and down a fair amount, but it's confusing because they're so close together."

With clarification from the interviewer, they (P4) were able to quickly grasp the differences and make sense of the rest of the rendering. As we witnessed with P1 and P4, sufficient training with the renderings should address any stark differences in comfort based on the existing knowledge.

Theme Summary: This theme highlighted the importance of considering the variability of previous experiences when designing products and experiences. Specifically, designing training modules that cater to a diverse population with varying tactile and cognitive abilities may help mitigate issues encountered with differences in exposure to Braille and sonification. Additionally, user-configurable parameters should be investigated to address the variance identified in this theme.

4.2.6 Discussion

In this study, our objective was to investigate the effectiveness of different rendering modes, including audio-only, audio with FF, and audio with tactile perception, to convey spatial map information to users. We explored the advantages of each rendering strategy for communicating map information.

Additionally, we aimed to understand how multimodal renderings compared to unimodal renderings. Through careful observation and analysis of participants' interactions with these different renderings, we sought to gain a comprehensive understanding of the benefits and limitations of each modality and device. Our objective was to improve our rendering strategies and gain direct user feedback to improve the design of harmonized multimodal interactions that effectively communicate spatial information.

Based on our observations, we identified several takeaways that could be categorized into the following:

1. Improvements and modifications to the initial rendering strategies
2. Improving communication of rendering strategies through targeted training
3. Considerations for the design of audio-haptic renderings

4.2.6.1 Improvements and modifications to the initial rendering strategies

The rendering strategies used in this study were designed to convey map information on using three modalities: audio, FF, and tactile haptics (refreshable pin array). The audio renderings used sonification to trace the path of a transit line from west to east, stopping to announce intersection points. The FF renderings also followed the contour of a transit line, while simultaneously allowing users to explore at their own pace. The Dot Pad renderings provided users with greater control and flexibility in exploration, as they could use both hands to explore the map at their own pace.

When examining how to improve these renderings, Don Norman's concept of visceral design provides a useful framework for understanding how initial and lasting impressions are formed. Visceral design describes how users form quick, often subconscious, impressions of a product that can have a lasting impact on their overall experience [106].

Negative feedback on the audio quality, such as tonality and volume, had a significant impact on the user's initial impression of the rendering and highlighted the importance of investigating certain sonifications that elicited this outcome. Improvement in audio quality seems to be an important aspect to consider in the design stage to enable improved visceral design. Specifically, parameters such as clarity, volume, and appropriateness remain important for sustained positive interactions between the user and rendering. Similarly, with the 2DIY, ensuring stability is essential for creating positive initial and lasting impressions. This can be achieved by modifications to the control algorithm or through mechanical changes. By focusing on stability, designers can create a more seamless user experience that will leave a lasting positive impression on the user. Looking at the performance scores to gauge user understanding of renderings of each modality, we observed that the Dot Pad was able to best convey both the shape of the transit line and facilitate the identification of intersection points.

For intersection identification, both the Dot Pad and audio yielded similar results. Since participants were provided with audio before their interactions with the Dot Pad, this result was unsurprising as the audio-only condition performed well. However, we also observed that participants had increased difficulty understanding the shape of the transit line in the audio-only condition as compared to the Dot Pad and the FF guidance. From the user feedback, we conjecture that one way to improve user understanding is by bolstering the user control for parameters such as speed and level of information. P6 referred to existing audio solutions such as Job Access with Speech (JAWS) (a popular screen reader), where the user can control and customize many aspects of the audio, including the playback rate. On a related note, several users reported experiencing problems with speech interruptions during the announcement of intersections. This is likely due to interruptions that cause a rapid shift in attention. Previous research has shown that sudden shifts in attention can have an adverse effect on the encoding and retrieval of a stimulus in memory [107]. One potential solution for addressing the feedback regarding speech interruptions during intersection announcements is to offer users the ability to customize the level of detail in the initial audio rendering. By providing this option, users can select the level of detail that best suits their needs and preferences, which may help reduce interruptions and improve comprehension. Additionally, offering a high degree of customization may be particularly important given the reported challenges with understanding the renderings, as it may allow users to tailor the experience to their individual needs and improve overall usability.

Our observations of the participants revealed that they specifically experienced difficulties with the north-south spatialization, which was mapped using pitch. This was likely due to the fact that pitch discrimination deteriorates with age [108, 109]. Furthermore, recent statistics in the BLV community [110] suggest that this population tends to skew older, as many of the conditions that lead to blindness occur later in life. Given that all of our participants were over 50 years of age, the use of pitch as a means of conveying spatial information can pose challenges.

Looking at the user performance on the 2DIY, most participants were able to describe the shape of the transit line better through their experience with the 2DIY in haptic guidance mode compared to the audio-only condition. However, participants gained a better understanding of both the shape and intersections with the audio-only condition compared with the 2DIY in the exploration mode. Examining individual interactions more closely, we conjecture that since the 2DIY only had a single point of contact, participants took longer to explore. During their exploration, users were required to remember all the previous intersection points, greatly increasing their cognitive load. Additionally, we noticed that participants often either ended their exploration prematurely or stagnated in their progression, both of which contributed to poor intersection identification scores.

As outlined in Sec. 4.2.1, our participant pool consisted of individuals aged 50 and above, a crucial demographic parameter that requires careful consideration in interpreting the results of our study.

It is well established that proprioceptive acuity decreases with age [111], affecting the ability of older adults to perceive and interact with FF devices. This may partially explain the challenges that some participants experienced when using the 2DIY in exploration mode. One potential solution to improve the usability of the 2DIY for older adults and individuals with reduced tactile sensitivity is to incorporate additional hardware that allows users to rest their hand on the device, providing a more stable reference point for their interactions. However, it is important to keep in mind that any modifications to the form factor of the 2DIY may increase its production costs, affecting its affordability and accessibility to a wider range of users. Considering that the low cost of the 2DIY was one of the main reasons for selecting the device, it is essential to balance the need for improved usability, general affordability, and accessibility. Therefore, alternative solutions should be explored to address the challenges faced by older adults and individuals with reduced tactile sensitivity, without significantly increasing the cost of the device. Moreover, the findings from the thematic analysis highlight several issues that can be attributed to the relatively lower fidelity of the current 2DIY device. However, the hardware improvements discussed above should address stability concerns in future versions of the device.

4.2.6.2 Improving communication of rendering strategies through targeted training

Another important takeaway from the analysis was the need for more targeted training sessions or examples. It was clear from the results that the Dot Pad was much more intuitive for users. However, it is widely accepted in the human-computer interaction (HCI) community that intuitiveness is a meaningless target, as it usually speaks more to the user's familiarity with the components of the system rather than the efficiency of the design [112].

This was particularly noticeable with the audio renderings, as participants had to experience them repeatedly to understand, suggesting a need for practice.

Even within the time frame of the experiment (approximately two hours), participants seemed to gain comfort with the renderings as the experiment progressed.

As noted above, there is a great variance in the needs and capacities of the BLV population relating to how and when they acquired their vision loss. Therefore, it is important to consider these differences when designing tutorials. As revealed in Themes 1, 5, and 6, a few users possessed a high level of comfort with sonifications and would benefit from having their training geared towards delving deeper into the technical aspects of the renderings to further refine their skills.

One observation from the study was that haptic feedback not only reinforced information but also aided in the comprehension of audio renderings. With this in mind, incorporating the Dot Pad or a similar tactile device into future training or plans could be beneficial. However, due to the cost of the device, it may be more feasible to host the device in a public community space, such as a library, where users can have periodic access to improve their proficiency in using the audio renderings.

4.2.6.3 Considerations for the design of audio-haptic renderings.

The lessons learned from the feedback can be categorized into two main groups: temporal congruence and spatial congruence.

Temporal Congruence

Temporal congruence refers to the degree of synchronization in time between different modalities. This concept is particularly relevant to FF interactions, as the timing of haptic and audio cues can greatly impact the user's understanding and perception of the rendering.

In this study, participants indicated that they felt a discrepancy in the duration of the audio and the haptic guidance. It should be noted that the audio and haptic stimuli were presented sequentially, which may have contributed to the perceived discrepancy. If both were presented simultaneously, the discrepancy in temporal congruence may have been even more pronounced. Previous studies have suggested that the JND for asynchrony of audio-haptic interactions can be as low as 5 to 8 milliseconds [113].

To the best of our knowledge, there has been limited research on establishing tolerances for temporal asynchrony in FF-based audio-haptic interactions.

As discussed in Section 4.1.2.3, synchronizing premade audio files with force guidance is challenging, especially due to the limitations of the 2DIY's resolution, which can make it difficult to achieve fine motor control. This difficulty is compounded by variations in grip and weight applied to the EE, altering the applied friction. To overcome these challenges, more advanced calibration and control algorithms are required to ensure that the pace of the guidance aligns with that of the audio file. However, there is potential for improvement with the latest version of the 2DIY, which features enhanced motors and encoders. These improvements may alleviate resolution problems by improving the accuracy of position tracking, which could make it easier to achieve more precise control.

Another possible solution is to use real-time audio synthesis, taking advantage of the 2DIY's existing position-tracking capabilities. However, generating dynamic audio in real-time can be computationally demanding. This is particularly challenging as the control of the 2DIY is implemented on the computer rather than the device firmware, greatly taxing the computational load on the users' devices.

Spatial Congruence

Spatial congruence is related to the match in spatial mapping between the modalities. In Theme 2, it was observed that there was a significant difference in how participants perceived the magnitude of values in the vertical plane between audio and haptic modalities. This difference may be attributed to the vertical-horizontal illusion, a well-known phenomenon in haptic literature. The phenomenon suggests that people tend to perceive changes in the vertical direction to be longer than the same change in the horizontal plane. The vertical-horizontal illusion was first documented as a visual illusion by Wundt et al. [114] in 1862, and has since been extensively studied and observed in both visual [115, 116] and haptic [117, 118] modalities.

Renier et al. [119] hypothesize that the origins of the phenomenon are from the visual modality, and it, therefore, has no effect on congenitally blind individuals who have never been introduced to the visual illusion. They studied this by using an auditory sensory substitution device on both the sighted and congenitally blind. The researchers found that the illusion was ineffective the congenitally blind and only had a moderate effect on the sighted.

However, in our study, we observed a high incidence of this discrepancy even in congenitally blind individuals, suggesting that our findings may be more related to our rendering strategy than to the illusion itself. Specifically, our use of nonpure tones in the audio renderings may have increased the difficulty in identifying pitch changes. It is, therefore, difficult to conclude to which level the horizontal-vertical illusion contributes to our observations of a spatial discrepancy.

4.2.6.4 Limitations

Our study provided us with valuable information from the user population on the benefits and limitations of our rendering strategies to convey map information to BLV individuals. However, due to the qualitative nature of this study, there are some inherent limitations. As our study was primarily exploratory, we did not prioritize evaluating the performance of our prototypes using quantitative measures. However, as we refine our prototypes in the future, we intend to measure the accuracy of the shapes and points conveyed through our renderings using quantitative metrics.

Since most of our team speaks English as their primary language, we faced limitations in the number of French-speaking participants that we could comfortably recruit for our study. Furthermore, since we conducted the study in both languages, there were slight variations in the renderings due to differences in the TTS voice used for each language. Another limitation of conducting a bilingual study was the thematic analysis, as not all team members were fluent in French and therefore had to rely on translation software or other team members for accurate translations.

Our study also had certain limitations due to the small sample size and participant pool, primarily consisting of people over the age of 50 years, half of whom reported congenital blindness. As a result, it is worth noting that the generalizability of our findings to other populations may be limited.

4.3 Developments on Interactive Tactile Maps

To improve the functionality of the Dot Pad, we integrated position tracking, enabling interactive multimodal applications. Participants identified that this added functionality would be a significant improvement, as it would increase the multimodal congruence by allowing the audio to dynamically adapt to the touch inputs. In this section, we delve into the design of a prototyping environment for crafting multimodal interactions on the Dot Pad. Next, we showcase an example rendering created using this tool. Finally, we conclude with a discussion of the challenges we faced during the process and outline potential directions for future research.

4.3.1 Prototyping Environment

Following the realization that our initial prototyping required extensive technical effort, preventing us from conducting rapid iterations based on user feedback, we opted to design a low-fidelity prototyping environment. This allowed us to test concepts more rapidly within the team and with potential end-users. In this prototyping environment, we designed two distinct operating modes: edit and playback, to frequently assess and adapt designs.

In the editing mode, the designer could manually control the status (raised or lowered) of any pin on the Dot Pad, and they could associate audio with selected pins. Audio events were triggered using specified gestures such as "tap" and "swipe". We intended to make this environment flexible and versatile for testing various concepts quickly and efficiently.

We used the data obtained from the IMAGE extension to ensure that the prototypes generated were realizable in the IMAGE extension. These data were used as a starting point, and the designer selectively modified the status of the pins and audio associations.

4.3.1.1 Dynamic Audio

As the Dot Pad lacks textures, we heavily rely on audio to convey intricate and nuanced information. However, the current audio synthesis pipeline in the IMAGE project often experiences bottlenecks due to a limited number of audio developers and the complexity of the system. To address this issue, we explored two methods for incorporating dynamic audio capabilities into the renderings: dynamic playback and dynamic synthesis. The first approach, dynamic playback, involved selectively playing premade audio files, whereas the second, dynamic synthesis, entailed dynamically generating simple sonifications.

Using dynamic playback, we integrated high-quality renderings generated using the IMAGE architecture into our prototyping environment by using audio sprites. However, we opted not to pursue this approach due to the considerable time required to accumulate a sufficient database of sounds for productive renderings. Additionally, this method would not solve the bottleneck issue as renderings still require significant effort and time to create.

Rather than relying on premade audio files or audio sprites, we decided to use the *Web Audio API* to generate rudimentary sonifications. This approach enabled us to build upon the strategy used in our initial prototypes and expand to more complex graphics, such as street maps. To achieve this, we employed frequency modulation synthesis methods to create a few base sonifications that could be reused in the prototypes.

We used a popular open source option, the *Web Speech API* for our TTS. However, a major limitation of this API was that it did not offer any spatialization capabilities. Where possible, we recommend that future studies incorporate spatialized audio to improve audio-haptic congruence.

4.3.2 Preplanning Example

As an illustration of the prototyping environment detailed above, we present an example rendering created using the tool.

In the multimodal study, a user explained how it was frequently challenging for them to understand the surroundings of a target destination ahead of time. Since many BLV individuals rely on paratransit systems, they are often in a situation where they need to wait around in a safe location. We wanted to design an interaction where the user inputs a target destination (e.g., McGill University) and would be given information about what is around this destination.

We used data from an IMAGE preprocessor that can communicate with OSM, an open-source database maintained by volunteers. Using this preprocessor, we can query a specific latitude and longitude and obtain information about the streets and landmarks within a certain bounding box, which we can then use to create renderings.

Although these data serve as our starting point for the interaction, there is an abundance of street data available from OSM that cannot be displayed on a lower-resolution device such as the Dot Pad. Therefore, we needed to selectively filter the data to not overwhelm the user (Fig. 4.10).

For this scenario, we chose not to use any sonifications as we found that they did not add any meaningful information. Instead, we used TTS to speak the names of streets, intersections, and landmarks. Additionally, the user can get a readout of the distance between two points. In future prototypes, we would like to investigate the use of earcons and other soundscape techniques to communicate information about landmarks.

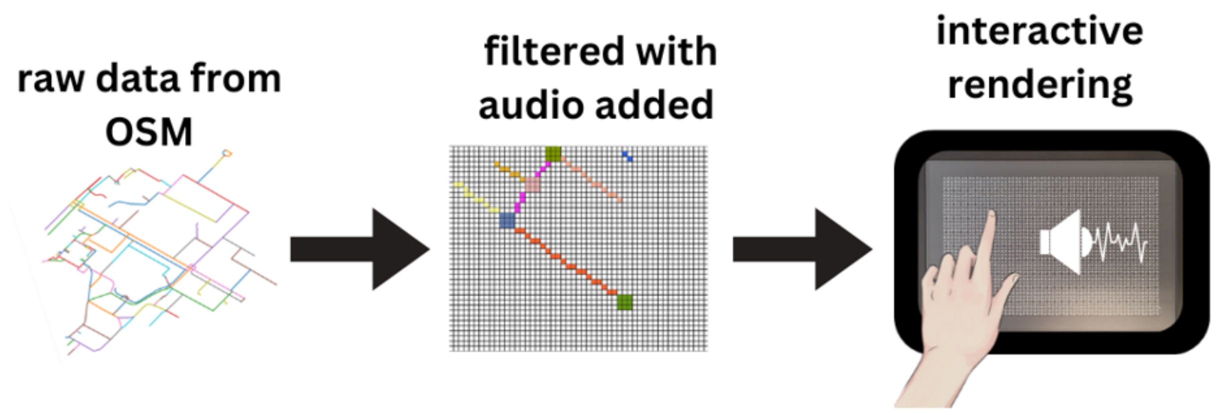


Fig. 4.10: The prototyping process used for generating audio-haptic street maps on the Dot Pad using OSM data.

4.3.3 Challenges and Future Directions

While the work discussed in this chapter endeavours to enhance accessibility for understanding graphical content, several challenges persist, impeding the widespread adoption of these techniques and devices for delivering map information.

- Open-source projects, like OSM, provide convenient access to precise data. However, since they rely on volunteer contributions, significant regions remain unmapped.
- Considering the vast amount of data that we can potentially obtain from these databases, it is necessary to conduct more focused research to identify the most crucial information

in street maps for users. This research could even result in the development of a data hierarchy.

- The multimodal study and ongoing prototype development assume users' capability to set up and operate devices for daily use. Nevertheless, it is crucial to recognize the numerous technical challenges linked to the operation of haptic devices, such as the Dotpad. Therefore, additional research is warranted to investigate effective training methods, enhancing users' autonomy and comfort in using these devices in their existing configuration.

Chapter 5

Conclusion

The sense of touch is a powerful tool that significantly enhances how we interact with technology. Specifically, the use of haptics in developing accessible maps for BLV individuals is increasingly promising. By providing haptic feedback, users can gain a better understanding of the physical layout of the map and the location, enabling them to interact with media that is typically presented visually. However, to design and develop effective haptic interactions, we require a thorough understanding of the underlying mechanisms of haptic perception as well as the knowledge of how to leverage other modalities, such as audio, to improve the transfer of information. As such, motivated by the growing popularity of wrist-mounted wearable devices, we conducted a study to investigate the effect of distance on haptic perception through postural changes. Although our findings were not statistically significant, they suggested that frequency discrimination was better at greater spatiotopic distances. This trend was consistent with similar studies, indicating a potential benefit in using more complex vibrotactile renderings based on frequency changes when designing devices in real-world settings and applications. However, it is important to note that the observed effect is likely to be very small, as indicated by the relatively small percentage difference between JNDs in our study.

By taking a multimodal approach and integrating haptic feedback with other sensory cues, we can create more accessible technology that improves the lives of individuals with blindness. To this effect, we generated prototypes on three separate modalities: audio, FF haptics, and tactile haptics. The design of these prototypes was informed by previous literature and early user testing. Since our focus was on communicating information about street shapes and landmarks, we used transit maps as our stimuli of choice.

Before investing in technical efforts to achieve higher fidelity renderings, we sought user feedback for our strategies. Through this study, we gained insights into what factors in the renderings are appealing to users and what is not. Additionally, our results highlighted some considerations for combining modalities effectively. Of note, we observed that aspects such as speech interruptions and force walls can act as distractions that affect the user's attention, leading to increased cognitive load. Furthermore, we realized that targeted training was essential to both the understanding and enjoyment of our rendering strategies. Using these insights, we iterated on the designs by prototyping interactive multimodal renderings. Due to notable issues on the 2DIY, we opted to focus these efforts on the Dot Pad. As the Dot Pad had not yet been incorporated into the IMAGE extension, we developed a prototyping environment outside of the extension that would allow for the testing of renderings with similar capabilities to the extension. A user scenario from the *Multimodal study* was developed into an example rendering using this prototyping environment.

From the *Multimodal study* we observed that the Dot Pad's unimodal renderings were already effective at conveying spatial information. We are interested in measuring the added value of interactive multimodal renderings in future studies. We also believe there is potential for more detailed audio-haptic representations of maps that include street sizes, crosswalk locations, and landmarks. Considering the current demand for more readily accessible haptic technology, we used a unique combination of qualitative responses and quantitative metrics. Our studies aim to highlight crucial factors in designing haptic devices, such as actuator distance and the influence of body position, while also exploring the effectiveness of multi-modal representations. The findings from our research contribute to the expanding body of work in this field, particularly enhancing our understanding of the technical and sensory considerations involved in haptic device design.

References

[1] R. B. Gillespie and A. M. Okamura, "Haptic interaction for hands-on learning in system dynamics and controls," *Control Systems Magazine*, 2008.

[2] P. Kortum, *HCI beyond the GUI: Design for haptic, speech, olfactory, and other nontraditional interfaces*. Elsevier, 2008.

[3] L. Jones, *Haptics*. MIT press, 2018.

[4] J. M. KENNEDY, "Haptics," in *Perceptual Coding*, pp. 289–318, Elsevier, 1978.

[5] M. A. Srinivasan, "What is haptics," *Laboratory for Human and Machine Haptics: The Touch Lab, Massachusetts Institute of Technology*, pp. 1–11, 1995.

[6] B. Roder, C. Spence, and F. Rasler, "Assessing the effect of posture change on tactile inhibition-of-return," *Experimental Brain Research*, vol. 143, pp. 453–462, Apr. 2002.

[7] S. Soto-Faraco, A. Ronald, and C. Spence, "Tactile selective attention and body posture: Assessing the multisensory contributions of vision and proprioception," *Perception & Psychophysics*, vol. 66, pp. 1077–1094, Oct. 2004.

[8] S. D. Novich and D. M. Eagleman, "Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput," *Experimental brain research*, vol. 233, pp. 2777–2788, 2015.

[9] K. E. MacLean and V. Hayward, "Do it yourself haptics: Part ii [tutorial]," *IEEE robotics & automation magazine*, vol. 15, no. 1, pp. 104–119, 2008.

[10] S. Ungar, M. Blades, and C. Spencer, "The role of tactile maps in mobility training," *British Journal of Visual Impairment*, vol. 11, pp. 59–61, July 1993. Publisher: SAGE Publications Ltd.

[11] A. Weill-Duflos, N. Ong, F. Desourdy, B. Delbos, S. Ding, and C. Gallacher, "Haply 2diy: an accessible haptic platform suitable for remote learning," in *Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 839–840, 2021.

[12] M. Hersh, "Route learning by blind and partially sighted people," *Journal of Blindness Innovation and Research*, vol. 10, no. 2, 2020.

[13] P. Chatzina and D. Gavalas, "Route Planning and Navigation Aid for Blind and Visually Impaired People," in *The 14th Pervasive Technologies Related to Assistive Environments Conference*, (Corfu Greece), pp. 439–445, ACM, June 2021.

[14] J. J. Gibson, *The Senses Considered as Perceptual Systems*. Houghton Mifflin, 1966.

[15] K. O. Johnson, T. Yoshioka, and F. Vega-Bermudez, "Tactile functions of mechanoreceptive afferents innervating the hand," *Journal of Clinical Neurophysiology*, vol. 17, no. 6, pp. 539–558, 2000.

[16] F. Mancini, A. Bauleo, J. Cole, F. Lui, C. A. Porro, P. Haggard, and G. D. Iannetti, "Whole-body mapping of spatial acuity for pain and touch," *Annals of Neurology*, vol. 75, no. 6, pp. 917–924, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ana.24179.

[17] C. Zeagler, "Where to wear it: functional, technical, and social considerations in on-body location for wearable technology 20 years of designing for wearability," in *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pp. 150–157, 2017.

[18] A. Farnè, C. Brozzoli, E. Làdavas, and T. Ro, "Investigating multisensory spatial cognition through the phenomenon of extinction," *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII*, pp. 183–206, 2008.

[19] S. Lakatos and R. N. Shepard, "Time-distance relations in shifting attention between locations on one's body," *Perception & Psychophysics*, vol. 59, pp. 557–566, June 1997.

[20] D. I. Shore, K. Gray, E. Spry, and C. Spence, "Spatial Modulation of Tactile Temporal-Order Judgments," *Perception*, vol. 34, pp. 1251–1262, Oct. 2005. Publisher: SAGE Publications Ltd STM.

[21] N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide*. New York: Psychology Press, 2 ed., Aug. 2004.

[22] E. H. Weber, E. H. Weber, E. H. Weber, H. E. Ross, D. J. Murray, and Experimental Psychology Society, *E.H. Weber on the tactile senses*. Hove: Erlbaum (UK) Taylor & Francis, 1996.

[23] E. C. Tolman, "Cognitive maps in rats and men.," *Psychological Review*, vol. 55, no. 4, p. 189, 1948.

[24] R. G. Golledge, "Human wayfinding and cognitive maps," in *Behavior and Environment: Psychological and Geographical Approaches*, pp. 5–45, Elsevier, 1999.

[25] A. W. Siegel and S. H. White, "The development of spatial representations of large-scale environments," *Advances in Child Development and Behavior*, vol. 10, pp. 9–55, 1975.

[26] P. W. Thorndyke and B. Hayes-Roth, "Differences in spatial knowledge acquired from maps and navigation," *Cognitive Psychology*, vol. 14, no. 4, pp. 560–589, 1982.

[27] D. R. Montello, D. Waller, M. Hegarty, and A. E. Richardson, "Spatial memory of real environments, virtual environments, and maps," in *Human Spatial Memory*, pp. 271–306, Psychology Press, 2004.

[28] R. Passingham, "The hippocampus as a cognitive map| j. o'keefe & l. nadel, oxford university press, oxford (1978). 570 pp.,£ 25.00," 1979.

[29] C. Thinus-Blanc and F. Gaunet, "Representation of space in blind persons: vision as a spatial sense?," *Psychological bulletin*, vol. 121, no. 1, p. 20, 1997.

[30] C. Rizzi, "Cognitive maps and spatial navigation in animals: a comparative perspective," *Journal of Experimental Biology*, vol. 220, no. 18, pp. 3193–3203, 2017.

[31] H. Burton, "Visual cortex activity in early and late blind people," *Journal of Neuroscience*, vol. 23, no. 10, pp. 4005–4011, 2003.

[32] G. Miletic, "Perspective taking: Knowledge of level 1 and level 2 rules by congenitally blind, low vision, and sighted children," *Journal of Visual Impairment & Blindness*, vol. 89, no. 6, pp. 514–523, 1995.

[33] M. Szubielska, E. Niestorowicz, and B. Marek, "Jak rysują osoby, które nigdy nie widziały? badania niewidomych uczniów/drawing without eyesight. evidence from congenitally blind learners," *Annals of Psychology*, no. 19, p. 4, 2016.

[34] S. de Castro Pitano and R. E. Noal, "Blindness and mental representation of knowledge by concepts: Comparison between congenital and acquired blind people," *Educação Unisinos*, vol. 22, no. 2, p. 128, 2018.

[35] D. Goldreich and I. M. Kanics, "Tactile acuity is enhanced in blindness," *Journal of Neuroscience*, vol. 23, no. 8, pp. 3439–3445, 2003.

[36] C. J. Sabourin, Y. Merrikhi, and S. G. Lomber, "Do blind people hear better?," *Trends in Cognitive Sciences*, 2022.

[37] D. Marr, "Visual information processing: The structure and creation of visual representations," *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 290, no. 1038, pp. 199–218, 1980.

[38] J. Pearson, "The human imagination: the cognitive neuroscience of visual mental imagery," *Nature Reviews Neuroscience*, vol. 20, no. 10, pp. 624–634, 2019.

[39] M. E. Wheeler, S. E. Petersen, and R. L. Buckner, "Memory's echo: vivid remembering reactivates sensory-specific cortex," *Proceedings of the National Academy of Sciences*, vol. 97, no. 20, pp. 11125–11129, 2000.

[40] C. Cornoldi and T. Vecchi, *Visuo-spatial working memory and individual differences*. Psychology Press, 2004.

[41] G. Lehnert and H. D. Zimmer, "Auditory and visual spatial working memory," *Memory & Cognition*, vol. 34, pp. 1080–1090, 2006.

[42] G. Lehnert and H. D. Zimmer, "Modality and domain specific components in auditory and visual working memory tasks," *Cognitive Processing*, vol. 9, pp. 53–61, 2008.

[43] T. Vecchi, "Visuo-spatial imagery in congenitally totally blind people," *Memory*, vol. 6, no. 1, pp. 91–102, 1998.

[44] T. Vecchi, C. Tinti, and C. Cornoldi, "Spatial memory and integration processes in congenital blindness," *Neuroreport*, vol. 15, no. 18, pp. 2787–2790, 2004.

[45] R. W. Cholewiak and A. A. Collins, "Vibrotactile localization on the arm: Effects of place, space, and age," *Perception & Psychophysics*, vol. 65, pp. 1058–1077, Oct. 2003.

[46] L. TamÃ¨, A. WÃ¼hle, C. D. Petri, F. Pavani, and C. Braun, "Concurrent use of somatotopic and external reference frames in a tactile mislocalization task," *Brain and Cognition*, vol. 111, pp. 25–33, Feb. 2017.

[47] V. A. Diwadkar and T. P. McNamara, "Viewpoint dependence in scene recognition," *Psychological Science*, vol. 8, no. 4, pp. 302–307, 1997.

[48] A. L. Shelton and T. P. McNamara, "Multiple views of spatial memory," *Psychonomic Bulletin & Review*, vol. 4, no. 1, pp. 102–106, 1997.

[49] B. Roskos-Ewoldsen, T. P. McNamara, A. L. Shelton, and W. Carr, "Mental representations of large and small spatial layouts are orientation dependent.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 24, no. 1, p. 215, 1998.

[50] W. Mou and T. P. McNamara, "Intrinsic frames of reference in spatial memory.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 28, no. 1, p. 162, 2002.

[51] A. Raz, H. Jacobson, and A. Krakowski, "The use of reference frames in spatial cognition in blind and sighted individuals," *Frontiers in Psychology*, vol. 5, p. 957, 2014.

[52] T. Iachini, G. Ruggiero, and F. Ruotolo, "Does blindness affect egocentric and allocentric frames of reference in small and large scale spaces?," *Behavioural Brain Research*, vol. 273, pp. 73–81, 2014.

[53] N. Burgess, "Spatial memory: How egocentric and allocentric combine," *Trends in Cognitive Sciences*, vol. 10, no. 12, pp. 551–557, 2006.

[54] H. Konig, J. Schneider, and T. Strothotte, "Haptic exploration of virtual buildings using non-realistic haptic rendering," in *Int. Conf. on Computers Helping People with Special Needs (ICCHP)*, pp. 377–384, Citeseer, 2000.

[55] L. Ran, S. Helal, and S. Moore, "Drishti: an integrated indoor/outdoor blind navigation system and service," in *Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the*, pp. 23–30, Mar. 2004.

[56] B. N. Walker and J. Lindsay, "Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 48, pp. 265–278, June 2006.

[57] T. V. Tran, T. Letowski, and K. S. Abouchacra, "Evaluation of acoustic beacon characteristics for navigation tasks," *Ergonomics*, vol. 43, pp. 807–827, June 2000.

[58] E. Björk, "Laboratory annoyance and skin conductance responses to some natural sounds," *Journal of Sound and Vibration*, vol. 109, no. 2, pp. 339–345, 1986.

[59] I. Hertrich, S. Dietrich, A. Moos, J. Trouvain, and H. Ackermann, "Enhanced speech perception capabilities in a blind listener are associated with activation of fusiform gyrus and primary visual cortex," *Neurocase*, vol. 15, no. 2, pp. 163–170, 2009.

[60] P. B. Nelson, S.-H. Jin, and A. E. Carney, "Understanding speech in modulated interference: cochlear implant users and normal-hearing listeners," *Journal of the Acoustical Society of America*, vol. 113, no. 2, pp. 961–968, 2003.

[61] W. W. Gaver, "What in the world do we hear? an ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[62] T. Hermann, A. Hunt, and J. Neuhoff, *The Sonification Handbook*. Logos Verlag Berlin GmbH, 2011.

[63] F. L. Wightman and D. J. Kistler, "Resolution of front-back ambiguity in spatial hearing by listener and source movement," *The Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2841–2853, 1999.

[64] J. C. Middlebrooks and D. M. Green, "Sound localization by human listeners," *Annual Review of Psychology*, vol. 42, no. 1, pp. 135–159, 1991.

[65] P. Majdak, M. J. Goupell, and B. Laback, "3-d localization of virtual sound sources: Effects of visual environment, pointing method, and training," *Attention, Perception, & Psychophysics*, vol. 72, no. 2, pp. 454–469, 2010.

[66] J. B. Krygier, "Sound and geographic visualization," in *Modern cartography series*, vol. 2, pp. 149–166, Elsevier, 1994.

[67] P. Voss, F. Lepore, F. Gougoux, R. J. Zatorre, and M. Lassonde, "Differential occipital responses in early-and late-blind individuals during a sound-source discrimination task," *NeuroImage*, vol. 23, no. 1, pp. 364–373, 2004.

[68] B. Röder, W. Teder-Sälejärvi, A. Sterr, F. Rosler, and S. A. Hillyard, "Improved auditory spatial tuning in blind humans," *Nature*, vol. 400, no. 6740, pp. 162–166, 1999.

[69] D. Josselin, U. ESPACE, and F. AVIGNON, "Spatial analysis using sonorous cartography. some propositions.," *Proceedings of ICC'2011, paris, 3-8 july 2011*, 2011.

[70] I. Miller, A. Pather, J. Milbury, L. Hasty, A. O'Day, D. Spence, and S. Osterhaus, "Guidelines and standards for tactile graphics," *The Braille Authority of North America*, 2010.

[71] M. Espinosa, S. Ungar, E. OchaÄÌta, M. Blades, and C. Spencer, "Comparing Methods for Introducing Blind and Visually Impaired People to Unfamiliar Urban Environments," *Journal of Environmental Psychology*, vol. 18, pp. 277–287, Sept. 1998.

[72] R. G. Golledge, "Tactual strip maps as navigational aids," *Journal of Visual Impairment & Blindness*, vol. 85, no. 7, pp. 296–301, 1991.

[73] S. Ungar, M. Blades, and C. Spencer, "Strategies for knowledge acquisition from cartographic maps by blind and visually impaired adults," *The Cartographic Journal*, vol. 34, no. 2, pp. 93–110, 1997.

[74] L. Holloway, K. Marriott, and M. Butler, "Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, (Montreal QC Canada), pp. 1–13, ACM, Apr. 2018.

[75] S. Giraud, A. M. Brock, M. J.-M. Macé, and C. Jouffrais, "Map learning with a 3d printed interactive small-scale model: Improvement of space and text memorization in visually impaired students," *Frontiers in Psychology*, vol. 8, p. 930, 2017.

[76] M. Simonnet, "Smart map augmented: Exploring and learning maritime audio-tactile maps without vision: The issue of finger or marker tracking," *Multimodal Technologies and Interaction*, vol. 6, no. 8, p. 66, 2022.

[77] B. Taylor, A. Dey, D. Siewiorek, and A. Smailagic, "Customizable 3d printed tactile maps as interactive overlays," in *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 71–79, 2016.

[78] J. Brus, R. Barvir, and A. Vondrakova, "Interactive 3d printed haptic maps-touchit3d," *CARTOGRAPHY FOR SPECIFIC USERS*, p. 23, 2019.

[79] M. Kuribayashi, S. Kayukawa, H. Takagi, C. Asakawa, and S. Morishima, "Linechaser: A smartphone-based navigation system for blind people to stand in lines," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.

[80] S. Azenkot, R. E. Ladner, and J. O. Wobbrock, "Smartphone haptic feedback for nonvisual wayfinding," in *The proceedings of the 13th international ACM SIGACCESS conference on Computers and Accessibility*, pp. 281–282, 2011.

[81] P. Wacker, C. Wacharamanotham, D. Spelmezan, J. Thar, D. A. Sánchez, R. Bohne, and J. Borchers, "Vibrovision: an on-body tactile image guide for the blind," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 3788–3791, 2016.

[82] O. Ozioko and R. Dahiya, "Smart tactile gloves for haptic interaction, communication, and rehabilitation," *Advanced Intelligent Systems*, vol. 4, no. 2, p. 2100091, 2022.

[83] S. Kammoun, C. Jouffrais, T. Guerreiro, H. Nicolau, and J. Jorge, "Guiding blind people with haptic feedback," *Frontiers in Accessibility for Pervasive Computing (Pervasive 2012)*, vol. 3, 2012.

[84] G. Flores, S. Kurniawan, R. Manduchi, E. Martinson, L. M. Morales, and E. A. Sisbot, "Vibrotactile guidance for wayfinding of blind walkers," *IEEE Transactions on Haptics*, vol. 8, no. 3, pp. 306–317, 2015.

[85] V. Hayward, O. R. Astley, M. Cruz-Hernandez, D. Grant, and G. Robles-De-La-Torre, "Haptic interfaces and devices," *Sensor review*, vol. 24, no. 1, pp. 16–29, 2004.

[86] B. Tornil and N. Baptiste-Jessel, "Use of force feedback pointing devices for blind users," in *User-Centered Interaction Paradigms for Universal Access in the Information Society: 8th ERCIM Workshop on User Interfaces for All, Vienna, Austria, June 28-29, 2004, Revised Selected Papers 8*, pp. 479–485, Springer, 2004.

[87] K. Rassmus-Grohn, C. Magnusson, and H. Eftring, "AHEAD - Audio-Haptic Drawing Editor And Explorer for Education," in *2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games*, pp. 62–66, Oct. 2007.

[88] H. Z. Tan, S. Choi, F. W. Lau, and F. Abnousi, "Methodology for maximizing information transmission of haptic devices: A survey," *Proceedings of the IEEE*, vol. 108, no. 6, pp. 945–965, 2020.

[89] N. G. Forta, M. J. Griffin, and M. Morioka, "Vibrotactile difference thresholds: Effects of vibration frequency, vibration magnitude, contact area, and body location," *Somatosensory & Motor Research*, vol. 29, no. 1, pp. 28–37, 2012.

[90] B. R. Glasberg and B. C. Moore, "Psychoacoustic abilities associated with music perception," in *Music Perception*, pp. 87–102, Springer, 2002.

[91] G. Wetherill, H. Chen, and R. Vasudeva, "Sequential estimation of quantal response curves: A new method of estimation," *Biometrika*, vol. 53, no. 3-4, pp. 439–454, 1966.

[92] B. Berglund, U. Berglund, and G. Ekman, "Temporal integration of vibrotactile stimulation," *Perceptual and Motor Skills*, vol. 25, no. 2, pp. 549–560, 1967.

[93] K. Johnson, "Neural basis of haptic perception," in *Steven's handbook of experimental psychology: Sensation and perception, Vol. 1, 3rd ed*, pp. 537–583, Hoboken, NJ, US: John Wiley & Sons Inc, 2002.

[94] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

[95] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *arXiv preprint arXiv:1406.5823*, 2014.

[96] C. Germann, R. Sutter, and D. Nanz, "Novel observations of Pacinian corpuscle distribution in the hands and feet based on high-resolution 7-T MRI in healthy volunteers," *Skeletal Radiology*, vol. 50, pp. 1249–1255, June 2021.

[97] J. Regimbal, J. R. Blum, and J. R. Cooperstock, "Image: a deployment framework for creating multimodal experiences of web graphics," in *Proceedings of the 19th International Web for All Conference*, pp. 1–5, 2022.

[98] J. D. Novak and A. J. Cañas, "The theory underlying concept maps and how to construct and use them," 2008.

[99] E. Tanhua-Piiroinen and R. Raisamo, "Micole: A multimodal learning environment for visually impaired and sighted children," in *Disability Rehabilitation Management Through ICT*, pp. 159–169, River Publishers, 2022.

[100] I. Xenakis, "Elements of stochastic music," *Gravesaner*, vol. 18, pp. 84–105, 1960.

[101] B. Bernschütz, "A spherical far field hrir/hrtf compilation of the neumann ku 100," in *Proceedings of the 40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (DAGA) Conference on Acoustics*, p. 29, German Acoustical Society (DEGA) Berlin, 2013.

[102] C. Hatzfeld and T. A. Kern, *Engineering haptic devices*. Springer, 2016.

[103] V. Braun and V. Clarke, *Thematic analysis.* American Psychological Association, 2012.

[104] R. Rampin and V. Rampin, "Taguette: open-source qualitative data analysis," *Journal of Open Source Software*, vol. 6, no. 68, p. 3522, 2021.

[105] V. Hayward and K. E. MacLean, "Do it yourself haptics: part i," *IEEE Robotics & Automation Magazine*, vol. 14, no. 4, pp. 88–104, 2007.

[106] D. A. Norman, *Emotional design: Why we love (or hate) everyday things*. New York: Basic Books, 2004.

[107] C. Hautekiet, S. Verschooren, N. Langerock, and E. Vergauwe, "Attentional switching between perception and memory: Examining asymmetrical switch costs," *Attention, Perception, & Psychophysics*, pp. 1–11, 2023.

[108] C. Bunch and T. Raiford, "Race and sex variations in auditory acuity," *Archives of Otolaryngology*, vol. 13, no. 3, pp. 423–434, 1931.

[109] J. C. Steinberg, H. C. Montgomery, and M. B. Gardner, "Results of the world's fair hearing tests," *The Bell System Technical Journal*, vol. 19, no. 4, pp. 533–562, 1940.

[110] N. F. of the Blind, "Blindness statistics." `https://nfb.org/resources/blindness-statistics`, n.d. Accessed on April 10, 2023.

[111] J. C. Stevens, E. Foulke, and M. Q. Patterson, "Tactile acuity, aging, and braille reading in long-term blindness.," *Journal of Experimental Psychology: Applied*, vol. 2, no. 2, p. 91, 1996.

[112] D. A. Norman, "The problem with intuitive design," *Dr. Dobb's Journal*, vol. 15, no. 11, pp. 32–36, 1990.

[113] B. D. Adelstein, D. R. Begault, M. R. Anderson, and E. M. Wenzel, "Sensitivity to haptic-audio asynchrony," in *Proceedings of the 5th international conference on Multimodal interfaces*, pp. 73–76, 2003.

[114] W. M. Wundt, *Contributions to the theory of sense perception*. CF Winter, 1862.

[115] B. Wallace, "Creation of the horizontal-vertical illusion through imagery," *Bulletin of the Psychonomic Society*, vol. 22, no. 1, pp. 9–11, 1984.

[116] D. Pearce and L. Matin, "Variation of the magnitude of the horizontal-vertical illusion with retinal eccentricity," *Perception & Psychophysics*, vol. 6, no. 4, pp. 241–243, 1969.

[117] M. A. Heller, J. A. Calcaterra, L. L. Burson, and S. L. Green, "The tactual horizontal-vertical illusion depends on radial motion of the entire arm," *Perception & Psychophysics*, vol. 59, pp. 1297–1311, 1997.

[118] M. A. Heller, A. Smith, R. Schnarr, J. Larson, and S. Ballesteros, "The horizontal–vertical curvature illusion in touch is present in three-dimensional objects and raised lines," *The American Journal of Psychology*, vol. 126, no. 1, pp. 67–80, 2013.

[119] L. Renier, R. Bruyer, and A. G. De Volder, "Vertical-horizontal illusion present for sighted but not early blind humans using auditory substitution of vision," *Perception & Psychophysics*, vol. 68, pp. 535–542, 2006.