Perception of audio-haptic textures for new touchscreen interactions

Corentin Bernard

AIX-MARSEILLE UNIVERSITÉ

Doctoral School: Sciences pour l'ingénieur

PRISM (UMR 7061)

A thesis submitted in fulfillment of the requirements for the degree of Doctor at Aix-Marseille Université

Perception of audio-haptic textures for new touchscreen interactions

---

Corentin Bernard

Defended on 22th of February 2022 before committee:

Laurent GRISONI	University of Lille	Reviewer
Davide ROCCHESSO	University of Palermo	Reviewer
Anne KAVOUNOUDIAS	Aix-Marseille University	Examiner
Vincent LÉVESQUE	École de technologie supérieure, Montréal	Examiner
Jocelyn MONNOYER	Stellantis	Industrial supervisor
Michaël WIERTLEWSKI	TU Delft	Supervisor
Sølvi YSTAD	PRISM, CNRS, AMU	Supervisor

Declaration of Authorship

I, Corentin Bernard, declare that this thesis titled, "Perception of audio-haptic textures for new touchscreen interactions" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Corentin Bernard

Date: 16/02/2022

"Les hasards, les rencontres forgent une destinée."

Otis

Abstract

Perception of audio-haptic textures for new touchscreen interactions

An ever-increasing number of human-machine interfaces have embraced touchscreens as their central component, such as in smartphones, laptops, terminals, etc. Their success has been particularly noticeable in the automotive industry, where physical buttons have been replaced by touchscreens to handle multiple elements of the driving environment. However, contrary to physical buttons, these interfaces do not possess any tangible elements that allows the user to feel where the commands are. Without tactile feedback, users have to rely on visual cues and simple adjustment tasks become significant distractions that may lead to dangerous situations while driving.

Recently, haptic touchscreens have emerged to restore tangibility to these interfaces, by rendering the sensation of feeling textures and shapes through friction modulation. However, we still do not have a good understanding of how these synthetic textures are perceived by humans, which is crucial to design meaningful and intuitive haptic interfaces. In this thesis, I first show that the perception thresholds of friction modulated textures are similar to vibrotactile thresholds. Then, I investigate the perception of haptic gradients, i.e., textures whose spatial frequency gradually changes. Hence, a law that describes the minimal exploration distance to perceive a given gradient is deduced. This law is similar to the auditory perception of rhythm variations, which suggests that there are common mechanisms between the two modalities. Finally, I demonstrate that gradient haptic feedback can guide a user to adjust a setting on an interface without vision.

The findings shed new light on the understanding of haptic perception and its multisensory interactions and open up new possibilities in terms of human-machine interaction.

Keywords: Haptics, Acoustics, HMI, Perception, Psychophysics, Ultrasonic Friction Modulation, Multisensory Integration.

Résumé

Perception des textures audio-haptiques pour de nouvelles interactions avec un écran tactile

Les écrans tactiles ont envahi notre quotidien et sont maintenant présents dans un grand nombre d'appareils tels que les téléphones, les ordinateurs, les bornes de commande, etc. Leur succès a aussi gagné l'industrie automobile, où les écrans ont remplacé les boutons physiques pour gérer les paramètres de l'environnement de conduite. Cependant, contrairement aux boutons physiques, les interfaces sur écran tactile ne possèdent pas d'éléments tangibles permettant à l'utilisateur de sentir les commandes. Sans retours tactiles, les utilisateurs ne peuvent se fier qu'aux indications visuelles et de simples tâches de réglage peuvent devenir des distractions importantes au volant.

Récemment, les écrans haptiques ont fait leur apparition pour redonner une certaine tangibilité à ces interfaces en permettant de procurer des sensations de texture et de relief à leur surface grâce à la modulation de frottement. Cependant, nous ne comprenons toujours pas bien comment ces textures synthétiques sont perçues par l'homme, ce qui est une question cruciale pour concevoir des interfaces haptiques pertinentes et intuitives. Dans cette thèse, je montre tout d'abord que les seuils de perception des textures par modulation de frottement sont similaires aux seuils de perception des vibrations. Ensuite, j'étudie la perception des gradients haptiques, des textures dont la fréquence spatiale évolue progressivement. Il en ressort une loi décrivant la distance minimale d'exploration pour percevoir un gradient donné. Cette loi se révèle similaire à la perception auditive des variations de rythme, ce qui suggère qu'il existe des mécanismes communs entre les deux modalités. Enfin, je démontre que les retours haptiques de type gradient peuvent guider un utilisateur pour lui permettre d'ajuster un paramètre sur une interface sans recourir à la vision.

Ces résultats apportent un éclairage nouveau sur la compréhension de la perception haptique et ses liens avec la perception auditive, et ouvrent de nouvelles possibilités en termes d'interaction homme-machine.

Mots clés : Haptique, Acoustique, IHM, Perception, Psychophysique, Modulation de Frottement Ultrasonore, Intégration Multimodale.

List of publications

Bernard, C., Monnoyer, J., Ystad, S., & Wiertlewski, M. (2022). Eyes-Off Your Fingers: Gradual Surface Haptic Feedback Improves Eyes-Free Touchscreen Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2022.*

Bernard, C., Monnoyer, J., Wiertlewski, M., & Ystad, S. (2022). Rhythm perception is shared between audio and haptics. *Scientific Reports.*

Bernard, C., Ystad, S., Monnoyer, J., & Wiertlewski, M. (2020). Detection of friction-modulated textures is limited by vibrotactile sensitivity. *IEEE transactions on haptics.*

Bernard, C., Monnoyer, J., & Wiertlewski, M. (2018). Harmonious textures: The perceptual dimensions of synthetic sinusoidal gratings. *In International Conference on Human Haptic Sensing and Touch Enabled Computer Applications 2018. Nominated for Best Student Paper.*

Huloux, N., **Bernard, C.**, & Wiertlewski, M. (2020). Estimation of the modulation of friction from the mechanical impedance variations. *IEEE Transactions on Haptics.*

Bernard, C., Monnoyer, J., Ystad, S., & Wiertlewski, M. (2019). Interferometric Tribometer for Wide-Range/High-Bandwidth Measurement of Tactile Force Interaction. *In IEEE World Haptics Conference.*

Bernard, C., Monnoyer, J., Denjean, S., Wiertlewski, M., & Ystad, S. (2019). Sound and Texture Synthesizer. *In International Workshop on Haptic and Audio Interaction Design-HAID.*

Kanzari, K., **Bernard, C.**, Monnoyer, J., Denjean, S., Wiertlewski, M.,& Ystad, S. (2019). A Real-time Synthesizer of Naturalistic Congruent Audio-Haptic Textures. *In CMMR 2019: 14th International Symposium on Computer Music Multidisciplinary Research.*

Fery, M., **Bernard, C.**, Thoret, E., Kronland-Martinet, R., Ystad, S. (2021). Audio-Tactile Perception of Roughness. *Computer Music Multidisciplinary Research 2021.*

Bernard, C., Monnoyer, J., & Wiertlewski, M. (2020). Procédé et appareil de calibration d'une paroi haptique associée à un capteur de force. *French Patent n° FR3077131.*

Bernard, C., Monnoyer, J., Denjean, S., & Wiertlewski, M., Ystad, S. (2022). Procédé et dispositif de contrôle d'un système de retour haptique. *French Patent, submitted*

Acknowledgements

Je tiens en premier lieu à remercier mes directeurs de thèse. Merci Sølvi d'avoir été présente au jour le jour et disponible dès que j'en avais besoin. Grâce à toi je n'ai pas oublié les sons ! Merci Michaël de m'avoir partagé ton goût pour la recherche dès les débuts de mon stage, de m'avoir donné envie de poursuivre en thèse et d'avoir continué à suivre mes travaux depuis Delft. Merci à vous deux de m'avoir confronté aux exigences scientifiques en me poussant à remettre régulièrement en question mes problématiques de recherche et mes protocoles, ce qui a permis à ma thèse d'avancer dans la bonne direction.

Merci Jocelyn de m'avoir fait confiance dès le début de cette aventure et de m'avoir soutenu dans les différentes directions que j'avais envie de prendre.

Mes remerciements vont également à Richard et Stéphane pour m'avoir accueilli dans leurs laboratoires respectifs.

A Sébastien, qui a suivi cette thèse depuis Paris, merci d'avoir été disponible quand j'avais besoin de tes conseils.

Un grand merci à mes acolytes Laurence et Nicolas pour tous nos échanges, les relectures d'articles et de présentations, et d'avoir ouvert la voie quelque temps avant moi. Ces années de thèse n'auraient pas été les mêmes sans vous, entre Japon, Pays-Bas, Italie et Méditerranée.

Merci à l'équipe des thésards (et ingénieur !) de PRISM, Simon, Théophile, Antoine, Thomas, Adrien, Pierre et Samuel pour la bonne ambiance, les p'tits cartons et les bracos, aussi bien que pour vos retours avisés en réunion d'équipe et votre aide pour organiser la soutenance. Merci Étienne pour tes précieux conseils, tes encouragements et ton inspirante persévérance.

Merci aussi aux autres membres de PRISM, notamment Mitsuko, Jocelyn et Salomée, ainsi que tous ceux que j'ai eu l'occasion d'y croiser, en particulier Khoubeib, Matis et Madeline.

Merci à tous ceux passés par l'équipe biorob' pour les bons moments passés ensemble, festifs et sportifs, en particulier à Anna, Xi, Evandro, Aimie, Léo, Illya, Jean-Marc, Océane, Léandres, Valentin, les Julien et Marc.

Je remercie également l'équipe de TU Delft pour nos échanges.

Je remercie ma famille, mes parents, ma sœur et mes frères pour leur soutien. Je profite que vous soyez public captif pendant 45 minutes aujourd'hui pour qu'enfin je puisse vous expliquer les tenants et les aboutissants de ma thèse en toute simplicité !

Merci à ma mamie présente aujourd'hui, toi tu seras dispensée de l'évaluation finale. Pour Hubert, c'est l'occasion de préouïr quelques nouveaux mots.

Merci Marie d'être là, super sympa !

Une pensée à ceux qui ont forgé ma culture scientifique étant plus jeune, M. Gérard, mon professeur de physique au collège qui m'a partagé sa passion et sa rigueur, mon oncle Eric qui m'a fait découvrir l'ambiance de recherche internationale au CERN, sans oublier Fred, Jamy et Marcel !

Merci à tous mes amis, ceux de l'Ain, ceux de Marseille, mes anciens colocs et camarades de promo ainsi que toutes les personnes qui ont fait le déplacement aujourd'hui pour me soutenir.

Contents

List of Figures

Chapter 1

Introduction

Contents

Problem statement

Machines are omnipresent in our society. We therefore had to find the most effective ways of using them. Since electronic devices are based on switches, press buttons naturally appeared as the basic units for interacting with these machines. On modern computers, a panel of buttons in the form of the keyboard and mouse pair has been widely adopted as the most efficient interface. Simple spring and stop systems have been developed to provide haptic feedback when pressing buttons, confirming that the action has been completed successfully. The haptic feedback of the keys makes the interaction more intuitive, and has been proven to reduce typing errors [Ma et al., 2015].

Recently, touchscreens have emerged and are now common everywhere. They have replaced many physical interfaces, such as telephone keypads, vending machine buttons and even keyboards on tablets. The success of touchscreens mainly come from their high level of reconfigurability and versatility. Since the visual interface and the touch inputs can be easily switched between menus, the same compact device can be used to control many various parameters. Despite their assets, touchscreens present a major drawback. These flat surfaces deprive the user of any tactile feedback, causing certain tasks to be more complex and less intuitive.

This issue is particularly relevant in the automotive context. Indeed, the design of efficient on-board interfaces is essential since the driver needs to concentrate on the main task of driving and cannot devote too much attention to the control of in-car parameters. Yet car dashboards have not escaped the general trend and physical buttons have also been mostly replaced by touchscreens.

Without tactile feedback, simple setting tasks that could even be performed without vision, such as air conditioning setting, now require much more visual attention. This is a major concern since 80% of car accidents are caused by driver inattention [Dingus et al., 2006], and especially by distraction that comes from phones and screens [Lee et al., 2008]. A simple task such as dialling a phone number, for example, increases the braking reaction time of about 0.5 s [Lamble et al., 1999]. Some works have already suggested implementing tactile feedback on car touchscreens. They showed a strong user preference for tactile feedback in addition to the visual interface, but the performance gains were not significant [Pitts et al., 2012] [Serafin et al., 2007]. However, those studies rendered basic tactile feedback using simple lateral vibrations.

By acting on friction, new haptic touchscreens can now provide more precise and robust haptic feedback. In this thesis, the haptic technology used is the ultrasonic friction modulation. Ultrasonic vibrations of the surface cause the finger to slightly levitate on a thin film of air, strongly reducing the friction between the finger and the screen. Modulation of the vibration amplitude produces friction changes that are perceived as synthetic textures or shapes and can be used as haptic feedback.

However, we still do not fully understand how friction-modulated textures are perceived by humans. Although a wide variety of tactile stimuli can be rendered, they do not feel like the familiar stimuli we encounter every day, such as real textures or vibrations.

The first objective of the thesis is to improve our understanding of the perception of this type of synthetic textures. The design of the study was inspired by experiments on auditory perception already well documented in psychoacoustics.

Based on the perceptive results, the second objective of the thesis is to propose new human-computer interactions that exploit the full potential of haptic feedback to substitute for the visual modality. Since the addition of tactile and auditory feedback showed promising results both in terms of performance gain and user preference on car interfaces [Serafin et al., 2007] [Gaffary and Lécuyer, 2018], I also investigate which kind of sounds could be combined with friction modulated textures to assist the interaction.

A better understanding of synthetic textures perception and the demonstration of their application potential could help, in the near future, to design efficient and tangible interfaces.

Thesis overview

The thesis is organized as follows:

Chapter 2 provides a literature review covering the scope of the thesis. First, I introduce the auditory and tactile sensory systems. Then, I detail how the two sensory systems respond to basic stimuli and compare their detection capabilities. The next section describes the different approaches of tactile and auditory synthesis. The state of the art then focuses on haptic surfaces. I explain how these technologies function and I present previous works that explored how humans perceive synthetic textures. Next, I show how haptic and auditory modalities can influence each other. Finally, I present how audio and haptic feedback can guide the user in his interaction with an interface.

Chapter 3 refines our understanding of the perception of friction-modulated textures by investigating the minimal amplitude of friction variations that humans can perceive. It raises the question of whether synthesized textures are perceptually closer to vibrations or to real textures. With a psychophysical experiment, I measure the detection thresholds of sinusoidal gratings at various spatial frequencies for two finger velocity conditions, and compare the results with thresholds from the literature on real reliefs and vibrotactile perception. The resulting threshold curves are also helpful to design perfectly calibrated haptic stimuli.

Chapter 4 takes the study of haptic perception a step further by considering an unfamiliar type of texture: textures that evolve continuously while being explored by the finger. With a psychophysical experiment, I first measure the exploration distance needed to perceive a certain evolution rate defined by a spatial frequency gradient. The results are compared with the auditory literature on rhythmic changes and explained by a multimodal model of rhythm perception, revealing strong similarities between the two modalities. A second experiment highlights audio-haptic sensory integration in the perception of evolving textures.

Chapter 5 proposes a new human-computer interaction using an evolving haptic feedback based on the previous results. The interaction is applied to the use case of setting a parameter on a touchscreen. I present an experiment that demonstrates that, after learning, this interaction is suitable to guide the user in performing the setting task. I also propose and compare different learning strategies that uses other modalities, with visual or auditory feedback, in order to find the most suitable training method depending on the context.

Chapter 6 summarizes the main contribution of this thesis and discusses the results and their limitations. I also present new research questions arising from the experiments that may lead to future work.

Chapter 2

State of the art

Contents

Preface to Chapter 2

This state of the art is constructed to highlight parallels and differences between the senses of hearing and touch. Firstly, I describe the biological functioning of the auditory and tactile sensory systems. Then, I present how both senses react to elementary stimuli and how we can compare these two modalities. Next, I focus on signal synthesis. Whereas sophisticated algorithms can synthesize very realistic sounds, an equivalence of recordings and reproductions through the microphone-loudspeaker chain still does not exist for haptics. Current research on haptic synthesis involves both software and hardware solutions. The next section describes the different haptic surface technologies with a focus on ultrasonic friction modulation and how this kind of haptic feedback is perceived. Then, several studies which demonstrate how auditory and haptic senses influence each other are presented. Finally, I will show how multimodal feedback can guide a user on an interface.

2.1 Senses of hearing and touch

Before exploring human perception in greater depth, this first section will describe how the auditory and tactile sensory systems operate from a biological point of view.

2.1.1 Auditory system

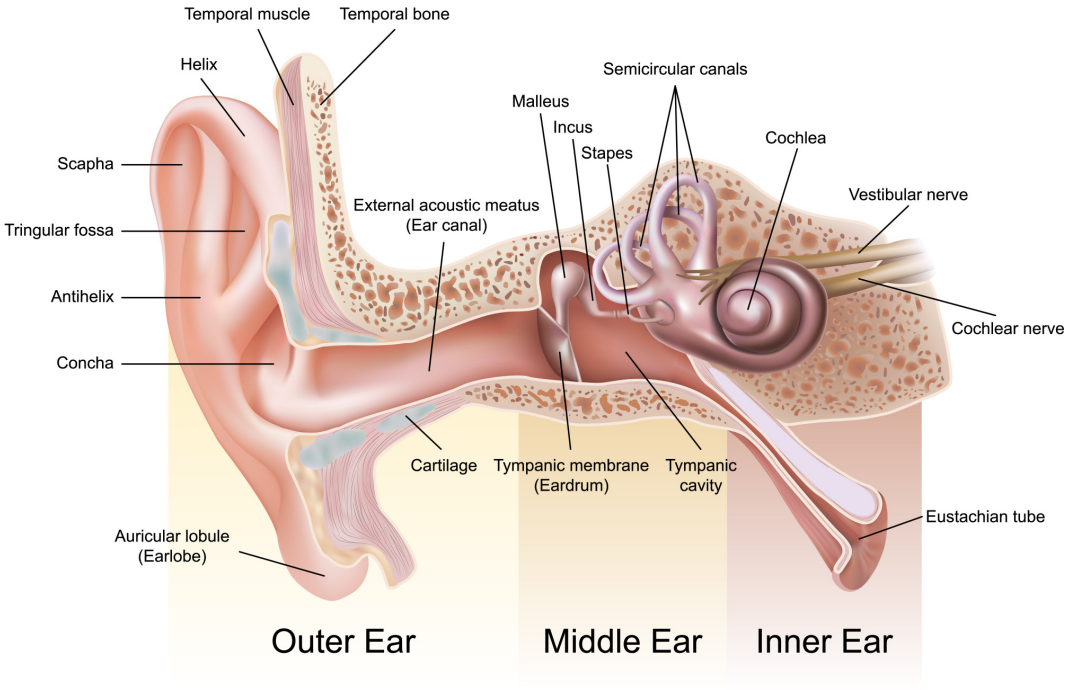

FIGURE 2.1: Schematic of the auditory system. Adapted from *Thompsons Road Physiotherapy*

The auditory system, responsible for the sense of hearing, is capable of perceiving sounds over a very wide range of frequencies, from approximately 20 Hz to 20 kHz, and a wide range of intensities, from the detection threshold of approximately 10^{-12} W/m^2 up to the pain threshold of approximately 1 W/m^2. It is composed of three parts presented in Figure 2.1: the outer ear, the middle ear and the inner ear.

The outer ear

The outer ear is the visible part of the auditory system, made up of the auricular pinna, which captures acoustic waves thanks to its shape and concentrates them towards the ear canal. The auricular pinna amplifies frequencies of approximately 2 kHz by a few decibels. The ear canal, in which waves are traveling, acts as a resonant pipe that amplifies frequencies between 2 and 5 kHz, the bandwidth for which the auditory system is most sensitive.

The middle ear

The middle ear consists of the eardrum and the ossicular chain, a chain of three small bones called the malleus, incus, and stapes (also called hammer, anvil, and stirrup). Acoustic waves transmitted through the ear canal create vibrations at the eardrum. The eardrum is a membrane which converts acoustic waves in the air into mechanical vibrations. The ossicles amplify these vibrations and transmit them to the inner ear. The stape muscle can contract to damp excessive pressure, thus protecting the inner ear. Simultaneously, the Eustachian tube opens periodically to equalize the pressure on both sides of the eardrum.

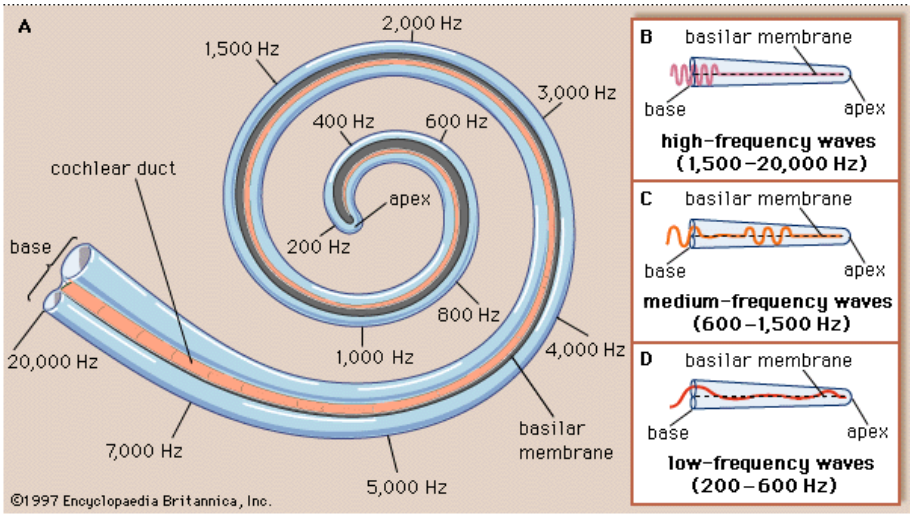

FIGURE 2.2: Schematic of the cochlea and description of acoustic waves propagation according to their frequency. *Encyclopedia Britannica*

The inner ear

The stape strikes the oval window and thus transmits the acoustic wave to the cochlea. The cochlea is a hollow fluid-filled organ consisting of three spirally coiled tubes. The cochlear duct is surrounded by the vestibular duct and the tympanic duct, separated by the Reissner's membrane and the basilar membrane, respectively. Vibrations propagate through these tubes and make the basilar membrane move, which in turn causes movements of inner and outer haircells. The basilar membrane, which is thin and taut at the base, continuously changes its mechanical properties up to its thick and loose end (the Apex). This means that depending on the input frequency, different regions of the basilar membrane will be more or less excited. As shown in Figure 2.2, the base of the cochlea is excited by high frequency sound waves, while the apex is excited by low frequencies. Thus, depending on their position on the basilar membrane, the hair cells respond to a particular excitation frequency. They perform the mechanoelectrical transduction, transforming a movement of their cilia into nerve signals that are further transmitted to the brain

via the auditory nerve. The auditory system therefore operates by decomposing sounds into a multitude of narrow frequency bands, in some ways comparable to a Fourier transform. There are other models of the auditory system based on temporal variations.

2.1.2 Tactile sensory system

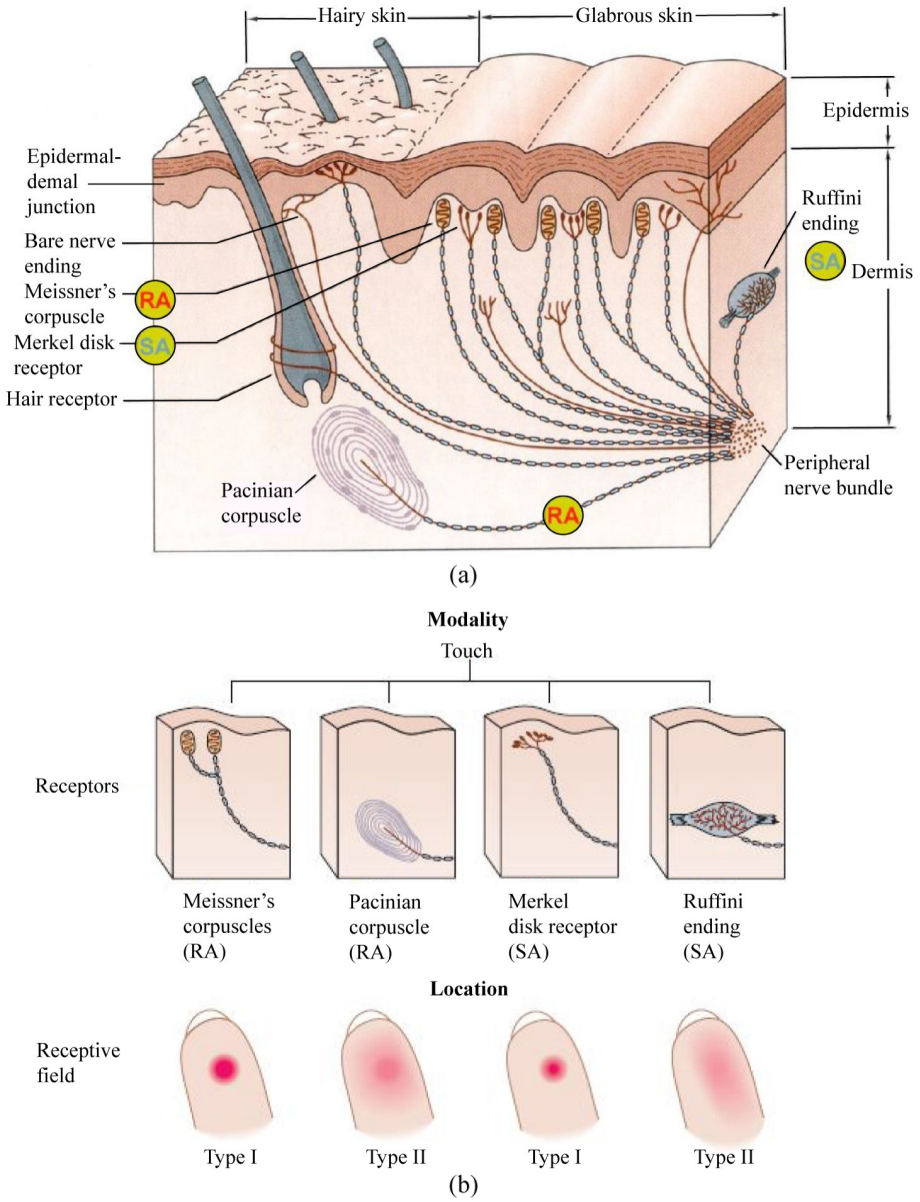

FIGURE 2.3: Schematic of skin mechanoreceptors and their receptive field sizes [Ding and Bhushan, 2016].

The sense of touch is a complex system that combines different sensory receptors connected by nerves to the somaesthetic areas of the brain. Its precise mechanism is not yet well understood. Fingertips are particularly sensitive areas, with a high density of receptors [Johansson and Vallbo, 1979]. For the area of interest in the present

work, the glabrous skin (without hair), 4 types of mechanoreceptors are observed which are presented in Figure 2.3

Each type of mechanoreceptor is associated with a type of nerve ending [Bolanowski Jr et al., 1988], characterized by its receptive field size (small = type I or large = type II), and by its adaptation speed (Slow-Adapting = SA or Fast-Adapting = FA). The following table 2.1 summarizes these categories.

Adaptation speed	Receptive field size	
	Small	Large
Slow	Slow-adapting type (I SA I) Merkel ($< 5Hz$)	Slow-adapting type II (SA II) Ruffini
Fast	Fast-adapting type I (FA I) Meissner (approx. 5 to $40Hz$)	Fast-adapting type II (FA II) Pacini (approx. 40 to $400Hz$)

TABLE 2.1: Classification of mechanoreceptors [Lederman and Klatzky, 2009]. The nerve ending associated with each mechanoreceptor type is detailed as well as its frequency bandwidth.

The receptive field size describes the ability of the mechanoreceptor to spatially segregate nearby stimuli (type I) or to sum them up (type II). As we can see in Figure 2.3, type I receptors are located near the surface of the skin, whereas type II receptors are located more deeply, which impacts their spatial resolution. The adaptation speed characterizes the ability to respond to fast strain variations of the skin (RA), or rather to detect sustained deformations due to constant pressures (SA). Each type of mechanoreceptor therefore has its own particularities, listed below:

Merkel disk receptors - SA I

Merkel disk receptors are slowly adapting mechanoreceptors at the base of the epidermis. Their function is still not well understood. They capture information about pressure, position and static characteristics, such as shapes and edges. They are the receptors that are most sensitive to low frequency vibrations, below 5Hz.

Ruffini endings - SA II

Ruffini endings detect continuous pressure and stretching of the skin. They are slowly adapting fibers with a larger receptive field than Merkel disk receptors. There is much debate about their role. Perhaps they are involved in the position control and movement of the finger. They also act as thermoreceptors.

Meissner's corpuscles - FA I

Meissner's corpuscle are located in the upper part of the dermis. They react to light stimulation and pressure, this is why they are concentrated in areas that are sensitive to gentle contact, such as the fingers or lips. They can determine where and when

the skin has been touched. Thanks to these properties, they are the mechanoreceptors that are used to read Braille. Meissner's corpuscle are fast-adapting receptors sensitive to vibrations between 5 and 40 Hz. .

Pacinian corpuscles - FA II

Pacinian corpuscles are located in the lower layer of the dermis. They are sensitive to vibrations and also play a role in the perception of pain. They are very rapidly adapting receptors. They only respond to sudden disturbances and not to continuous stresses. They are sensitive to vibrations from 40 to 400 Hz, with an optimal sensitivity at 250 Hz. While actively exploring a texture, Pacinian and Meissner's corpuscles are the most important mechanoreceptors to perceive its characteristics (roughness for example).

Mechanoreceptor activation

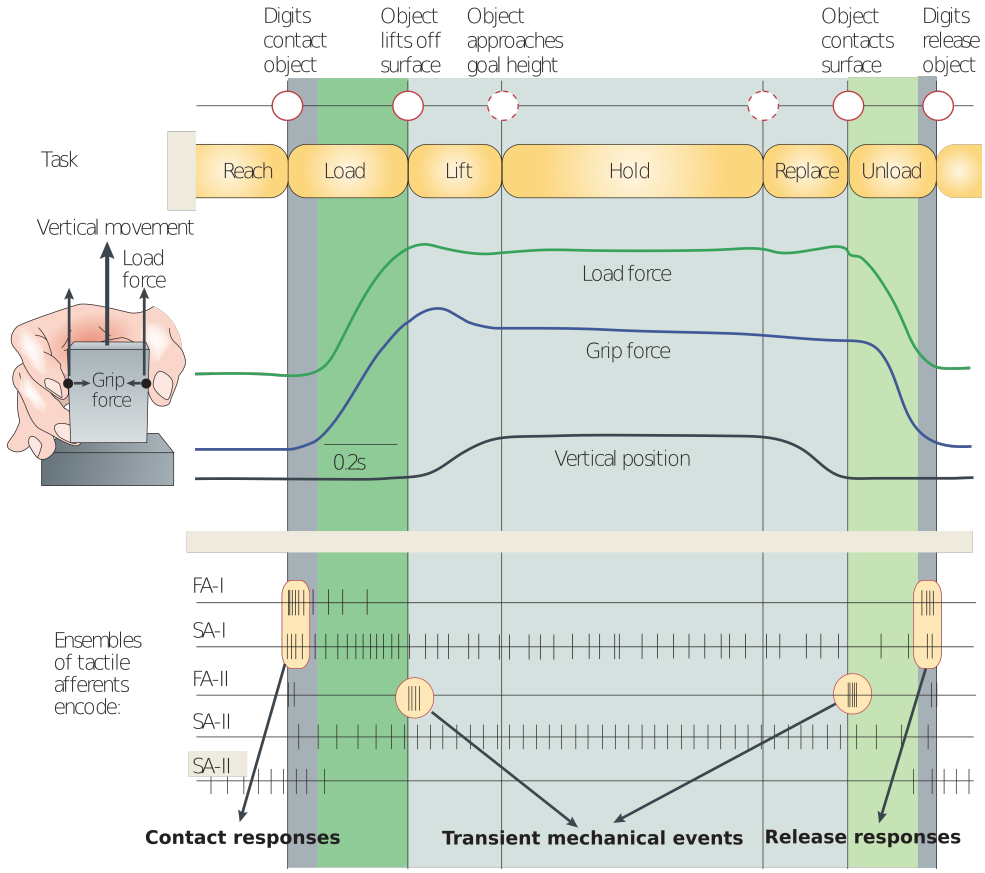

FIGURE 2.4: Activation of the 4 types of mechanoreceptors while a subject grasps an object, lifts it up and replaces it. Figure from [Johansson and Flanagan, 2009].

A technique called microneurography uses electrodes inserted into the skin to measure the nerve fiber activity and evaluate the afferent firing rates

[Vallbo and Hagbarth, 1968]. Figure 2.4 is a typical illustration of the mechanoreceptors activation during a classical object manipulation task. The subject grasps an object, lifts it up and replaces it. It shows that the FA fibers are activated during transition between the different phases of the task, when strains change, whereas the SA fibers respond continuously during the stimulation. For texture encoding, the involvement of each afferent varies with the size of the asperities and the speed of exploration [Weber et al., 2013].

Type C fibers are others noteworthy afferents. They optimally respond to gentle skin strokes by objects around 32°C, like hand caress [Ackerley et al., 2014]. They are associated with pleasant stimuli, and are of great importance in affective and social touch.

2.1.3 Haptics

Haptics is the combination of two perceptual systems: the cutaneous perception, or tactile perception, presented above and the kinesthetic perception [Lederman and Klatzky, 2009]. Cutaneous perception provides information on the deformations of the upper layer of the skin, while kinesthetic perception provides information about our spatial environment and our body position. It results from active exploratory movements of our hand in contact with objects. It gathers information from the skin, but also from all the muscles, joints and tendons of the hand and arm.

As we can see in Figure 2.5, the combination of cutaneous and kinesthetic perceptions makes it possible to get a lot of information about the properties of an object being held, such as its weight, volume, shape, hardness, temperature, and the characteristic that will interest us most in this thesis: its texture.

2.2 Audio and tactile perception of elementary stimuli

I previously explained how human auditory and tactile sensory systems capture and process stimulations from the environment. I will now describe how these stimulations are perceived. Psychophysics quantitatively investigates the relationship between physical stimuli and the perceptions they produce. In this research field, the human being is seen as a black box that we try to model by linking properties of the stimulus to the subject's responses. This section presents parallels and differences between psychoacoustics and psychohaptics, through psychophysical experiments on hearing and touch.

2.2.1 Perception of pure tones

Auditory perception threshold

Primary psychoacoustic studies investigated the perception of pure tones: sinusoidal signals at given frequencies and amplitudes. First, we need to know the

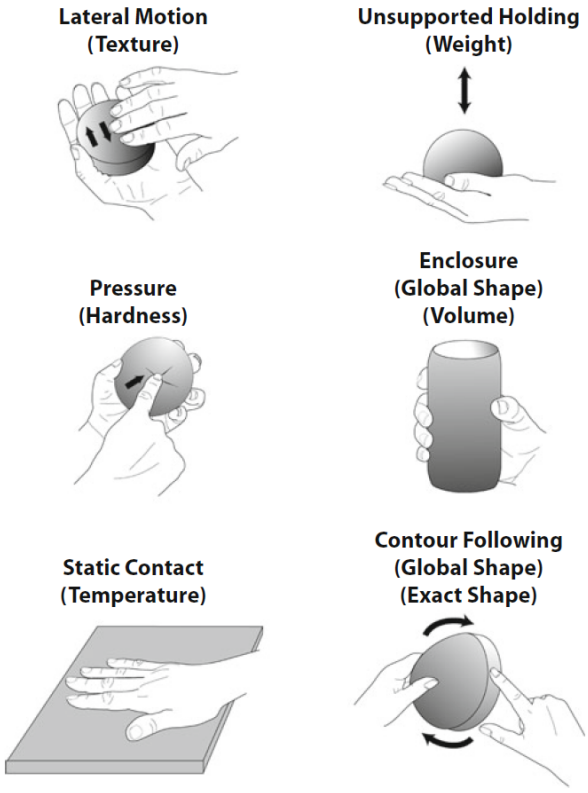

FIGURE 2.5: Description of the six haptic properties of an object associated with their manual exploration procedures [Lederman and Klatzky, 2009].

detection threshold linked to the intensity of the audio signal, i.e. the amplitude level above which an audio signal is perceived. This amplitude threshold varies according to the frequency. We can therefore construct the auditory threshold curve, presented in dashed lines in Figure 2.6 which is averaged over the normal hearing population. Sounds below this curve are too weak to be noticeable. The minimum threshold level of the curve is at approximately 3500 Hz, and corresponds to the frequencies for which the ear is the most sensitive. The detection thresholds increase at the extremities of the audible frequency bandwidth, i.e. around 20 Hz and 20 kHz. Auditory threshold measurements are for example used clinically to characterize a person's hearing loss. Audiograms are often performed during medical check-ups

Equal-loudness contour

For sounds above the threshold of hearing, it is interesting to know how strongly they are perceived. This subjective quantity is called the loudness. The unit of loudness is the phon. By definition, 1 phon corresponds to the perceived loudness of a reference sound at 1000 Hz and 40 dB SPL in free field. A 2 phon sound will be perceived as twice as loud as this reference sound. Like auditory thresholds, the loudness varies as a function of the frequency. We can then measure the equal-loudness contours shown in Figure 2.6. Two points on the same curve will be perceived with

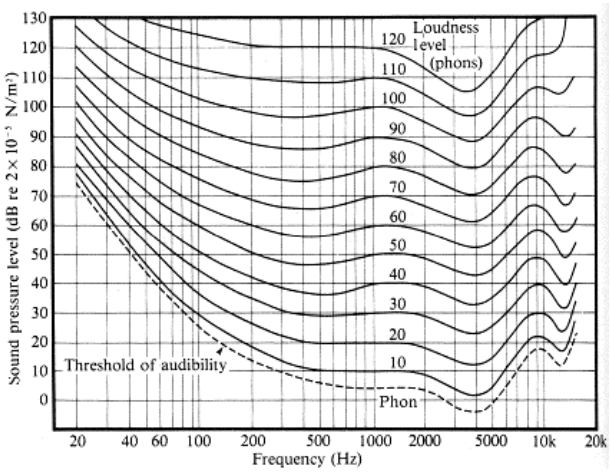

FIGURE 2.6: Auditory threshold curve and equal-loudness contours of Fletcher and Munson [Fletcher and Munson, 1933]

the same loudness even if they differ by their amplitude and frequency. While the loudness of pure tones is easily predictable, the loudness of more complex sounds is still not well understood [Moore et al., 1997].

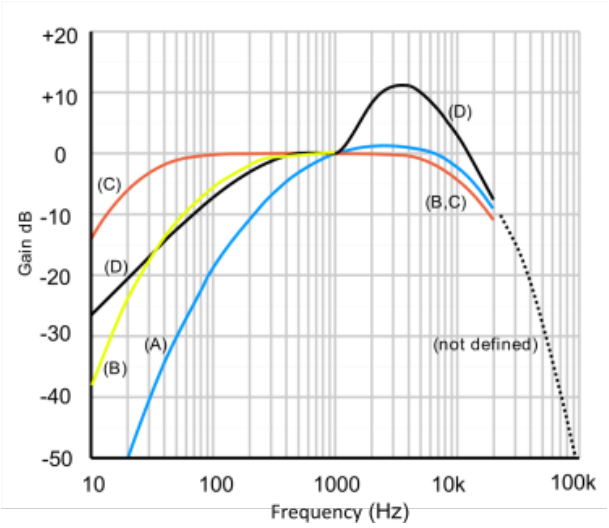

FIGURE 2.7: A-, B-, C- and D-weightings curves across the audible frequency range defined by international standards. Adapted from *Wikipedia*.

Weightings curves

When measuring sound levels, we obtain sound pressures expressed in dB. However, as presented previously, we are not equally sensitive to all frequencies. Equal-loudness contour can be inverted to define weighting curves, presented in Figure 2.7. The weighting curves are applied to the instrument-measured sounds to take into account human hearing sensitivity. They are used, for example, in measurements of

environmental and industrial noise to assess potential hearing damage or discomfort. Historically, the first weighting was the A-weighting (dBA). Today, international standards recommend the use of other weights for louder sounds.

Similar corrections will be used for haptic stimuli in the experiments presented in this thesis.

2.2.2 Tactile perception of vibrations

Studies have evaluated perceived intensities in the case of touch, also measuring threshold curves and iso-intensity curves [Verrillo et al., 1969]. These measurements presented in Figure 2.8 were performed with a vibrator in contact with the finger that excited the skin surface at various frequencies and amplitudes. The tactile curves present the same shape as the audio ones, with a much more limited frequency range. The perception of vibration is optimal around 250 Hz and decreases until it reaches its limits around 20 Hz and 800 Hz. Tactile iso-intensity curves follow the same trend.

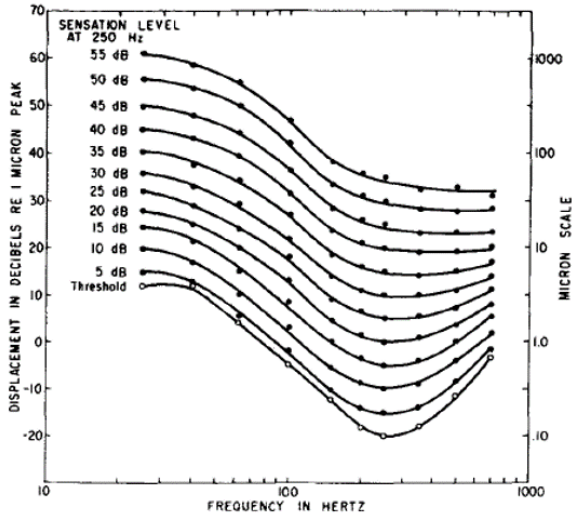

FIGURE 2.8: Detection thresholds curve and iso-intensity curves for vibrations on the finger [Verrillo et al., 1969].

2.2.3 Audio and tactile detection thresholds comparison

Now that the detection thresholds of auditory and tactile stimuli at various frequencies have been described, it would be interesting to compare the sensitivities of the two modalities [Merchel and Altinsoy, 2019]. The comparison is not straightforward because the two sensitivities are not measured with the same physical quantities. Auditory thresholds are evaluated in terms of acoustic sound pressure (in Pa) whereas tactile thresholds are assessed in terms of skin displacement in presence of a vibrator (in μm).

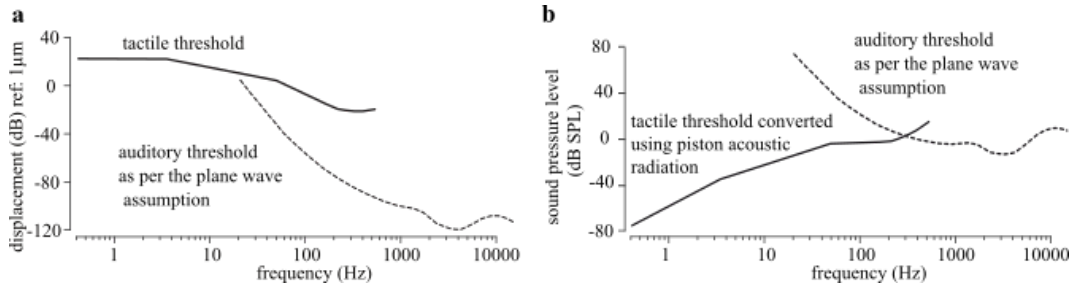

FIGURE 2.9: Comparison of the detection thresholds of auditory and tactile stimuli according to two comparative methods [Hudin and Hayward, 2020].

A first comparison method consists in transforming the sound pressure of the auditory thresholds into eardrum displacements. Thanks to Euler's equation, we can relate the pressures of the acoustic wave to the displacements of the air molecules, which are considered identical to the eardrum displacements. This leads to the threshold curves presented in Figure 2.9.a. Unsurprisingly, we observe that the eardrum is largely more sensitive to the amplitude of the vibrations than the skin. A second method [Hudin and Hayward, 2020] evaluates the acoustic radiation of a piston of a given diameter that vibrates on the skin. The displacement of the vibrator can hereby be linked to the acoustic pressure that propagates to the ear when the vibrator is assumed to be 50 cm away. The result is presented in Figure 2.9.b. With this comparison method, touch appears as more sensitive than the ear in the low frequency range. This can be illustrated by the fact that small objects (like bugs) that move slowly on our arm are detected mostly through touch.

2.2.4 Auditory and tactile perception at different scales

An intriguing phenomenon in sound perception is the transition from rhythmic perception to pitch perception. If we listen to a sound consisting of a pulse train at low frequencies, below 24 pulses per second, we hear the pulses independently of each other. Above about 30 pulses per second, we begin to perceive a pitch. At more than 150 Hz, the perception of the rhythm fades out and we only perceive a pitch, with a certain roughness. Thus there is a transient frequency bandwidth, below which we perceive discrete events independently, and above which this succession of close events is perceived as a continuous stimulus.

This phenomenon is also found in tactile perception. The large asperities are perceived as independent events (relief), while the finer asperities are perceived as a continuous texture.

This duality could be explained by the *duplex theory of tactile perception* [Lederman, 1974]. It states that the perception of fine textures (high frequencies) would be more specifically due to the vibrations created by the finger exploration, i.e. temporal cues [Cascio and Sathian, 2001], while coarse textures (low frequencies) would be perceived rather by their geometric properties that

induce deformations of the skin, i.e. spatial clues. This theory is supported, for example, by the fact that the finger, when stationary, can feel the relief but not the roughness [Hollins and Risner, 2000, Fagiani and Barbieri, 2016]. We would then have two types of texture encoding, one spatial for relief with Merkel corpuscles (SA I) and one temporal for fine structures with Pacinians (FA II). However, more recent studies focusing on mechanoreceptors activation have shown that spatial and temporal encoding cannot be so easily segregated [Weber et al., 2013].

2.3 Audio and tactile synthesis

The previous section described how auditory and tactile sensory systems perceive very basic stimuli. The perception of our environment is however much more complex. A theory, the ecological approach to perception, stipulates that we perceive our environment through invariant structures contained in the stimuli, and that we can recognize stimuli thanks to these invariants. This approach was firstly proposed by [Gibson, 1979] for the sense of vision, and was later extended to the sense of hearing by [Warren and Verbrugge, 1984] and [McAdams and Bigand, 1993]. Two categories of invariants were defined: *structural invariants* that characterize the object and its physical properties, and *transformational invariants* that describe the action exerted on this object that produces the sound, such as scratching on a wood plate. [Gaver, 1993] extended this theory by defining environment listening as the experience of listening to sound-producing events rather than sounds themselves. Thus, we naturally listen to sounds in order to identify the underlying objects and actions that interact. This concept is suitable to describe audio-tactile perception when we explore an object, because sounds and vibrations arise from the same source caused by the finger movement. Therefore, interaction sound and vibration synthesis is also related to this approach. In the present section, I will present various works that propose synthesis methods of sound and touch.

2.3.1 Interaction sound synthesis

Interaction sounds are sounds produced by applying an action with a certain object on another object. Such sounds are particularly interesting for this thesis because they include sounds that can be generated by a finger rubbing a surface. To create realistic sounds, two main synthesis approaches exist. Synthesis by physical modelling consists in modelling and simulating the physical behaviour of an object, usually a musical instrument, to generate the waveform of the sounds it produces. Another approach is synthesis by signal modelling which consists in imitating the recorded signal through an analysis-synthesis approach. First the sounds that we want to imitate are recorded. Then, the important temporal or spectral structures can be identified through signal analysis. Listening tests further enable

to identify perceptually relevant sound structures responsible for the evocations induced by these sounds. These signal structures that are responsible for the evocation (for instance the metallic aspect or the scraping action) are considered as invariants. To render interaction sounds, the analysis-synthesis approach was mainly applied. [Conan et al., 2012] proposed to render rubbing and scratching sounds using a subtractive synthesis based on noise with a certain impact density, low-pass filtered with a cutoff frequency proportional to the finger velocity. Perceptive experiments revealed that a noise with a high impact probability generates rubbing sounds, whereas more sparse impacts produce scratching sounds. These signals can be processed by resonant filter banks to render the illusion of scratching or rubbing on a plate of a certain material, like wood or metal [Conan et al., 2014].

A similar sound synthesis method was developed with a different noise generation by [Van Den Doel et al., 2001]. They noticed that interaction sound spectra generally present a fractal noise, a noise whose power spectrum linearly decreases as the frequency increases (on a log scale) with a $1/f^{\alpha}$ relationship. For example, pink noise is typical fractal noise for $\alpha = 1$. We will see later that fractal noise is also an important characteristic of the vibrations that propagate into the finger during tactile interaction.

2.3.2 Sound texture synthesis

Sound textures are the collective result of many similar acoustic events. They are stationary and are often encountered in the environment. A texture sound analysis and synthesis algorithm was developed by [McDermott and Simoncelli, 2011] based on the cochlea processing. The analysis consists in decomposing a typical recorded sound into 30 frequency bandwidths, mimicking the critical bands in the cochlea [Patterson et al., 1992]. Signal envelopes are then processed by a second filter bank, to perform several statistical measurements: the marginal moments of a signal and correlations between two filtered signals. The synthesis consists in starting from white noise and iteratively imposing the previously measured statistics of the noise. This procedure works well to synthesize a large variety of sound textures, like wind, waves, insects, applause, fire... What is interesting about this approach is that it also provides a better understanding of how the auditory system works. It suggests that sound texture perception is mediated by relatively simple statistics of early auditory representations.

2.3.3 Tactile synthesis via a stylus

Usually, we explore objects in our environment directly with our hands and fingers. Sometimes, we touch things through tools. For example, we can use a fork or a spoon to get clues about the cooking of a food. When exploring a surface with a tool, we feel the surface and not the tool [Katz, 2013]. Many studies have been performed to investigate in depth how texture characteristics, like roughness, are

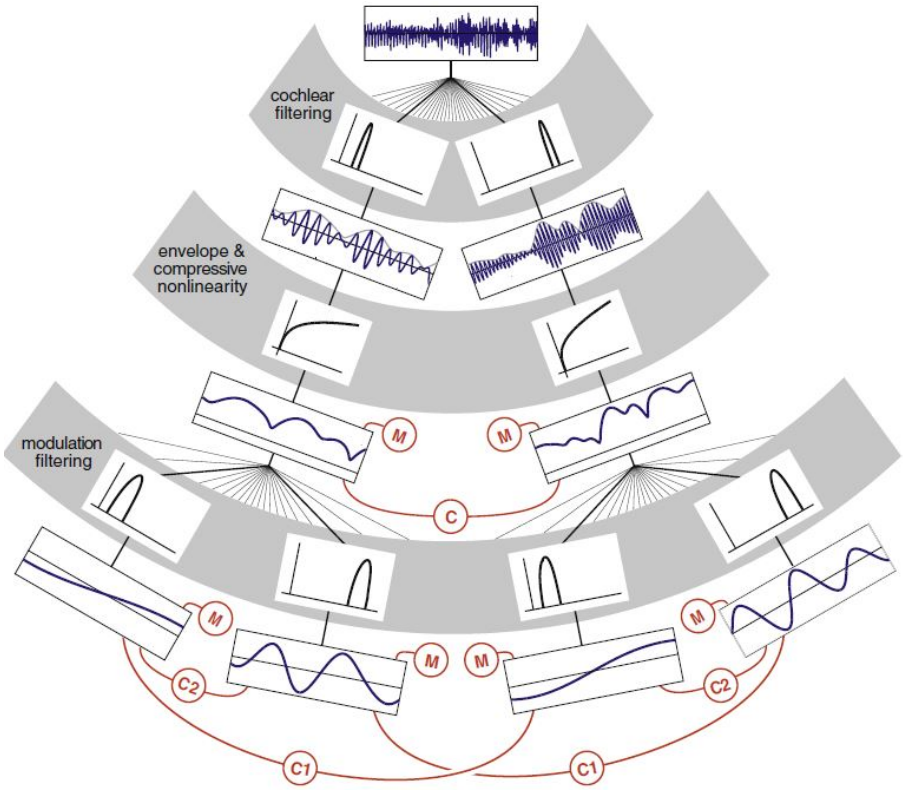

FIGURE 2.10: Cochlear filter bank for sound analysis and synthesis [McDermott and Simoncelli, 2011]

perceived through a probe and how the probe properties influence their perceptions [Klatzky et al., 2003, Klatzky and Lederman, 2002, Lawrence et al., 2007].

Tools also appear as an efficient way to recreate realistic textures. Texture vibrations can easily be recorded while the user drags an instrumented stylus on a surface with a certain force and displacement speed. The use of an intermediate tool avoids the need of recreating the fine skin deformations during tactile exploration, which is still a technological issue. A large collection of 100 various real textures (such as paper, plastic, fabric, metal...) has been recorded [Culbertson et al., 2014]. After the recording, the same stylus equipped with an actuator, can render the texture signal. The challenge is to map the vibrations recorded for certain force and speed conditions to the user's actual force and speed in real time [Romano and Kuchenbecker, 2011] [Culbertson et al., 2012]. Whereas the velocity-dependence of the vibrations is crucial, the consideration of force may not affect the realism of the rendered texture [Culbertson and Kuchenbecker, 2015].

2.3.4 Audio-haptic synthesis

Exploring a texture with a stylus also produces sounds. In order to enhance the tactile rendering presented above, [Lu et al., 2020] proposed a method to record and resynthesize sounds produced by tool-surface interaction on various materials

(Video). The challenge is still the same: the exploratory movements during the rendering are not the same as those for which the sounds were recorded. Their method decomposes the signal in short time windows during which the stylus velocity is constant. The windows are then analyzed using wavelet transform to exhibit a set of parameters. During the rendering, the parameters corresponding to the current velocity are processed to reconstruct grains of sound using inverse wavelet transform that are then played successively.

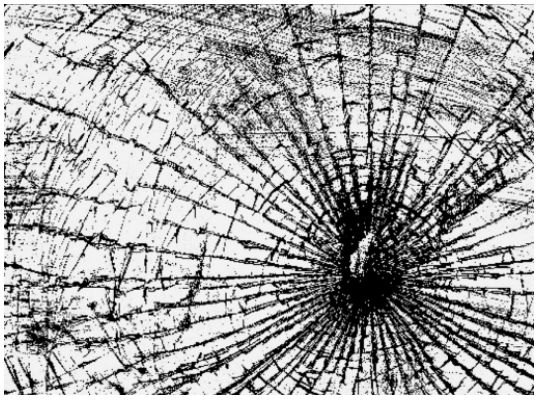

FIGURE 2.11: Example of texture map to synthesize audio and haptic signals, [Rocchesso et al., 2016].

Instead of recording and replaying sounds and vibrations, another approach of audio-haptic synthesis consist in creating signals from a texture map, like the greyscale image in Figure 2.11. [Del Piccolo et al., 2015] proposed an algorithm, called Sketch-a-Scratch, that uses a micro-impact synthesizer based on a physical model to generate sounds and vibration from a graphical map in response to exploratory gestures (Video). [Chan et al., 2021] developed an audio-haptic synthesizer with a similar method, presented in Figure 2.12. Virtual objects are encoded at different scales. At the macro level, their shapes are presented by a polygonal mesh. Meso features, like surface height, appear in the displacement map. Micro features are encoded as fractal noises in the roughness map. When the object is explored, a real-time, rigid body physical simulation based on the displacement and roughness maps synthesizes the sounds and vibrations. This kind of audio-haptic synthesis is of great interest when interacting with the environment in virtual reality.

2.4 Texture rendering on haptic surfaces

This section presents the different haptic surface technologies and explains in details the ultrasonic friction modulation. I also describe several studies on the perception of friction-modulated textures.

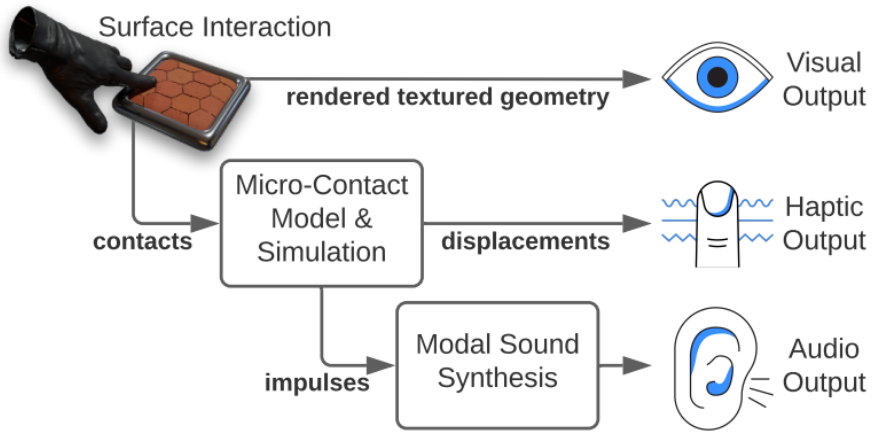

FIGURE 2.12: Audio-haptic synthesis principle for texture interactions in virtual reality proposed by [Chan et al., 2021].

2.4.1 Lateral force illusion

While exploring an object with our finger, information from the mechanoreceptors is processed by the brain. The final percept can be biased at this stage. It is indeed possible to use some properties of haptic perception to recreate the illusion of touching a relief on a flat surface. [Minsky et al., 1990] were the first to show that variations in lateral forces collinear with the movement can give the subject the illusion of exploring bumps and holes. This was demonstrated with a force feedback joystick by applying this principle: the user experiences a force opposing his movement when moving up the bump, and a force helping his movement when moving down. In general, it is possible to create an illusion of relief by applying a lateral force proportional to the height gradient of the relief. It has been shown that the lateral force is even more important perceptually than the shape [Robles-De-La-Torre and Hayward, 2001]. This method also recreates the illusion of exploring rough textures, seen as a succession of small bumps.

In this thesis, I am particularly interested in tactile synthesis technologies that allow direct exploration of texture by the fingertip, without intermediate items. This field of research is called surface haptics. It is possible to apply the above principle to haptic surfaces by varying the frictional force, i.e. the lateral force that opposes the movement of the finger in contact with the plate, to recreate the illusion of touching reliefs or textures. In this manuscript, I will often refer to the dynamic coefficient of friction, noted μ, defined as the ratio between the tangential force, opposed to the displacement when the finger is steadily sliding, and the normal force at the contact: $\mu = F_t / F_n$.

2.4.2 Haptic surface technologies

Various technologies make it possible to modify the friction on a surface to provide haptic stimuli. Such technologies are described below:

Vibrotactile technology

Vibrotactile technologies are the most widely used to provide tactile feedback. They consist in making objects vibrate to create wave propagation through the finger, at frequencies that are well perceived by the user, i.e. between 50 and 500 Hz. These vibrations are usually generated by vibrating motors such as the *Eccentric Rotating Mass vibration motor (ERM)*, which puts an unbalanced mass into rotation, or the *Linear Resonant Actuator (LRA)*, which works like a loudspeaker and uses the resonance of a mass-spring system to provide high vibration amplitude with low energy. Vibrations can also be generated by means of piezoelectric ceramics, thanks to their ability of deformation under an electric field.

Vibrotactile feedback are found in many everyday objects, especially in smartphones [Poupyrev and Maruyama, 2003, Chen et al., 2011]. They are usually very basic, but they can still recreate fine and precise haptic sensations with well-designed vibrotactile interfaces [Visell et al., 2008]. Indeed, friction modulations can be recreated by producing lateral vibrations of a plate explored by the finger. When the plate moves in the opposite direction to the finger movement, this causes a brief increase of the coefficient of friction, and vice versa. By linking vibrations to the finger position, it is possible to render sensations of textures [Wiertlewski et al., 2011b]. This technology is limited to high-frequencies and cannot to provide a constant friction level, which is for example necessary to create large bumps.

Friction modulation with ultrasonic vibrations

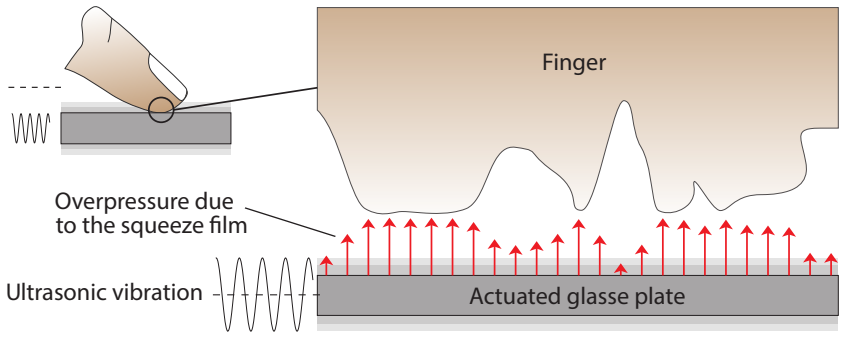

FIGURE 2.13: Illustration of ultrasonic friction modulation. The overpressure of the squeeze film created by the vibrations lifts the finger asperities of the plate and reduces the friction. From [Huloux, 2021]

Ultrasonic friction modulation is the technology that I used in the work presented in this thesis. By means of piezoelectric ceramics, the glass plate above the

screen vibrates at one of the frequencies of its natural resonance modes. The frequency used is higher than the audible range, generally in the order of 25 to 45 kHz depending on the dimensions of the plate and the desired resonance mode. The maximum peak-to-peak vibration amplitudes at the antinodes are typically from 2 to 6 μm. These vibrations are not perceptible by the finger because the frequencies are too high to be detected by the mechanoreceptors of the skin, but they cause a drastic reduction in the frictional forces between the skin and the glass.

Two hypotheses have been emitted to explain this phenomenon. The first work which highlighted it [Watanabe and Fukui, 1995] suggested that the vibrations generate a *squeeze air film effect*. An air cushion would appear between the finger and the plate, reducing the contact area and thus the friction. Another possible explanation was the intermittent contact theory [Vezzoli et al., 2017]. Due to the vibrations, the fingertip would bounce off the glass, reducing the effective contact duration and thus the friction. It has since been demonstrated by using stroboscopic imagery that the reality is a combination of these two theories: the finger bounces on a cushion of air [Wiertlewski et al., 2016], as presented in figure 2.13.

Many studies have shown how it was possible to apply ultrasonic friction modulation to touchscreens to enhance them with haptic feedback [Biet et al., 2007, Winfield et al., 2007]. The start-up hap2U develops haptic touchscreens with this technology.

Friction modulation with electroadhesion

Another technology to modify the coefficient of friction is based on electroadhesion. It consists of generating an electrostatic attraction force on the fingertip which increases the coefficient of friction [Meyer et al., 2013]. This is accomplished by coating a glass plate with a thin electrically conductive layer and then an insulating layer. This combination is used in conventional capacitive touchscreens to measure the finger position. In the case of electroadhesion, a high voltage is sent to the conductive layer. In the vicinity of the charged layer, the fingertip, which is separated from the charged layer by the insulating layer, becomes polarized. The generated electrostatic force attracts the finger towards the screen, increasing the frictional forces. Friction can therefore be controlled by modulating the voltage to create various haptic effects [Shultz et al., 2018]. The startup Tanvas develops touchscreens with this technology.

A more detailed comparison of the haptic surface technologies can be found in the review [Basdogan et al., 2020].

2.4.3 Texture recording and rendering on haptic surface

The different technologies presented previously provide friction forces modulations. A first step in texture reproduction is therefore to study how these lateral forces vary when we explore real textures. By measuring the lateral vibrations and the finger

position, it is possible to reconstruct spatial signals of the textures, whose spectrum can be analyzed.

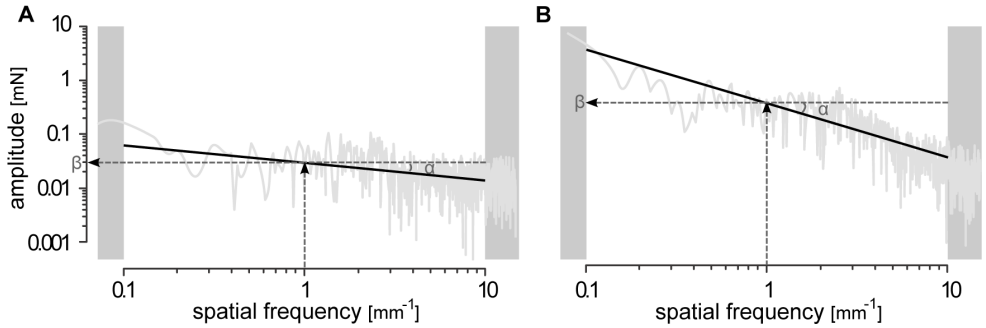

FIGURE 2.14: Spatial spectra of paper (A) and sandpaper (B) from [Klöcker et al., 2013]. Both spectra present fractal noise with different parameters.

In line with interaction sounds, all the texture spectra present a background fractal noise: a noise whose spectral density decreases with the frequency according to the law: $S = \frac{\beta}{f^{\alpha}}$ with β and α characteristic parameters of the texture [Wiertlewski et al., 2011a] as presented in figure 2.14. These parameters also appear to be correlated with texture pleasantness judgments [Klöcker et al., 2013]. In addition to the fractal noise, the spatial spectrum of the texture can also present harmonics which correspond to the texture regularities [Janko et al., 2015].

Once the lateral force signals have been recorded, it is possible to replay them on a haptic interface [Wiertlewski et al., 2011b, Messaoud et al., 2016b, Meyer et al., 2016]. The synthetic textures created in this manner are recognizable but still remain different from real textures. These technologies do not allow to modify the compressibility and the thermal properties of the surface. Rendering realistic textures remains a challenge.

2.4.4 Technological challenges within surface haptics

To increase the realism of texture rendering on haptic surfaces, one research area focuses on the precise control of the frictional forces. While exploring a surface with the finger, many factors indeed influence the mechanical behavior of the contact. They cause a lot of friction variability, even for the same subject trial after trial. The main factors influencing friction are the normal force of the finger [Tomlinson et al., 2009], the touch angle, the velocity [Pasumarty et al., 2011], as well as the temperature [Choi et al., 2021] and the moisture [Li et al., 2020].

A method used to overcome friction variability is to perform closed-loop control of the lateral force [Huloux et al., 2018]. The friction coefficient is measured and the feedback loop regulates the vibration amplitude in real time. It makes it possible to render texture signals with a high accuracy [Grigorii et al., 2019]. This method however requires high-precision force sensors and can only be achieved while the

finger is sliding and producing lateral frictional forces. Yet, the initial movement, when the finger starts sliding, is crucial for texture perception [Grigorii et al., 2019]. It could also be possible to develop control strategies based on the contact area between the finger and the glass plate (acquired with fast imaging) or based on the contact impedance (measured through the perturbations of the plate vibration) because these physical parameters are closely related to friction [Huloux et al., 2021a, Huloux et al., 2021b].

Another challenge in surface haptics is the ability to produce localized stimuli. The aim is to independently stimulate multiple fingers in contact with the surface. Many technological approaches have been developed. [Hudin et al., 2015] proposed to create localized tactile stimuli through time-reversal wave focusing. [Dhiab and Hudin, 2019] showed that it was possible to spatially confine vibrotactile stimuli in narrow plates for vibration frequencies below a certain cutoff frequency. [Pantera and Hudin, 2019, Pantera and Hudin, 2020] proposed to produce multitouch vibrotactile feedback using the inverse filter technique. The glass plate is first characterized by measuring the transfer functions between each actuator and each possible finger position. The matrix of transfer functions is then inverted to address each finger position independently. This method is capable of delivering different stimuli to 6 fingers at the same time to enable visually impaired people to perceive messages through a type of Braille [Pantera et al., 2021]. Another promising method uses metamaterials, such as phononic crystals, to create vibration boundaries and wave guides in the plate [Daunizeau et al., 2021]. This technology can render different ultrasonic friction modulated stimuli on distinct areas of the surface.

2.4.5 Perception of friction

The friction coefficient typically varies from 0.1 for very slippery surfaces, to around 2 for very sticky textures.

Haptic surface technologies provide an easy way to generate various friction levels on the same surface, from 0.1 for very slippery surfaces, to approximately 2. This makes it a great tool to study human perception of friction. In addition, we need a detailed understanding of friction perception to design accurate feedback on haptic touchscreens.

This section describes the perception of smooth surfaces, with an approximately constant coefficient of friction, as well as the perception of the transitions between two friction levels. Two distinct exploration procedures are considered, the dynamic exploration when the finger moves on the surface, and the static exploration when the finger is still.

Dynamic perception of friction

First, we want to know the accuracy of the human tactile sensory system in judging that a surface is more or less slippery while exploring it actively with the finger.

Using ultrasonic friction modulation to recreate many friction levels on the same surface, it has been demonstrated by [Samur et al., 2009] that the perception of friction follows a Weber law with a Weber fraction of 18%. This means that on the whole friction coefficient scale (from 0.2 to 1 in this study), a human being is capable of perceiving a difference between two friction levels if they differ by more than 18%. For example, a friction coefficient of at least 0.236 will be distinguished from a friction coefficient of 0.2, and a friction coefficient of at least 1.18 at least will be distinguished from a friction coefficient of 1.

Dynamic perception of friction changes

The previous paragraph described the perceived differences due to various friction levels between two surfaces that we explore successively. We can also wonder how friction changes on the same surface are perceived, for example for surfaces for which one part is more slippery than the other and that can be explored in one movement. In this case, the Weber fraction drops to 11% [Gueorguiev et al., 2017], suggesting that we are more sensitive to friction transitions than to friction differences between two successive explorations. This type of sharp change in friction creates an additional sensation that can be considered as a haptic boundary feedback. The perceived intensity of such haptic feedback depends on the friction gap and the sharpness of the transition, but also on the normal force and the finger velocity [Gueorguiev et al., 2019, Saleem et al., 2019, Messaoud et al., 2016a].

Static perception of friction changes

Surprisingly, it has been shown, thanks to ultrasonic friction modulation, that it is also possible to feel a friction change even when the finger is not moving but remains static [Monnoyer et al., 2016, Monnoyer et al., 2017]. With fast imaging focused on the contact between the finger and the glass, the precise moment when the actuation makes the plate more slippery can be observed. It reveals that the fingertip stresses are suddenly released, causing micro-movements of the skin that are perceptible. Friction changes as a function of the pressure force can render haptic click feedback, as pressing a button.

Static perception of friction

Another noticeable ability of the human sensory system is linked to grabbing and lifting an object. In such cases, we apply more or less strong grip force depending on whether the object is more or less slippery, and this even before the loading [Johansson and Westling, 1984, Cadoret and Smith, 1996]. This means that skin receptors are able to evaluate the frictional state of the object's surface without any relative movement between the finger and the plate. Static friction perception could be due to patterns of finger skin deformations present from the initial

contact [Willemet et al., 2021]. Furthermore, incipient slippage could be processed by efficient encoding on characteristic strain patterns [Delhaye et al., 2016].

2.4.6 Perception of haptic textures

After we have seen how constant friction levels are perceived, we will now consider stimuli where the friction varies according to a waveform. A series of psychophysical studies specifically investigated the perception of haptic waveforms in order to better understand how this kind of haptic feedback is experienced. Amplitude detection and discrimination thresholds have been measured [Bau et al., 2010] for electroadhesion, as well as iso-intensity curves [Wijekoon et al., 2012]. Because they didn't control frictional forces, these experiments remain technology dependent and are hardly generalizable. Studied more carefully [Vardar et al., 2017a], detection threshold curves of sine and square waves follow similar trends as vibration thresholds presented in Figure 2.8.

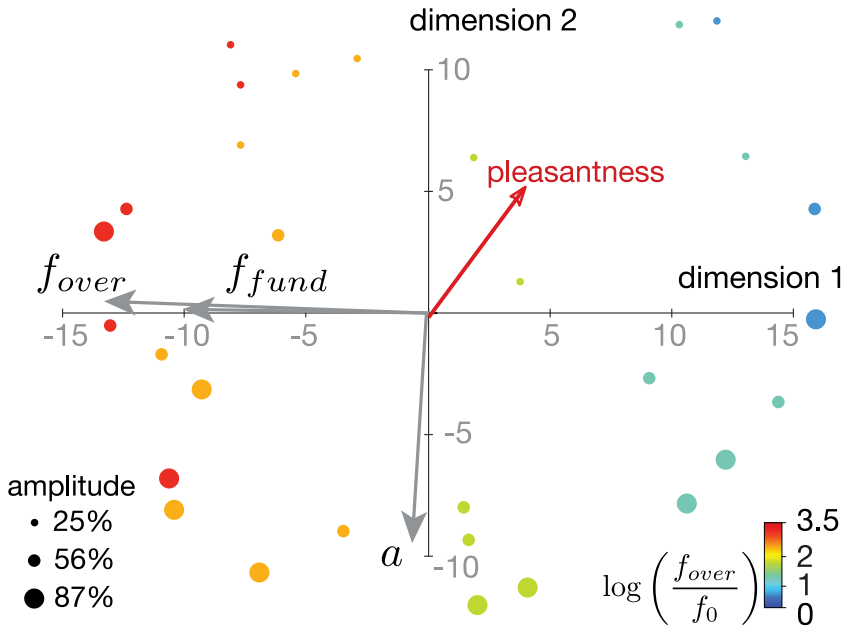

FIGURE 2.15: Representation of the perceptive space for 30 haptic textures composed of two harmonic frequencies, f_{fund} and f_{over} with various amplitudes. Two textures (displayed by color dots) that are close in this space are felt as similar. This space was obtained by a free-sorting experiment and a multidimensional scaling presented in [Bernard et al., 2018]. It was combined with pleasantness ratings that are displayed as a red arrow.

The perception of more complex haptic textures composed of many harmonics are still not well understood. Using ultrasonic friction modulation, one of my previous experiments [Bernard et al., 2018] presented in Figure 2.15 demonstrated that the highest-frequency component was perceptually prevailing in a two-harmonic grating. In the same line, the study of [Friesen et al., 2018] showed that a two spatial frequency components grating can be perceptually matched to a grating with

only one component in the mid-frequency band. They therefore proposed to design textures based on 3 parameters : spatial frequency, amplitude and irregularity. These 3 features appear as perceptually relevant to render a wide range of textures [Friesen et al., 2021]. It has also been established that the spatial spectrogram is an accurate perceptive representation to encode a haptic texture [Meyer et al., 2015].

The concept of roughness, which has been extensively studied for real textures, has also been explored for synthetic textures. Several intersecting studies [Bodas et al., 2019, Isleyen et al., 2019, Vardar et al., 2017b] revealed that the sensation of tactile roughness can be influenced by many synthesis parameters. Roughness sensation increases when the spatial period of the haptic signals decreases, when the amplitude increases, as well as when noise is added.

2.5 Multisensory integration

Human beings perceive the world through their senses (such as sight, smell, touch, taste, hearing, self-motion...) which collect and process information about their surrounding environment. Multimodal integration is the study of how information from different sensory modalities are integrated by the nervous system. This section presents first the basic principles of multisensory integration, and then focuses on on how haptics and audition interact.

2.5.1 Multisensory integration principles and illusions

When exploring our environment, plenty of objects are experienced by our senses. Therefore, the brain receives plenty of stimulation from different senses and has to decide which stimulations come from the same object or event. This categorization is based on four main cues : the modality, the intensity, the location, and the duration of the stimulus. For example, a flash of light, an intense burst of sound and vibrations of a choc wave that are perceived simultaneously and that come from the same direction are highly likely to originate from the same source and event. However, this process is not error-free and the brain can be fooled. Many biases are in favour of visual capture, i.e the visual modality dominates over the other senses. Visual capture is well illustrated by the ventriloquist effect [Alais and Burr, 2004]. Since the dummy's mouth and not the speaker's is moving, the illusion occurs and we perceive the speech sound as originating from the dummy. In this case, the visual and auditory stimuli are not congruent and the speaker's voice is captured by the vision so that the sound appears as emerging from another source. In some cases of multisensory incongruence, however, both modalities affect the final percept. This can be demonstrated with the McGurk effect [McGurk and MacDonald, 1976]. We can experience this stunning illusion by listening to a sound recording of a phoneme and, at the same time, watching a video of someone saying a different phoneme. We usually perceive another phoneme, intermediate between the two. For example, if the phoneme "ba-ba" are spoken as we see the lip movements of "ga-ga", we perceive the

phoneme "da-da" (Video). In this case, since the discrepancy between auditory and visual information remains weak, both stimuli are integrated into a single percept that is a compromise between the two perceived phonemes.

External stimuli can even alter the perception of our own body. The rubber hand illusion [Botvinick and Cohen, 1998][Ehrsson et al., 2004] provokes an illusion of body transfer toward an artificial object (Video). A subject sitting at a table places his own hand hidden from view next to a realistic rubber hand. The experimenter strokes both hands simultaneously and identically with a paintbrush. This illusion makes the subject feel that touch stimulations from his hand are coming from the rubber hand. He can even consider the rubber hand as his own hand. This effect can be demonstrated by the reaction of surprise of the subject when the fake hand is being struck. This body ownership illusion reveals interconnections in sensory processing between touch, vision and proprioception.

2.5.2 Bayesian integration

When contradictory information is provided to two senses, the observer migh either be aware of this contradiction or might experience one unified impression, usually according to the degree of discrepancy. In the second case, as described in the previous illusions, the brain either relies on one of the modalities or combines the two to create a single percept. An interesting question is related to the dominance of one sense on another and the influence of each modality when there is a compromise. These issues were investigated in the case of the perception of an object's size, through the visual and tactile modalities, by seeing it or by grasping it with the hand. The first experiments revealed that vision is strongly dominant in shape estimation in standard conditions [Rock and Victor, 1964]. However, when the visual stimulation is degraded, the unified impression appears to be more in favor of the haptic modality. Those observations led to the development of a model, called Bayesian integration, based on maximum likelihood estimation, to predict the multisensory integration outcome [Ernst and Bülthoff, 2004].

The perception of the property of an object in our environment is described by the equation:

$$\hat{S} = f(S) \tag{2.1}$$

with S the physical property being estimated, $\hat{S}$ the estimate of the property by one sensory system (vision or touch for example) and f the operation by which the nervous system does the estimation. Since the human sensory system is not 100% accurate, $\hat{S}$ is corrupted by noise. Noise is assumed to be Gaussian with a variance σ^2. This means that if a subject was asked to judge the property of the same object several times, the judgments would not be exactly the same, but would slightly vary around the average estimate. This would lead to the type of probability distributions plotted in dashed lines in Figure 2.16. In this example, a subject experiences a bar in a discrepancy situation where the height of the visual stimulus S_V is greater than

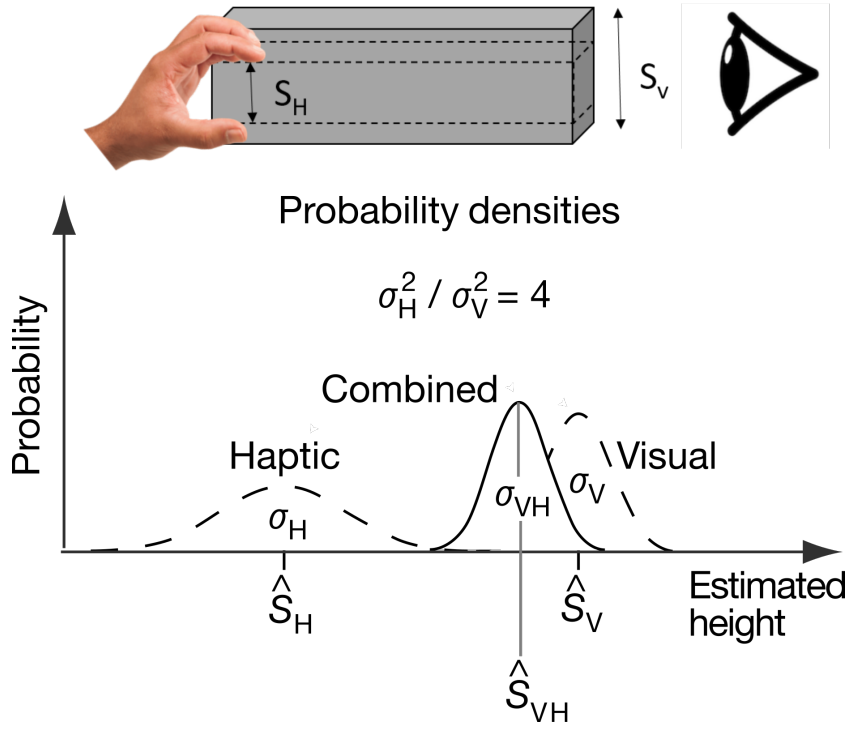

FIGURE 2.16: Description of Bayesian integration. The height of a physical object is estimated by both vision and proprioception. (Adapted from [Ernst and Banks, 2002]).

the height of the haptic stimulus S_H. Here, the visual sense is more accurate than the haptic sense, which leads to a higher haptic variance σ_H^2 and then a flatter Gaussian distribution for the haptic estimation. According to the variances, the Bayesian integration model associates weight w to each of the modalities defined by the following equations :

$$\frac{w_V}{w_H} = \frac{\sigma_H^2}{\sigma_V^2} \quad \text{and} \quad w_V + w_H = 1 \tag{2.2}$$

$$=> \quad w_H = \frac{\frac{1}{\sigma_H^2}}{\frac{1}{\sigma_H^2} + \frac{1}{\sigma_V^2}} \quad \text{and} \quad w_V = \frac{\frac{1}{\sigma_V^2}}{\frac{1}{\sigma_H^2} + \frac{1}{\sigma_V^2}} \tag{2.3}$$

These weights represent the importance of each sense in the final estimate $\hat{S}_{VH}$, given by the Maximum-likelihood estimate:

$$\hat{S}_{VH} = w_H \hat{S}_H + w_V \hat{S}_V \tag{2.4}$$

In Figure 2.16, since the variance of the visual modality is lower than the variance of the haptic modality, its weight is higher. The distribution of the visuo-haptic estimate is then shifted towards the visual estimate. This means that in the case of a discrepancy, the human sensory system performs a compromise according to the reliability of each sense. The variance of the visuo-haptic estimate is given by:

$$\sigma_{VH}^2 = \frac{\sigma_V^2 \sigma_H^2}{\sigma_V^2 + \sigma_H^2} \tag{2.5}$$

By construction, the final estimate has always a lower variance than both the visual and the haptic estimate, revealing that the combination of several sources of information provides a more precise and reliable percept. The Bayesian model of multisensory integration is relevant in many situations, also with other senses.

2.5.3 Visuo-haptic integration

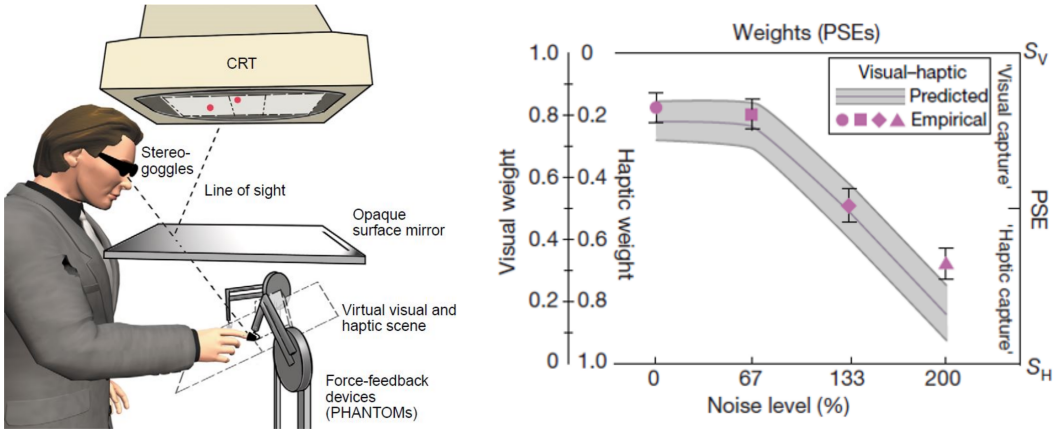

FIGURE 2.17: Illustration of the experimental setup used in [Ernst and Banks, 2002] and graph of the principal result: weights of the two modalities with respect to the visual noise level.

The Bayesian model presented previously was first demonstrated for visuo-haptic integration [Ernst and Banks, 2002]. To address the complexity of creating well-controlled non-congruent stimuli, an experimental set-up was constructed by combining a 3D optical system to render the visual stimulus and force feedback devices to haptically render virtual shapes, as presented in Figure 2.17. Subjects were asked to judge the size of a bar whose visual height was different from the haptic height, as presented in Figure 2.16. This setup allows the experimenters to blur the visual modality by adding noise to the visual background. They showed that without noise, the estimated height was more in favor of the visual height, but as the visual noise level increased, the estimated height became more in favor of the haptic height. The weights associated to each modality appear as well predicted by the theory of Bayesian Integration.

Visual and haptic sensory systems also interact in the perception of textures. [Heller, 1982] demonstrated that touch and vision alone provide comparable levels of performance in the evaluation of a texture roughness (selection of the smoothest surface), but the combination of both visual and tactual input led to greater accuracy.

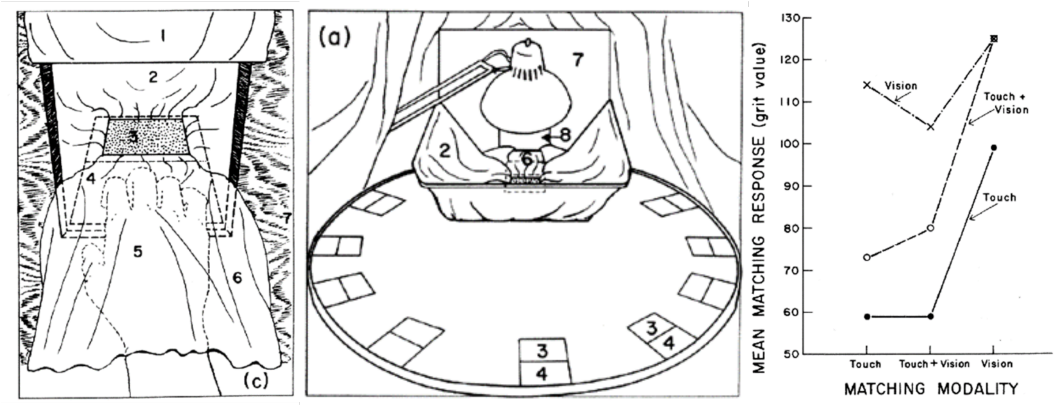

FIGURE 2.18: Experimental setup to render visuo-tactile discrepancy used in [Lederman and Abbott, 1981] and results.

Delivering dissimilar visual and tactile stimuli during the exploration of a texture with the finger is a challenging issue. [Lederman and Abbott, 1981] proposed an experimental setup where subjects could observe half of the texture while exploring the other half with their hand under a curtain, as presented in Figure 2.18. It enables the presentation of slightly different textures between the two halves without the subject noticing it. In this experiment, subjects were asked to match an incongruent standard (a sandpaper with grit value of 60 for the tactile part and 120 for the visual part) using either vision, touch or both modalities with one of the sandpapers of various grit values placed on a rotating table that allows quick switch between stimuli. The results in the case of visuo-tactile stimulation showed that the bimodal percept was a compromise equally influenced by vision and touch. This study demonstrates that the multimodal experience of exploring textures with the finger can also be described by the theory of Bayesian integration.

2.5.4 Audio-Haptic integration

Exploring a texture with the finger produces an interaction sound that is inherently congruent with the tactile stimulation. [Lederman, 1979] showed that we are able to judge the roughness of a texture only by hearing that sound. Auditory roughness estimations were similar, but not identical to judgments performed with only haptic cues. When the texture was experienced both with touch and audition, roughness estimations tended to be similar to haptic judgments, showing the preponderant weight of tactile cues in the bimodal integration in this case. However, exploring a texture not directly with the finger but with a stylus led to different results [Lederman et al., 2002]. In this case, the audio-haptic estimation of the roughness was a compromise with a weight of 62% for touch and 38% for audition. These results highlight that when the haptic sensation is degraded, with only vibrations and no skin deformations, humans integrate the touch produced auditory information in their perception of roughness.

The previous findings were exhibited for touch produced sounds that are naturally congruent with the haptic sensation. But what happens if the sounds are incongruent? [Jousmäki and Hari, 1998] showed that an alteration of the sound can strongly change the tactile perception. When rubbing our hands together, a modification of the frequency contents of the produced sound obtained by damping or amplifying frequencies above 2 kHz, produces a decrease or an increase in the perceived roughness snesation of the hands. Since amplification of high audio frequencies makes the skin of the palm feel drier, almost like a parchment paper, this effect was called the "Parchment-skin illusion".

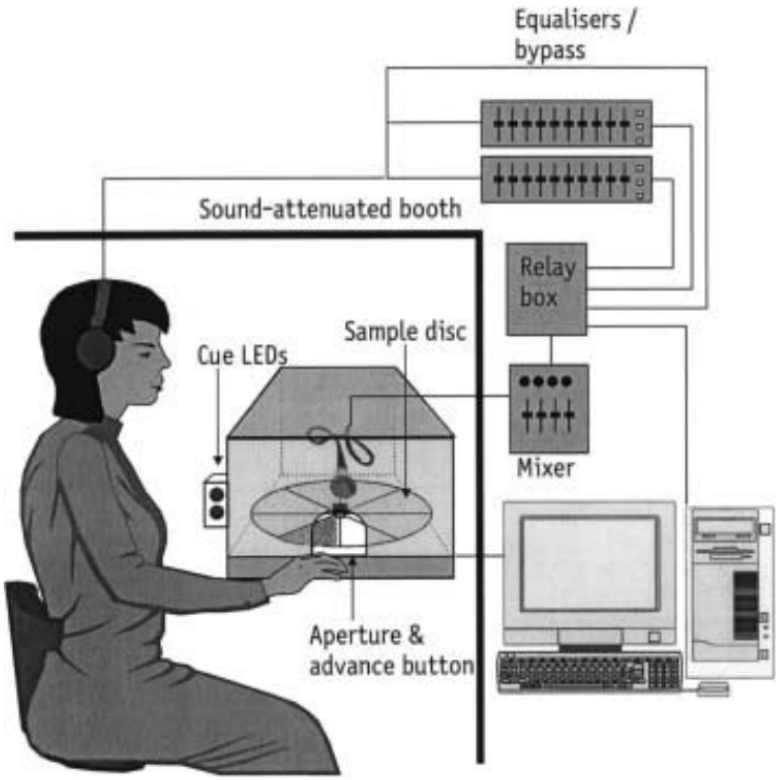

FIGURE 2.19: Experimental setup in which the frequency contents of the rubbing sounds is modified in real time to demonstrate audio-tactile interaction [Guest et al., 2002].

However, the experimental procedure was based on roughness judgements that could have been biased by the experimenter's expectations. This is the reason why [Guest et al., 2002] proposed to demonstrate this effect more properly. They used a speeded, forced-choice discrimination task. The participants were first trained to recognize two abrasive papers of different grit values, one smooth and one rough. The experimental setup, presented in Figure 2.19, randomly presents one of the two stimuli and the subject had to explore the texture and quickly identify it as the smooth one or the rough one. During the exploration, touch produced sounds were recorded and played back with earphones, while their frequency content was

modified in real time. The interaction between sensory modalities was revealed by variations in discrimination accuracy. The smooth sample was accurately classified as the smooth one more often in presence of a sound which high frequencies were attenuated, and was classified as the rough one more often with a sound which high frequencies were amplified. This experiment confirmed the trends exhibited with the "parchment skin illusion".

Audio frequency contents appear to have a strong impact on tactile perception, but other aspects of the sound also could be involved. [Suzuki et al., 2008] showed that the "Parchment-skin" effect still works if touch produced sounds are replaced by noise which gain is synchronised with finger movements, but not with a pure sound. Moreover, the effect disappears if the discrepancy between the auditory and haptic stimulations is too large in terms of intensity [Suzuki and Gyoba, 2009] and spatial localization [Gyoba and Suzuki, 2011]. Those studies demonstrate that sounds need to be congruent with the haptic stimuli to influence the tactile perception.

Multisensory integration appears to be variable between individuals. The "Parchment skin illusion" for example did not work for all the subjects. Besides, multisensory integration can evolve with aging. For example, older adults do not perform as well as young adults in segregating irrelevant auditory information in a roughness discrimination task with auditory distractors [Landelle et al., 2021].

Many studies have shown that the combination of auditory and tactile cues enhance the perception compared to conditions involving only audition or touch. For example, such combinations lead to better discrimination of materials, whether the subject are exploring the surface [Chan et al., 2021], feeling the vibrations of bouncing events on the object [De Pra et al., 2020] or even walking on the material [Giordano et al., 2012].

2.6 Auditory and haptic feedback for user guidance

Haptic and auditory feedback appears as a promising solution to improve human-computer interaction and to guide the user. User guidance is considered as means used to advise, orientate, inform and lead the user when interacting with an interface. This section presents previous works that implemented auditory or haptic feedback in an interface.

2.6.1 Performance metrics

The quantitative evaluation and comparison of different interfaces and types of feedback is not a straightforward exercise. We can measure many physical properties (like the latency or the precision), but it will not fully determine how the user handles the interface. A classical approach is to assess the performances of the interface by observing the way users execute tasks on it [Samur, 2012].

Objective metrics of user performances can be exhibited with the paradigm of Fitts' law [Fitts, 1954], a predictive model of human movement extensively used in

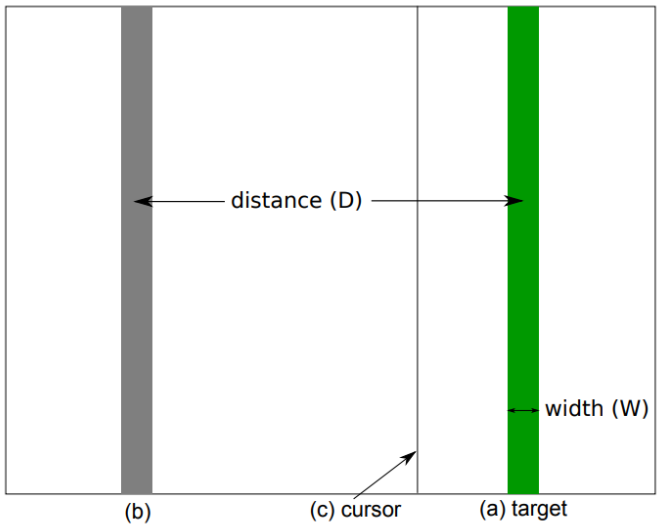

FIGURE 2.20: Visual interface for the pointing task experiment described by Fitts' Law. Subjects have to select the cursor in the grey starting position and rapidly drag it to the green target. Adapted from [Casiez et al., 2011]

human–computer interaction, ergonomics and motor control. This model describes the trade-off between precision and rapidity that humans perform when reaching a target, either to interact with the environment or with an interface. Fitts' law predicts the time required to perform pointing tasks, i.e. to move along one dimension to reach a visual target as quickly and accurately as possible, as presented in Figure 2.20. First, the complexity of the task is described by the index of difficulty (ID) as the ratio between the distance from the starting point to the target (D) and the width of the target (W):

$$ID = \log_2\left(\frac{2D}{W}\right) \tag{2.6}$$

Then, Fitts' law predicts the average time to complete the movement (MT) as :

$$MT = a + b \times ID = a + b \times \log_2\left(\frac{2D}{W}\right) \tag{2.7}$$

with a and b constants that depend on the interface. This means that the task duration increases linearly with the index of difficulty. Many gestures in our daily life can be described with Fitts' law, such as reaching a doorbell button or reaching an icon with the computer mouse. In HCI, the index of difficulty is most frequently expressed using the Shannon formulation [MacKenzie, 1992], which yields to the equation :

$$MT = a + b \times ID = a + b \times \log_2\left(\frac{D}{W} + 1\right) \tag{2.8}$$

This standardized formulation is preferred to reflect the information transmitted by performing the task (in units of bits).

This discrete model was extended to the continuous case in two dimensions by the steering law [Accot and Zhai, 1997]. As shown in Figure 2.21.a, this model predicts the time required (T) to navigate, or steer, through a 2-dimensional tunnel of variable width :

$$T = a + b \int_C \frac{ds}{W(s)} \tag{2.9}$$

with C the path parameterized by s and $W(s)$ the width of the path at s. Again the constants a and b depend on the context. The steering law is an abstraction that well describes the navigation in a computer menu with the mouse, as illustrated in Figure 2.21.b.

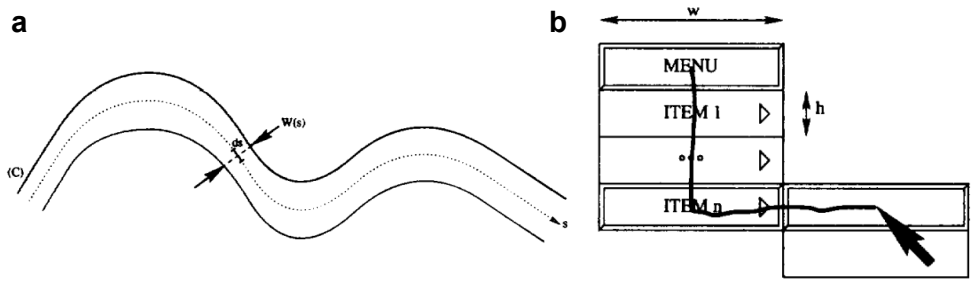

FIGURE 2.21: a. Navigation through a 2-dimensional path described by the steering law. b Example of application: navigation through a menu. Adapted from [Accot and Zhai, 1997]

.

As they depend on the interface and/or the type of feedback, the parameters a and b can stand as performance metrics. They can be measured experimentally for each condition by performing linear regressions of the completion time for various indices of difficulty. In some cases, when the task is not exactly a pointing or a steering task, the completion time and the precision error can be directly used as performance metrics.

The performance metrics presented in this section describe the entire interaction of a user operating on an interface and can be used to evaluate the benefits of providing audio or haptic feedback.

2.6.2 Haptic guidance

On touchscreens, [Levesque et al., 2011] evaluated the performance gain of adding friction-modulated haptic feedback. The haptic feedback was constructed with a low friction level at the beginning and a high friction level when the finger reached the target. Compared to control conditions with constant low or high friction levels, friction changes led to a reduction of completion time, especially for thin targets, i.e with a high index of difficulty. Linear regressions in Figure 2.22 demonstrate that results are in accordance with Fitts' Law predictions, and that the haptic feedback improves the performances.

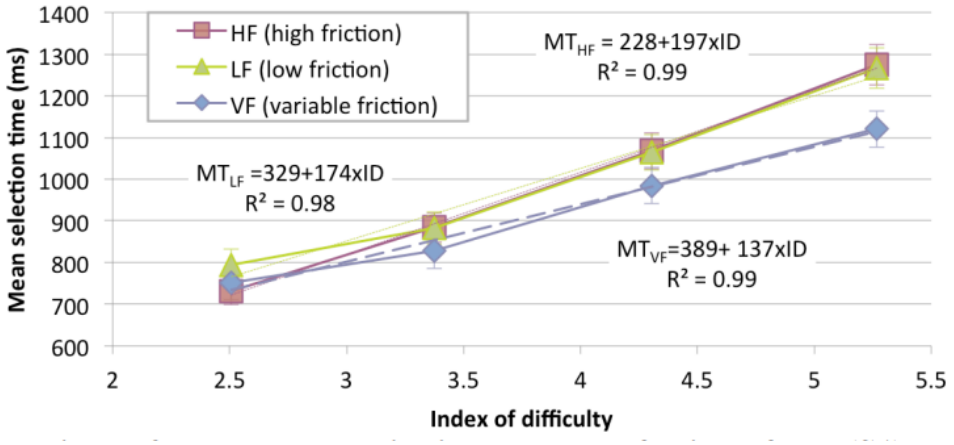

FIGURE 2.22: Results of the pointing experiment with haptic feedback and Fitts' law models, from [Levesque et al., 2011]. Pointing tasks are achieved either with constant friction (HF and LF), or with a friction increase on the target (VF).

Similar results have been reported in [Casiez et al., 2011], which also showed that haptic feedback in the form of step profiles (with sharp friction transitions) were more efficient than Gaussian-shaped friction profiles. Moreover the presence of distractors, i.e intermediate targets placed before the principal target also enhanced with friction changes, did not affect the performances.

[Zhang and Harrison, 2015] investigated whether pointing performances could be further improved with other types of haptic feedback, presented in Figure 2.23. The haptic feedback from the previous experiments ("Fill") is compared to haptic feedback with more punctual friction changes that are placed either at the entrance of the target ("Line Leading Edge"), in the middle of the target ("Line Center"), or between the targets ("Line Background"). They found that the "Fill" condition led to better performances both in terms of movement time and error rate. In this experiment, friction modulated feedback was rendered with electradhesion technology. [Kalantari et al., 2018] investigated whether these results could be extended to friction modulation with ultrasonic levitation. They performed the same experiment with this technology. Surprisingly, the findings were not the same and the "Line Center" condition provided the best speed-accuracy trade-off. Altogether, these results show that the hardware configuration (position sensing system and actuation technology) should be considered while designing haptic feedback. Nevertheless, performances were better with haptic feedback than without in all cases.

In these experiments, subjects could feel the haptic feedback and then intentionally stop their finger on the target. However, tactile cues can also unconsciously influence movement. [Bianchi et al., 2017] demonstrated that the ridges orientation of a texture could bias the finger movement in a slightly different direction than the target.

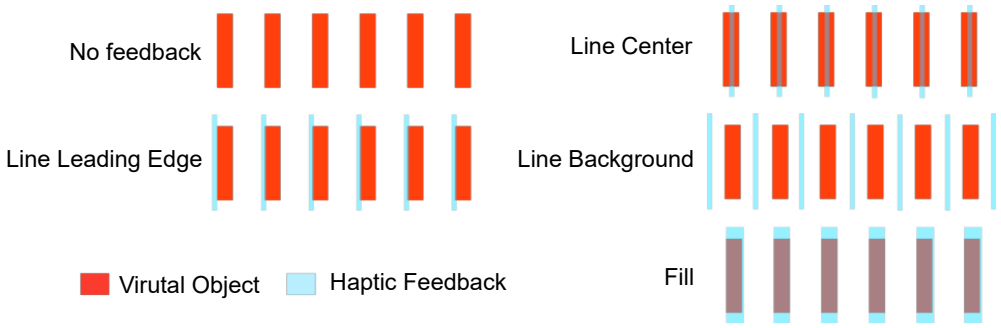

FIGURE 2.23: Different types of haptic feedback that can be used to guide the user in performing a pointing task, from [Zhang and Harrison, 2015].

2.6.3 Audio guidance

Auditory feedback is extensively present in our daily interfaces. It mostly consists of simple transitory sounds ("beep" or "click") that validate certain actions. These relatively short abstract sound patterns are called *earcons* [Blattner et al., 1989]. On a computer, it has been shown that the navigation trough folders and files was improved by the implementation of *earcons* that share similar sound attributes when clicking on similar types of folders/files [Brewster et al., 1993]. To guide the user also during his/her movement, more complex auditory feedback solutions have been proposed with sonification, defined as the use of non-speech auditory feedback to convey information to the user [Barrass, 1995]. A famous example is the car's reversing radar, whose beeps help the driver to park. Sonification has proved to assist the user in many other applications, such as in pedestrian navigation [Wilson et al., 2007], in surgery [Wegner, 1998] or in rehabilitation [Danna et al., 2013]. It is also of interest in sport to learn new skills like golf putting [O'Brien et al., 2017] or improve the performances, for instance within bicycling [Vidal et al., 2020] or rowing [Dubus, 2012]. In the automotive context, sonification of electrical motors has shown to improve drivers' perception of vehicle dynamics [Denjean et al., 2019].

[Parseihian et al., 2013] proposed different sonification strategies to guide the user towards a target. The distance to the target was mapped to different sound attributes, i.e. pitch, loudness, tempo, and timbre (brightness, inharmonicity and roughness). To reach the target, the subjects had to find the position which corresponded to an extremum of the attribute (the lowest pitch or the fastest tempo, for example). The experiment showed that participants were faster and more precise with sonification strategies based on pitch or tempo variations and by combination of these strategies [Parseihian et al., 2016].

[Bressolette et al., 2021] designed a car interface based on a virtual object that can be manipulated by mid-air gestures enhanced with auditory feedback. For example an increase in the air conditioning temperature is performed with a swipe gesture

that corresponds to throwing the virtual object toward a specific location. A further throw leads to higher temperature increase, as presented in Figure 2.24. The virtual object is considered as a ball that rolls away on a slope and comes back. In [Bressolette et al., 2018], many sonification strategies of the virtual object movement were compared in terms of temperature setting accuracy. The sonification strategies with pitch and brightness appeared as the most relevant for this interaction.

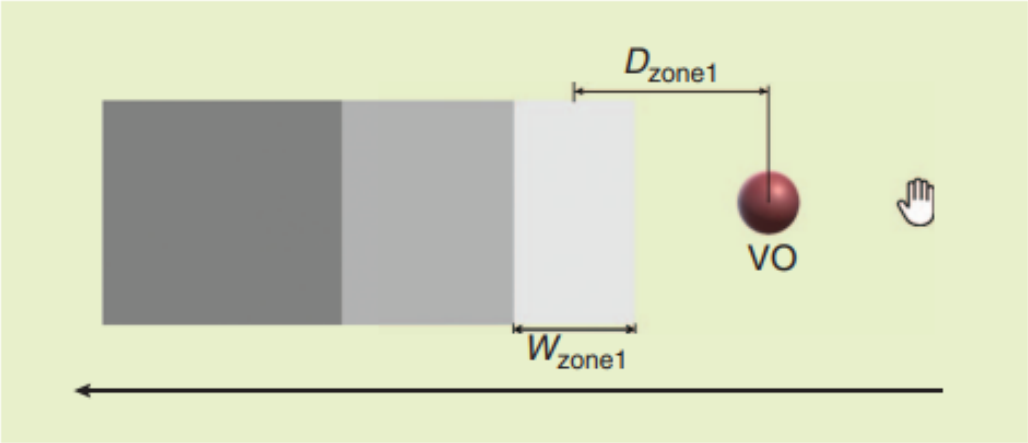

FIGURE 2.24: Illustration of the task in [Bressolette et al., 2018]. With a swipe gesture, subjects had to throw the virtual object (VO) toward a target area. Different sonification strategies of the virtual object movement were compared.

It has been demonstrated that auditory feedback can also unconsciously influence movement. [Thoret et al., 2016] showed that if subjects are asked to draw a circle but hear the sound of an ellipse being drawn, they tend to draw an ellipse.

2.6.4 Audio-haptic guidance

Few studies have compared the benefit of auditory and tactile feedback to guide a user. In [Rocchesso et al., 2016], the authors used the audio-tactile synthesizer they developed to guide the user through a path, made of a succession of rectangular bumps. Subjects could feel the vibrations or hear the sounds produced by the stylus passing over the bumps, confirming that they were on the path. The experiment with steering tasks showed that visual feedback is prevailing in this task and is not affected when combined with auditory or vibratory feedback. Moreover, vibrations, audio and audio + vibrational feedback appeared as equivalent since they led to similar movement times. In [Del Piccolo et al., 2018], subjects had to navigate through more complex paths, relying only on auditory or tactile cues. They could feel a sound or vibration when they were on the path, and the feedback stopped when they were outside. The experiment showed that the accuracy was not affected by the feedback modality, but the completion time was lower with auditory feedback.

In summary, the compilation of studies on interactions with auditory and haptic feedback shows that performance gains are rather dependent on the context and the interface. No clear trend reveals an ideal type of feedback that works in all cases. Yet the presence of auditory or haptic feedback is always more beneficial than none.

Chapter 3

Haptic perception of uniform textures

Contents

Adapted from: Corentin Bernard, Sølvi Ystad, Jocelyn Monnoyer, and Michael Wiertlewski. "Detection of friction-modulated textures is limited by vibrotactile sensitivity". IEEE Transactions on Haptics, 2020.

Preface to Chapter 3

As discussed in the state of the art, haptics still needs technological developments to precisely record and recreate textures. My first approach aimed at mimicking a microphone by building a device to record haptic textures. We designed a force sensor based on an interferometer capable of measuring friction variations with a better sensitivity than a human finger, for the entire frequency bandwidth of tactile sensitivity from continuous friction to 800 Hz vibrations. Its functioning and performances are described in detail in [Bernard et al., 2019].
The force sensor was used in the psychophysical experiment presented in this chapter. The aim of this study was to measure the detection thresholds of elementary sine waves, as the curves presented in the state of the art (chapter 2.2), specifically applied to the perception of haptic surfaces with ultrasonic friction modulation. As this technology produces sensation of reliefs and textures using force variations, the idea was to investigate whether the thresholds of friction modulated gratings were similar to real relief thresholds or to vibration thresholds. A precise understanding of these elementary thresholds is required to further synthesize more complex stimuli. The experiment was published in [Bernard et al., 2020].

3.1 Abstract

Modulation of the frictional force of a fingertip sliding over a surface-haptic device can produce compelling sensations of texture and relief. The virtual sensation is particularly apparent and feels as fixed in space if the stimulus is rigorously correlated with the displacement of the finger. While frictional textures tactually resemble their real counterparts, some exploratory conditions under which the sharpness of the texture declines exist. We postulate that this decline in sharpness is caused by the perceptual limitation of the attempt to interpret the variation in friction as an out-of-plane sinusoidal topography. To investigate these questions, we measured the detection thresholds of sinusoidal friction-modulated gratings for a wide range of spatial periods explored at two different speeds. We compared the results with the detection thresholds, reported in the literature, of real gratings and vibrotactile stimuli.

We found that the detection of spatial friction-modulated textures does not follow the same trend as that of real textures but is more similar to the vibrotactile rendering, which is strongly influenced by the exploratory speed. This study provides a better understanding of the perception of friction-modulated textures and provides insight into how to design impactful stimuli on surface-haptic devices.

3.2 Introduction

Surface haptics is a promising way to enhance the human-machine interaction by providing localized sensations on a touchscreen. These technologies lend themselves to a large number of applications such as consumer electronics and the automotive industry, with the promise of decreasing visual distraction. Owing to the fine control over the frictional force on the skin, these devices also provide an opportunity for furthering the understanding of human tactile perception.

Changes in friction between the user finger and the glass plate can provide the sensation of touching shapes that protrude from the plate or even fine textures. The mechanism behind the perceptual integration of friction-modulated spatial patterns into believable tactile textures is still an open question. Friction changes affect the entire contact surface and can vary only over time; yet, the perceptual experience is convincingly similar to touching a physical relief, the features of which are distributed in space.

For the relief to be perceived in a coherent way, it must be precisely localized. Therefore, friction is usually modulated as a function of the finger position, i.e., each finger position corresponds to one friction level. The illusion breaks down if the presentation of the stimulus lags behind the user's motion [Okamoto et al., 2009].

Since the frictional force can be modulated only over the whole fingertip, creating synthetic curvatures smaller than the area of contact should be impossible. However, friction variations below this limit produce vibratory oscillations that give the

illusion of touching fine textures in a fashion similar to that of vibrotactile stimuli [Wiertlewski et al., 2011b].

For virtual as well as real textures, the frequency of the vibrations produced by the skin-surface contact that propagate at the surface of the skin shifts according to the exploration velocity. Vibrations are perceptually integrated with exploratory motion as a unique and invariant percept. These vibrations are believed to be one of the main factors that mediate texture perception [Lederman et al., 1982, Bensmaia and Hollins, 2003, Bochereau et al., 2018]. Inversely, when proprioceptive cues are absent, the scanning speed can be perceived by the frequency content [Dépeault et al., 2008, Delhaye et al., 2019].

The sensitivity to vibrations follows a U-curve from 10 Hz up to 800 Hz, with an optimal sensitivity at 250 Hz [Verrillo et al., 1969, Verrillo, 1985]. This sensitivity curve was found when subjects were static and the apparatus presented a fixed stimulus. In contrast, when touching real or surface-haptic textures, the observer has to conduct an active exploration to acquire relevant informations. In this context the skin is in relative motion with the display, causing a fundamentally different mechanical interaction.

Natural textures, in contrast, have a much richer interaction than fixed vibrotactile stimuli. Louw et al [Louw et al., 2000] showed that the detection thresholds for real Gaussian bumps follow a linear trend with respect to the spatial period of the features. The results were subsequently extended to sinusoidal gratings [Nefs et al., 2001]. The authors found that the preponderant factor in the determination of the detection threshold was the slope of the relief. The minimal perceptible gradient is 1.3 µm/mm over a large range of spatial scales from a few hundred microns to several centimetres. In these experiments, subjects were free to explore the textures without any speed restriction. It is thus not possible to deduce the spectral content of the vibrations that propagated in the finger from this study.

These results lead us to question which mechanism is involved in the perception of friction-modulated gratings: Are spatially defined frictional textures perceptually integrated as a spatial topography or as time-varying vibrotactile signals?

To investigate this question, the present chapter studies the detection thresholds of spatially defined sinusoidal gratings of various spatial periods under two exploration velocity conditions using ultrasonic friction modulation technology. For the same spatial period λ, the induced vibration frequency f is affected by the scanning speed v_{finger}, following the relation $f = v_{finger}/\lambda$. The faster the exploration speed is, the higher the vibration frequency. Our assumption is that if these textures are spatially integrated, the detection threshold of a grating at a given spatial period λ will not be affected by the scanning speed, as presented in Fig. 3.1.a (left). Consequently, a difference occurs between the two finger velocities if we compare the detection thresholds of the induced vibration frequencies $f = v_{finger}/\lambda$, as presented in Fig. 3.1.a (right). In the opposite case, if the textures are temporally integrated, the detection thresholds of the stimuli should be guided by the vibration frequencies f

produced during the exploration, as presented in Fig. 3.1.b (right). In this case, the thresholds are invariant with finger velocity in the frequency domain, which implies that the detection threshold of a grating at a given spatial period λ will vary with the finger velocity v, as presented in Fig. 3.1.b (left).

To measure the friction variations induced by the stimuli around perception thresholds, we designed a custom force sensor sensitive to stimuli four times smaller than the detection threshold of a human observer. In addition, we instrumented physical quantities that are relevant to texture perception, such as the amplitude of the plate vibration, the vibration of the skin and the subject's perception. The setup assesses all the physical quantities involved in the perception of friction-modulated texture.

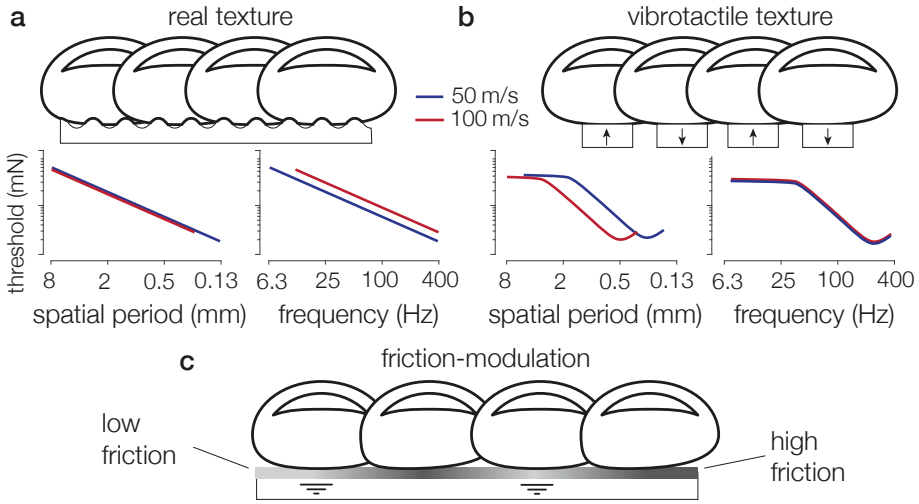

FIGURE 3.1: **a**. Detection threshold curve with a spatially based perception model in the spatial and frequency domains. Data extracted from [Louw et al., 2000, Nefs et al., 2001]. **b**. Detection threshold curve with a frequency based perception model in the spatial and frequency domains. Data extracted from [Verrillo et al., 1969]. 50 mm/s and 100 mm/s velocities are shown in blue and red respectively. **c**. Interaction with a virtual texture on a surface-haptic device.

3.3 Materials and methods

3.3.1 Design motivation

Experimental setup

The experimental setup was designed with the objective of providing a highly consistent signal by ensuring a spatial and temporal resolution of the stimulus orders-of-magnitude higher than the perceptual limits. In addition, the quantification of the output signal was sufficiently fine to avoid any artefacts. The spatial stimuli were created by designing a friction map; a function of the desired friction level according

to the finger position. Every 200 µs, the controller polled the finger position and sent the actuation command.

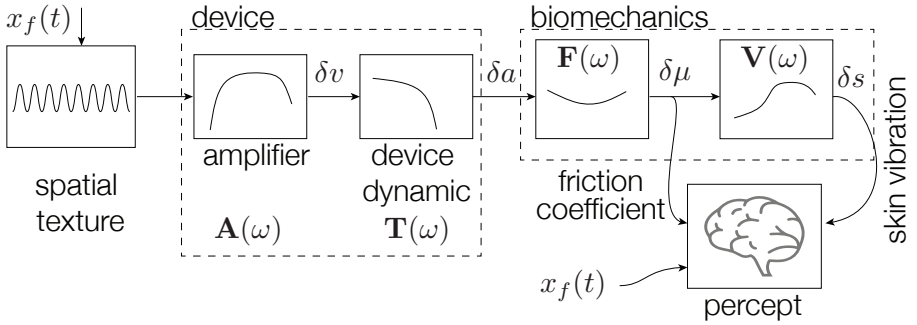

FIGURE 3.2: The signal is generated from a lookup table indexed as a function of the user's position $x_f(t)$. The signal is then amplified and the voltage modulation δv is sent to the piezoelectric actuators. Changes in the ultrasonic vibrations of the plate δa result in changes in the frictional force $\delta\mu$, which creates skin vibrations, captured by an accelerometer on the skin surface on the first phalanx of the finger δs.

Psychophysics and perceptual transfer functions

Fig. 3.2 summarizes the physical quantities that are related to the perception of frictional texture, from the actuation to the perception. The algorithm used in the threshold experiment presents a stimulus defined as an amplitude modulation command. This command is converted into a voltage modulation that mixes with the carrier frequency and is provided to the actuators on the glass plate. The effective vibration modulation of the plate, which depends on its dynamics, induces squeeze-film levitation which in turn modulates the frictional forces between the finger and the plate. Friction changes at the contact between the skin and the glass plate produce mechanical vibrations that propagate into the whole finger. Since the precise relevance of friction force and vibration of the skin in the perceptual experience of tactile texture is unclear, we measured both variables during the psychophysical experiment. Fig. 3.3 shows the experimental setup, and the technical details are described in the following section.

3.3.2 Texture production on surface-haptic devices

The finger position is tracked with a small ring attached to the participant's finger. The ring is connected to a pulley-encoder system that measures unidirectional finger displacements along the length of the glass plate. The precision of this system is approximately 0.01 mm and can be captured at 4000 samples/sec without any significant latency. A microcontroller (Teensy 3.5) reads the encoder and outputs a modulating signal according to a friction map encoded in memory on a fixed internal real-time timer. The carrier signal, a 35 kHz sine wave, is created by a function generator (BK Precision 4052) and amplitude modulated by the analog signal coming

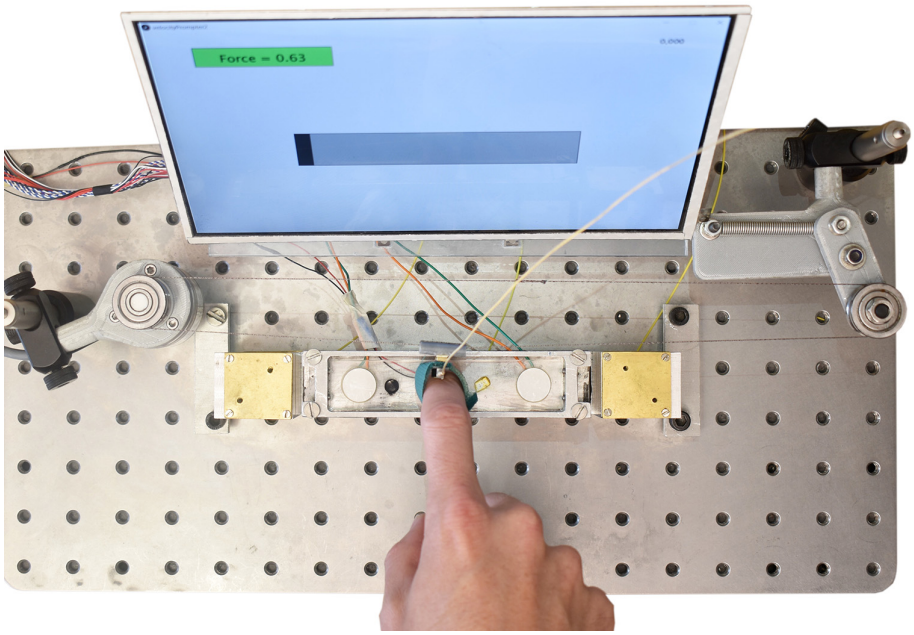

FIGURE 3.3: Experimental setup. The subject touches the actuated glass plate on top of the force sensor. The finger is linked to a pulley-encoder system for position measurement and an accelerometer is placed on the first phalanx. The screen shows a cursor imposing the velocity and provides feedback on the normal force.

from the microcontroller. The resulting signal is then amplified 20-fold (WMA-100, Falco Systems) to drive two piezoelectric actuators glued on a $105 \times 22 \times 3.3$ mm glass plate. Modulation of the amplitude of vibration of the glass plate induces friction variations during exploration.

3.3.3 High-precision force sensor

To measure the interaction force acting on the participants' finger, we designed a custom sensor able to measure forces with a better sensitivity than that of the human sensory system, over its entire sensitive frequency range [Verrillo et al., 1969]. The force sensor is based on a rigid elastic structure in which nanometre-scale deformation is measured via a Fabry-Perot interferometer. The sensor is optimized to cover a frequency bandwidth that spans continuous forces to stimulation up to 800 Hz, and is able to resolve forces with amplitudes lower than 1 mN.

Mechanical structure

The structure of the force sensor is presented in Fig. 3.4.a. The actuated glass plate is fixed to the top of an aluminium support with a lightweight yet stiff honeycomb structure. The aluminium support is suspended by three flexures that allow for minute lateral displacements but are virtually infinitely stiff in the other directions. This structure guides the deformation along the sensing axis of the transducer. The lateral force sensor is then suspended above the ground with two brass four-bar

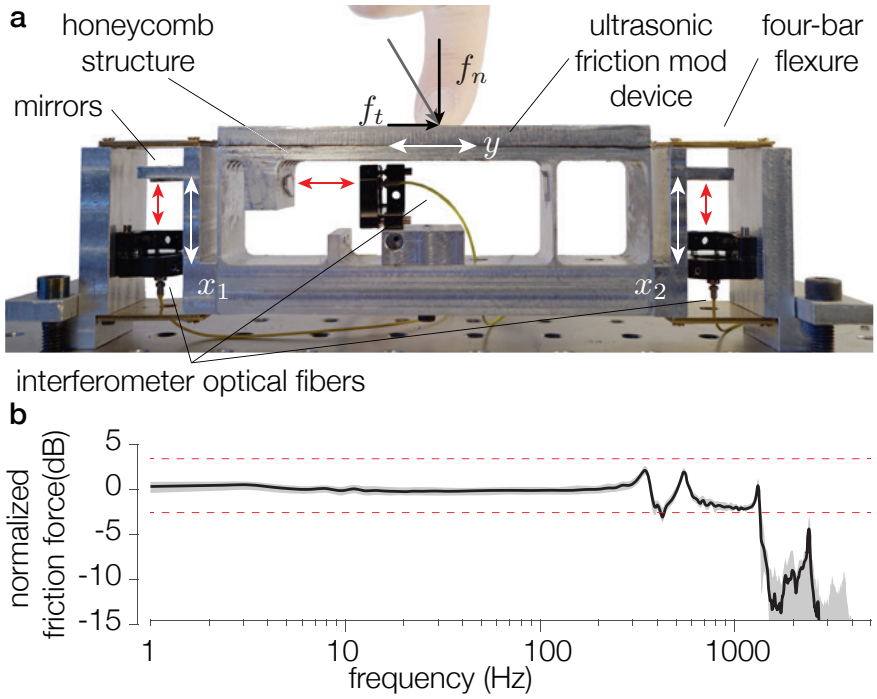

FIGURE 3.4: **a** Picture of the tribometer. The glass plate is mounted on a honeycomb structure that provides light and stiff support. The test sample is suspended by a set of three four-bar flexures whose deflection is measured with an interferometer via fibre optics. Red arrows show the laser path between the fibre optics and the mirrors. **b** Frequency response of the sensor. The median of ten normalized lateral displacement amplitudes is printed in black and its first and third quartiles are shaded in grey. Red dashed lines show the $\pm$ 3 dB range.

flexure linkages on each side. Similar to the lateral flexures, these horizontal brass flexures allow for vertical movements while limiting motion in the other degrees of freedom.

Deformation measurement and calibration

The deformation of the structure is measured by a 3-axis Fabry-Perot interferometer (IDS 3010, Attocube). The laser is guided by an optical fibre and focused by a lens (D4/F1, Attocube), which is mounted on the base of the structure. The laser is then reflected to the lens by a mirror, mounted on the flexible part of the structure. Micrometric adjustment screws enable alignment of the laser path. The picometer-sized displacements of the three axes (2 normal x_1 and x_2 and one tangential to the fingertip exploration y) are then converted into an analog output refreshed at 10 MHz. The calibration is performed by first applying a known force to the upper part of the sensor with a mass and then by using the same weight with a string and pulley system to apply a tangential load.

Performance

The frequency response of the structure to a tangential impact on the top plate is shown in Fig. 3.4.b. The frequency response shows a main cut-off frequency of approximately 1300 Hz, with some normal modes between 300 Hz and 600 Hz. The measured noise floor of the presented structure and interferometer system is 0.4 mN.

3.3.4 Additional sensors

In addition to the 2-axis force sensor, an accelerometer (Model 2250A / AM1-10, Meggitt) with a flat response over a frequency bandwidth spanning from 2 to 15 000 Hz, is attached to the plastic ring in contact with the skin of the first phalanx to measure the propagation of vibrationsinto the surface of the skin of the index finger. Indeed the vibrations from the tactile interaction propagate in the whole hand [Shao et al., 2016]. The finger vibrations, contact forces and finger positions are recorded with an acquisition card (USB X Series Multifunction DAQ, National Instruments) at a 10 kHz sampling rate preceded by an antialiasing filter. To measure the vibration of the plate, a third piezoelectric ceramic acting as a sensor is glued to the glass plate. The input voltage applied to the piezoelectric actuator and the output voltage from the piezoelectric sensor are recorded by a dedicated acquisition card (NI USB-6211, National Instruments) at a 100 kHz sampling rate, to provide enough resolution for demodulation of the ultrasonic wave. The output signal is calibrated with an interferometer (IDS 3010, Attocube) to obtain the glass plate vibration in micrometres. Before and after each session, the moisture on the subjects' index finger is assessed with a dedicated device (gpskin Barrier pro). This measurement does not provide any insight into the results and will therefore not be discussed.

3.3.5 Psychophysical experiment

Participants

17 volunteers, 5 females and 12 males, 15 right-handed and 2 left-handed, ranging from 22 to 42 years old (mean 28.2) participated in the study. They were naive to the aims of the study and none of them reported having any skin concerns. The study was approved by the Ethical Committee of Aix-Marseille University. The participants gave their informed consent before the experiment. They were paid for their participation. They washed and dried their hands and the glass plate was cleaned with an alcoholic solution before the experiment. The results of two subjects were discarded because of technical issues.

Stimuli

We investigated the subjects' detection thresholds for friction-modulated sine waves spatially encoded by the position of the finger. The experiment was divided into two parts corresponding to the two finger velocity conditions (50 mm/s and 100 mm/s).

The presentation order of the velocity condition was alternated between subjects. For the 50 mm/s velocity, haptic stimuli were rendered for 7 spatial periods: 0.125, 0.25, 0.5, 1, 2, 4 and 8 mm. For the 100 mm/s velocity, haptic stimuli were rendered for 6 spatial periods: 0.25, 0.5, 1, 2, 4 and 8 mm. According to the 2 finger velocities, the vibrations transmitted to the finger varied from 6.25 Hz to 400 Hz. The 0.125 mm condition was not presented for the high velocity since it would produce a fundamental frequency of 800 Hz, which could not be rendered by the glass plate. The spatial period sessions were presented in random order. The experiment lasted for approximately 2 hours.

Psychophysics procedure

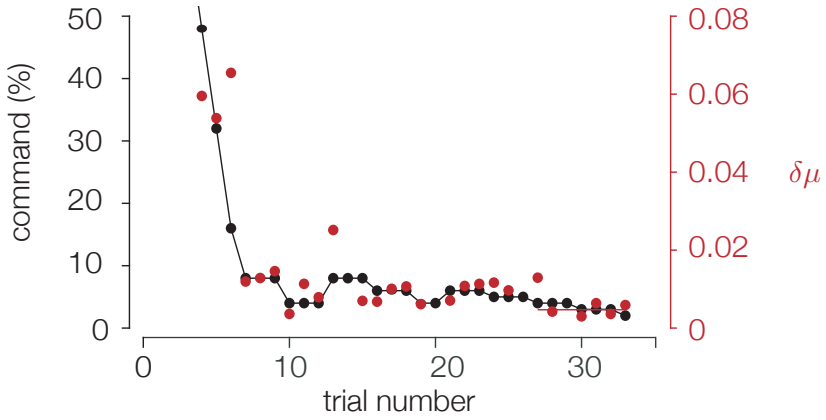

FIGURE 3.5: Typical 3-down/1-up adaptive staircase procedure. The desired friction modulation amplitude is plotted in black and the resulting friction variation amplitude $\delta\mu$ is plotted in red. The red line reports the mean value of the last 6 trials.

Participants sat in a chair in front of the experimental desk and wore headphones with pink noise to prevent any auditory cues from the device. They put their right index finger into the position-tracking apparatus. They were asked to continuously explore the glass plate with their finger by moving back and forth from left to right while synchronizing their movement with a cursor presented on a screen that imposed the finger velocity. After each trial, feedback on the normal force applied to the glass plate was displayed on-screen. They were asked to keep the normal force between 0.4 and 0.8 N. For each trial, they had to perform 2 successive explorations on the glass plate (2 back-and-forth movements). One of the explorations randomly contained the haptic stimulus (modulated friction) whereas the other one was smooth (constant friction) acting as a reference. Both the stimulus and the reference had the same average friction level. The psychophysical method was a two-alternative forced choice (2AFC): the participants had to report which of the two explorations presented the most irregular texture.

They could answer via an interface on a laptop situated to their left. For each spatial period, the detection threshold, i.e., the minimal perceptible amplitude, was

evaluated according to a 3-down/1-up adaptive staircase procedure. After 3 successive correct answers, the amplitude of the stimulus decreased, and after one wrong answer, the amplitude of the stimulus increased. This algorithm converges to the detection threshold. In our experiment, the procedure was stopped after 6 reversals. After every second reversal, the step size was divided by two. To quicken the convergence, the procedure started with a simple 1-down/1-up process until the first reversal. A typical procedure is presented in Fig. 3.5.

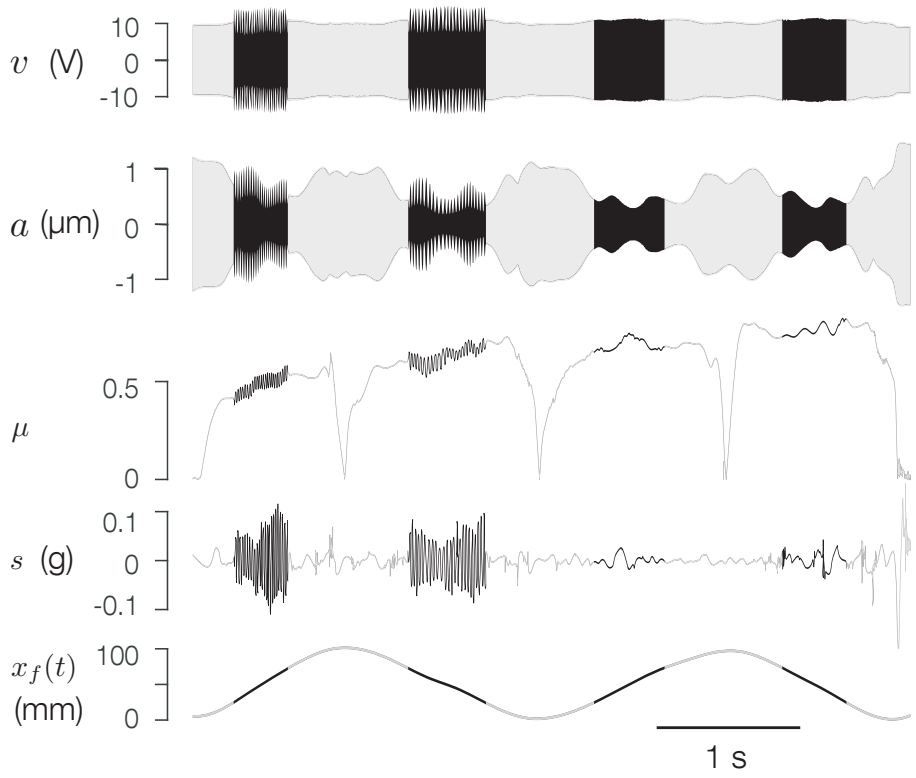

FIGURE 3.6: Raw data measured during one trial. The subject explores the virtual texture with 2 successive lateral back-and-forth movements across the plate. In this trial, the stimulus was presented during the first lateral movement. During the other one, friction reduction was constant. v is the input voltage provided to the piezoelectric actuators, a is the glass plate vibration and μ is the friction coefficient between the plate and the finger. s is the vibration of the finger, assessed on the first phalanx. x_f is the position of the finger. Darkened parts are the selections for which the variations of these variables are calculated

3.3.6 Data analysis

Signal processing

The friction coefficient was computed from the absolute value of the tangential force divided by the normal force which was filtered with a 100 Hz low-pass filter to remove fluctuations. For each trial, the finger position, input voltage to the piezoelectric actuators, glass plate vibrations, friction coefficient and vibrations of the finger

were recorded. Fig. 3.6 presents typical curves of these measurements. Data are first partitioned to isolate the part that corresponds to the 48 mm length area at the centre of the plate, where the stimulus is (or is not) presented. The finger velocity is assumed to be constant in this part. For each trial, we thus obtain 4 samples: 2 explorations with the haptic stimulus, one in the left-to-right direction and the other in the right-to-left direction, and 2 explorations with a constant actuation which will be used as a references. Calculation methods presented in the next sections are also performed on these references to measure the noise floor. This noise is partly caused by the sensor variations and mainly produced by fluctuations of the frictional force during the exploration of the unactuated surface.

Transfer functions of the block diagram Fig. 3.2 are calculated by taking the ratio between the mean value of two successive variables for each stimulus frequency. They show how each variable affects the next one as a function the stimulus frequency.

Estimation of friction variations

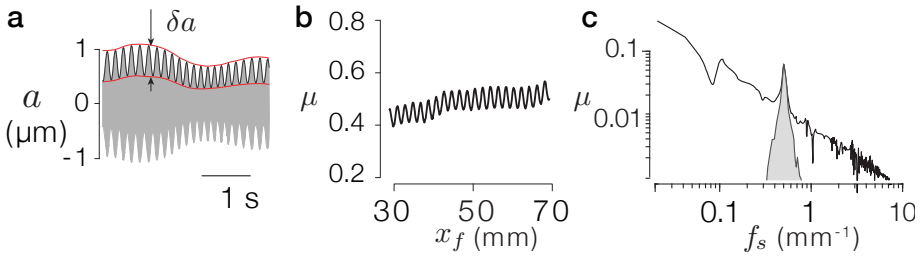

FIGURE 3.7: **a**. Vibration of the glass plate as a function of time. The envelope of the carrier is in black. The glass plate vibration modulation δa is measured by averaging the difference between the upper and lower red envelopes. **b**. Friction coefficient as a function of the finger position. **c**. Spatial spectrum of the friction signal. A bandpass filter centred around the spatial frequency of the stimulus is applied to the spatial signal to extract the energy of the signal in the frequency band of interest.

The friction coefficient is interpolated on regularly sampled spatial coordinates system (0.01 mm sampling interval) to obtain the signal as a function of the finger position, as presented in Fig. 3.7.b. As seen in Fig 3.7.c, the friction coefficient is entailed with a large $1/f$ background noise, which is induced by the stochastic interaction between the finger and the plate [Wiertlewski et al., 2011a]. Because the friction variations at some thresholds are very close to the noise floor and do not stand out using envelope detection algorithms, we used the spatial frequency domain method illustrated in Fig. 3.7.c. The spatial signal μ_f is filtered with a bandpass filter around the spatial frequency of the stimulus (2nd-order Butterworth with cut-off frequencies of 0.7 and 1.3 times the stimulus spatial frequency). The energy is then computed with $E_\mu = \int |\mu_f(x)|^2 dx$ of the peak provided by the stimulus. The friction variation amplitude is then assessed by calculating the theoretical amplitude

A of a sinewave with the same energy $\delta\mu = \sqrt{2\,E_h/L}$, with $L = 48$ mm being the length of the signal. We selected only the maximal value $\delta\mu$ between the left-to-right direction and right-to-left direction. Indeed, many subjects reported that they sometimes felt the stimulus in only one direction. For each subject and each condition, the smallest perceptible friction variation was defined as the average of $\delta\mu$ over the last 6 trials.

Skin vibration induced by friction fluctuations

Vibrations that propagates on the surface of the finger skin, are measured by the an accelerometer and are processed in a way similar to that for the friction coefficient. Partitioned data are bandpass-filtered with the centre frequency corresponding to the resultant frequency of the stimulus according to the velocity condition. The vibration energy in the finger is thus given by $E_v = \int |a_f(t)|^2 dt$.

For the input voltage of the piezoelectric actuator and the glass plate vibration, the envelope of each signal is computed using the Hilbert transform to isolate the modulation from the 35 kHz carrier signal, as shown in Fig. 3.7.a.

3.4 Results

3.4.1 Subject performances

Explorations were performed with a normal force of a mean value and standard deviation of 0.68 ± 0.19 N, which is in line with the requested range of 0.4 to 0.8 N. We noticed that subjects tended to increase their normal exploration force as the stimuli became more subtle. The effective velocity of the finger was 55 ± 10.2 mm/s for the low velocity condition and 113.3 ± 24.3 mm/s for the high velocity condition. The spatial periods of the stimuli (0.125, 0.25, 0.5, 1, 2, 4 and 8 mm) led to effective frequencies of 440, 220, 110, 55, 27.5, 13.8 and 6.9 Hz on average for the low velocity condition and to 906.4, 453.2, 226.6, 113.3, 56.7, 28.3 and 14.2 Hz on average for the high velocity condition. Each session lasted between 15 and 42 trials (mean 25.4) until the subject converged to their detection threshold.

3.4.2 Tactile thresholds

For each session, the friction variation $\delta\mu$ is averaged over the last 6 trials of the session to obtain the subjects' friction variation detection threshold for the spatial period. We can thus reconstruct the *haptogram* –analogous to an audiogram for hearing– of a subject, a curve showing the tactile threshold of the subject at each spatial period as presented in Fig. 3.8. We decided to show these results in the spatial domain to compare the stimuli and in the frequency domain by taking into account the velocity of the finger.

Statistical analysis was performed on each of these four variables both in the spatial and frequency domains to investigate the effect of the velocity. We performed

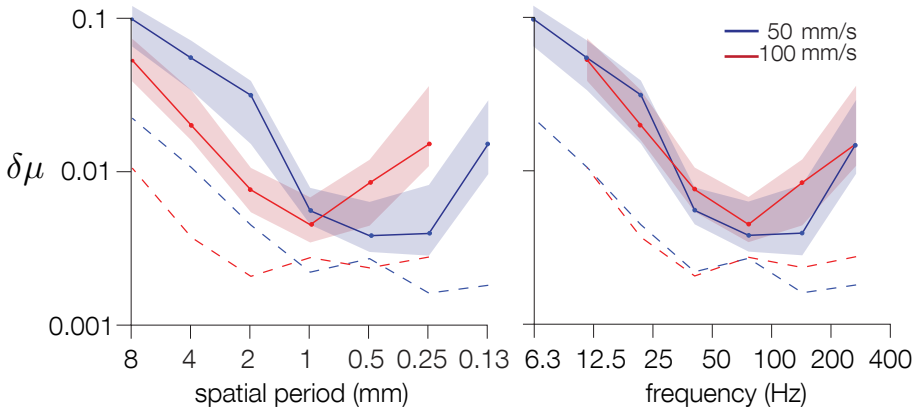

FIGURE 3.8: Friction variation detection thresholds. Curves are presented twice, one in the spatial domain, where the abscissa corresponds to the spatial period of the stimulus (the axis is inverted for comparison), and the other in the frequency domain, where the abscissa corresponds to the frequency given by the ratio between the finger velocity and the spatial period. The solid lines represent the median and the shaded zones represent the first and third quartiles. The dashed lines represent the noise floor calculated on the reference exploration without a stimulus.

two-way ANOVAs with the finger velocity and spatial period or frequency as the factors. In the spatial domain, the thresholds are compared according to their spatial period conditions. In the frequency domain, thresholds that share the same frequency during their exploration are compared.

The results showed that the velocity had a significant effect ($\alpha = 0.05$) on the friction variation thresholds in the spatial domain ($F_{1,5} = 10.43$, $p = 0.0015$) but not in the frequency domain ($F_{1,5} = 0.13$, $p = 0.72$).

In the same way, plate vibrations δa and finger vibrations δs were averaged over the last 6 trials of each session to obtain an average value at the subjects' detection thresholds. The results are summarized in Fig. 3.9. The same trend as for the friction variation threshold curve was observed but with less significance ($\alpha = 0.07$) for the plate vibrations. We observed an effect of the finger velocity on the plate vibration modulation at the threshold ($F_{1,5} = 3.4$, $p = 0.06$) in the spatial domain and no effect ($F_{1,5} = 0.003$, $p = 0.95$) in the frequency domain. There was no effect of finger velocity on finger vibrations at the thresholds in both the spatial ($F_{1,5} = 0.45$, $p = 0.5$) and temporal ($F_{1,5} = 1.88$, $p = 0.17$) domains.

To investigate the relevance of each variable to human perception, we measured their inter-subject variability. We assumed that a variable with less variability among participants would be the most relevant descriptor of tactile perception. We measured for each variable the coefficient of variation c averaged over all velocities and spatial period conditions. We found that the lowest variability was obtained for friction variations $c_{\delta f} = 0.42$ whereas the variability for plate vibration modulation was $c_{\delta a} = 0.73$ and $c_{\delta s} = 0.55$ for finger vibrations. We assume that these data demonstrate that friction variation is the most relevant variable encoding human tactile

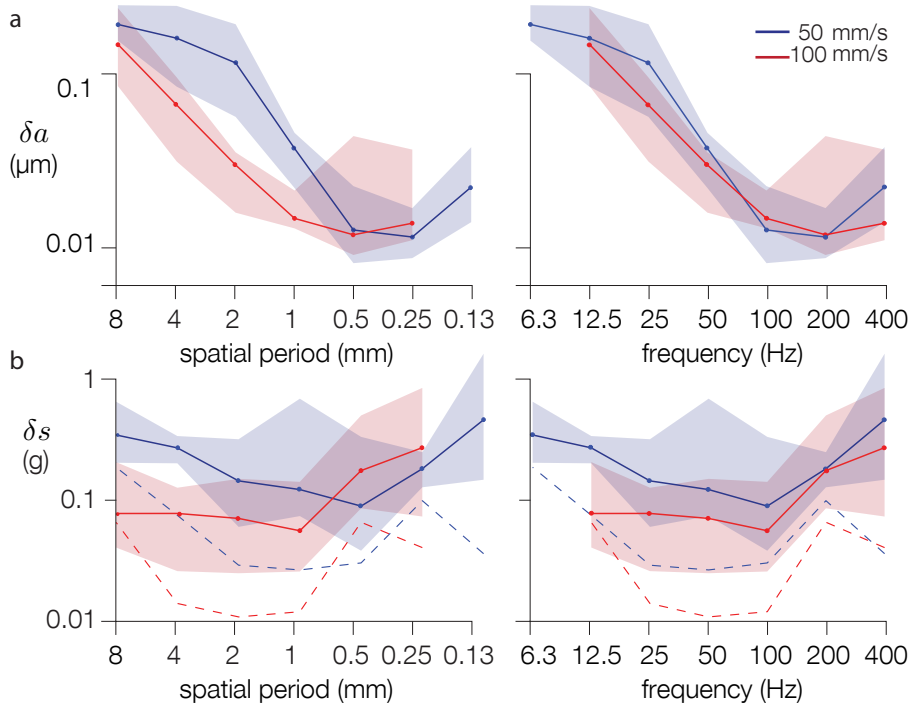

FIGURE 3.9: Values of plate vibration modulation and finger vibration at the detection thresholds in the spatial frequency domains. The solid lines represent the median and the shaded zones represent the first and third quartiles. Upon finger vibration, the dashed lines represent the noise floor calculated on the reference exploration without a stimulus.

sensitivity.

3.4.3 Transfer functions

As mentioned in the introduction, the rendering of a haptic stimulus follows several steps, summarized in the block diagram in Fig. 3.2. Each transfer function is calculated by averaging the data of all the trials, not only at the threshold, as their relationship is assumed to be linearly dependent on the input amplitude. Knowledge on these transfer functions provides a clearer picture of the role of each element involved in the rendering of surface-haptic stimuli.

The transfer function **T** (Fig. 3.10.a) shows how the deformations of the piezoelectric actuators make the glass plate vibrate. It reflects the frequency bandwidth of the glass plate vibration modulation. It has been noted before that due to the resonance of the glass plate, the amplitude modulation of the carrier wave will be attenuated for high frequencies. This effect is most noticeable when a step modulation function produces an exponential ring down of the plate oscillations [Meyer et al., 2014]. The attenuation is affected by the plate material and acts as a low-pass filter with a cut-off frequency of 100 Hz in this particular setup.

The transfer function **F** (Fig. 3.10.b) illustrates how the glass plate vibration affects friction between the finger and the glass plate. **F** characterizes the biomechanics of the skin-glass contact. The squeeze film effect at the origin of this

phenomenon has been explained and modelled [Wiertlewski et al., 2016], but the relationship between the modulation and the friction during an exploration is still subject to questioning. Interestingly, the transfer function **F** of this device is not constant. It shows that high and low modulation frequencies of the glass plate vibration are more effective to render strong friction variations.

The last transfer function **V** (Fig. 3.10.c) reflects how friction variations propagate into the finger as mechanical vibrations. The curve presents a U-shape with a peak transmissibility of the vibration around 100 Hz, in line with previous studies [Tanaka et al., 2012].

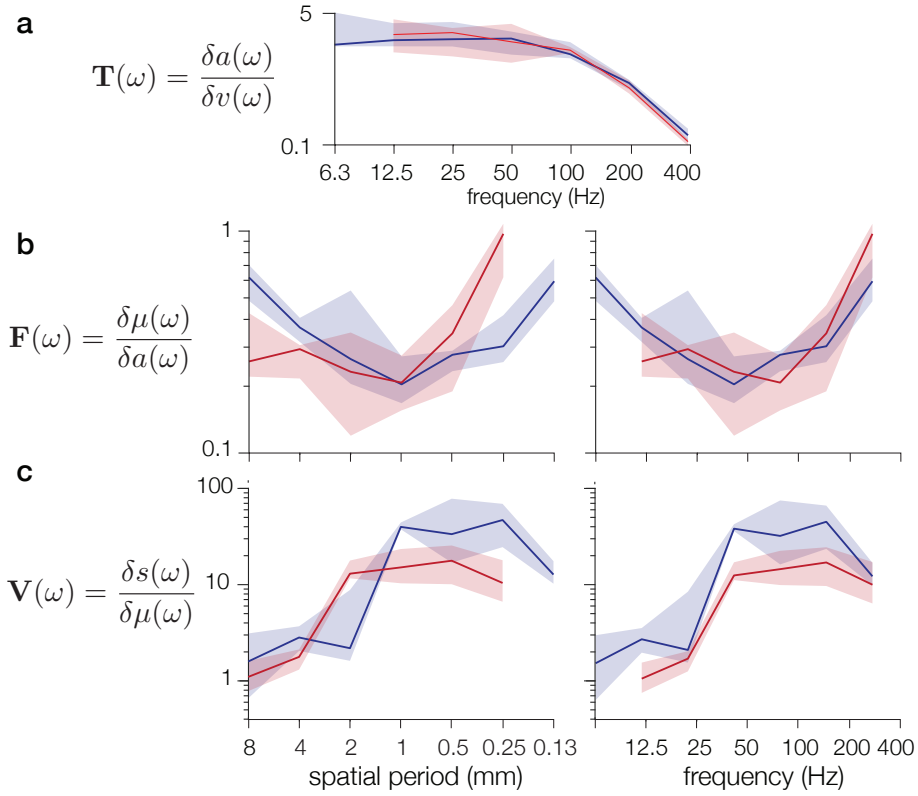

FIGURE 3.10: Transfer function between the different variables from actuation to perception. Data are measured on all trials. The solid lines represent the median and the shaded zone represents the first and third quartiles. **a** Transfer function between the input voltage modulation of the piezoelectric actuator and the actual vibration variations of the plate. **b** Effect of the modulation of the ultrasonic amplitude on the variation of friction, represented in the spatial and frequency domains. **c** Vibrations measured at the skin level of a given change in frictional force.

3.5 Discussion

Two main results emerge from the experiments presented Fig. 3.8. The finger velocity influences the detection thresholds of spatially based gratings, but no significant effect appears if we compare these thresholds according to the induced frequency. In

addition, threshold distributions across frequencies are not linear and show an optimum sensitivity between 100 and 200 Hz, with a decrease in sensitivity for higher frequencies. This observation suggests that friction-modulated textures are temporally integrated, since the detection threshold is not affected by the scanning speed, meaning that the frequency model (Fig. 3.1.b) is preferred over the spatial model hypothesis (Fig. 3.1.a). Friction-modulated grating detection thresholds follow similar trends to those of vibrotactile detection, and therefore, the vast literature on vibrotactile perception can support the design of meaningful friction-modulated stimuli. Pacinians are probably the mechanoreceptors that are the most involved in the detection of friction-modulated textures since their optimal sensitivity (250 Hz) is quite similar to the detection threshold curves (Fig. 3.8).

An interesting consequence of this result is that it exists an optimal combination of the spatial period and scanning velocity that enhances the tactile sensation of a friction-modulated texture. Small-scale synthetic textures are better perceived with low-velocity exploration whereas large relief perception is improved by faster explorations. A friction-modulated grating with a given spatial period can have the same detection threshold as that of another grating with a spatial period that is half the size if explored twice as fast.

It stands out that the stimuli used in the experiment may not be significantly different, from a perceptual point of view, from spatially encoded vibrotactile feedback with the same envelope. One major limitation of this result is that, as we investigated detection thresholds, the haptic effects were very subtle. This could be the reason why they are perceived more like vibrations than reliefs. This phenomenon might not be the same for gratings with higher amplitude.

The detection thresholds could have been altered by fatigue. It has been shown that tactile sensitivity is affected by the stimulation duration [Hahn, 1966]. We mitigated this issue by asking subjects to remove their finger from the glass plate between each trial, and imposing breaks.

Since the inter-subject variability is the lowest for the friction-variation variable, this variable seems to be the most relevant for controlling perception, in line with previous studies [Smith et al., 2002]. For example, some subjects needed higher plate vibration amplitudes at their detection threshold than other subjects, while the friction variation thresholds were similar between subjects. This is reflected in the relatively large variability of the transfer function F (Fig. 3.10.b) when compared to the vibratory transfer function V. The variability between plate vibration and friction reduction efficiency reflects differences in the bio-mechanical properties of the finger skin, which were also observed in [Kaci et al., 2019] and have been shown to significantly affect subjects' perception [Monnoyer et al., 2017]. Therefore, stimuli generation on surface haptics should focus on controlling the frictional force and friction coefficient rather than the ultrasonic amplitudes, as proposed in [Huloux et al., 2018, Messaoud et al., 2015].

Furthermore, the ratio between the perceptual thresholds and the amplitude of

the noise floor while sliding (acquired when no stimuli are presented) is relatively constant, at least for low frequencies. Thus, the frictional noise due to contact could be the limiting factor in the estimation of friction-modulated textures.

The finger vibrations at the detection thresholds presented in Fig. 3.9 show considerable variability. This is probably because the measured signals are close to the noise floor. Vibrations are attenuated by the skin during their propagation from the contact area to the sensor. The transfer function $\mathbf{V}(\omega)$ between the friction variations and the finger vibration (Fig. 3.10.c) is more reliable because it is measured for all trials and not only with low amplitudes at the threshold. Vibrations are maximally produced by friction changes between 100 and 200 Hz, which could explain the increased sensitivity in this frequency range.

The transfer function $\mathbf{F}(\omega)$, shown in Fig. 3.10.b, reveals that a given amplitude of ultrasonic vibration modulates the frictional force more effectively at high frequencies. This boost in high frequencies naturally and conveniently compensates the attenuation of the amplitude of the plate vibration at high modulation frequencies. The attenuation of the modulation of the ultrasonic vibration amplitude is due to the resonant behavior of the glass plate [Meyer et al., 2014] and is clearly illustrated in Fig. 3.10.a. The combination of both effects might explain why it is possible to render perceptible friction modulation feedback at frequencies above the cut-off frequency of the glass plate.

3.6 Conclusions

The objective of the present study was to investigate the detection thresholds of spatially defined fiction-modulated textures and to examine whether these textures were perceived more like real relief or like vibrations. We investigated the detection thresholds of spatial friction-modulated sinusoidal gratings at 7 spatial periods and for two exploratory velocity conditions because the scanning speed determines the frequency of the generated vibrations.

We found that the velocity had no effect on the thresholds if we compare the stimuli with their resulting vibration frequencies given by the ratio between the velocity and the spatial signal period. The detection threshold curves resulted in a U-curve with an optimal sensitivity between 100 and 200 Hz. These results demonstrate similar detection thresholds between ultrasonic friction-modulated gratings and vibrotactile ones.

Friction variations were measured with a one-of-a-kind force sensor using interferometry to provide unmatched precision over a frequency range spanning continuous forces up to kilohertz force fluctuations. Analyses of the transfer function between the different relevant variables provided some clues that could explain the curve shape for the thresholds. This study improves our understanding of the impact of the exploration speed on the perception of virtual features on haptic touchscreens.

Chapter 4

Haptic and audio perception of evolving textures

Contents

Adapted from: Corentin Bernard, Jocelyn Monnoyer, Michael Wiertlewski, and Sølvi Ystad. "Rhythm perception is shared between audio and haptics". Scientific Reports, under review.

Preface to Chapter 4

The previous chapter examined the perception of uniform gratings at various spatial frequencies. In the current chapter, I investigate the perception of evolving textures. This type of synthetic textures is constructed with a gradient of spatial frequency. By consequence, the ridge density of the texture gradually increases or decreases during the tactile exploration according to the gradient value. By touching such stimuli, we experience textures that become finer or coarser. The first psychophysical experiment investigates the conditions under which the evolution of the gradient can be perceived. Since the results reveal striking similarities with the auditory perception of rhythmic gradients, I further explore the interaction between haptic and auditory modalities. Sonification principles from the literature are applied to construct a congruent sound mapped to the relief of the texture and to the finger movement. I show that, combined with haptic feedback, this auditory feedback improves the perception of gradients.

The results from the previous experiment were used in this chapter to achieve an equalization of the perceived intensity over the spatial frequency range. For this purpose, we assumed that the iso-intensity curves follow the same trend as the threshold curve, as with audition (see Figure 2.6). It provides a corrective factor that amplifies the weakly perceived spatial frequencies and attenuates the stronger ones. This process ensures that subjects base their judgments of the gradient direction on spatial frequency changes and not on intensity cues.

4.1 Abstract

A surface texture is perceived through both the sounds and vibrations produced while being explored by our fingers. Because of the vibrations' common origin, both modalities have a strong influence on each other, particularly at higher frequencies (above 60 Hz), for which vibrotactile perception and pitch perception share common neural processes. However, whether the sensation of rhythm is shared between audio and haptic perception is still an open question. Here we show striking similarities between the audio and haptic perception of rhythmic changes, and that both modalities interact in this frequency range (below 60 Hz). Using a new surface-haptic device to synthesize arbitrary audio-haptic textures, psychophysical experiments demonstrate that the perception threshold curves of audio and haptic rhythmic gradients are the same. Moreover, multimodal integration occurs when audio and haptic rhythmic gradients are congruent. We propose a multimodal model of rhythm perception to explain these observations. These findings suggest that audio and haptic signals are likely to be processed by common neural mechanisms also for the perception of rhythm. They provide a framework for audio-haptic stimulus generation that is beneficial for nonverbal communication or modern human-machine interfaces.

4.2 Introduction

When we explore a texture with our fingers, the interaction between the skin and the surface produces vibrations that propagate both through the air, up to our ears, and through our skin, down to our mechanoreceptors. Both sensory channels contribute to the perception of the texture properties [Lederman, 1979]. These audio and tactile vibrations emanating from the same source are perceptually merged into a single amodal percept, creating a mental image of the surface [Lederman et al., 2002]. As both stimuli share the same origin, the two modalities greatly influence each other. Altering the frequency content of the touch-produced sound can bias the perception of tactile roughness [Guest et al., 2002, Suzuki et al., 2008]. This effect, that can be produced when we rub our hands together, is known as the parchment skin illusion [Jousmäki and Hari, 1998]. While psychophysical experiments demonstrate high-level interactions between audio and tactile sensory systems [Crommett et al., 2017, Crommett et al., 2019, Yau et al., 2009, Fery et al., 2021], neuroimaging studies suggest that these interactions also occur in early sensory areas [Kayser et al., 2005, Schürmann et al., 2006, Caetano and Jousmäki, 2006]. These experiments reveal strong interactions and common neural processes for vibrotactile perception and pitch perception, for frequencies above 60 Hz. However, audio-tactile interactions with lower frequency content, associated with rhythm, in particular rhythmic changes, are rarely investigated.

In the present paper, we investigated the perception of audio and haptic stimuli in which the rhythm evolves continuously with time. We decided to use the

term *rhythm* that is here considered as the succession of events forming periodic patterns, which elements are distinguishable from each other, sticking to the definition given by Cooper et al. [Cooper et al., 1963]: "to experience rhythm is to group separated sounds into structured patterns". We use the term for beat rates up to 60 Hz [Ungan and Yagcioglu, 2014], frequency range that is more commonly characterized as flutter range in tactile perception [Talbot et al., 1968]. Whether the sensation of accelerating or decelerating rhythms is shared between audio and haptic perception remains unknown. In audio, these evolving stimuli are better known as accelerando or decelerando, in the case of tempo increase or decrease. In touch, it has been shown that a 10% variation in the ridge density can be detected [Nefs et al., 2001, Nefs et al., 2002]. Here, we generated haptic stimuli whose spatial frequency gradually evolves during exploration by a finger on a glass plate actuated with ultrasonic friction modulation. This method uses ultrasonic levitation to change the friction between the finger and the glass [Wiertlewski et al., 2016]. Modulating friction in reaction to users' exploratory motion produces sensations of texture, shape and relief on a flat surface [Winfield et al., 2007, Biet et al., 2007, Bernard et al., 2018]. In addition, the use of synthesized stimuli makes it possible to freely combine auditory and haptic stimuli. A similar setup has already been used to show audio-haptic perception changes with aging [Landelle et al., 2021].

In the present study, we modulated the friction with respect to the position of the user' finger. The modulation is a spatial sinusoidal wave, which spatial frequency gradually increases or decreases, becoming finer or coarser. This process is illustrated in Fig. 4.4b. Touching these haptic stimuli produces the sensation of bumps that becomes closer or more distant from each other, like accelerating or decelerating rhythmic patterns.

The perception of these haptic gradients is here investigated by a psychophysical experiment, whose results are compared with the literature on auditory perception. We further explain these observations with a multimodal model of rhythm perception. This model predicts similar auditory and haptic mechanism in the perception of rhythmic gradients, confirmed by a final multimodal experiment that demonstrates interaction between the two modalities.

4.3 Materials and methods

4.3.1 Haptic gradient construction

Haptic gradients are spatially encoded signals, in which the spatial frequency ν (in mm^{-1}) evolves as a finger explores a surface. The spatial frequency can be considered as the number of ridges per millimeter. For the gradient evolution to be perceived equally along the frequency range, we adapted the spatial frequency to the Weber Law. According to this law, the just noticeable difference (JND) $\delta\nu$ of the frequency is proportional to the initial frequency ν multiplied by a constant of

proportionality g:

$$\delta\nu = g\nu \tag{4.1}$$

We called g the gradient value (in mm^{-1}). The instantaneous spatial frequency, or localized-in-space frequency, of the grating ν obtained by integrating (4.1) varies as a function of the finger position x (in mm) according to the following relationship:

$$\nu(x) = \nu_0 \exp(g(x - x_0)) \tag{4.2}$$

where $\nu_0 = 0.5$ mm^{-1} is the central spatial frequency, which is the same for all stimuli, x is the finger position and $x_0 = 50$ mm is the center of the glass plate. The sine wave gradient y_g that oscillates at the instantaneous frequency ν was then defined as follows:

$$y_g(x) = \cos\left(2\pi \int_0^x \nu(u)\, du\right) \tag{4.3}$$

Thus, to avoid potential influences due to perceived intensity variations, the stimulus amplitude was adjusted under the 50 mm/s finger velocity condition according to data from a previous experiment [Bernard et al., 2020] that provided frequency-dependent intensity judgments obtained with the same haptic device as in the current study. This adjustment represents a corrective factor $a_c(\nu) \in [0.5,\ 1]$, which attenuates the signal in the mid-frequency bandwidth. A maximum attenuation of 0.5 was hereby applied at $\nu = 2$ mm^{-1}. The intensity of the stimulus was therefore perceived as constant along the gradient.

The windowing function ϕ for a given window size w (in mm) was defined as the difference between two sigmoidal functions:

$$\phi(x) = \frac{1}{1 + e^{-5(x - x_0 + w/2)}} - \frac{1}{1 + e^{-5(x - x_0 - w/2)}} \tag{4.4}$$

The windowing function was also centered on x_0. Finally, the resulting signal of the stimulus $A(x)$ with respect to the finger position was given by:

$$A(x) = \frac{1}{2} + \frac{1}{2}\phi(x) a_c y_g(x) \tag{4.5}$$

A(x) is the modulating signal (in %) encoding the friction, which is electronically multiplied by the ultrasonic carrier signal. It is presented for the 4 gradient value conditions and the 60 mm window size condition in Fig. 4.1.

4.3.2 Apparatus

In this setup, ultrasonic friction modulation is achieved on a $105 \times 22 \times 3.3$ mm glass plate. To track the finger position of the subject, a small ring is attached on the first phalanx of his/her index finger, which is connected to a pulley-encoder system. It measures unidirectional displacements along the length of the glass plate with an accuracy of approximately 0.01 mm and a refresh rate of 4 kHz without any significant

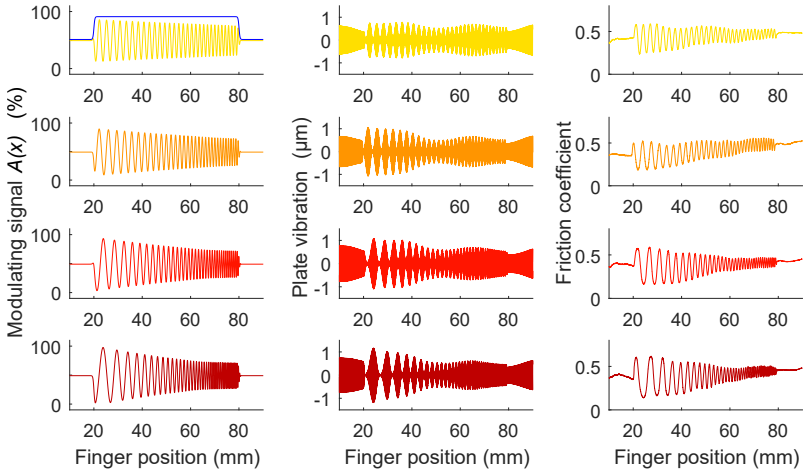

FIGURE 4.1: Presentation of the stimuli for the 4 gradient values (from top to bottom of the figure) under the 60 mm window condition with 3 types of measures. The left figures present the amplitude command of the modulating signal, the center present the measured vibration of the glass plate and the right present the measured friction coefficient between the finger and the surface. The amplitude is attenuated for higher frequencies to equalize the perceived intensity of the stimuli. Friction measurements are performed with a sensor described in a previous work [Bernard et al., 2020].

latency. A microcontroller (Teensy 3.5) reads the encoder and outputs a modulating signal (first column of Fig. 4.1) according to a friction map corresponding to the haptic stimulus. The carrier signal, a 35 kHz sine wave, is created by a function generator (BK Precision 4052) and amplitude-modulated by the analog signal provided by the microcontroller. The resulting signal is then amplified 20-fold (WMA-100, Falco Systems) to drive two piezoelectric actuators glued to the glass plate. Modulation of the amplitude of vibration of the glass plate (second column of Fig. 4.1) induces friction variations between the finger and the plate during the tactile exploration (third column of Fig. 4.1).

The graphical interface of the experiment, made with Max/MSP, is hosted an a computer connected to the microcontroller with serial communication. It handles the subjects' responses and audio feedback for the second experiment by receiving the finger velocity v_{finger} (approximately 50 mm/s), spatial frequency ν and windowing values in real time. The audio feedback is constructed as follows: filtered white noise (Butterworth 2nd-order bandpass filter between 400 and 800 Hz) is modulated from 0 to 100% by an oscillator at a frequency $f = \nu * v_{finger}$. Then, the windowing value acts as a gain from 0 when the finger is outside the window to 1 when the finger is on the haptic stimulus. Sounds are played through headphones (Sennheiser HD 26 Pro).

4.3.3 Protocol

Participants sat in a chair in front of the experimental desk and attached the ring connected to the position-tracking apparatus to their right index finger. Headphones

prevented any external auditory cues from the device. In each trial, the participants were asked to explore the actuated glass plate by sliding their finger from left to right and to synchronize their movement with a cursor presented on a screen in front of them that imposed a finger velocity of 50 mm/s, as presented in Fig 4.2. Participants could explore the stimulus only once. They were then asked to determine whether they felt that the ridge density increased (toward a "finer" texture) or decreased (toward a "coarser" texture) via a keyboard on the left-hand side of the setup. A training session familiarized the subjects with the terms and the corresponding stimuli.

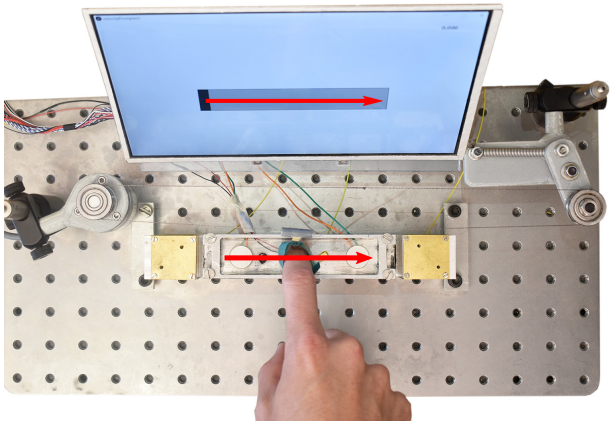

FIGURE 4.2: Experimental setup. The subject touches the actuated glass plate from left to right. The finger is linked to a pulley-encoder system for position measurement. The screen shows a cursor imposing the finger velocity.

In the first experiment, gradient perception was investigated following the method of constant stimuli for 4 value and 2 direction conditions, i.e., $g = \pm 0.015, \pm 0.025, \pm 0.035, \pm 0.045$ mm^{-1}, 6 window size conditions, i.e., $w =$ 10, 20, 30, 40, 50, 60 mm, and 8 repetitions, which led to $4 \times 2 \times 6 \times 8 = 384$ trials. Stimuli were presented in random order. Other parameters, such as the finger velocity and the central frequency of the gradient, were set as constants.

During the second experiment on multimodal perception, audio feedback was delivered through the headphones. The protocol was almost the same: 3 modality conditions, i.e., haptic only, audio only, and audio-haptic, for one gradient value and 2 directions, i.e., -0.025 and +0.025 mm^{-1}, 6 window size conditions, i.e., 10, 20, 30, 40, 50, 60 mm, and 8 repetitions, which led to $3 \times 2 \times 6 \times 8 = 288$ trials. The whole experiment lasted for approximately 90 minutes.

4.3.4 Subjects

Twenty-one subjects (5 females), 20 right-handed and 1 left-handed, ranging from 19 to 52 years old (mean 29) participated in the study. All the subjects participated in both experiments. They gave their informed consent before the experiment. All procedures were approved by the Ethical Committee of Aix-Marseille University and the experiment was carried out according to the relevant guidelines and regulations

expressed in the 1964 Declaration of Helsinki. They were paid for their participation. They washed and dried their hands before the experiment, and the glass plate was regularly cleaned with an alcoholic solution. Three subjects showed incoherent results, either due to technical issues or misunderstanding of the task. We defined a criterion for subject exclusion based on the percentage of correct answers under the easiest conditions: $g = 0.045$ mm^{-1} and $w = 50$ and 60 mm. The three subjects appeared as outliers according to the Tukey Fences method applied to these criteria, and their results were therefore discarded. Concerning the multimodal experiment, the same criterion was applied to the audio condition, and another subject was classified as an outlier. This subject's results were discarded from the multimodal analysis only.

4.3.5 Threshold measurement and statistical analysis

For each window size and gradient value condition, the answers from all the subjects were gathered to calculate the proportion of trials in which the stimulus was felt as becoming finer or coarser. The results, presented in Fig. 4.3a, reveal that for the smallest window size condition (10 mm), the subjects were not able to feel the difference between increasing and decreasing gradients, but this difference became more perceptible as the window size increased.

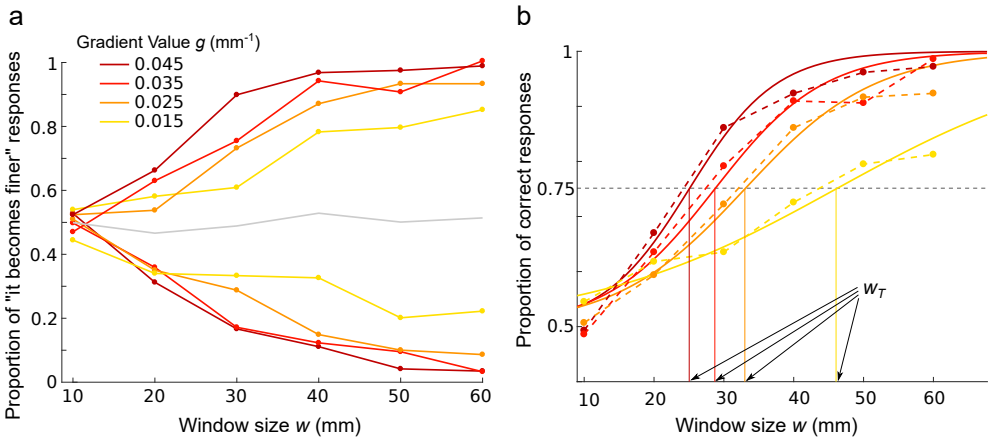

FIGURE 4.3: **a.** Proportion of "it becomes finer" against "it becomes coarser" responses for all subjects. Each point is the result of 144 trials (18 subjects $\times$ 8 repetitions) for one condition of the gradient value and window size. Positive value conditions (increasing gradients) appear in the upper part, and negative value conditions (decreasing gradients) appear in the lower part. The gray line shows the mean of the proportions across gradient value conditions to control a possible bias toward one type of response. **b.** Proportion of correct answers of all subjects (dashed lines) fitted by psychometric curves (solid lines). Window size thresholds w_T are measured at the level of 75% correct answers.

Since the mean of the proportions across gradient value conditions (gray line) remained at approximately 0.5, there was no bias toward one type of response. It

was therefore possible to average the increasing and decreasing conditions to obtain a common (direction-independent) gradient value. The proportion of correct answers calculated accordingly is presented in Fig. 4.3b. Similarly, we can observe that for small window size conditions, the proportion of correct answers was around the level of pure chance (50%), and as the window size increased, subjects tended to achieve correct answers 100% of the time, with slight variations within gradient value conditions. These data were fitted by psychometric curves given by the sigmoid function with the parameters γ and β:

$$sig(w) = 0.5 + \frac{0.5}{1 + e^{-\gamma(w-\beta)}} \tag{4.6}$$

The psychometric curves enabled us to measure the exploration distance thresholds required to perceive the gradient, i.e., the minimal exploration windows w_T to obtain 75% correct answers. w_T for the 4 gradient value conditions was calculated from the results of all 18 subjects. To measure if the effect of the gradient value was significant, we performed a method based on the jackknife resampling technique [Miller, 1974] used in [Kee et al., 2006] and [Micoulaud-Franchi et al., 2011]. This method, also called "leave-one-out", consists of running the analysis repeatedly while excluding one of the 18 subjects for each run, which means that the operation was repeated 18 times. A nonparametric Kruskal–Wallis test was performed on the 18 samples with the gradient value as factor. The test revealed a significant effect (at $\alpha = 0.05$) of the gradient value ($\chi^2_3 = 67, p < 0.001$). For the second experiment on multimodal integration, we ran exact same analysis procedure.

4.4 Results

4.4.1 Haptic detection of periodicity changes

The first experiment investigated how the exploration distance, constrained by an exploration window w, influences the detection of gradient g. The experimental design draws inspiration from studies on auditory perception, which explore the minimal duration needed to perceive a frequency or tempo change at a given rate of change [Madison, 2004, Yanagida et al., 2016, Sapp, 2006, Hart et al., 2006, Sergeant and Harris, 1962, Klatt, 1973, Pollack, 1968, Schouten, 1985, Rossi, 1971].

The detection thresholds were investigated for 4 gradient value conditions ($g =$ 0.015, 0.025, 0.035, 0.045 mm^{-1}) and 2 directions (increasing or decreasing) with 6 window sizes ($w =$ 10, 20, 30, 40, 50, 60 mm). A stimulus corresponding to the increasing direction is presented in Fig. 4.4a. In each trial, subjects had to explore the stimulus once and report if they felt that the stimulus "became finer" or "became coarser", which corresponded to increasing or decreasing spatial frequencies, respectively. Subjects' responses and the related analyses are presented in Fig. 4.3a and 4.3b. The percentages of correct answers for all subjects and for each condition are fitted with psychometric curves to obtain the window size thresholds w_T. The

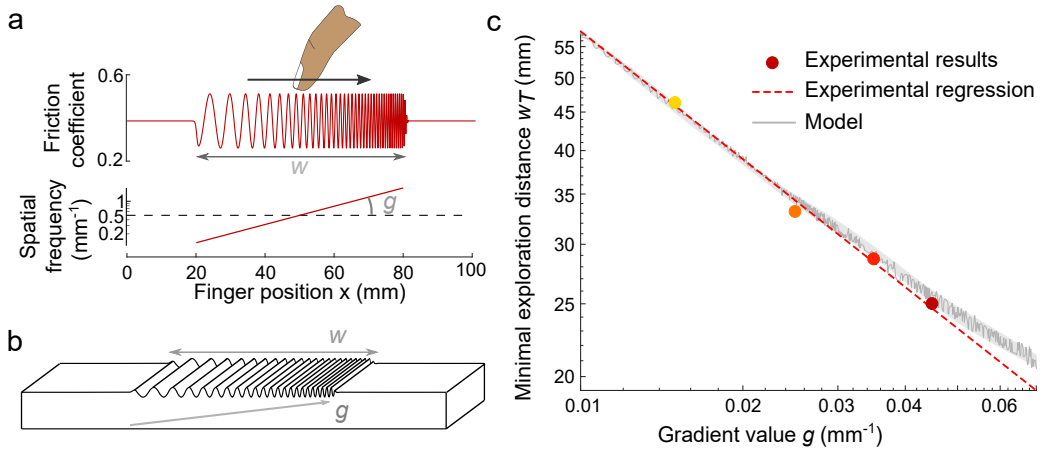

FIGURE 4.4: Overview of the experiment on haptic gradient detection thresholds. Subjects explore a synthetic sinusoidal grating whose spatial frequency evolves with respect to the finger position. **a.** Illustration of the haptic stimulus for the gradient value condition $g = 0.045$ mm^{-1} and the window size $w = 60$ mm. **b.** Illustration of the illusion produced by modulating fingertip friction with the haptic interface. **c.** Minimal exploration distances w_T necessary to detect the variation in ridge density, shown as dots for various gradient values g. Logarithmic regression, shown by the dashed red line, leads to a goodness of fit $R^2 = 0.997$. The model predictions are shown in dark gray, and the margin of error is shown in light gray.

minimal exploration distances to perceive a change in the gradient value g = 0.015, 0.025, 0.035 and 0.045 mm^{-1} were found to be w_t = 46.2, 33.1, 28.7 and 25 mm, respectively. The thresholds w_T decrease as the gradient value g increases with a linear dependency on a logarithmic scale, as illustrated Fig. 4.4c. A logarithmic regression relveals a significant correlation (p=0.004) between the window size threshold and the gradient value such that $\log(w_T) = 1.50 - 0.55\log(g)$, which can also be written as $w_T \times g^{0.55} = 4.48$.

4.4.2 Comparison of audio and haptic thresholds

The exploration distance w_T and the gradient value g are not proportional, but follow a power law with an exponent of 0.55. To compare the results of this experiment with data from the literature on tempo and frequency gradients in auditory stimuli [Madison, 2004, Yanagida et al., 2016, Sapp, 2006, Hart et al., 2006, Sergeant and Harris, 1962, Klatt, 1973, Pollack, 1968, Schouten, 1985, Rossi, 1971], the exploration distances w (in mm) and gradient value g (in mm^{-1}) were converted into stimulus durations $\Delta T = w/v_{finger}$ (in s) and frequency rates $r = g \times v_{finger}$ (in s^{-1}), respectively, using the finger velocity $v_{finger} = 59.6 \pm 9.7$ mm/s. Participants were asked to explore the stimuli with a constant speed by synchronizing their finger movement with a cursor on a visual display. Fig. 4.5 provides a comparison between our results and the literature data. The haptic gradient threshold curves

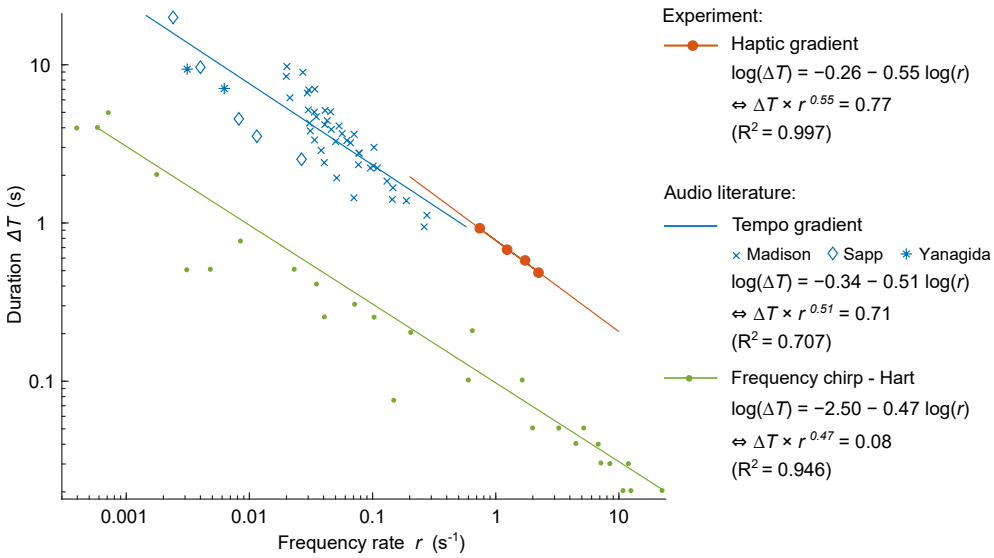

FIGURE 4.5: Comparison of the experimental results with the literature. Data points are fitted with logarithmic regressions (displayed as solid lines). The regression equations are presented in two forms with their goodness of fit. Thresholds for the haptic gradients, converted to time with the finger velocity value, are shown in red. Thresholds from the literature on the perception of tempo gradients (accelerando) [Madison, 2004, Yanagida et al., 2016, Sapp, 2006] are displayed in blue. Thresholds from the meta-analysis of Hart [Hart et al., 2006], who gathered data from many studies on the perception of frequency chirps (glissando), are shown in green [Sergeant and Harris, 1962, Klatt, 1973, Pollack, 1968, Schouten, 1985, Rossi, 1971].

strongly resemble the audio tempo gradient threshold curves, with the only difference being that the haptic thresholds are presented for shorter durations. Textures of a few centimeters explored at a velocity of approximately 50 mm/s typically lasted approximately 1 s, which is indeed below the usual tempo durations for audio stimuli. The graph also shows that the slope distribution of the tempo and haptic gradients is close to that obtained for frequency chirps.

Stimuli	Logarithmic regression	Equivalent expression
Haptic gradient	$\log(\Delta T) = -0.26 - 0.55\log(r)$	$\Delta T \times r^{0.55} = 0.77$
Audio tempo variation	$\log(\Delta T) = -0.34 - 0.51\log(r)$	$\Delta T \times r^{0.51} = 0.71$
Audio frequency chirp	$\log(\Delta T) = -2.5 - 0.47\log(r)$	$\Delta T \times r^{0.47} = 0.08$

TABLE 4.1: Logarithmic regressions of the threshold curves from the experiment (haptic gradient) and the audio literature.

To numerically investigate these similarities, we performed logarithmic regressions which yielded the equations in Table 4.1. We can compare the values of the exponent e and the constant c of the threshold laws $\Delta T \times r^e = c$. This analysis confirms that the 3 exponent values are in the same range and, most importantly, that the haptic and auditory tempo values differ by only 7.3% and 7.8% (relative error) for e and c, respectively.

In summary, haptic gradient thresholds follow the same law as rhythmic gradients. This suggests that similar mechanisms are activated in the two modalities for low-frequency gradient perception.

4.4.3 Perceptual model of audio-haptic rhythmic gradients

To investigate these mechanisms, we adapted a model from the literature on the perception of irregular rhythmic patterns based on the work of Schulze [Schulze, 1978]. Three theories compete to explain the encoding of tempo perception. *Successive interval discrimination* theory proposes that each interval between two beats is compared with the previous interval. When a difference exceeds a given threshold, an irregularity is perceived. *Comparison with an internal rhythm* theory states that the first beats are internalized and used as a rhythmic reference. When a beat differs by more than a threshold from the reference, an irregularity is perceived. *Comparison with internal interval* theory is quite similar to that with internal rhythms but uses the interval rather than the rhythmic difference. It postulates that the first interval is internalized and used as a reference. When a duration difference between one interval and the reference exceeds a certain threshold, an irregularity is perceived.

These theories were tested by Schulze on beat sequences that contained carefully chosen irregularities. The results of his study revealed that the *internal rhythm* theory was a good predictor but that the results were also in agreement with the internal interval model predictions. The experiment was reproduced by Keele et al. [Keele et al., 1989], who concluded that the *comparison with the internal interval* theory was more likely to predict the perceived rhythm. A generalization of Schulze's model was later proposed to take into account the influence of the initial pace [Vos et al., 1997]. Investigating the perception of linear tempo gradients [Madison, 2004], Madison explained his results using models of previous studies with the principle of accretion, which considers that the accumulation of small differences reinforces the global difference.

The haptic frequency gradient perception model, derived from its audio counterpart, is illustrated in Fig. 4.6. First, the haptic signal encoding the friction is converted into a pulse train, where each pulse corresponds to the local maximum of the virtual shape. The pulse train signal mimics the response of the Pacinian channels to sinusoidal stimulation [Bell et al., 1994]. The duration between two pulses τ is then computed. The probability of perceiving 2 intervals of duration τ_1 and τ_2 as identical is described by the probability distribution $P(\alpha)$. We assume that the probability of perceiving a difference in successive intervals depends on the duration ratio $\alpha = \tau_2/\tau_1$ and follows a log-normal function as presented in Fig. 4.6a. The standard deviation of logarithmic values σ is the only parameter that is fixed in the model. We compared three theories of tempo perception (see Fig. 4.8 in the supplementary materials). Among the three theories, the internal interval with accretion theory is the best predictor of the observed results. In this model, the first interval is internalized and used as a reference, and then each interval duration is compared

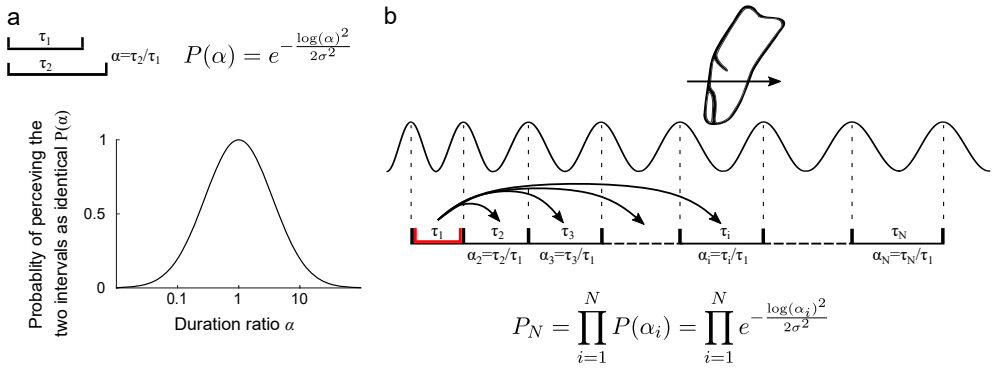

FIGURE 4.6: Model of tempo perception applied to haptics. **a.** Formula and basic principle of the model: the probability P of perceiving two intervals as having the same duration follows a log-Gaussian function with respect to the ratio between the two interval durations α. **b.** Application of the model: the haptic stimulus is converted into a pulse train. The theory of *comparison with an internal interval with accretion* is then applied to the intervals τ_i to calculate the probability of perceiving the gradient.

to the reference duration to calculate the probability P of perceiving no difference. The small imperceptible variations are compounded using the accretion principle, which stipulates that the accumulation of small differences reinforces the global difference. The overall probability of perceiving no change in the stimulus P_N is then the product of all the previous probabilities $P(\alpha_i)$. The final probability of perceiving the change in frequency is given by $P_{g,w} = 1 - P_N$.

In line with the experimental design, this procedure is applied to all gradient magnitude g and window size w conditions. The theoretical thresholds are calculated by performing the same analysis with psychometric curve fittings (see Fig. 4.9). These thresholds are defined as the critical window sizes w_T that yield a 50% chance of perceiving the irregularity ($P_{g,w_T} = 0.5$). To minimize the error between the four w_T values of the model and of the experiment, we optimize the parameter σ of the log-normal distribution. We find that $\sigma = 1.153$ leads to the best predictions of the observed data with a mean relative error of 1.74%. The proposed model can also extrapolate the experimental thresholds to a broader range of values, as presented in Fig. 4.4c.

4.4.4 Audio-haptic interaction

The previous results hint at a shared process between the haptic and the audio perception of rhythmic gradients. To test whether both modalities do indeed influence each other, we measured their influence on the overall detection threshold when both modalities were present. This methodology was successfully used in previous studies to unravel the interaction between haptics and other modalities

[Ernst and Banks, 2002, Ro et al., 2009, Lederman and Abbott, 1981]. To find a multimodal interaction, haptic stimuli were enhanced with congruent auditory stimuli. Audio feedback was synthesized from filtered white noise, which evokes a natural interaction sound, such as rubbing [Conan et al., 2012]. These signals were amplitude-modulated by an oscillator whose frequency matches that of the haptic stimuli used to render the virtual shape. The auditory stimuli were hereby perceived as successive beats with increasing or decreasing tempos perfectly synchronized with the haptic stimuli both in terms of modulation and time window, as illustrated in Fig 4.7a. Since the auditory stimuli were amplitude modulated signals derived from filtered white noise, no noticeable pitch was perceived, even for frequencies above 30 Hz.

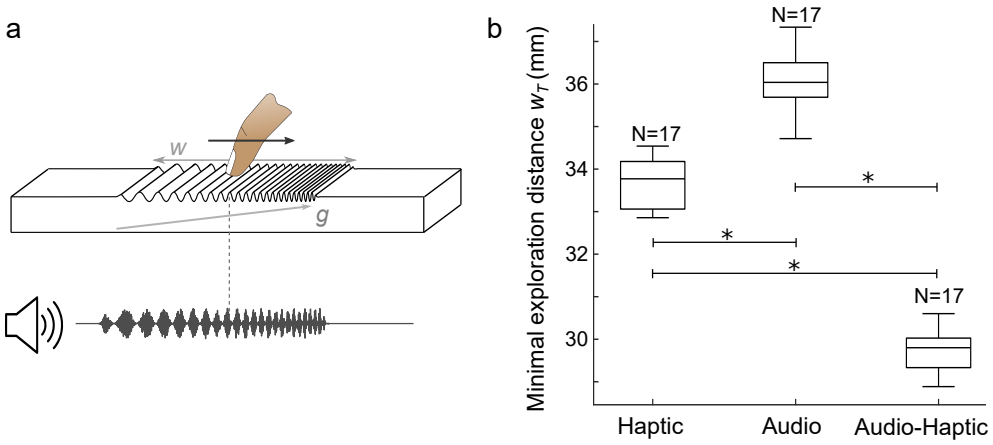

FIGURE 4.7: Audio-haptic interaction. **a.** Illustration of the multimodal experiment with audio feedback derived from the haptic texture. **b.** Perception threshold for the unimodal and multimodal conditions. The thresholds are represented by the minimal exploration distance to perceive a difference in frequency when the gradient value $g = 0.025$ mm^{-1}.

In a second experiment, we investigated whether the addition of audio feedback could improve the gradient perception and thus lower the minimal exploration distance. Thresholds were investigated for the 0.025 mm^{-1} gradient value condition with the same 6 window sizes (from 10 to 60 mm) for 3 modality conditions: haptic only, audio only and bimodal audio-haptic. The analysis of the subject responses was achieved with the same algorithm as in the first experiment. We used the jackknife resampling method for the statistical analysis. This method is presented in the materials and methods section, and the thresholds are presented in the boxplot in Fig. 4.7b.

A non-parametric Kruskal–Wallis test was performed with the sensory modality as a factor (haptic, audio or audio-haptic) on the 17 samples from the jackknife method. The test shows a significant effect (at $\alpha = 0.05$) of the modality ($\chi^2_2 = 44, p = 2.2 \times 10^{-10}$). This result was validated by post hoc Nemenyi tests, which presented significant differences between the 3 modality conditions (H-A: $p = 2.5 \times 10^{-3}$, H-AH: $p = 2.5 \times 10^{-3}$, A-AH: $p = 7.8 \times 10^{-11}$).

Haptic-only and audio-only thresholds fall into the same order of magnitude, as we expected from the previous comparison in Fig. 4.5. However, when both modalities are combined, the detection threshold is significantly reduced. Both modalities contribute to heightening of the sensitivity, and the results indicate a multimodal integration of rhythm.

4.5 Discussion

Audio and haptic perception are known to interact for pitches above 100 Hz, and in this series of experiments we demonstrated that this interaction extends to the perception of rhythm and its temporal evolution. When exploring with our bare finger a simulated surface, rhythm was haptically perceived after a minimal exploration distance. This distance followed a power law with the rate of change in the frequency, with an exponent of $\approx$ 0.5, matching the known behavior of auditory signal perception when the tempo increases or decreases. A model of audio tempo perception, based on an internal interval theory, accurately predicted the results of the haptic experiment.

Adding congruent audio feedback to the haptic sensation resulted in a significant interaction between the two modalities. When the auditory variations of the tempo followed the shape of the haptic signal, the participant needed 12% less distance to achieve an accurate detection. These results attest to a bimodal integration of audio and haptic stimuli, which suggests that the perception of the energy envelope of audio and haptic signals shares common perceptual mechanisms.

The model predicts the probability of perceiving both audio and haptic rhythmic variations on a wide range of variation rates and durations; however, its applicability has limits. If we take a closer look at the probability distribution shown in Fig 4.6a, it is difficult to interpret the value of the parameter $\sigma = 1.153$. This value would lead to a just noticeable difference between two interval durations of 289% (P(0.389)=0.5), which is not coherent with the just noticeable difference of $\approx$ 10% reported in previous works [Drake and Botte, 1993].

Because of the protocol design, some stimuli in our experiment tended to exceed the flutter range (<60 Hz) at their extremity, making the interpretation more complex. However, this issue concerns only the stimuli with the longest time windows; the stimuli at the thresholds all remain within 17 to 52 Hz, corresponding to the flutter range. This limit comes from the fact that two tactile events need to be separated by at least a certain duration to be perceived independently. Considering two successive isolated pulses, the minimal interval is about 40 ms [Pastor et al., 2004], whereas with pulse trains, the limit between the flutter range (discrete events) and the continuous vibration range was evaluated at 60 Hz ($\sim$15 ms intervals) [Mountcastle et al., 1967, Talbot et al., 1968]. This value has been verified using both periodic [Birznieks and Vickery, 2017, Ng et al., 2020] and aperiodic [Ng et al., 2018] stimuli: the authors showed that the tactile sensation of

frequency is determined by the duration of the silence intervals between the pulses. Intervals longer than 15 ms are crucial whereas shorter intervals have only a limited effect on the frequency evaluation, exhibiting the limit of the discrete perception below this value. Haptic signals with higher frequencies are felt more as continuous textures. This principle also occurs in audition: audio beating progressively turns into sound roughness from 30 to 100 Hz, and then to pitch. The present model is limited to a specific frequency bandwidth, and does not take into account phenomena that occur outside this band. However, the comparison principle and accretion of the probabilities are likely to be effective for the perception of higher frequency changes as well, as seen in Fig. 4.5. In the future, this limit of the model could be overcome by applying weights to each interval with respect to its length to minimize the effect of small indistinguishable intervals, as the approach discussed previously [Birznieks and Vickery, 2017, Ng et al., 2018].

The findings extend previous works that showed similarities between the perception of tactile vibration and auditory pitch to the perception of discrete, dynamic low-frequency stimuli: time-varying haptic gratings and audio pulse trains. In addition, the results are of interest in the field of human-machine interfaces for the design of textural haptic feedback to guide the user on touchscreens without requiring visual attention. Other studies have shown that haptic gradients with intensity variations are promising for guiding exploratory motion [Klatzky et al., 2017, Bodas et al., 2019]. Our findings extend this promise by showing an unambiguous relationship between the exploration window and the magnitude of the gradient ($w \times g^{0.55} > 4.48$) to create salient stimuli. The perception of these stimuli can be further enhanced by adding congruent audio feedback.

In addition, these results open up new perspectives related to nonverbal communication and sensory substitution. In speech, for instance, frequency transitions are central in conveying information [Schouten, 1985]. The emotional aspect of speech is strongly conveyed through fundamental frequency changes (f0 trajectory) of voiced vowels (parent-child communication) [Mithen et al., 2006]. It has been shown that downward pitches are often associated with cold or anger and rising pitches with fear, surprise and happiness [Bänziger and Scherer, 2005]. Both temporal and frequency variations (portamento, accelerando/ritardando) are also extremely important for conveying emotions through music [Schubert and Wolfe, 2013]. Hence, although only time-varying pulses have been explored in the present study, perceived glissandi in the haptic domain could help produce emotional reactions in line with musical stimuli, since similar mechanisms are activated in the tactile and auditory domains.

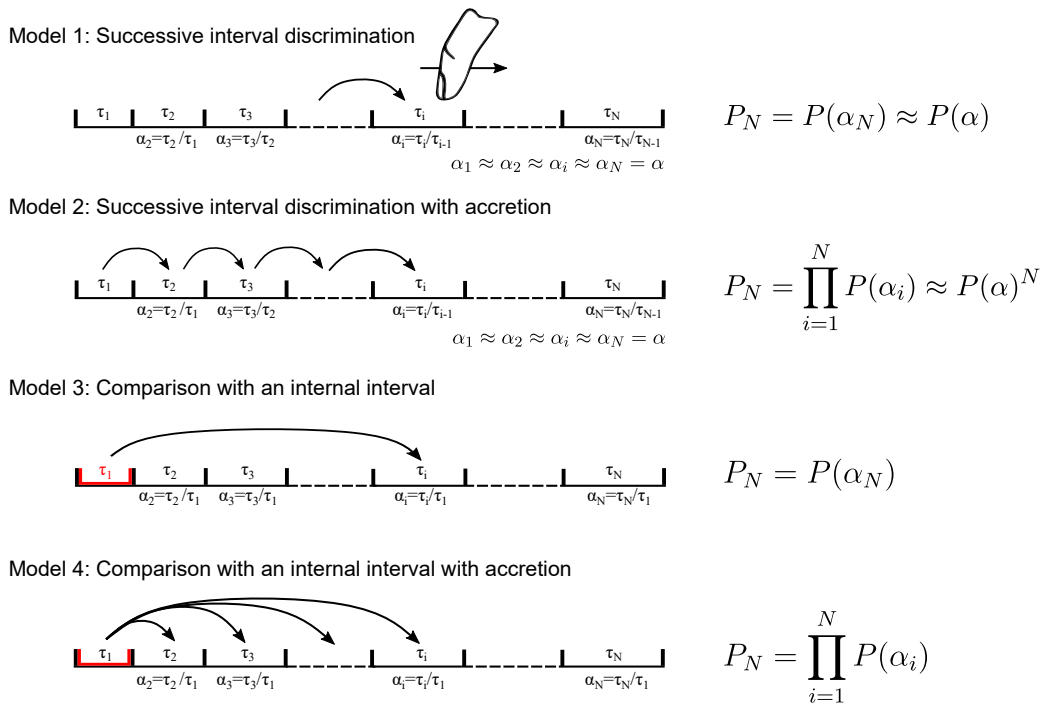

FIGURE 4.8: Presentation of the models performed with 4 theories on tempo perception adapted from the literature.

4.6 Supplementary materials: comparison of different theories

The model we proposed was also evaluated with alternative theories of tempo perception derived from the literature, as presented in Fig. 4.8.

1. Successive interval discrimination: each interval is compared with the previous interval.

2. Successive interval discrimination with accretion: each interval is compared with the previous interval and previous comparisons are kept in mind.

3. Comparison with an internal interval: the first interval is internalized, and each interval is compared with this reference.

4. Comparison with an internal interval with accretion: the first interval is internalized, and each interval is compared with this reference. Previous comparisons are kept in mind. This is the model presented in the core of the article.

Because the stimuli of the experiment present intervals that are either all increasing or all decreasing, there are no irregularities due to variation in the direction. Hence, we do not evaluate the theory of comparison with an internal rhythm from [Schulze, 1978]. In our case, this is equivalent to the theory of comparison with an internal interval.

	σ_{opt}	RMSE	error %
Model 1	/	/	/
Model 2	0.118	90.06	30.32
Model 3	0.539	13.71	11.04
Model 4	1.153	0.412	1.74

TABLE 4.2: Results of the optimization of σ on the root-mean-square error (RMSE) between the predicted thresholds and the thresholds from the experiment for the 4 models. The error is also expressed as the percentage error for visualization. Optimization with model 1 is not possible.

For a given gradient value g and window size w, the corresponding haptic stimulus is converted into a pulse train, where each pulse coincides with a maximal friction value. According to the interval duration of the pulse train, the probability of not perceiving any change P_N is computed. These probabilities are plotted in Fig. 4.9. The value $1 - P_N/2$ is preferred to match the experiment, which is designed with a two stimuli-one interval paradigm. The analysis is then the same as for the experimental results. Data are fitted with psychometric curves to measure the window size which gives $1 - P_N/2 = 0.75$.

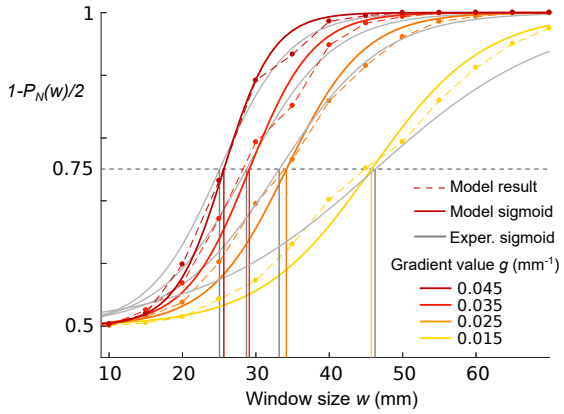

FIGURE 4.9: Comparison between the prediction of Model 4 and the experimental results. Probabilities are calculated with $\sigma_{opt} = 1.153$. The probability of perceiving the gradient variation is plotted for window sizes every 5 mm for the 4 gradient value conditions. These data are fitted by sigmoid curves to find the window size thresholds predicted by the model. Thresholds from the experiment are plotted in gray.

For each model, an optimization of the parameter of the probability distribution function σ (see Fig. 4.6a) is performed. A gradient descent algorithm finds the value of σ_{opt} that minimizes the root-mean-square error between the predicted thresholds w_T from the model and those from the experiment. Table 4.2 compares the accuracy of the models.

Because P_N does not change with the number of intervals N for model 1, since the window size has no influence on P_N, optimization is not possible. This model is thus discarded. Comparing the errors, it appears unequivocally that model 4 has the

best prediction performance based on the experimental results. This model is also the one that stands out from the literature.

Chapter 5

Evolving textures as feedback for gesture guidance

Contents

Adapted from: Corentin Bernard, Jocelyn Monnoyer, Sølvi Ystad, and Michael Wiertlewski. "Eyes-Off Your Fingers: Gradual Surface Haptic Feedback Improves Eyes-Free Touchscreen Interaction", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2022, under review.

Preface to Chapter 5

In the previous chapter, I presented a specific type of evolving haptic textures: the spatial frequency gradients, and studied how they are perceived. In this chapter, I show that this type of haptic feedback can be implemented in a human-computer interface and I demonstrate that this haptic interaction is appropriate to guide the user. The study focuses on the simple task of setting a parameter on a touchscreen, here performed without vision. Since the interaction is not straightforward, different training strategies are proposed to assimilate the haptic feedback with the aid of other modalities.

Previous findings on the minimal exploration distance to perceive a texture with a given spatial frequency gradient are used here to design the haptic stimulus. The threshold law ensures that the direction of the gradient is perceived with a single finger slide on the interface. Moreover, the previous conclusions that highlighted similarities in the processing of rhythm between auditory and haptic modalities are here taken into account to propose congruent auditory feedback to assist the learning of the haptic interaction.

5.1 Abstract

Moving a slider to set the music volume or control the air conditioning is a familiar task that requires little attention. However, adjusting a virtual slider on a featureless touchscreen is much more demanding and can be dangerous in situations such as driving. Here, we study how a gradual tactile feedback, provided by a haptic touchscreen, can replace visual cues. As users adjust a setting with their finger, they feel a continuously changing texture, which spatial frequency correlates to the value of the setting. We demonstrate that, after training with visual and auditory feedback, users are able to adjust a setting on a haptic touchscreen without looking at the screen, thereby reducing visual distraction. Every learning strategy yielded similar performance, suggesting an amodal integration. This study shows that surface haptics can provide intuitive and precise tuning possibilities for tangible interfaces on touchscreens.

5.2 Introduction

Interacting with machines often consist in selecting programs or adjusting parameters, such as turning a knob to increase the volume, or moving a slider to set a program on a dishwasher. The physicality of these primitive interfaces is slowly disappearing in favor of versatile touchscreens, which can combine thousands of operations in a single unit. Touchscreens have the advantage of allowing reprogrammable tactile inputs and co-localized visuals. However, the generalization of such interfaces also come with an important trade-off on the type of feedback they provide to the user. A well-designed mechanical knob provides distinctive haptic responses, via its 3 dimensional shape when searched for, and via its mechanical impedance –the motional response to dynamic forces– when operated, which is pivotal to guide the user into reaching and operating the knob. On the other hand, the featureless flat pane of glass of a touchscreen provides no haptic cues on the location of the interface, nor on its operation. As a consequence, interface designers provide visual cues – progress bars, numerical values, or sliders– or auditory cues –clicks– as a way to indicate the progression of a setting change in response to the user's motion. However, relying on visual or auditory modalities is often not suitable or even possible. For example, while driving, the use of dashboard screens causes visual distraction, and in dark environments visual cues might not even be available. Auditory feedback can also be degraded due to noisy environments.

Surface haptics has emerged as one potential solution to restore tangibility to touchscreens. It creates haptic sensations directly onto the user's fingertip by modifying the friction between the skin and the glass plate. Controlling friction can be achieved either using ultrasonic vibration [Biet et al., 2007, Wiertlewski et al., 2014, Wiertlewski et al., 2016] or electroadhesion [Bau et al., 2010, Shultz et al., 2015, Vardar et al., 2017a], and by modulating the friction force as a function of the position on the screen, sensations of textures and shapes can be

recreated [Friesen et al., 2021, Messaoud et al., 2016b, Bernard et al., 2018]. Creating zones of low friction can also help in guiding the user to a given target point, improving the information capacity of the user [Casiez et al., 2011, Levesque et al., 2011].

Here, we show that in the particular interaction of tuning an integer value, surface haptics can significantly improve the interaction. Such adjustment tasks are found when tuning the ventilation power, changing a sound level, raising the height of window blinds, and so on. In this chapter, the controlled parameter is referred to as the *setting*, noted γ, which can take a value between 0% and 100%. The interaction consists of adjusting the setting by sliding the finger up or down, which is made possible through clutching movements, where the finger hovers over the plate back to the initial position. As the setting changes, the user feels a synthetic texture on the haptic surface, whose spatial frequency (or number of ridges per millimeters) is proportional to the value. Through learning and practice, the user can learn how to associate the setting to the state of the texture, and use this information to adjust the value on an absolute scale.

We propose different learning strategies with visual and auditory stimuli to assess how subjects adapt to haptic feedback with the help of other modalities: a visual feedback rendered as a slider, and an auditory feedback for which the pitch and beat are tied to the value. To study the effect of the guidance hypothesis which states that learning with extra cues might lead to a poorer transfer to the actual task [Sigrist et al., 2013], we added two more conditions where the visual or the auditory feedback slowly fades out from trial to trial. Lastly, we added a fifth control condition, where only haptic feedback was present.

To evaluate the proposed interaction, we constructed an experiment in which subjects are asked to adjust the setting to a target value as quickly and precisely as possible, with first a learning session and then an evaluation session. The performance of the subjects are measured in terms of *completion time* and *adjustment error*. Based on these metrics, we compare the efficiency of the five different learning strategies in helping the subject to retain the interaction during the evaluation phase. We found that, after training, the subjects improved their setting performance, regardless of the learning strategies.

5.3 Related works

5.3.1 Eyes-free interaction

Haptic feedback of everyday objects reduces the need for visual attention during interface control. Using force feedback knobs and sliders, Snibbe et al. [Snibbe et al., 2001] introduced multiple rendering strategies to provide physical feedback to continuous interactions, for example while manipulating media such as video, audio or computer graphics. Haptic event count is a common communication strategy. This interaction is for example experienced in mouse wheels. Applying this approach to interactions with touchscreens proves effective to help

users set a value or select a menu by counting the number of vibration ticks they can feel from a vibrotactile surface [Liao et al., 2017] or with actuators directly on the wrist [Youn et al., 2021]. While these methods are based on our temporal perception and segregation of tactile events, other strategies exploit our tactile spatial acuity [Gupta et al., 2016] by conveying information through the localization of vibrations on the body. The perceptual motion of a vibration around the wrist enables the user to perform pointing, selection and drag-and-drop tasks without vision. In the present approach, users actively explore a texture with evolving spatial periods to set a given value. The access to the haptic information is directly linked to their movement, mimicking real world sensorimotor interactions.

5.3.2 Haptic feedback for pointing tasks

Setting a slider can be considered as pointing to the desired value. Human performance when reaching a target is well quantified by Fitts' law [Fitts, 1954]. This seminal finding predicts the time required to rapidly move to a target as a function of the size and the distance of the target, which together define the index of difficulty. This discrete model was extended to the continuous case of following a tunnel, by the steering law [Accot and Zhai, 1997]. These models initially considered the integration of visual cues to guide participants' motion. Recently, it has been shown that the addition of haptic feedback can also significantly reduce the time it takes, effectively illustrating that the difficulty of the task is reduced in the presence of an additional type of feedback [Casiez et al., 2011, Levesque et al., 2011, Kalantari et al., 2018]. These studies considered the simple case of a binary feedback, where friction was low everywhere but high on the target, or vice-versa. However, binary feedback only guides the user when a specific target is reached and not during the entire motion. Therefore, this kind of feedback is ill-suited to enhance tuning tasks, since the value of the setting is not known in advance by the rendering algorithm. For this purpose, continuous feedback, where the amplitude or the frequency evolves continuously, has shown potential for guiding users [Klatzky et al., 2017].

5.3.3 Auditory guidance using sonification

Other forms of feedback have also shown promising results for guiding the user on a touchscreen, in particular auditory feedback via sonification. Sonification is the use of non-speech auditory feedback to convey information to the user [Barrass, 1995]. It can be relevant in interface manipulation because the auditory system performs well at analyzing dynamic information and recognizing temporal and frequency patterns. It has even been shown that gesture sonification can bias our movements [Thoret et al., 2016]. The main perceptual attributes of a sound, i.e. pitch, loudness, tempo, and timbre (brightness, inharmonicity and roughness), have been compared in their capability of guiding the user gesture in a target experiment [Parseihian et al., 2016]. These attributes were

used to sonify the distance to the target, which means that subjects had to find the position of the lowest pitch or the fastest tempo, for example. Participants were faster and more precise with sonification strategies that varied in pitch and tempo. Pitch sonification is also the preferred strategy for mid-air gesture interfaces [Bressolette et al., 2018]. Moreover, the performance at two-dimensional path following tasks with audio and haptic guidance are similar across modalities, suggesting that the presence of any type of feedback is enough to guide the control of motor commands [Rocchesso et al., 2016, Del Piccolo et al., 2018].

5.3.4 Skill acquisition with multimodal feedback

Learning a new interaction with an interface is closely related to learning a new gesture, since a subtle coordination between sensory input and muscle output has to be found. We drew inspiration from the literature on motor learning to design our training method, with a particular attention towards methods related to practice and refinements of new skills [Nieuwboer et al., 2009]. Augmented feedback, through the use of visual, auditory or haptic modalities, has been shown to effectively enhances motor learning [Sigrist et al., 2013]. Two types of feedback are considered: the *concurrent feedback*, which is a real-time feedback presented during the task execution that helps to perform the gesture, and the *terminal feedback*, which present the performances after the completion of the task. This feedback can be a display of the error, the completion time or the trajectory. Concurrent feedback and terminal feedback are most effective during the early learning phase. However, to develop a persistent internal representation of the task, feedback frequency and intensity should be progressively decreased [Kovacs and Shea, 2011]. Indeed, the sudden removal of a feedback can cause a significant degradation of performance [Huang et al., 2006]. In [Robin et al., 2005], authors compared the effect of no feedback, full visual feedback or weak visual concurrent feedback (with low contrast), during the acquisition phase. They found that the latter leads to better pointing performance in the transfer phase when no more visual feedback is provided. By mitigating the appropriation of the guidance dynamics into the task dynamics, the assistance can progressively decrease during the training session [Patoglu et al., 2009].

5.4 Interface and haptic signal

5.4.1 Interaction design

The interaction that we propose is performed on an inclined touchpad presented in Figure 5.1.a. It mimics new interfaces that emerge in automobile cockpits, with a touchpad on the driver's right armrest. The touchpad provides control of the graphical dashboard display by sliding gestures. The interaction focuses on the use case of an absolute setting of a value, i.e. reaching a desired value, as opposed to a relative setting that consists of increasing or decreasing the value by a certain amount.

Clutching interaction is implemented to virtually increase the physical space of the touchpad and extend the setting range without loss of adjustment precision. Clutching has been shown to provide better performances than continuous gestures in some cases [Nancel et al., 2015]. The sliding gestures are performed in the vertical direction, from top-to-bottom to increase the setting value and from bottom-to-top to decrease it, as if interacting with a vertical carousel whose numbers increase in the upward direction. We chose to use ultrasonic friction-modulation rather than vibrotactile rendering for its more precise sensation delivery. Indeed, the just-noticeable difference of vibration frequencies (i.e. the lowest frequency difference to perceive a change in vibration) is between 18% and 20% [Pongrac, 2008] whereas it is between 6% and 11% for spatial frequencies of gratings [Nefs et al., 2001]. An illustration of the illusion that is produced by the haptic feedback is shown in Figure 5.2.a.

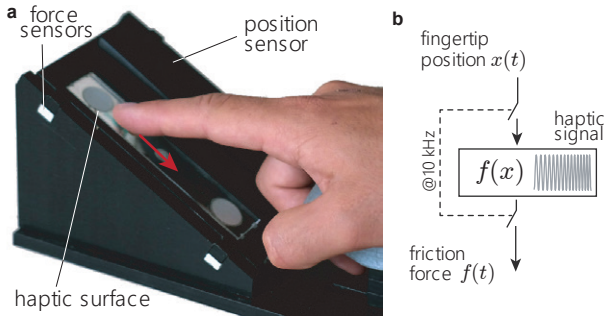

FIGURE 5.1: (a): Bespoke haptic touchpad used in the study. Users can explore the vertical surface and feel bumps and holes via ultrasonic friction-modulation. (b): The finger position is tracked in real time to update the frictional force according to the haptic signal.

5.4.2 Surface haptic display

The touchpad, presented in Figure 5.1.a, is a custom-made haptic surface actuated with ultrasonic friction modulation. The workspace is a $105 \times 22 \times 3.3$ mm glass plate. This size was chosen because it is technically easier to render powerful feedback on a plate of such dimensions rather than on a touchpad with conventional width, and it does not disturb one-dimensional vertical gestures. The position of the finger on the plate is measured at 400 Hz with an optical sensor (NNAMC1221PC01 Neonode Touch Sensor Module) connected to a microcontroller (Teensy 3.5). The microcontroller extrapolates the finger position at a frequency of 10 kHz and outputs a modulating signal according to the finger position and the friction map of the haptic stimulus, as presented in Figure 5.1.b. The carrier signal, a 35 kHz sine wave, is created by a function generator (BK Precision 4052) and amplitude-modulated by the analog signal provided by the microcontroller. The resulting signal is then amplified 20-fold (WMA-100, Falco Systems) to drive two piezoelectric actuators glued to the glass plate. Modulation of the amplitude of vibration of the glass plate induces friction variations between the finger and the plate during the tactile exploration as

presented in Figure 5.2.d. The presence of the finger on the plate is detected with 2 capacitive force sensors (CS8-1N, Cal 8mm diameter, 1N/0.22lb force, SingleTact). The microcontroller is connected to the graphical interface with USB serial communication. The setting γ, the finger position, the finger velocity v_{fing}, and the spatial frequency of the haptic feedback ν are thus transmitted in real time and recorded.

5.4.3 Haptic signal

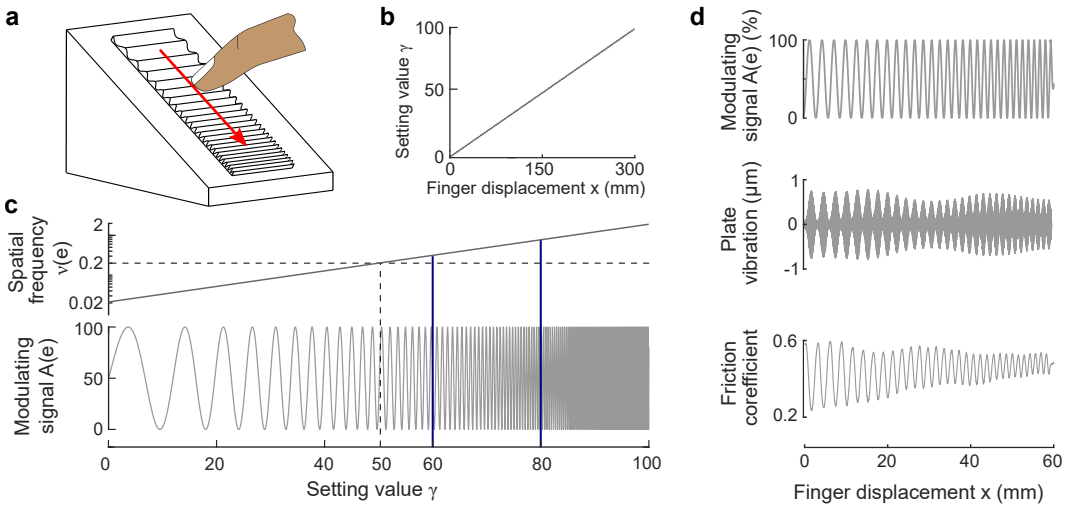

FIGURE 5.2: Description of the haptic feedback. (a): Illustration of the illusion produced by modulating fingertip friction on the haptic touchpad. (b): Relationship between the finger displacement and the setting. Theoretical representation with the whole range of variation (as if the whole range could be reached in one single movement). (c): Haptic feedback and the evolution of the spatial signal frequency with respect to the setting. The two target values are shown in blue lines. (d): Example of the actuation during one typical 60 mm finger slide, from $\gamma = 40$ to $\gamma = 60$. The modulating signal (upper plot) pilots the plate vibration envelope (middle plot) that produces friction variations (lower plot). Friction measurements were made with an interferometric force sensor described in [Bernard et al., 2020].

The haptic feedback consists in a sine wave of friction modulation, which spatial frequency evolves with the setting. First, the setting γ evolves with the finger swipes, and the variations of the setting $\Delta\gamma$ are proportional to the finger displacement Δx (in mm) by a factor $\alpha = 1/3$ mm^{-1}:

$$\Delta\gamma = \alpha\Delta x \tag{5.1}$$

This relationship is illustrated in Figure 5.2.b. A 3 mm displacement of the finger on the interface induces a change of the setting value of 1. The haptic gradient feedback is a spatially encoded signal, in which the spatial frequency ν (in mm^{-1}) evolves with the setting γ as the finger explores the interface. The spatial frequency can be considered as the number of ridges per millimeter. For the gradient evolution to be perceived equally along the frequency range, we adapted the spatial frequency

to the Weber Law. According to this law, the just noticeable difference (JND) $\delta\nu$ of the frequency is proportional to the initial frequency ν multiplied by a constant of proportionality g:

$$\delta\nu = g\nu \tag{5.2}$$

The constant g is the gradient value (in mm^{-1}) and represents how fast the spatial frequency changes when the finger moves. The variations of the spatial frequency of the grating ν are obtained by integrating (5.2) with respect to the setting:

$$\nu(\gamma) = \nu_0 \exp\left(g\frac{\gamma - \gamma_0}{\alpha}\right) \tag{5.3}$$

where $\nu_0 = 0.2$ mm^{-1} is the central spatial frequency which is reached when the setting is at half range, i.e. $\gamma = \gamma_0 = 50$.

The modulating signal A that oscillates at the instantaneous frequency ν is then defined as follows:

$$A(\gamma) = \cos\left(2\pi \int_0^{\gamma} \nu(u)\, du\right) \tag{5.4}$$

The gradient value was fixed to $g = -0.015$ mm^{-1}, an appropriate trade-off to ensure that users can feel the spatial frequency change over the distance of a typical finger slide (>50 mm) [Bernard et al.,], while allowing for fine-tuning of the setting value and respecting the limits of the device in terms of spatial frequency bandwidth ($\approx$ 0.02 to 2 mm^{-1}). The resulting haptic signal is presented in Figure 5.2.c.

5.5 User study: interaction learning

Since the interaction with haptic feedback requires training, we performed an experiment to investigate how well subjects can learn the task. We measured their performances after a training session, in which additional visual or auditory feedbacks helped the learning process.

5.5.1 Task

Both the training session and the learning session were conducted on the same task. At each trial, the setting was randomly initialized at one of the five initial values $\gamma_i = 0, 10, 20, 30, 40$. Subjects were not explicitly informed of the initial value, but they could sense it from the feedback. It prevents subjects from basing their actions on proprioceptive cues only (such as the finger displacement and number of movements) that would bias the experiment. The target value, either $\gamma_t = 60$ or $\gamma_t = 80$, is displayed on the screen. By successive slides of the index finger on the touchpad, the subjects were asked to increase the setting until it reaches the target value as quickly and precisely as possible, as presented in Figure 5.3. They selected the final value by lifting the finger for more than two seconds. To achieve the task, subjects need to mentally associate the two target values to two spatial frequencies of the texture.

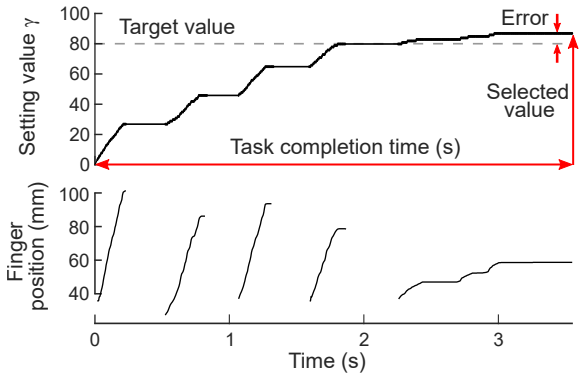

FIGURE 5.3: Setting and finger position variations for a typical trial (initial value $\gamma_i = 0$, target value $\gamma_t = 80$, haptic feedback only). The setting is increased by successive slides of the finger on the touchpad with clutching. The finger position is not measured during clutches. The 3 performance metrics are also presented.

The target values $\gamma_t = 60$ and $\gamma_t = 80$ correspond to gratings of spatial frequency $\nu_t = 0.31$ and $\nu_t = 0.77$ mm^{-1}, respectively. Subjects were asked to avoid reaching a too high value (above 100) to prevent any other proprioceptive biases. The 5 initial values and the 2 target values led to 10 task conditions. During the experiment, the task conditions were presented in randomized orders per block of 10.

5.5.2 Protocol and learning strategies

Participants sat in a chair in front of the computer on the experimental desk and wore headphones. After being informed about the experimental procedure, they placed their right hand on the wrist rest placed on their right-hand side. They were asked to interact with the touchpad with their right index finger. Before the experiment, subjects could interact with the touchpad without any feedback during 1 min to become familiar with the gesture. Then the experiment started with the learning session (or acquisition phase). This session was composed of 200 trials (20 blocks) of the task described in the previous section. The touchpad was actuated with the concurrent haptic feedback. At the end of each trial, subjects were informed about their performance by a digital display of the selected value on the graphical interface. This terminal feedback helped them correct themselves and improve their performance on a trial by trial basis. The haptic feedback can be enhanced with additional, auditory or visual, feedback. The five learning strategies were investigated with five groups of subjects following a between-subjects experiment, as presented in Figure 5.4.

The experiments ended with the evaluation session (or transfer phase) which was the same for each subject. It was composed of 40 trials (4 blocks), with the same task as in the learning session, with the haptic feedback only. No additional feedback or terminal feedback were presented. The whole experiment lasted for about 50 minutes.

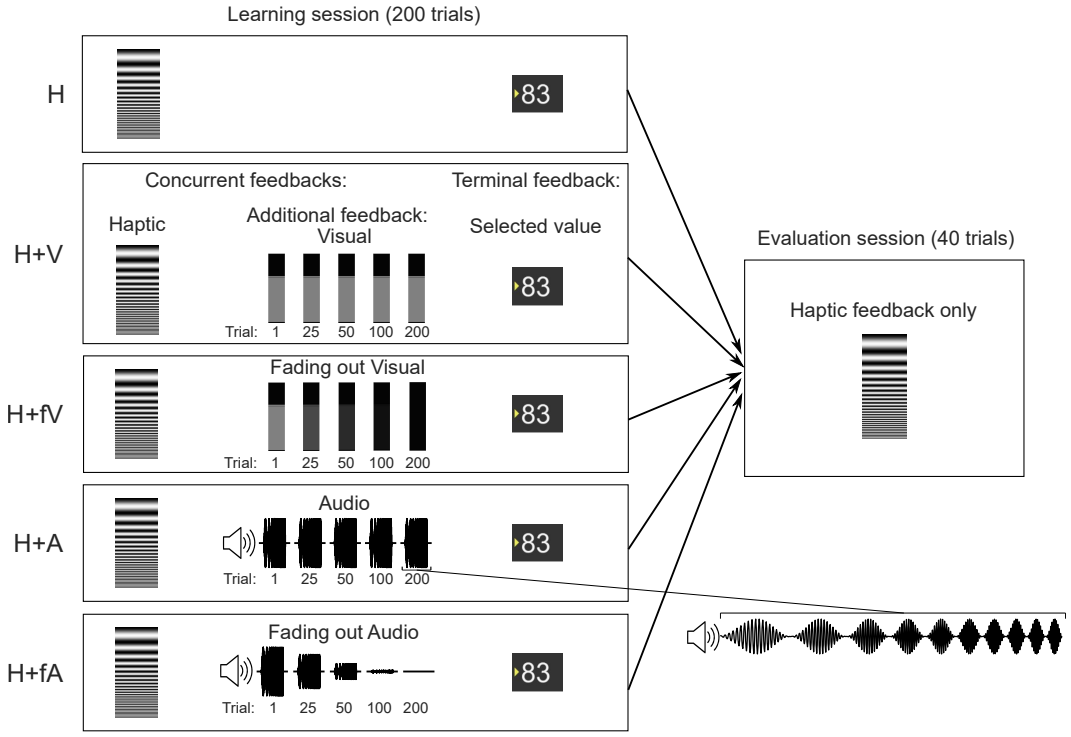

FIGURE 5.4: Illustration of the experimental protocol and the five learning strategies experienced by 5 groups of 8 subjects. Group conditions only differ by the additional concurrent feedback (auditory or visual) during the learning session. Haptic feedback is a gradient of spatial frequencies rendered with ultrasonic friction modulation technology. For the *Visual* condition, a slider is displayed on the graphical interface. For the *fading out Visual* condition, the slider contrast decreases until it becomes invisible around trial 100. For the *Audio* condition, the pitch and the modulation frequency of the sound are mapped to the setting value and to the haptic friction variations. For the *fading out Audio* condition, the sound level decreases until it becomes inaudible around trial 100.

5.5.3 Additional feedback

The additional visual feedback consists of a vertical slider on the graphical interface that moves up and down as the subject adjusts the setting value. It is similar to what we experience with everyday interfaces. For the *fading out Visual* condition, the contrast of the slider decreases progressively so that it becomes less and less visible and then disappears, as presented in Figure 5.4.

The additional auditory feedback is constructed by sonification (both pitch and tempo related) of the setting, the haptic friction signal and the finger velocity. The setting is sonified by a monochromatic tone whose frequency varies between 100 and 600 Hz (for $\gamma = 0$ to 100). The frequency is logarithmically mapped to the setting. The target values $\gamma_t = 60$ and $\gamma_t = 80$ correspond to pitch frequencies of 293 and 419 Hz, respectively. The sonification of the haptic friction signal is tempo-related and made by amplitude-modulations whose modulation frequency $f_{mod} = \nu * v_{finger}$ matches those of the haptic stimuli to ensure multisensory congruency between the

two stimuli [Bernard et al.,]. The resulting signal is illustrated in Figure 5.4. For the *fading out Audio* condition, the sound level decreases progressively and becomes less and less audible. Informal pre-tests measured the perceptual thresholds of the audio and visual stimuli to ensure that additional feedback became unnoticeable about halfway through the training session (trial number 100) under the fading out conditions.

The Max/MSP interface generates the visual and auditory feedback. Sounds were played through noise-cancelling headphones (3M PELTOR ProTac III) that also prevented any external auditory cue from the device.

5.5.4 Subjects

40 subjects, 14 females and 26 males, 34 right-handed and 6 left-handed, ranging from 20 to 54 years old (mean 27.9) participated in the study. They gave their informed consent before the experiment and the study was approved by the Ethical Committee. They were paid for their participation. Before the experiment, subjects washed and dried their hands and the touchpad was cleaned with an alcoholic solution. They were randomly assigned to one of the five learning strategy groups.

5.6 Results

Task performances are assessed using three metrics: the final value, the error and the completion time, as presented in Figure 5.3.

5.6.1 Overall results

After training, participants were able to use the haptic cues to adjust the setting to $\gamma_t = 60$ or $\gamma_t = 80$. Boxplots of the distribution of the selected values obtained during each trial are shown in Figure 5.5a. The clear difference in the median of the two conditions demonstrates that subjects were successful in performing the task. A two-sample Student-Welch t-test with unequal variances validates the significant difference between the two target value conditions (t=-27.5, df=1591, p=$3.8 \cdot 10^{-136}$). We can observe that the target value at 80 was reached more accurately (median selected value: 80) than the target value at 60 (median selected value: 62.5), with the same variability (interquartile range: 13).

Figure 5.5b presents the effect of the initial value on the setting error in order to control any possible proprioceptive biases. If the subjects had relied purely on the estimation of the distance covered by the finger on the interface, the selected value would have been underestimated for the lowest initial value and overestimated for the highest initial values, as illustrated by the grey line in Figure 5.5.b. Here, even if the initial value shows a significant effect on the setting error (one-way ANOVA: $F_4 = 19.9$, p=$5.4 \cdot 10^{-16}$), due to an overestimation for the $\gamma_i = 30$ and 40 initial

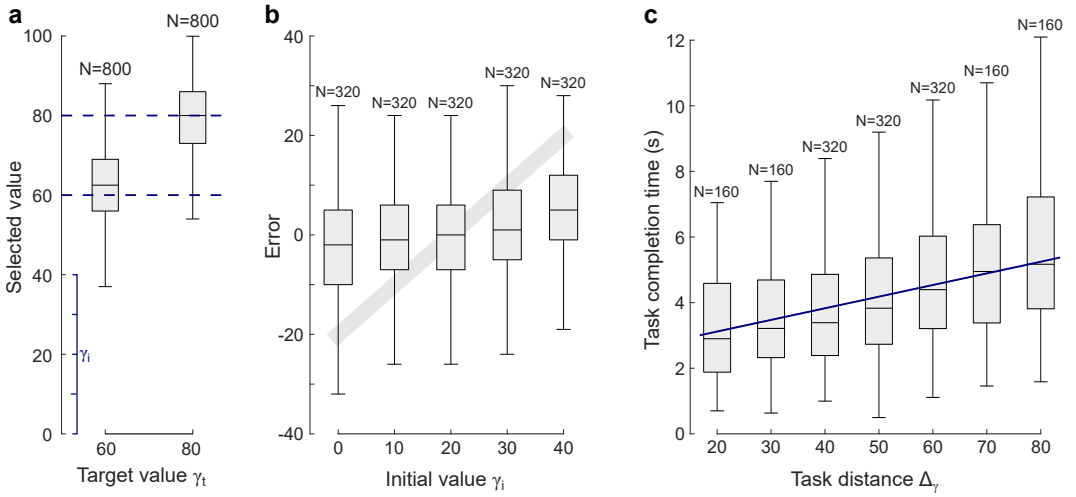

FIGURE 5.5: Results of the evaluation session. (a): Boxplot of the selected value for the two target value conditions during the evaluation session for all subjects. The targets are also displayed in dashed lines. The initial value range is plotted in blue. (b): Boxplot of the error on the selected value for the five initial value conditions during the evaluation session for all subjects. If responses were only based on proprioceptive cues, the error would follow the trend shown by the grey line in the background. (c): Boxplot of the task completion time as a function of the task distance ($\Delta_\gamma = \gamma_t - \gamma_i$) during the evaluation session for all subjects. The mean completion time (with outliers exclusion) increases linearly with respect to the task distance, as proven by the linear regression presented in blue line (R^2=0.99).

values, this bias is much lower than the theoretical bias with proprioception only. It demonstrates that subjects succeeded in integrating tactile cues in their selection.

The task completion time is presented in Figure 5.5.c as a function of the setting distance between the initial and the target value $\Delta_\gamma = \gamma_t - \gamma_i$. The task distance shows a significant effect on the completion time (one-way ANOVA: $F_6 = 12.8$, p=$3.3 \cdot 10^{-14}$). Subjects spent more time to adjust the setting when the target value was much greater than the initial value, as exhibited by the linear regression performed on the mean completion time $T = 0.035 \times \Delta_\gamma + 2.42$ ($R^2 = 0.993$).

In summary, the results allow us to quantify the user performance. With haptic feedback only, subjects performed the task on average in 4.1 s with 8.1% of error (mean of the absolute value of the error and mean of the completion time on all subjects during the evaluation session, outlier trials excluded).

5.6.2 Learning strategy comparisons

We can compare the five learning strategies by looking in detail at the learning process. Figure 5.6 presents the learning curves for the five strategies, in terms of error and completion time. Trials are gathered by blocks of 10, with one of each target value and initial value condition in each block. Outlier trials are excluded thanks to the quartiles method (more than 1.5 interquartile ranges above the upper quartile or below the lower quartile) before averaging the trials' metrics in each block. The

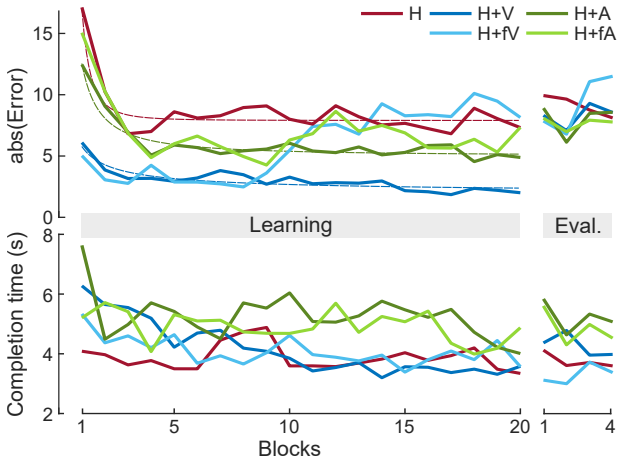

FIGURE 5.6: Learning curves for the five learning strategies. Performance metrics are averaged by blocks of 10 trials. The 20 first blocks are for the learning session, the 4 last blocks are for the evaluation session. Learning curves for strategies with permanent additional feedback are fitted with power law functions as displayed in dashed lines.

learning curves show how performances evolve during the training with respect to the learning strategy.

To characterize the evolution of the error, learning curves are fitted with power law functions $\gamma(k) = a \times k^{-b} + c$ that are typical of skill acquisition [Newell and Rosenbloom, 1981]. These functions are calculated for the 3 conditions with permanent feedback only. The parameter $a + c$ shows the error level at the beginning of the training and c the error level at the plateau. The parameter b reflects how quickly subjects improve their performances.

Subjects with haptic feedback only and additional auditory feedback start with a high error ($a + c = 17.1$ and 12.6 respectively) whereas it is much lower for subjects with visual feedback ($a + c = 5.8$). However, the error drops quickly for haptic and auditory strategies ($b = 2.7$ and 1.3 respectively) while it only slightly decreases for visual strategies ($b = 0.7$). After the initial improvement, the error remains stable for the 3 learning strategies with permanent feedback, with an error level of $c = 8$, 5 and 1.9 for the *Haptic*, *Audio* and *Visual* conditions, respectively. In contrast, for the learning strategies with feedback that fades out with time, the error increases at the middle of the training session when audio and visual stimuli become imperceptible, and reaches the error level of the haptics only condition. The transition to the evaluation session causes a significant gap for the visual and audio conditions, but the error finally appears to be approximately at the same level regardless of the training strategy.

For the task completion time, we globally observe a slight decrease during the training session. Subjects appear to be slower in the auditory learning conditions, even when the feedback disappears.

We can quantify the efficiency of the five learning strategies concerning the mastering of the interaction. Figure 5.7 shows the performances during the evaluation

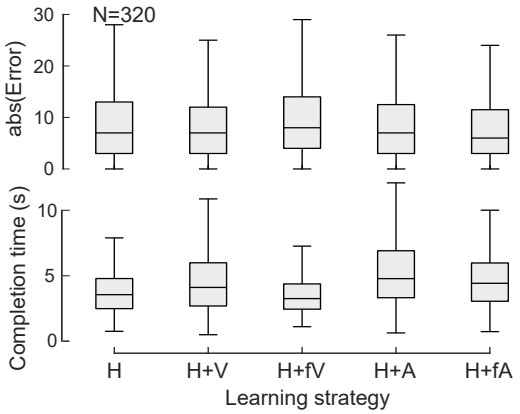

FIGURE 5.7: Performances during the evaluation session for the five groups of subjects associated to the five learning strategies.

session with respect to the learning strategy group. Statistical analysis exhibits a significant effect (at $\alpha = 0.05$) of the learning strategy on the absolute value of the error (one-way ANOVA: $F_4 = 221$, p=0.011). This effect appears to be only due to a slight difference between the *fading out Audio* and *fading out Visual* condition (Post hoc Tukey's test: $p_{fV-fA} = 0.012$)

Learning strategies also have a significant effect (at $\alpha = 0.05$) on the completion time during the evaluation session (one-way ANOVA: $F_4 = 13.66$, p=$5.9 \cdot 10^{-11}$). The post hoc Tukey's test reveals that this effect is due to the *Visual*, *Audio* and *fading out Audio* conditions that lead to a significantly higher completion time.

As the error disparities are low and the p-value is still close to the level of significance, we assume that there are no major differences between the learning strategy groups during the evaluation session in terms of task error and that statistical artifacts could be due to the high subject variability.

Both the learning curves and the boxplot show that the task completion time is longer for conditions with audio feedback. As audio stimuli are based on the sonification of the haptic signal, whose temporal variations depend on the finger velocity, it is possible that subjects adjusted their exploratory motion to optimize the information acquisition from the sound. Remarkably, subjects maintain this slower exploratory motion even when the audio feedback is removed, for the *fading out* condition and for the evaluation.

5.6.3 Inter-subject variability

We investigated inter-subject variability to better understand the previous results. Figure 5.8 presents the subjects' performances in a two-dimensional space, with mean error and completion time as dimensions (outlier trials excluded). Intra-subject variability is also displayed through the error bars. The bottom-left corner corresponds to subjects with the best performances, i.e. the lowest error and the fastest completion time. Since subjects adopted various approaches during the task performance, their performances vary greatly. The graph shows large disparities

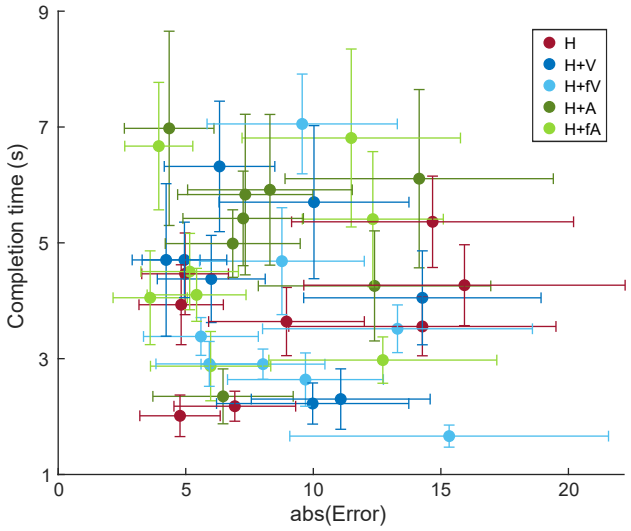

FIGURE 5.8: Illustration of the inter-subject variability for the task with haptic feedback only. Each point corresponds to one subject, characterized by its performances: the mean completion time and the mean absolute value of the error on all trials during the evaluation session. Error bars show the standard deviations, which reflect the intra-subject variability. The 5 learning strategies are presented with different colors.

between the subjects, which makes it difficult for a trend in learning strategy to emerge. We can still observe that the subjects with audio learning conditions (in green) are mainly in the upper part of the graph, as they spend more time performing the task.

5.7 Discussion

Participants were able to adjust a setting on a 0 to 100 scale with an error of 8.1%, using a progressive haptic feedback only. In comparison, for the same task with visual feedback in the form of a slider on a screen, the setting error was approximately at 2.2% (mean of the last 100 trials of the H+V condition). On our custom-made interface, such setting errors correspond to finger position shifts of 24.3 mm with haptic feedback and 6.6 mm with vision, both are within the same order of magnitude as the size of the contact made by the fingertip. The results reveal that setting a value on a screen using only haptics leads to degraded results compared to using a screen, but still within an acceptable level of performance. This discrepancy should be taken into account while designing haptic interfaces by adding a margin of error or dead-zones. Alternatively, the progressive haptic feedback interaction without vision can be used in applications that require a low precision, such as window controls. In the experiment, subjects had to adjust the setting to reach 2 different target values only. However, the gradual haptic feedback does not present any specific cues at these locations. In practice, gradual feedback allows the users to reach any setting value. Therefore, this interaction provides an absolute adjustment of the setting, whereas

conventional binary haptic feedback only provides relative adjustment to change the setting by a certain amount.

Furthermore, across conditions and subjects, the task is performed within an average of 4.1 s using haptic feedback, which is comparable with the visual feedback condition of 3.5 s. The completion time increases with larger distances, similarly to Fitts' law. However, we found that the relationship between distance and time is linear and not logarithmic, perhaps due to the nature of the interaction, which differs from a pointing task and thereby involves other control mechanisms.

The learning curve during the training sessions with only haptic feedback demonstrates that subjects needed about 35 trials to reach 95% of the final error level, which shows that the interaction is mastered remarkably quickly. The initial trials with only haptic feedback were two times more laborious than the steady state performance and about three to four times longer than with visual feedback, suggesting that the task with only haptic feedback is initially cognitively demanding. However, after even a short training session, the differences in performance vanishes, which is a promising result for enabling eyes-free control of sliders.

We found that the type of training feedback had little effect on the final performance during the evaluation session. However, these results were obtained from 8 subjects by groups. Since the inter-subject variance is large, the absence of a larger effect could be due to a too small sample size. By design, participants were aware of the absence of feedback during the evaluation session before starting the experiment. As a consequence, some of them reported that they had focused on the haptic feedback and paid less attention to the visual or auditory feedback. This might explain why the performances were the same with additional concurrent feedback presented during the whole training session, whereas this type of strategy was reported in the literature to be less effective for retention [Robin et al., 2005, Sigrist et al., 2013]. Finally, given the relatively low impact of congruent feedback, we can conclude that the terminal feedback, given on a digital display, seems to be the essential factor enabling the retention process.

Since retention after training did not depend on the learning modality, the learning method can be adapted to different contexts and to the users' preferences. Visual feedback would be selected preferentially for its simplicity and rapidity of training, but auditory feedback can complement vision. Auditory feedback still requires more time for learning, perhaps because users are less familiar with this kind of feedback since it is scarcely implemented. Additionally, the learning curves show a plateau after about 40 trials, or four blocks, so in practice the training session could probably be significantly shortened.

Only short-term retention has been studied, with the evaluation performed after the training. Future research will investigate to which extent users can capitalize on their learning over a long period. In addition, since the interaction was learned in laboratory conditions, it would be interesting to contextualize the experiment to real-life scenarios and evaluate the device and the learning procedure

in real situations, as in a car cockpit [Mullenbach et al., 2013] or even in the subway [Hoggan et al., 2008].

In realistic cases, we cannot expect the visual and auditory feedback to be always present. However we expect that the current method based on haptic feedback will facilitate the operation of the slider without having to look at it or in a noisy environment. With time and practice, users will naturally rely less on vision and more on haptic perception. For example, in [Bressolette et al., 2021] authors showed that using a gesture based sonified interface was challenging at first, but after training, users preferred it to a conventional touch screen in a driving situation as it distracted them less from monitoring the road.

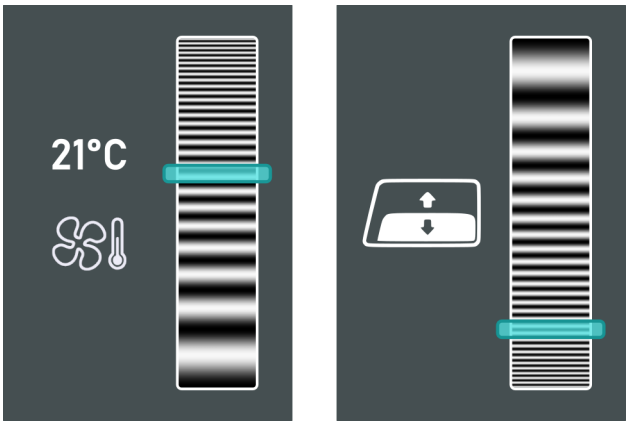

FIGURE 5.9: Example of two use cases with graphical display. The gradual haptic feedback illustrated in grey scale enables the user to control the temperature and the opening of a window without vision.

A vertical one-dimensional touchpad was used in these experiments, but the interaction can trivially be extended to two-dimensional touchpads or touchscreens equipped with haptic feedback. Moreover, we used only one type of haptic feedback, a sine wave with a spatial frequency gradient. To recognize the parameters in a complex interface for example, we can imagine other types of haptic stimulations, such as random irregularities or amplitude gradients [Bodas et al., 2019]. As illustrated in Figure 5.9, the interaction with progressive haptic feedback can be implemented in a wide range of human-machine interactions, such as the setting of the temperature level, the strength of ventilation or the music level, the selection of a radio station, the opening of windows or stores.

5.8 Conclusions

We propose a new control device that allows setting a value without looking at it, using a new haptic touchpad. By feeling a texture that evolves continuously, the control parameters of the interface can be adjusted in an absolute way, a task that until now seemed impossible without vision. After learning, users succeeded in performing the task with 8.1% of error and in 4.1 s on average.

This performance comes after a short learning period, during which any kind of feedback seems to help improve the skills. This new guidance method which leverages a continuous mapping of the haptic cues to an abstract value has only marginal degradation of performance compared to the use of visual cues. Providing this feedback on touchscreens can reduce distraction and improve intuitiveness of the interaction.

Chapter 6

Conclusion

Contents

6.1 Summary of the contributions

This thesis first proposes a literature review covering the broad fields of hearing, touch, multisensory integration, haptic technologies and human-computer interfaces.

Chapter 3 describes a perceptual threshold test on haptic stimuli that provides perceived intensity curves with a maximum sensibility around 100 to 200Hz. It shows that the perception thresholds of synthetic friction modulated textures is closer to the perception of vibrations than of real textures. The resulting threshold curves with respect to the spatial period were used in the following experiments to construct haptic stimuli in which the spatial period gradually evolves. Chapter 4 provides a law that describes the minimal finger exploration distance required to detect the evolution of a given haptic gradient. This relationship showed similarities with the auditory perception of tempo changes, which can be explained by a multimodal model of rhythmic variations perception. Based on these results I suggest that a congruent sound combined with the haptic feedback can improve the perception of the gradient, hereby demonstrating audio-haptic integration. Spatial frequency gradients were finally applied in chapter 5 as haptic feedback to control a human-computer interface. I demonstrated that users can integrate haptic gradient information to perform the absolute adjustment of a setting without vision with an acceptable level of performance. Various strategies with the use of visual or auditory modalities proposed different training processes that all led to successful

learning of the adjustment task. This type of interaction with haptic feedback could be implemented in car dashboards to reduce visual distraction.

6.2 Future directions

6.2.1 Audio-haptic integration

The combined synthesis of sounds and textures is still a challenge. For the two stimuli to be perceived as a unique percept, a certain congruence between the sounds and textures is required. One possible solution is to design sounds that are correlated with the synthetic textures so that both stimuli appear to originate from the same source. To do so, the approach developed in this thesis used the friction map of the haptic texture as a signal that modulates the amplitude of the sound texture. That way, the sounds and the textures share the same fluctuations during the exploration. The sound texture, however, can take different forms. In chapter 4, a filtered white noise was used to mimic frictional sounds and increase the realism. In chapter 5, it was a monochromatic tone whose frequency was mapped to the setting in order to bring supplementary information. In auxiliary works [Kanzari et al., 2019], we proposed to use sound textures with timbre of different materials (wood, stone, metal...) from the interaction sound synthesizer developed in [Conan et al., 2014]. However, the addition of these sounds wasn't enough to render the illusion of touching different materials with the finger, and this objective still requires further research.

In the light of the outcomes of Chapter 4, we may wonder whether the Bayesian model of multisensory integration could by applied to multisensory perception of evolving textures. The second experiment was first conducted with another experimental condition, with incongruent auditory and haptic stimuli. To explore whether subjects would base their judgments more on the auditory or haptic cues, the auditory and haptic gradients were of equal value but presented in opposite directions. However, this type of stimulus resulted in such strong differences that subjects reported that they could feel the discrepancy. For this condition, the results showed that subjects responded globally at random (with about 50% preference for each modality). Therefore, we could not carry the Bayesian analysis any further. To explore the application of the Bayesian model to the audio-haptic perception of evolving textures, another experiment could be imagined with smaller differences between audio and haptic stimuli by choosing slightly different gradients that evolve in the same direction.

Another interesting parallel between audition and touch appears in perceived roughness. For touch, texture roughness is well known as one of the main perceptive attributes when exploring a surface with the finger [Tiest and Kappers, 2006]. Auditory roughness refers to another perceptual attribute at the basis of phenomena such as consonance and dissonance in music. It can be described as the perception of very fast fluctuations in sounds. For stimuli composed of two monochromatic tones $s(t) = sin(2\pi f_1 t) + sin(2\pi f_2 t)$, the sensation of roughness is driven by the space

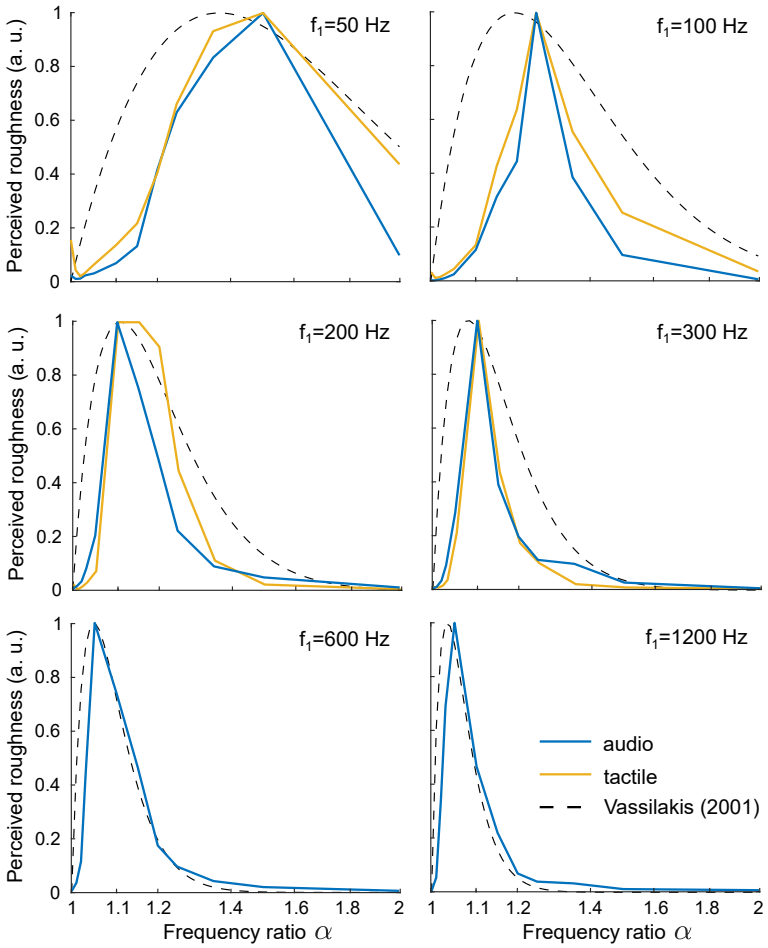

FIGURE 6.1: Auditory roughness curves and their tactile counterparts. The two modalities elicit similar roughness maxima and are coherent with the theoretical auditory roughness model proposed by [Vassilakis, 2001] plotted in dashed lines. The stimuli with f_1= 600 and 1200 Hz were presented only for the audio condition because they are beyond the frequency bandwidth of tactile perception.

between the two frequency components, defined by the frequency ratio $\alpha = f_2/f_1$. Such a phenomenon reveals the existence of auditory critical bands, a fundamental characteristic of auditory filters [Terhardt, 1974]. While the auditory perception of such phenomena has been largely studied, other modalities such as touch elicit similar behaviors [Makous et al., 1995]. Interestingly, it is possible to produce similar stimuli as auditory roughness with vibrotactile actuators. We investigated this question in [Fery et al., 2021]. In a pairwise comparison, subjects were asked to judge which tone combination (or tactile stimulation) was the most "granular" (granuleux in French) between to frequency ratio conditions. The results presented in Figure 6.1 show similar roughness curves between the two modalities. This study extends the results presented in Chapter 4, suggesting that auditory and tactile modalities share common principles also in the perception of roughness and beatings. It would be interesting to investigate more in detail the similarities between these two senses.

6.2.2 Human Computer Interfaces

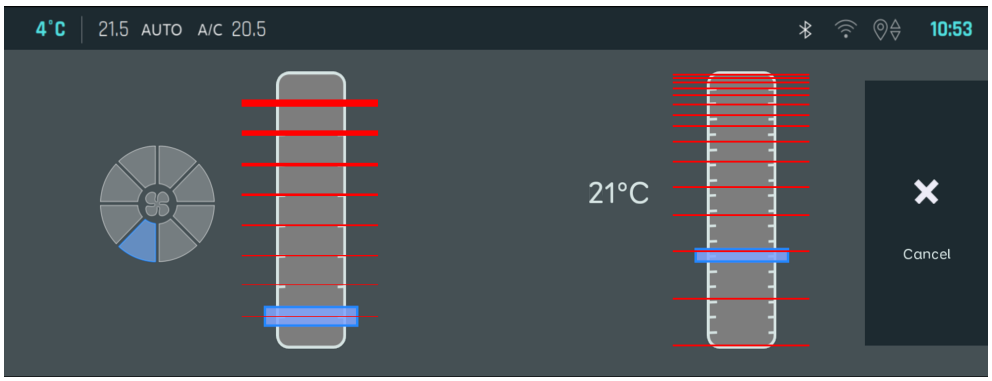

FIGURE 6.2: Example of an interface with evolving textures: amplitude gradient for the ventilation strength and a spatial frequency gradient for the air temperature. Haptic feedback is illustrated in red.

Chapter 5 presented an interaction with haptic feedback for the absolute adjustment of a setting on an interface. Figure 6.2 shows an example of an interface using this type of interaction. The visual display is controlled by a haptic touchpad located near the driver's right armrest. Up/down swipes on the left part of the touchpad enable the user to control the ventilation strength and up/down swipes on the right part control the temperature. To distinguish the two settings, two types of haptic feedback are proposed: a gradient of spatial frequency, as studied previously, and a gradient of amplitude. They can be combined with congruent auditory feedback developed in this thesis to enhance the task.

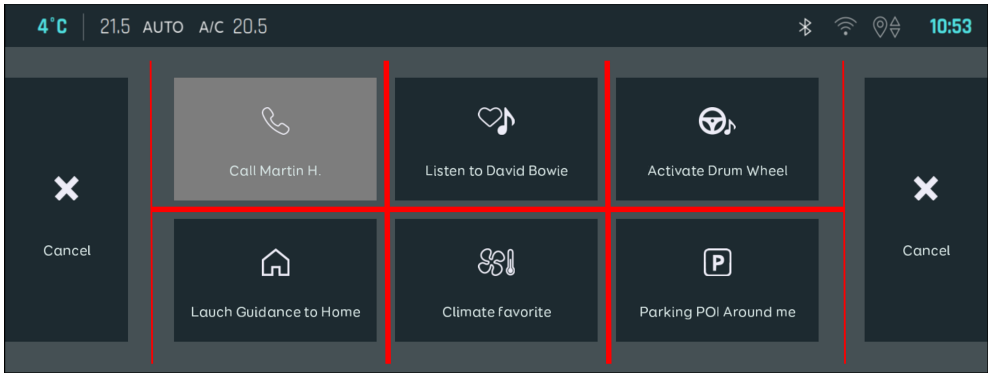

FIGURE 6.3: Example of an interface with haptic boundaries for shortcuts selection. Haptic feedback is illustrated in red.

Other types of interaction could be implemented on interfaces with haptic feedback. Figure 6.3 presents another demonstrator that we developed. Here, the touchpad is used to rapidly select one of six predefined actions. Once in this menu, the user can slide his finger from one cell to another and release when on the selected one. Here, we proposed to implement haptic feedback as sharp friction increases to

render the sensations of feeling boundaries between two cells. This makes the interface more intuitive and user-friendly. This type of haptic feedback can be enhanced with classical "click" sounds.

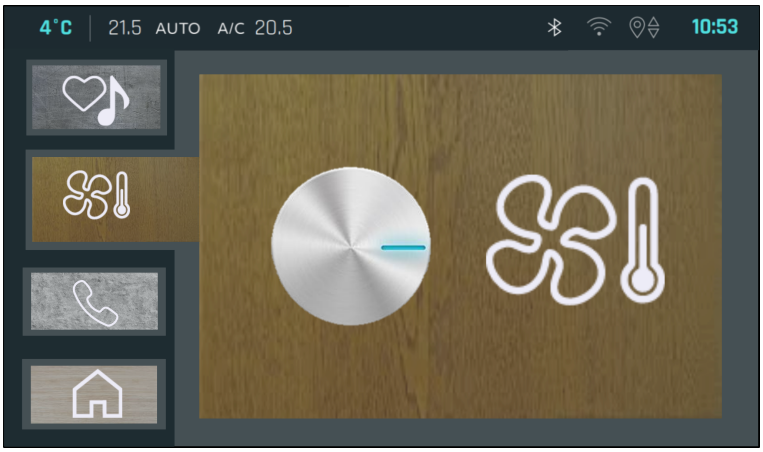

FIGURE 6.4: Example of interface with a menu of audio-haptic textures.

With the interaction sound synthesizer connected to the haptic surface, we can also imagine other original interfaces. Figures 6.4 shows a possible implementation of audio-haptic textures to navigate through menus. Each interface panel is associated with a background texture that users can recognize by touching and hearing it. Here for example, the air conditioning menu is linked to a "wood" texture that produces the sound of a finger rubbing on a wood plate when the screen is explored. With learning, users should be able to assess the actual menu and change it without having to look at it. The carousel in Figure 6.5 is a great example where users already have to slide their finger to switch between menus. This interaction could naturally be enhanced with audio-haptic feedback by the presented approach. A huge advantage of sound and texture synthesis is that feedback can be fully personalized to adapt to the user's sensitivity and preferences.

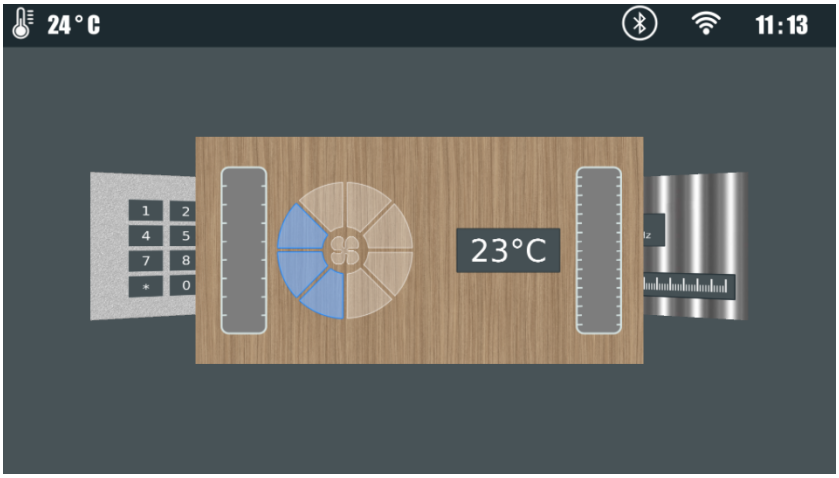

FIGURE 6.5: Example of interface with a carousel of audio-haptic textures.

In summary, the synthesis of high-quality sounds and textures is a great tool that can change the way we think about human-computer interfaces. In this thesis we proposed a new interaction and demonstrated its efficiency but, with these new opportunities, a multitude of other possibilities can be imagined.

Bibliography

[Accot and Zhai, 1997] Accot, J. and Zhai, S. (1997). Beyond fitts' law: models for trajectory-based hci tasks. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, pages 295–302.

[Ackerley et al., 2014] Ackerley, R., Wasling, H. B., Liljencrantz, J., Olausson, H., Johnson, R. D., and Wessberg, J. (2014). Human c-tactile afferents are tuned to the temperature of a skin-stroking caress. *Journal of Neuroscience*, 34(8):2879–2883.

[Alais and Burr, 2004] Alais, D. and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, 14(3):257–262.

[Bänziger and Scherer, 2005] Bänziger, T. and Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech communication*, 46(3-4):252–267.

[Barrass, 1995] Barrass, S. (1995). Personify: a toolkit for perceptually meaningful sonification. In *ACMA*. Citeseer.

[Basdogan et al., 2020] Basdogan, C., Giraud, F., Levesque, V., and Choi, S. (2020). A review of surface haptics: enabling tactile effects on touch surfaces. *IEEE transactions on haptics*, 13(3):450–470.

[Bau et al., 2010] Bau, O., Poupyrev, I., Israr, A., and Harrison, C. (2010). Teslatouch: electrovibration for touch surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 283–292. ACM.

[Bell et al., 1994] Bell, J., Bolanowski, S., and Holmes, M. H. (1994). The structure and function of pacinian corpuscles: a review. *Progress in neurobiology*, 42(1):79–128.

[Bensmaia and Hollins, 2003] Bensmaia, S. J. and Hollins, M. (2003). The vibrations of texture. *Somatosensory & motor research*, 20(1):33–43.

[Bernard et al., 2018] Bernard, C., Monnoyer, J., and Wiertlewski, M. (2018). Harmonious textures: the perceptual dimensions of synthetic sinusoidal gratings. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 685–695. Springer.

[Bernard et al.,] Bernard, C., Monnoyer, J., Wiertlewski, M., and Ystad, S. Rhythm perception is shared between audio and haptics. *Scientific Reports, under review*.

[Bernard et al., 2019] Bernard, C., Monnoyer, J., Ystad, S., and Wiertlewski, M. (2019). Interferometric tribometer for wide-range/high-bandwidth measurement of tactile force interaction. In *IEEE World Haptics Conference 2019*.

[Bernard et al., 2020] Bernard, C., Ystad, S., Monnoyer, J., and Wiertlewski, M. (2020). Detection of friction-modulated textures is limited by vibrotactile sensitivity. *IEEE transactions on haptics*, 13(3):542–551.

[Bianchi et al., 2017] Bianchi, M., Moscatelli, A., Ciotti, S., Bettelani, G. C., Fioretti, F., Lacquaniti, F., and Bicchi, A. (2017). Tactile slip and hand displacement: Bending hand motion with tactile illusions. In *2017 IEEE World Haptics Conference (WHC)*, pages 96–100. IEEE.

[Biet et al., 2007] Biet, M., Giraud, F., and Lemaire-Semail, B. (2007). Squeeze film effect for the design of an ultrasonic tactile plate. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 54(12):2678–2688.

[Birznieks and Vickery, 2017] Birznieks, I. and Vickery, R. M. (2017). Spike timing matters in novel neuronal code involved in vibrotactile frequency perception. *Current Biology*, 27(10):1485–1490.

[Blattner et al., 1989] Blattner, M. M., Sumikawa, D. A., and Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, 4(1):11–44.

[Bochereau et al., 2018] Bochereau, S., Sinclair, S., and Hayward, V. (2018). Perceptual constancy in the reproduction of virtual tactile textures with surface displays. *ACM Transactions on Applied Perception (TAP)*, 15(2):10.

[Bodas et al., 2019] Bodas, P., Friesen, R. F., Nayak, A., Tan, H. Z., and Klatzky, R. (2019). Roughness rendering by sinusoidal friction modulation: Perceived intensity and gradient discrimination. In *2019 IEEE World Haptics Conference (WHC)*, pages 443–448. IEEE.

[Bolanowski Jr et al., 1988] Bolanowski Jr, S. J., Gescheider, G. A., Verrillo, R. T., and Checkosky, C. M. (1988). Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical society of America*, 84(5):1680–1694.

[Botvinick and Cohen, 1998] Botvinick, M. and Cohen, J. (1998). Rubber hands 'feel'touch that eyes see. *Nature*, 391(6669):756–756.

[Bressolette et al., 2018] Bressolette, B., Denjean, S., Roussarie, V., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2018). Harnessing audio in auto control: The challenge of sonifying virtual objects for gesture control of cars. *IEEE Consumer Electronics Magazine*, 7(2):91–100.

[Bressolette et al., 2021] Bressolette, B., Denjean, S., Roussarie, V., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2021). Movecho: A gesture-sound interface allowing blind manipulations in a driving context. *ACM Transactions on Applied Perception (TAP)*, 18(3):1–19.

[Brewster et al., 1993] Brewster, S. A., Wright, P. C., and Edwards, A. D. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 222–227.

[Cadoret and Smith, 1996] Cadoret, G. and Smith, A. M. (1996). Friction, not texture, dictates grip forces used during object manipulation. *Journal of neurophysiology*, 75(5):1963–1969.

[Caetano and Jousmäki, 2006] Caetano, G. and Jousmäki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *Neuroimage*, 29(1):15–28.

[Cascio and Sathian, 2001] Cascio, C. J. and Sathian, K. (2001). Temporal cues contribute to tactile perception of roughness. *Journal of Neuroscience*, 21(14):5289–5296.

[Casiez et al., 2011] Casiez, G., Roussel, N., Vanbelleghem, R., and Giraud, F. (2011). Surfpad: riding towards targets on a squeeze film effect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2491–2500.

[Chan et al., 2021] Chan, S., Tymms, C., and Colonnese, N. (2021). Hasti: Haptic and audio synthesis for texture interactions. In *2021 IEEE World Haptics Conference (WHC)*, pages 733–738. IEEE.

[Chen et al., 2011] Chen, H.-Y., Park, J., Dai, S., and Tan, H. Z. (2011). Design and evaluation of identifiable key-click signals for mobile devices. *IEEE Transactions on Haptics*, 4(4):229–241.

[Choi et al., 2021] Choi, C., Ma, Y., Sequeira, S., Chatterjee, S., Li, X., Felts, J. R., and Hipwell, M. C. (2021). Effect of surface temperature on finger friction and perception in electroadhesion. In *2021 IEEE World Haptics Conference (WHC)*, pages 680–684. IEEE.

[Conan et al., 2012] Conan, S., Aramaki, M., Kronland-Martinet, R., Thoret, E., and Ystad, S. (2012). Perceptual differences between sounds produced by different continuous interactions. In *Acoustics 2012*.

[Conan et al., 2014] Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Ystad, S., and Kronland-Martinet, R. (2014). An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling. *Computer Music Journal*, 38(4):24–37.

[Cooper et al., 1963] Cooper, G. W., Cooper, G., and Meyer, L. B. (1963). *The rhythmic structure of music*. University of Chicago press.

[Crommett et al., 2019] Crommett, L. E., Madala, D., and Yau, J. M. (2019). Multisensory perceptual interactions between higher-order temporal frequency signals. *Journal of Experimental Psychology: General*, 148(7):1124.

[Crommett et al., 2017] Crommett, L. E., Pérez-Bellido, A., and Yau, J. M. (2017). Auditory adaptation improves tactile frequency perception. *Journal of neurophysiology*, 117(3):1352–1362.

[Culbertson and Kuchenbecker, 2015] Culbertson, H. and Kuchenbecker, K. J. (2015). Should haptic texture vibrations respond to user force and speed? In *2015 IEEE World Haptics Conference (WHC)*, pages 106–112. IEEE.

[Culbertson et al., 2014] Culbertson, H., Lopez Delgado, J. J., and Kuchenbecker, K. J. (2014). The penn haptic texture toolkit for modeling, rendering, and evaluating haptic virtual textures.

[Culbertson et al., 2012] Culbertson, H., Romano, J. M., Castillo, P., Mintz, M., and Kuchenbecker, K. J. (2012). Refined methods for creating realistic haptic virtual textures from tool-mediated contact acceleration data. In *2012 IEEE Haptics Symposium (HAPTICS)*, pages 385–391. IEEE.

[Danna et al., 2013] Danna, J., Velay, J.-L., Paz-Villagrán, V., Capel, A., Petroz, C., Gondre, C., Thoret, E., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2013). Handwriting movement sonification for the rehabilitation of dysgraphia. In *10th International Symposium on Computer Music Multidisciplinary Research (CMMR)-Sound, Music & Motion-15-18 oct. 2013-Marseille, France*, pages 200–208.

[Daunizeau et al., 2021] Daunizeau, T., Gueorguiev, D., Haliyo, S., and Hayward, V. (2021). Phononic crystals applied to localised surface haptics. *IEEE Transactions on Haptics*.

[De Pra et al., 2020] De Pra, Y., Fontana, F., Järveläinen, H., Papetti, S., and Simonato, M. (2020). Does it ping or pong? auditory and tactile classification of materials by bouncing events. *ACM Transactions on Applied Perception (TAP)*, 17(2):1–17.

[Del Piccolo et al., 2015] Del Piccolo, A., Delle Monache, S., Rocchesso, D., Papetti, S., and Mauro, D. A. (2015). To" sketch-a-scratch".

[Del Piccolo et al., 2018] Del Piccolo, A., Rocchesso, D., and Papetti, S. (2018). Path following in non-visual conditions. *IEEE transactions on haptics*, 12(1):56–67.

[Delhaye et al., 2016] Delhaye, B., Barrea, A., Edin, B. B., Lefèvre, P., and Thonnard, J.-L. (2016). Surface strain measurements of fingertip skin under shearing. *Journal of the Royal Society Interface*, 13(115):20150874.

[Delhaye et al., 2019] Delhaye, B. P., O'Donnell, M. K., Lieber, J. D., McLellan, K. R., and Bensmaia, S. J. (2019). Feeling fooled: Texture contaminates the neural code for tactile speed. *PLoS biology*, 17(8).

[Denjean et al., 2019] Denjean, S., Kronland-Martinet, R., Roussarie, V., and Ystad, S. (2019). Zero-emission vehicles sonification strategy based on shepard-risset glissando. In *14th International Symposium on Computer Music Multidisciplinary Research*.

[Dépeault et al., 2008] Dépeault, A., Meftah, E.-M., and Chapman, C. E. (2008). Tactile speed scaling: contributions of time and space. *Journal of neurophysiology*, 99(3):1422–1434.

[Dhiab and Hudin, 2019] Dhiab, A. B. and Hudin, C. (2019). Confinement of vibrotactile stimuli in narrow plates. In *2019 IEEE World Haptics Conference (WHC)*, pages 431–436. IEEE.

[Ding and Bhushan, 2016] Ding, S. and Bhushan, B. (2016). Tactile perception of skin and skin cream by friction induced vibrations. *Journal of colloid and interface science*, 481:131–143.

[Dingus et al., 2006] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., et al. (2006). The 100-car naturalistic driving study, phase ii-results of the 100-car field experiment. Technical report, United States. Department of Transportation. National Highway Traffic Safety ....

[Drake and Botte, 1993] Drake, C. and Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & psychophysics*, 54(3):277–286.

[Dubus, 2012] Dubus, G. (2012). Evaluation of four models for the sonification of elite rowing. *Journal on Multimodal User Interfaces*, 5(3):143–156.

[Ehrsson et al., 2004] Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! activity in premotor cortex reflects feeling of ownership of a limb. *Science*, 305(5685):875–877.

[Ernst and Banks, 2002] Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433.

[Ernst and Bülthoff, 2004] Ernst, M. O. and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, 8(4):162–169.

[Fagiani and Barbieri, 2016] Fagiani, R. and Barbieri, M. (2016). A contact mechanics interpretation of the duplex theory of tactile texture perception. *Tribology International*, 101:49–58.

[Fery et al., 2021] Fery, M., Bernard, C., Thoret, E., Kronland-Martinet, R., and Ystad, S. (2021). Audio-tactile perception of roughness. In *Computer Music Multidisciplinary Research 2021, in press*.

[Fitts, 1954] Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381.

[Fletcher and Munson, 1933] Fletcher, H. and Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4):377–430.

[Friesen et al., 2021] Friesen, R., Klatzky, R. L., Peshkin, M. A., and Colgate, E. (2021). Building a navigable fine texture design space. *IEEE Transactions on Haptics*.

[Friesen et al., 2018] Friesen, R. F., Klatzky, R. L., Peshkin, M. A., and Colgate, J. E. (2018). Single pitch perception of multi-frequency textures. In *2018 IEEE Haptics Symposium (HAPTICS)*, pages 290–295. IEEE.

[Gaffary and Lécuyer, 2018] Gaffary, Y. and Lécuyer, A. (2018). The use of haptic and tactile information in the car to improve driving safety: A review of current technologies. *Frontiers in ICT*, 5:5.

[Gaver, 1993] Gaver, W. W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological psychology*, 5(1):1–29.

[Gibson, 1979] Gibson, J. J. (1979). The ecological approach to visual perception.

[Giordano et al., 2012] Giordano, B. L., Visell, Y., Yao, H.-Y., Hayward, V., Cooperstock, J. R., and McAdams, S. (2012). Identification of walked-upon materials in auditory, kinesthetic, haptic, and audio-haptic conditions. *The Journal of the Acoustical Society of America*, 131(5):4002–4012.

[Grigorii et al., 2019] Grigorii, R. V., Peshkin, M. A., and Colgate, J. E. (2019). Stiction rendering in touch. In *2019 IEEE World Haptics Conference (WHC)*, pages 13–18. IEEE.

[Gueorguiev et al., 2017] Gueorguiev, D., Vezzoli, E., Mouraux, A., Lemaire-Semail, B., and Thonnard, J.-L. (2017). The tactile perception of transient changes in friction. *Journal of The Royal Society Interface*, 14(137):20170641.

[Gueorguiev et al., 2019] Gueorguiev, D., Vezzoli, E., Sednaoui, T., Grisoni, L., and Lemaire-Semail, B. (2019). The perception of ultrasonic square reductions of friction with variable sharpness and duration. *IEEE transactions on haptics*.

[Guest et al., 2002] Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, 146(2):161–171.

[Gupta et al., 2016] Gupta, A., Pietrzak, T., Roussel, N., and Balakrishnan, R. (2016). Direct manipulation in tactile displays. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 3683–3693.

[Gyoba and Suzuki, 2011] Gyoba, J. and Suzuki, Y. (2011). Effects of sound on the tactile perception of roughness in peri-head space. *Seeing and perceiving*, 24(5):471–483.

[Hahn, 1966] Hahn, J. (1966). Vibrotactile adaptation and recovery measured by two methods. *Journal of Experimental Psychology*, 71(5):655.

[Hart et al., 2006] Hart, J., Collier, R., and Cohen, A. (2006). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge University Press.

[Heller, 1982] Heller, M. A. (1982). Visual and tactual texture perception: Intersensory cooperation. *Perception & psychophysics*, 31(4):339–344.

[Hoggan et al., 2008] Hoggan, E., Brewster, S. A., and Johnston, J. (2008). Investigating the effectiveness of tactile feedback for mobile touchscreens. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1573–1582.

[Hollins and Risner, 2000] Hollins, M. and Risner, S. R. (2000). Evidence for the duplex theory of tactile texture perception. *Perception & psychophysics*, 62(4):695–705.

[Huang et al., 2006] Huang, F. C., Gillespie, R. B., and Kuo, A. D. (2006). Human adaptation to interaction forces in visuo-motor coordination. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(3):390–397.

[Hudin and Hayward, 2020] Hudin, C. and Hayward, V. (2020). When hearing defers to touch. *arXiv preprint arXiv:2004.13462*.

[Hudin et al., 2015] Hudin, C., Lozada, J., and Hayward, V. (2015). Localized tactile feedback on a transparent surface through time-reversal wave focusing. *IEEE transactions on haptics*, 8(2):188–198.

[Huloux, 2021] Huloux, N. (2021). *Étude de la réflexion et de l'absorption des ondes ultrasonores par le doigt : application aux surfaces haptiques*. Theses, AMU - Aix Marseille Université.

[Huloux et al., 2021a] Huloux, N., Bernard, C., and Wiertlewski, M. (2021a). Estimating friction modulation from the ultrasonic mechanical impedance. *IEEE transactions on haptics*, 14(2):409–420.

[Huloux et al., 2018] Huloux, N., Monnoyer, J., Boyron, M., and Wiertlewski, M. (2018). Overcoming the variability of fingertip friction with surface-haptic force-feedback. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 326–337. Springer.

[Huloux et al., 2021b] Huloux, N., Willemet, L., and Wiertlewski, M. (2021b). How to measure the area of real contact of skin on glass. *IEEE Transactions on Haptics*.

[Isleyen et al., 2019] Isleyen, A., Vardar, Y., and Basdogan, C. (2019). Tactile roughness perception of virtual gratings by electrovibration. *IEEE Transactions on Haptics*.

[Janko et al., 2015] Janko, M., Primerano, R., and Visell, Y. (2015). On frictional forces between the finger and a textured surface during active touch. *IEEE transactions on haptics*, 9(2):221–232.

[Johansson and Flanagan, 2009] Johansson, R. S. and Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience*, 10(5):345–359.

[Johansson and Vallbo, 1979] Johansson, R. S. and Vallbo, A. B. (1979). Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *The Journal of physiology*, 286(1):283–300.

[Johansson and Westling, 1984] Johansson, R. S. and Westling, G. (1984). Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental brain research*, 56(3):550–564.

[Jousmäki and Hari, 1998] Jousmäki, V. and Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Current biology*, 8(6):R190–R191.

[Kaci et al., 2019] Kaci, A., Torres, A., Giraud, F., Giraud-Audine, C., Amberg, M., and Lemaire-Semail, B. (2019). Fundamental acoustical finger force calculation for out-of-plane ultrasonic vibration and its correlation with friction reduction. In *2019 IEEE World Haptics Conference (WHC)*, pages 413–418. IEEE.

[Kalantari et al., 2018] Kalantari, F., Lank, E., Rekik, Y., Grisoni, L., and Giraud, F. (2018). Determining the haptic feedback position for optimizing the targeting performance on ultrasonic tactile displays. In *2018 IEEE Haptics Symposium (HAPTICS)*, pages 204–209. IEEE.

[Kanzari et al., 2019] Kanzari, K., Bernard, C., Monnoyer, J., Denjean, S., Wiertlewski, M., and Ystad, S. (2019). A real-time synthesizer of naturalistic congruent audio-haptic textures. In *CMMR 2019: 14th International Symposium on Computer Music Multidisciplinary Research*.

[Katz, 2013] Katz, D. (2013). *The world of touch*. Psychology press.

[Kayser et al., 2005] Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron*, 48(2):373–384.

[Kee et al., 2006] Kee, K. S., Horan, W. P., Wynn, J. K., Mintz, J., and Green, M. F. (2006). An analysis of categorical perception of facial emotion in schizophrenia. *Schizophrenia Research*, 87(1-3):228–237.

[Keele et al., 1989] Keele, S. W., Nicoletti, R., Ivry, R. I., and Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgements? *Psychological Research*, 50(4):251–256.

[Klatt, 1973] Klatt, D. H. (1973). Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *The Journal of the Acoustical Society of America*, 53(1):8–16.

[Klatzky et al., 2017] Klatzky, R. L., Adkins, S., Bodas, P., Osgouei, R. H., Choi, S., and Tan, H. Z. (2017). Perceiving texture gradients on an electrostatic friction display. In *2017 IEEE World Haptics Conference (WHC)*, pages 154–158. IEEE.

[Klatzky and Lederman, 2002] Klatzky, R. L. and Lederman, S. J. (2002). Perceiving texture through a probe. *Touch in virtual environments*, pages 180–193.

[Klatzky et al., 2003] Klatzky, R. L., Lederman, S. J., Hamilton, C., Grindley, M., and Swendsen, R. H. (2003). Feeling textures through a probe: Effects of probe and surface geometry and exploratory factors. *Perception & psychophysics*, 65(4):613–631.

[Klöcker et al., 2013] Klöcker, A., Wiertlewski, M., Théate, V., Hayward, V., and Thonnard, J.-L. (2013). Physical factors influencing pleasant touch during tactile exploration. *Plos one*, 8(11):e79085.

[Kovacs and Shea, 2011] Kovacs, A. J. and Shea, C. H. (2011). The learning of 90 continuous relative phase with and without lissajous feedback: external and internally generated bimanual coordination. *Acta psychologica*, 136(3):311–320.

[Lamble et al., 1999] Lamble, D., Kauranen, T., Laakso, M., and Summala, H. (1999). Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. *Accident Analysis & Prevention*, 31(6):617–623.

[Landelle et al., 2021] Landelle, C., Danna, J., Nazarian, B., Amberg, M., Giraud, F., Pruvost, L., Kronland-Martinet, R., Ystad, S., Aramaki, M., and Kavounoudias, A. (2021). The impact of movement sonification on haptic perception changes with aging. *Scientific Reports*, 11(1):1–12.

[Lawrence et al., 2007] Lawrence, M. A., Kitada, R., Klatzky, R. L., and Lederman, S. J. (2007). Haptic roughness perception of linear gratings via bare finger or rigid probe. *Perception*, 36(4):547–557.

[Lederman, 1974] Lederman, S. J. (1974). Tactile roughness of grooved surfaces: The touching process and effects of macro-and microsurface structure. *Perception & Psychophysics*, 16(2):385–395.

[Lederman, 1979] Lederman, S. J. (1979). Auditory texture perception. *Perception*, 8(1):93–103.

[Lederman and Abbott, 1981] Lederman, S. J. and Abbott, S. G. (1981). Texture perception: studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *Journal of Experimental Psychology: Human perception and performance*, 7(4):902.

[Lederman and Klatzky, 2009] Lederman, S. J. and Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71(7):1439–1459.

[Lederman et al., 2002] Lederman, S. J., Klatzky, R. L., Morgan, T., and Hamilton, C. (2002). Integrating multimodal information about surface texture via a probe: relative contributions of haptic and touch-produced sound sources. In *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002*, pages 97–104. IEEE.

[Lederman et al., 1982] Lederman, S. J., Loomis, J. M., and Williams, D. A. (1982). The role of vibration in the tactual perception of roughness. *Perception & Psychophysics*, 32(2):109–116.

[Lee et al., 2008] Lee, J. D., Young, K. L., and Regan, M. A. (2008). Defining driver distraction. *Driver distraction: Theory, effects, and mitigation*, 13(4):31–40.

[Levesque et al., 2011] Levesque, V., Oram, L., MacLean, K., Cockburn, A., Marchuk, N. D., Johnson, D., Colgate, J. E., and Peshkin, M. A. (2011). Enhancing physicality in touch interaction with programmable friction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2481–2490.

[Li et al., 2020] Li, X., Choi, C., Ma, Y., Boonpuek, P., Felts, J. R., Mullenbach, J., Shultz, C., Colgate, J. E., and Hipwell, M. C. (2020). Electrowetting: a consideration in electroadhesion. *IEEE transactions on haptics*, 13(3):522–529.

[Liao et al., 2017] Liao, Y.-C., Chen, Y.-C., Chan, L., and Chen, B.-Y. (2017). Dwell+ multi-level mode selection using vibrotactile cues. In *Proceedings of the 30th annual acm symposium on user interface software and technology*, pages 5–16.

[Louw et al., 2000] Louw, S., Kappers, A. M., and Koenderink, J. J. (2000). Haptic detection thresholds of gaussian profiles over the whole range of spatial scales. *Experimental brain research*, 132(3):369–374.

[Lu et al., 2020] Lu, S., Chen, Y., and Culbertson, H. (2020). Towards multisensory perception: Modeling and rendering sounds of tool-surface interactions. *IEEE transactions on haptics*, 13(1):94–101.

[Ma et al., 2015] Ma, Z., Edge, D., Findlater, L., and Tan, H. Z. (2015). Haptic keyclick feedback improves typing speed and reduces typing errors on a flat keyboard. In *2015 IEEE World Haptics Conference (WHC)*, pages 220–227. IEEE.

[MacKenzie, 1992] MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1):91–139.

[Madison, 2004] Madison, G. (2004). Detection of linear temporal drift in sound sequences: empirical data and modelling principles. *Acta psychologica*, 117(1):95–118.

[Makous et al., 1995] Makous, J. C., Friedman, R. M., and Vierck, C. J. (1995). A critical band filter in touch. *Journal of Neuroscience*, 15(4):2808–2818.

[McAdams and Bigand, 1993] McAdams, S. E. and Bigand, E. E. (1993). Thinking in sound: The cognitive psychology of human audition. In *Based on the fourth workshop in the Tutorial Workshop series organized by the Hearing Group of the French Acoustical Society.* Clarendon Press/Oxford University Press.

[McDermott and Simoncelli, 2011] McDermott, J. H. and Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron*, 71(5):926–940.

[McGurk and MacDonald, 1976] McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588):746–748.

[Merchel and Altinsoy, 2019] Merchel, S. and Altinsoy, M. E. (2019). Psychophysical comparison of the auditory and vibrotactile perception-absolute sensitivity. In *International Workshop on Haptic and Audio Interaction Design-HAID2019*.

[Messaoud et al., 2015] Messaoud, W. B., Amberg, M., Lemaire-Semail, B., Giraud, F., and Bueno, M.-A. (2015). High fidelity closed loop controlled friction in smart-tac tactile stimulator. In *2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe)*, pages 1–9. IEEE.

[Messaoud et al., 2016a] Messaoud, W. B., Bueno, M.-A., and Lemaire-Semail, B. (2016a). Relation between human perceived friction and finger friction characteristics. *Tribology International*, 98:261–269.

[Messaoud et al., 2016b] Messaoud, W. B., Bueno, M.-A., and Lemaire-Semail, B. (2016b). Textile fabrics' texture: from multi-level feature extraction to tactile simulation. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 294–303. Springer.

[Meyer et al., 2013] Meyer, D. J., Peshkin, M. A., and Colgate, J. E. (2013). Fingertip friction modulation due to electrostatic attraction. In *2013 world haptics conference (WHC)*, pages 43–48. IEEE.

[Meyer et al., 2015] Meyer, D. J., Peshkin, M. A., and Colgate, J. E. (2015). Modeling and synthesis of tactile texture with spatial spectrograms for display on variable friction surfaces. In *2015 IEEE World Haptics Conference (WHC)*, pages 125–130. IEEE.

[Meyer et al., 2016] Meyer, D. J., Peshkin, M. A., and Colgate, J. E. (2016). Tactile paintbrush: A procedural method for generating spatial haptic texture. In *2016 IEEE Haptics Symposium (HAPTICS)*, pages 259–264. IEEE.

[Meyer et al., 2014] Meyer, D. J., Wiertlewski, M., Peshkin, M. A., and Colgate, J. E. (2014). Dynamics of ultrasonic and electrostatic friction modulation for rendering texture on haptic surfaces. In *2014 IEEE Haptics Symposium (HAPTICS)*, pages 63–67. IEEE.

[Micoulaud-Franchi et al., 2011] Micoulaud-Franchi, J.-A., Aramaki, M., Merer, A., Cermolacce, M., Ystad, S., Kronland-Martinet, R., and Vion-Dury, J. (2011). Categorization and timbre perception of environmental sounds in schizophrenia. *Psychiatry research*, 189(1):149–152.

[Miller, 1974] Miller, R. G. (1974). The jackknife-a review. *Biometrika*, 61(1):1–15.

[Minsky et al., 1990] Minsky, M., Ming, O.-y., Steele, O., Brooks Jr, F. P., and Behensky, M. (1990). Feeling and seeing: issues in force display. In *Proceedings of the 1990 symposium on Interactive 3D graphics*, pages 235–241.

[Mithen et al., 2006] Mithen, S., Morley, I., Wray, A., Tallerman, M., and Gamble, C. (2006). The singing neanderthals: the origins of music, language, mind and body. *Cambridge Archaeological Journal*, 16(1):97–112.

[Monnoyer et al., 2016] Monnoyer, J., Diaz, E., Bourdin, C., and Wiertlewski, M. (2016). Ultrasonic friction modulation while pressing induces a tactile feedback. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 171–179. Springer.

[Monnoyer et al., 2017] Monnoyer, J., Diaz, E., Bourdin, C., and Wiertlewski, M. (2017). Optimal skin impedance promotes perception of ultrasonic switches. In *2017 IEEE World Haptics Conference (WHC)*, pages 130–135. IEEE.

[Moore et al., 1997] Moore, B. C., Glasberg, B. R., and Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4):224–240.

[Mountcastle et al., 1967] Mountcastle, V. B., Talbot, W. H., Darian-Smith, I., and Kornhuber, H. H. (1967). Neural basis of the sense of flutter-vibration. *Science*, 155(3762):597–600.

[Mullenbach et al., 2013] Mullenbach, J., Blommer, M., Colgate, J. E., and Peshkin, M. A. (2013). Reducing driver distraction with touchpad physics. *Master's thesis, Northwestern Univ., Evanston, IL, USA*.

[Nancel et al., 2015] Nancel, M., Vogel, D., and Lank, E. (2015). Clutching is not (necessarily) the enemy. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 4199–4202.

[Nefs et al., 2001] Nefs, H. T., Kappers, A. M., and Koenderink, J. J. (2001). Amplitude and spatial-period discrimination in sinusoidal gratings by dynamic touch. *Perception*, 30(10):1263–1274.

[Nefs et al., 2002] Nefs, H. T., Kappers, A. M., and Koenderink, J. J. (2002). Frequency discrimination between and within line gratings by dynamic touch. *Perception & psychophysics*, 64(6):969–980.

[Newell and Rosenbloom, 1981] Newell, A. and Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, 1(1981):1–55.

[Ng et al., 2018] Ng, K. K., Birznieks, I., Ian, T., Andersen, J., Nilsson, S., and Vickery, R. M. (2018). Perceived frequency of aperiodic vibrotactile stimuli depends on temporal encoding. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 199–208. Springer.

[Ng et al., 2020] Ng, K. K., Olausson, C., Vickery, R. M., and Birznieks, I. (2020). Temporal patterns in electrical nerve stimulation: burst gap code shapes tactile frequency perception. *Plos one*, 15(8):e0237440.

[Nieuwboer et al., 2009] Nieuwboer, A., Rochester, L., Müncks, L., and Swinnen, S. P. (2009). Motor learning in parkinson's disease: limitations and potential for rehabilitation. *Parkinsonism & related disorders*, 15:S53–S58.

[Okamoto et al., 2009] Okamoto, S., Konyo, M., Saga, S., and Tadokoro, S. (2009). Detectability and perceptual consequences of delayed feedback in a vibrotactile texture display. *IEEE Transactions on Haptics*, 2(2):73–84.

[O'Brien et al., 2017] O'Brien, B., Juhas, B., Bieńkiewicz, M., Pruvost, L., Buloup, F., Bringnoux, L., and Bourdin, C. (2017). Considerations for developing sound in golf putting experiments. In *International Symposium on Computer Music Multidisciplinary Research*, pages 338–358. Springer.

[Pantera and Hudin, 2019] Pantera, L. and Hudin, C. (2019). Sparse actuator array combined with inverse filter for multitouch vibrotactile stimulation. In *2019 IEEE World Haptics Conference (WHC)*, pages 19–24. IEEE.

[Pantera and Hudin, 2020] Pantera, L. and Hudin, C. (2020). Multitouch vibrotactile feedback on a tactile screen by the inverse filter technique: Vibration amplitude and spatial resolution. *IEEE transactions on haptics*, 13(3):493–503.

[Pantera et al., 2021] Pantera, L., Hudin, C., and Panëels, S. (2021). Lotusbraille: Localised multifinger feedback on a surface for reading braille letters. In *2021 IEEE World Haptics Conference (WHC)*, pages 973–978. IEEE.

[Parseihian et al., 2013] Parseihian, G., Gondre, C., Aramaki, M., Martinet, R. K., and Ystad, S. (2013). Exploring the usability of sound strategies for guiding task: toward a generalization of sonification design. In *Proc. of the 10th International Symposium on Computer Music Multidisciplinary Research*.

[Parseihian et al., 2016] Parseihian, G., Gondre, C., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2016). Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Transactions on Multimedia*, 18(4):674–686.

[Pastor et al., 2004] Pastor, M. A., Day, B. L., Macaluso, E., Friston, K. J., and Frackowiak, R. S. (2004). The functional neuroanatomy of temporal discrimination. *Journal of Neuroscience*, 24(10):2585–2591.

[Pasumarty et al., 2011] Pasumarty, S. M., Johnson, S. A., Watson, S. A., and Adams, M. J. (2011). Friction of the human finger pad: influence of moisture, occlusion and velocity. *Tribology Letters*, 44(2):117–137.

[Patoglu et al., 2009] Patoglu, V., Li, Y., and O'Malley, M. K. (2009). On the efficacy of haptic guidance schemes for human motor learning. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, pages 203–206. Springer.

[Patterson et al., 1992] Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (1992). Complex sounds and auditory images. In *Auditory physiology and perception*, pages 429–446. Elsevier.

[Pitts et al., 2012] Pitts, M. J., Skrypchuk, L., Wellings, T., Attridge, A., and Williams, M. A. (2012). Evaluating user response to in-car haptic feedback touchscreens using the lane change test. *Advances in human-computer interaction*, 2012.

[Pollack, 1968] Pollack, I. (1968). Detection of rate of change of auditory frequency. *Journal of experimental psychology*, 77(4):535.

[Pongrac, 2008] Pongrac, H. (2008). Vibrotactile perception: examining the coding of vibrations and the just noticeable difference under various conditions. *Multimedia systems*, 13(4):297–307.

[Poupyrev and Maruyama, 2003] Poupyrev, I. and Maruyama, S. (2003). Tactile interfaces for small touch screens. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 217–220. ACM.

[Ro et al., 2009] Ro, T., Hsu, J., Yasar, N. E., Elmore, L. C., and Beauchamp, M. S. (2009). Sound enhances touch perception. *Experimental brain research*, 195(1):135–143.

[Robin et al., 2005] Robin, C., Toussaint, L., Blandin, Y., and Proteau, L. (2005). Specificity of learning in a video-aiming task: Modifying the salience of dynamic visual cues. *Journal of Motor Behavior*, 37(5):367–376.

[Robles-De-La-Torre and Hayward, 2001] Robles-De-La-Torre, G. and Hayward, V. (2001). Force can overcome object geometry in the perception of shape through active touch. *Nature*, 412(6845):445–448.

[Rocchesso et al., 2016] Rocchesso, D., Delle Monache, S., and Papetti, S. (2016). Multisensory texture exploration at the tip of the pen. *International Journal of Human-Computer Studies*, 85:47–56.

[Rock and Victor, 1964] Rock, I. and Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143(3606):594–596.

[Romano and Kuchenbecker, 2011] Romano, J. M. and Kuchenbecker, K. J. (2011). Creating realistic virtual textures from contact acceleration data. *IEEE Transactions on haptics*, 5(2):109–119.

[Rossi, 1971] Rossi, M. (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica*, 23(1):1–33.

[Saleem et al., 2019] Saleem, M. K., Yilmaz, C., and Basdogan, C. (2019). Psychophysical evaluation of change in friction on an ultrasonically-actuated touchscreen. *IEEE transactions on haptics*.

[Samur, 2012] Samur, E. (2012). *Performance metrics for haptic interfaces*. Springer Science & Business Media.

[Samur et al., 2009] Samur, E., Colgate, J. E., and Peshkin, M. A. (2009). Psychophysical evaluation of a variable friction tactile interface. In *Human vision and electronic imaging XIV*, volume 7240, page 72400J. International Society for Optics and Photonics.

[Sapp, 2006] Sapp, C. (2006). The mazurka project, http://mazurka.org.uk/experiments/tempojnd.

[Schouten, 1985] Schouten, M. E. H. (1985). Identification and discrimination of sweep tones. *Perception & psychophysics*, 37(4):369–376.

[Schubert and Wolfe, 2013] Schubert, E. and Wolfe, J. (2013). The rise of fixed pitch systems and the slide of continuous pitch: A note for emotion in music research about portamento. *Journal of Interdisciplinary Music Studies*, 7(1-2):1–27.

[Schulze, 1978] Schulze, H.-H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research*, 40(2):173–181.

[Schürmann et al., 2006] Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., and Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage*, 30(4):1325–1331.

[Serafin et al., 2007] Serafin, C., Heers, R., Tschirhart, M., Ullrich, C., and Ramstein, C. (2007). User experience in the us and germany of in-vehicle touch screens with integrated haptic and auditory feedback. *SAE Transactions*, pages 357–364.

[Sergeant and Harris, 1962] Sergeant, R. L. and Harris, J. D. (1962). Sensitivity to unidirectional frequency modulation. *The Journal of the Acoustical Society of America*, 34(10):1625–1628.

[Shao et al., 2016] Shao, Y., Hayward, V., and Visell, Y. (2016). Spatial patterns of cutaneous vibration during whole-hand haptic interactions. *Proceedings of the National Academy of Sciences*, 113(15):4188–4193.

[Shultz et al., 2018] Shultz, C., Peshkin, M., and Colgate, J. E. (2018). The application of tactile, audible, and ultrasonic forces to human fingertips using broadband electroadhesion. *IEEE transactions on haptics*, 11(2):279–290.

[Shultz et al., 2015] Shultz, C. D., Peshkin, M. A., and Colgate, J. E. (2015). Surface haptics via electroadhesion: Expanding electrovibration with johnsen and rahbek. In *2015 ieee world haptics conference (whc)*, pages 57–62. IEEE.

[Sigrist et al., 2013] Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review*, 20(1):21–53.

[Smith et al., 2002] Smith, A. M., Chapman, C. E., Deslandes, M., Langlais, J.-S., and Thibodeau, M.-P. (2002). Role of friction and tangential force variation in the subjective scaling of tactile roughness. *Experimental brain research*, 144(2):211–223.

[Snibbe et al., 2001] Snibbe, S. S., MacLean, K. E., Shaw, R., Roderick, J., Verplank, W. L., and Scheeff, M. (2001). Haptic techniques for media control. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 199–208.

[Suzuki and Gyoba, 2009] Suzuki, Y. and Gyoba, J. (2009). Effects of sounds on tactile roughness depend on the congruency between modalities. In *World Haptics 2009-Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 150–153. IEEE.

[Suzuki et al., 2008] Suzuki, Y., Gyoba, J., and Sakamoto, S. (2008). Selective effects of auditory stimuli on tactile roughness perception. *Brain Research*, 1242:87–94.

[Talbot et al., 1968] Talbot, W. H., Darian-Smith, I., Kornhuber, H. H., and Mountcastle, V. B. (1968). The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. *Journal of neurophysiology*, 31(2):301–334.

[Tanaka et al., 2012] Tanaka, Y., Horita, Y., and Sano, A. (2012). Finger-mounted skin vibration sensor for active touch. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 169–174. Springer.

[Terhardt, 1974] Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). *Acta Acustica united with Acustica*, 30(4):201–213.

[Thoret et al., 2016] Thoret, E., Aramaki, M., Bringoux, L., Ystad, S., and Kronland-Martinet, R. (2016). Seeing circles and drawing ellipses: when sound biases reproduction of visual motion. *PLoS One*, 11(4):e0154475.

[Tiest and Kappers, 2006] Tiest, W. M. B. and Kappers, A. M. (2006). Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility. *Acta psychologica*, 121(1):1–20.

[Tomlinson et al., 2009] Tomlinson, S., Lewis, R., and Carré, M. (2009). The effect of normal force and roughness on friction in human finger contact. *Wear*, 267(5-8):1311–1318.

[Ungan and Yagcioglu, 2014] Ungan, P. and Yagcioglu, S. (2014). Significant variations in weber fraction for changes in inter-onset interval of a click train over the range of intervals between 5 and 300 ms. *Frontiers in psychology*, 5:1453.

[Vallbo and Hagbarth, 1968] Vallbo, Å. and Hagbarth, K.-E. (1968). Activity from skin mechanoreceptors recorded percutaneously in awake human subjects. *Experimental neurology*, 21(3):270–289.

[Van Den Doel et al., 2001] Van Den Doel, K., Kry, P. G., and Pai, D. K. (2001). Foleyautomatic: physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 537–544.

[Vardar et al., 2017a] Vardar, Y., Güçlü, B., and Basdogan, C. (2017a). Effect of waveform on tactile perception by electrovibration displayed on touch screens. *IEEE transactions on haptics*, 10(4):488–499.

[Vardar et al., 2017b] Vardar, Y., İşleyen, A., Saleem, M. K., and Basdogan, C. (2017b). Roughness perception of virtual textures displayed by electrovibration on touch screens. In *2017 IEEE World Haptics Conference (WHC)*, pages 263–268. IEEE.

[Vassilakis, 2001] Vassilakis, P. (2001). Auditory roughness estimation of complex spectra—roughness degrees and dissonance ratings of harmonic intervals revisited. *The Journal of the Acoustical Society of America*, 110(5):2755–2755.

[Verrillo, 1985] Verrillo, R. T. (1985). Psychophysics of vibrotactile stimulation. *The Journal of the Acoustical Society of America*, 77(1):225–232.

[Verrillo et al., 1969] Verrillo, R. T., Fraioli, A. J., and Smith, R. L. (1969). Sensation magnitude of vibrotactile stimuli. *Perception & Psychophysics*, 6(6):366–372.

[Vezzoli et al., 2017] Vezzoli, E., Vidrih, Z., Giamundo, V., Lemaire-Semail, B., Giraud, F., Rodic, T., Peric, D., and Adams, M. (2017). Friction reduction through ultrasonic vibration part 1: Modelling intermittent contact. *IEEE transactions on haptics*, 10(2):196–207.

[Vidal et al., 2020] Vidal, A., Bertin, D., Drouot, F., Kronland-Martinet, R., and Bourdin, C. (2020). Improving the pedal force effectiveness using real-time sonification. *IEEE Access*, 8:62912–62923.

[Visell et al., 2008] Visell, Y., Cooperstock, J. R., Giordano, B. L., Franinovic, K., Law, A., McAdams, S., Jathal, K., and Fontana, F. (2008). A vibrotactile device for display of virtual ground materials in walking. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 420–426. Springer.

[Vos et al., 1997] Vos, P. G., van Assen, M., and Fraňek, M. (1997). Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of schulze's (1978) internal beat model. *Psychological Research*, 59(4):240–247.

[Warren and Verbrugge, 1984] Warren, W. H. and Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: a case study in ecological acoustics. *Journal of Experimental Psychology: Human perception and performance*, 10(5):704.

[Watanabe and Fukui, 1995] Watanabe, T. and Fukui, S. (1995). A method for controlling tactile sensation of surface roughness using ultrasonic vibration. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 1, pages 1134–1139. IEEE.

[Weber et al., 2013] Weber, A. I., Saal, H. P., Lieber, J. D., Cheng, J.-W., Manfredi, L. R., Dammann, J. F., and Bensmaia, S. J. (2013). Spatial and temporal codes mediate the tactile perception of natural textures. *Proceedings of the National Academy of Sciences*, 110(42):17107–17112.

[Wegner, 1998] Wegner, K. (1998). Surgical navigation system and method using audio feedback. In *International Conference on Auditory Display'98*, pages 1–10.

[Wiertlewski et al., 2016] Wiertlewski, M., Friesen, R. F., and Colgate, J. E. (2016). Partial squeeze film levitation modulates fingertip friction. *Proceedings of the national academy of sciences*, 113(33):9210–9215.

[Wiertlewski et al., 2011a] Wiertlewski, M., Hudin, C., and Hayward, V. (2011a). On the 1/f noise and non-integer harmonic decay of the interaction of a finger sliding on flat and sinusoidal surfaces. In *2011 IEEE World Haptics Conference*, pages 25–30. IEEE.

[Wiertlewski et al., 2014] Wiertlewski, M., Leonardis, D., Meyer, D. J., Peshkin, M. A., and Colgate, J. E. (2014). A high-fidelity surface-haptic device for texture rendering on bare finger. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 241–248. Springer.

[Wiertlewski et al., 2011b] Wiertlewski, M., Lozada, J., and Hayward, V. (2011b). The spatial spectrum of tangential skin displacement can encode tactual texture. *IEEE Transactions on Robotics*, 27(3):461–472.

[Wijekoon et al., 2012] Wijekoon, D., Cecchinato, M. E., Hoggan, E., and Linjama, J. (2012). Electrostatic modulated friction as tactile feedback: Intensity perception. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 613–624. Springer.

[Willemet et al., 2021] Willemet, L., Kanzari, K., Monnoyer, J., Birznieks, I., and Wiertlewski, M. (2021). Initial contact shapes the perception of friction. *bioRxiv*.

[Wilson et al., 2007] Wilson, J., Walker, B. N., Lindsay, J., Cambias, C., and Dellaert, F. (2007). Swan: System for wearable audio navigation. In *2007 11th IEEE international symposium on wearable computers*, pages 91–98. IEEE.

[Winfield et al., 2007] Winfield, L., Glassmire, J., Colgate, J. E., and Peshkin, M. (2007). T-pad: Tactile pattern display through variable friction reduction. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*, pages 421–426. IEEE.

[Yanagida et al., 2016] Yanagida, M., Yamamoto, S., and Umata, I. (2016). Effects of the mode of tempo change on perception of tempo change. In *Proceedings of Meetings on Acoustics 22ICA*, volume 28, page 035007. Acoustical Society of America.

[Yau et al., 2009] Yau, J. M., Olenczak, J. B., Dammann, J. F., and Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. *Current biology*, 19(7):561–566.

[Youn et al., 2021] Youn, E., Lee, S., Kim, S., Shim, Y. A., Chan, L., and Lee, G. (2021). Wristdial: An eyes-free integer-value input method by quantizing the wrist rotation. *International Journal of Human–Computer Interaction*, pages 1–18.

[Zhang and Harrison, 2015] Zhang, Y. and Harrison, C. (2015). Quantifying the targeting performance benefit of electrostatic haptic feedback on touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, pages 43–46.