Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca
in Informatica
ciclo XXX

Tesi di Ricerca

Multisensory feedback for interactive surfaces

SSD: INF/01

Coordinatore del Dottorato
ch. prof. Riccardo Focardi

Supervisore
ch. prof. Davide Rocchesso

Dottorando
Alan Del Piccolo
Matricola 956199

Abstract (English)

Our experience of the surrounding environment involves all of our senses. While vision drives many of our everyday activities, the other senses contribute to complete the percept of the objects we interact with. Moreover, there are countless situations in which we react to what we hear and feel rather than to what we see. The interaction with digital devices rarely provides such complex, articulated experience: Conversely, it often focuses almost exclusively on visual information. The present thesis outlines the use of multisensory feedback in the context of interaction with digital devices. Haptic and audio feedback are analyzed first separately, then concerning their integration, both in the presence and absence of concurrent visual stimuli. Two experiments were conducted in the scope of two tasks concerning the interaction with two-dimensional surfaces: the multisensory exploration of virtual textures, and the following of a path by means of auditory and vibrotactile stimuli. The results of the experimentation stress the limitations of currently available technologies while asserting the effectiveness of a multisensory interaction, provided that its design is aware of human mechanisms of perception and processing of stimuli.

Abstract (Italiano)

L'esperienza dell'ambiente circostante coinvolge tutti i nostri sensi. La vista guida molte delle nostre attività quotidiane, tuttavia gli altri sensi contribuiscono a completare il percepito degli oggetti con cui interagiamo. Inoltre, in innumerevoli situazioni agiamo in risposta a ciò che sentiamo e tocchiamo, piuttosto che a ciò che vediamo. L'interazione con dispositivi digitali raramente fornisce un'esperienza così complessa e articolata: al contrario, spesso si concentra quasi esclusivamente sull'informazione visuale. La presente tesi delinea l'uso del feedback multisensoriale nel contesto dell'interazione con dispositivi digitali. Il feedback aptico e quello uditivo vengono analizzati prima separatamente, poi nella loro integrazione, sia in presenza che in assenza di stimoli visivi forniti in parallelo. Sono stati condotti due esperimenti nell'ambito di due attività riguardanti l'interazione con superfici bidimensionali: l'esplorazione multisensoriale di superfici virtuali e il percorrimento di un tracciato basandosi su stimoli uditivi e vibrotattili. I risultati della sperimentazione sottolineano i limiti delle tecnologie attualmente disponibili, affermando allo stesso tempo l'efficacia di una interazione multisensoriale, a patto che la sua progettazione tenga conto dei meccanismi di percezione ed elaborazione degli stimoli dell'essere umano.

Ringraziamenti

In chiusura del mio percorso dottorale non posso che trovarmi disorientato nel dover stilare un elenco delle persone a cui devo la mia gratitudine. Immagino che un buon modo per iniziare sia un ordine cronologico.

Premetto un caloroso ringraziamento al professor Federico Fontana dell'Università di Udine per avermi messo a conoscenza di questo dottorato, e in generale per avermi introdotto agli ambienti dell'informatica orientata al suono. Ambienti che ho trovato largamente più congeniali rispetto a quelli informatici "standard".

In questi tre anni la mia famiglia in campo accademico è stata lo SkAT-VG Lab all'interno dello IUAV di Venezia, a cui sono stato cortesemente prestato da Ca'Foscari. Vi ho trovato un gruppo di accademici talentuoso e variegato, che mi ha dimostrato come capacità intellettuali straordinarie possano accompagnarsi a doti umane altrettanto fuori dal comune.

Innanzitutto ringrazio il mio supervisore professor Davide Rocchesso, ora Professore Ordinario all'Università di Palermo, per avermi dato la possibilità di accedere al dottorato, aver diretto con (estrema) pazienza i miei passi e avermi fornito una quantità di spunti di ricerca e di opportunità formative. È stato e rimarrà per me un esempio di passione e curiosità scientifica.

Di seguito ringrazio il dottor Davide Andrea Mauro (Murivan), che ha abbandonato il Lab a metà corsa ma che ha fatto in tempo a darmi una grossa mano nel salto "al di là dello specchio". Il mio collega di ufficio dottor Stefano Delle Monache (SteM) è stato una costante fonte di consigli, supporto di ogni genere, e caffè. La mia gratitudine ed amicizia non si esauriranno certo con il dottorato. Lo stesso vale per il dottor Stefano Baldan (SteB), un'inesauribile fonte di aiuto e positività.

Un dovuto ringraziamento va al Collegio Dottorale di Ca'Foscari e al suo coordinatore professor Focardi, per il supporto puntuale ed amichevole.

Valicando le Alpi, il mio caro ringraziamento va al dottor Stefano Papetti (SteP) dell'Institute of Computer Music and Sound Technology (ICST) della Zürcher Hochschule der Künste di Zurigo, per l'amicizia, la collaborazione e l'aiuto, non ultimo nella preparazione alla trasferta nell'"esotica" (cit.)

Svizzera. Sempre all'interno dell'ICST, ringrazio i dottori Hanna Järveläinen, Martin Fröhlich e Sébastien Schiesser per il supporto, nonché il professor Germán Toro Pérez per aver reso possibile la mia permanenza.

Ogni tentativo di lista relativo ai professori, ricercatori e studenti che ho conosciuto durante le mie trasferte europee - Zurigo, Lille, Maynooth, Dresda - è destinato al naufragio. Spero di aver fatto tesoro dell'umiltà, dell'entusiasmo e del talento (il tutto in un ambito estremamente cordiale) di cui sono stato testimone.

Terminati i ringraziamenti in ambito accademico, rimangono quelli che sono gli stessi da una vita: ringrazio mia madre e la mia famiglia per il sostegno incondizionato, e gli amici vecchi e nuovi per continuare a volermi bene pur avendo rinunciato a capirmi (come dargli torto!). In ultimo, un ringraziamento a tutti quelli che hanno creduto in me, e a quelli che non l'hanno fatto: avevate tutti ragione.

Funding

Il mio dottorato è stato finanziato in buona parte dal progetto SkAT-VG facente parte del programma “Future and Emerging Technologies (FET) - Seventh Framework Programme for Research of the European Commission under FET-Open”, grant number 618067. Ho ricevuto un finanziamento aggiuntivo dall'ICST del ZHDK di Zurigo a supporto della mia permanenza in qualità di visiting student.

Contents

Chapter 1

Introduction

1.1 Motivation

Our experience of the surrounding environment involves all of our senses. Vision may well be assumed as the central sense driving many of our everyday actions, with the other senses complementing such primary information channel; Yet, there are countless situations in which we react to what we hear, feel or taste rather than to what we see. We feel our way to the light switch in a dark room, or reach for the mobile phone on the bedside table by locating the source of the ringing.

Moreover, vision can predict the qualities of an object in terms of other senses. By looking at a rough wall we can foresee how raspy it will feel at our fingertips. Of course, this happens with the other senses as well, with all sorts of combinations. By the looks and the smell of a dish we can foresee how exquisite, or horrible, it will taste in our mouth.

The interaction with digital devices rarely provides such complex, articulated experience. The currently available technology often reduces to visual perception the most part of it. Prior to that, the comprehension of the mechanisms ruling our perception is far from being fully unraveled. As such, the accurate multisensory rendition of physical objects or phenomena is not viable yet.

Nonetheless, when circumscribing the scope to a defined task, and analyzing the cognitive and perceptual processes that are involved in completing such task, we can at least approximate the information flow that would provide a guidance to our actions in a real world scenario. Such approximation requires the understanding of the involved perceptual mechanisms, the identification of stimuli that are able to activate them consistently with the real world events, and the practical re-creation of such stimuli. If proven correct, the resulting model can eventually reveal itself suitable for housing newer technologies instead of the current ones as they become available.

Considering the use of readily-available technologies for auditory and haptic feedback, what are currently the most prominent limitations to the simulation of real-life multisensory experiences? Can an acceptable rendition be achieved despite such limitations? Can the information usually delivered by means of vision be substituted by means of touch and hearing?

This work addresses such questions within the limited scope of specific tasks operated over interactive two-dimensional surfaces. Such surfaces are, to date, the most common type of digital artifact over which an interaction style based on direct manipulation has been implemented.

In the present thesis, tasks that - in a digital scenario - are usually performed under the guidance of a single sensory channel are proposed in a multisensory context, where visual information is either replaced or extended in a consistent manner. The underlying question refers to defining what is currently possible to achieve, what might be improved and what might not in terms of effectiveness of the stimuli.

1.2 Aim

The aim of this thesis is to analyze the aspects involved in the multisensory experience with digital devices, namely the interactions involving vision, hearing and touch. The limitations in the rendition of multisensory percepts as well as the resulting effectiveness concerning the execution of specific tasks are here evaluated.

Two experiments have been devised to investigate different tasks that commonly occur in our everyday interaction with the environment: The tool-mediated exploration of a surface, and the following of a path in two dimensions. While touch is the leading sense when performing the former task and the latter is driven by vision, the contribution of the other senses is hypothesized to be relevant in both cases. As a consequence, the effect of additional or alternative sensory channels to the interaction is evaluated.

1.3 Thesis structure

Chapter 1 presents the thesis' subject, aim, structure, and contributions.

Chapter 2 introduces the topic of touch. An overview of the related sensations as well as the neurophysical mechanisms that convey them is provided. Then, studies concerning the illusory alteration of such sensations are examined. Finally, current technologies for the exploitation of haptic feedback in digital devices are overviewed.

Chapter 3 introduces the topic of sound for interaction. Notions of sound ecology and of the related issues are provided. Then, the discipline of sonic interaction design is introduced. Tools for designing sounds for interaction are illustrated in the final part of the chapter.

Chapter 4 deals with multisensory feedback, including the requirements of an effective multisensory rendition of an event as well as the phenomena related to the topic. Several application scenarios are introduced.

Chapter 5 introduces the task of texture exploration and presents the design and the realization of an apparatus that affords texture exploration over an interactive surface.

Chapter 6 introduces the path following task. An overview of the related literature is provided. Then, an experiment of path following in non-visual conditions is illustrated.

Chapter 7 summarizes the results of the present research and of the related experimentation.

1.4 Contributions

The present thesis' contributions consist of a theoretical background for the design and the analysis of multisensory feedback in the interaction with digital devices, as well as the execution and the evaluation of two experiments involving the multisensory interaction with digital surfaces. The results provide a perspective on the current possibilities and limitations concerning the depicted scenarios.

1.5 Associated publications

- "To 'Sketch a Scratch' " (2015) A. Del Piccolo, S. Delle Monache, D. Rocchesso, S. Papetti, and D.A. Mauro. 12th Sound and Music Computing conference (SMC15), Maynooth University, Maynooth, Ireland.
- "Non-speech voice for sonic interaction: a catalogue" (2016) A. Del Piccolo, and D. Rocchesso. Journal on Multimodal User Interfaces · July 2016 DOI: 10.1007/s12193-016-0227-6

- “Non-visual path following” (2017) A. Del Piccolo, D. Rocchesso, and S. Papetti. *Under revision for IEEE Transactions on Haptics as of 02/12/2017*

Chapter 2

Haptics

The sense of touch is the first one that is developed in the human body, and the last one that ceases its functioning. It takes on great importance in the perception of the surrounding world, and of ourselves as well. The sensory system that is devoted to detecting tactile sensations is still matter of research, especially concerning the interactions among the receptive neurons of such system and the other receptive neurons, and the behaviour of the brain when processing touch-related stimuli. The emotional and social aspects of touch, or “affective touch”, have long been, and currently are, subjects of research as well[1].

A superset of the sense of touch is represented by haptics. Haptics encompasses both the sensations that are brought to one’s skin and those that are related to the perception of one’s limb position and movement, or “proprioception”. Albeit such functions are carried out by different sensory systems, their operations intermingle in several real-life situations, especially when the perception is related to a conscious action performed by the perceiver. In such cases, some kind of blending between tactile perception and proprioception takes place (both at sensory level and at brain level) to form a coherent percept.

The interaction with digital devices is often flawed by the minimal support to touch, and haptics in general, that is implemented. To date, the commonly found tactile feedback systems only provide vibratory stimuli, which are insufficient to render a plethora of tactile sensations. Concerning force feedback, the available systems are frequently costly and of impractical use. Nonetheless, novel technologies and approaches are being developed to fill in such gap for the sake of a more effective interaction.

In this chapter we will provide an overview of haptic sensations, of the

---

[1] International Association for the Study of Affective Touch, https://iasat.org

neurophysiological mechanisms underlying them, of the mental representation of such sensations, and of the technologies that are currently available to convey haptic sensations to the user of a digital device. Although the focus will be kept on the sensations generated by mechanical stimulations applied to the skin, we will show that proprioceptive and thermal sensations contribute to our percepts of the surrounding objects as well. As a consequence, proprioceptive sensations, also known as "kinaesthesia", and thermal sensations will be considered as well. Conversely, pain sensations are outside of the scope of this summary.

Being a vast and relatively unexplored topic, the brain processing of haptic sensations is outside of the scope of this document. Some related findings will be mentioned in no structured order whenever fitting.

The present chapter will focus on hand-based interaction and feedback, which is the most common in everyday interaction with digital devices. Nonetheless, research has been, and currently is, focused on providing haptic feedback to different body parts.

This chapter is organized as follows: First, we overview the basic haptic properties of objects that can be sensed by humans. Then, we briefly elucidate what sensory systems are able to sense such properties, and how they function. Then, we list the psychophysical factors that may alter haptic perception. Then, we move up to psychological level and discuss the haptic percepts, and how they can be deceived by means of sensory illusions. Finally, we overview the the currently available technologies for haptic feedback, and the research concerning haptic feedback on body parts other than the hands.

2.1 Haptic properties

Tactile sensations encompass those generated by mechanical stimulations, temperature changes, and all the external factors that cause pain in general. Different sensory systems sense and convey the different sensations, starting from different skin receptors. The brain processes such sensations differently as well: For instance, although pain and touch activate a network of similar brain regions, pain sensations are connoted by an intensity aspect, while touch sensations are not [156].

The sense of touch can be specified depending on the sensory inputs that are involved: "Cutaneous sense" is conveyed by the receptors that are located within the skin, while "kinaesthetic sense" is conveyed by the receptors within muscles, tendons and joints [90, 145]. Nonetheless, skin stretch receptors have been demonstrated to be relevant for kinaesthesia as well [38], especially for detecting weaker forces.

At the lowest level, haptic sensations refer to separate properties of the object the perceiver interacts with. Such properties may constitute a percept per se, or contribute as a part of the impression the perceivers create themselves about the object, together with other haptic properties and, most often, with the properties perceived by means of other senses.

In the present section we will overview a main categorization of haptic properties, regardless of the different sensory systems involved in the sensing process.

2.1.1 Material properties

The haptic perception of the properties of a material relies on both cutaneous and kinaesthetic sense, in that it encompasses surface qualities as well as structural and substantial qualities, or "bulk" properties. In general, the sensations that are related to the haptic perception of material properties have been summarized as follows [20]:

- Roughness, related to the height differences in the surface;
- Stickiness/slipperiness, related to the friction between surface and skin;
- Compliance, related to the material's elasticity or hardness;
- Coldness, related to the material's heat capacity and thermal conductivity.

While the first two refer to the surface qualities of an object, the other two refer to the bulk properties. Among the surface qualities, roughness is usually the most considered, since it apparently is the most important feature for discerning textures when exploring them haptically [101, 22]. Another distinction concerns the necessity of movement for the evaluation of a quality: While roughness and coldness can be perceived statically, compliance and slipperiness require movement (either of the object or of the skin) to be appreciated [20].

Interactions have been proven to exist among the sensations related to material perception. For instance, roughness and slipperiness seem to be correlated, being roughness a source for higher friction in a surface. Nonetheless, contrasting results have been found so far: Materials such as sandpapers have been found to cause an increase in friction [65], while grooved surfaces do not [219].

2.1.2 Surface properties

The capability of evaluating the tactile qualities of a surface has been investigated, resulting in the categorization of three different tactile abilities: tactile acuity, vibrotaction, and texture perception. Such abilities seem not to be related to one another [144], hence corroborating the hypothesis of different sensory systems and information channels for each of them.

Tactile acuity

Tactile acuity has been defined as "the ability to resolve spatial aspects of tactile stimulation, such as the orientation of a grating pressed against the skin" [100]. The classical experiment for assessing the tactile acuity consisted in the "two-point threshold": Two raised dots were moved progressively closer to each other until the perceiver could no longer feel the two dots as separate objects. Nonetheless, such measure was affected by factors such as the pressing force and the amount of skin that was stimulated. Later on, a more analytical measure was identified in the abovementioned "grating orientation task" (GOT), namely evaluating the orientation of a grate pattern carved on a surface being applied to the participant's skin for a limited time [45].

Vibrotaction

Vibrotaction has been defined as "the ability to detect and appreciate the properties (such as amplitude and frequency) of vibration imposed on or occurring in the skin" [100]. For technological reasons, vibrotaction is currently the most common sensory ability that is involved in the interaction with digital devices. An overview of the related technologies will follow. Selective sensory adaptation [94] is one phenomenon that was found to be connected to such ability: A low-frequency (10 Hz) adapting stimulus raised the detection threshold for a test stimulus at the same frequency, but had little effect on a high-frequency (200 Hz) threshold, and vice versa. Such effect supports the hypothesis of different populations of sensing neurons featuring different sensitivity to the frequency of vibration.

Texture perception

The feel of a surface is affected by sensations such as hardness, roughness, and slipperiness/stickiness, which are evoked by the microgeometry and the material properties of such surface. When sensing a surface, the skin is spatially deformed over time, i.e. locally stretched and indented: In this sense, texture perception has both a spatial and a temporal component.

The microgeometry, or the grain, of a texture determines two different perceptual strategies: Very fine surfaces (e.g. a center-to-center distance between texture elements below 200 μm) are sensed through a vibrotactile mechanism, that is the sensation originates from the high-frequency vibrations that are induced to the skin when scanning the surface. Such sensation is time-dependent, in that the scanning speed affects the sensation. Conversely, coarse and medium textures require a spatial code, that is to interpret the qualities of the individual texture elements that are scanned. This difference seems to be correlated with the involvement of different classes of sensing neurons, and with the execution of different algorithms at cortical level [237].

2.1.3 Force properties

The sensation of a force consists of two aspects: Magnitude and direction.

As overviewed in [21], magnitude is often evaluated either in the direction of gravity (that is, weight perception), or in other directions. The perceived weight of an object that is being held in one hand is shown to be susceptible to several factors, such as: The size of the object (as will be discussed later), the strength of the grasp exerted on it (if any), and whether the hand is still or moving. The area of the skin that is stimulated by the contact is relevant as well, but only when the holding hand is lying flat on a table (that is, when the cutaneous receptors are the only responsible for the sensation). Conversely, the capability of discriminating two weights improves considerably when the hand is lifted, thanks to the contribution of the kinaesthetic apparatus. Concerning the forces applied in other directions, the direction apparently affects the capability of discrimination between forces.

The perception of the direction of a force seemingly depends on the force, on the movement of the hand, and on the sensing area (whether a finger or the whole hand): The force direction is perceived with better precision with the whole hand and, within a certain range, when the magnitude is higher. A bias, namely a difference between the perceived and the actual direction, was shown to be subject-dependent.

2.1.4 Thermal properties

The normal range for skin temperature is 30-36 °C (the so-called “neutral thermal region”): Within such range we generally do not perceive warmth or coldness. Nonetheless, there can be a variation as high as 12 °C in normal individuals [165].

The skin temperature is usually higher than that of the objects we interact with. As an adaptation to this, the decrease in skin temperature is usually employed to identify the material an object is made of. Conversely, temperature increases are used to evaluate the object's temperature [108].

Although humans can detect fine temperature variations, the speed of such variations is relevant for the sensation that is produced: Slow variations (e.g. up to 0.5 °C per minute) can remain undetected as long as the temperature remains in the neutral thermal region, while fast variations (e.g. 0.1 °C/s) can produce startling sensations even within such range.

2.2 Haptic perception

Tactile sensations result from mechanical stimuli that are mostly detected by sensory receptors. After appropriate conversion and filtering, such stimuli are sent to the brain and processed in the somatosensory cortex.

The touch-related sensory neurons have been fully categorized depending on their "encoding properties": Once a stimulus is received, different classes of neurons have different sensitivities to the features of such stimulus, and convert them into electric signals accordingly. Nonetheless, real life tactile events are likely to involve all of the classes of neurons simultaneously. As such, their interaction has not been fully understood to date. Moreover, the elaboration that such stimuli undergo when reaching the brain is still matter of investigation as well.

Proprioceptive sensations rely on receptors located in muscles, tendons, and joints. The brain integrates such sensations with the information provided by the vestibular system, which is located in the inner ear and provides the sense of balance and of spatial orientation, to form the overall sense of body position, movement, and acceleration.

2.2.1 Mechanotransduction and spike propagation

The neurons that detect haptic stimuli are part of the somatosensory system. Such system includes the receptors for sensations of mechanical pressure or distortion (mechanoreceptors), pain (nociceptors), temperature (thermoreceptors), and limb position (proprioceptors).

The peripheral nerve terminals, or endings, of the cutaneous mechanoreceptors are in charge of transducing the stimuli generated by mechanical forces applied to the skin into electrical potential changes, or "spikes". Such process is named “mechanotransduction”. The spikes travel (see Figure 2.1) from the nerve endings to the dorsal root ganglia (DRG), which house the cell bodies of the sensory neurons, and then to the spinal cord through the afferent nerve fibers (ANF). The spinal cord then communicates such spikes to the brain cortex. Beside the cutaneous mechanoreceptors, other types of mechanoreceptors include the baroreceptors, which are excited by the stretch of blood vessels.

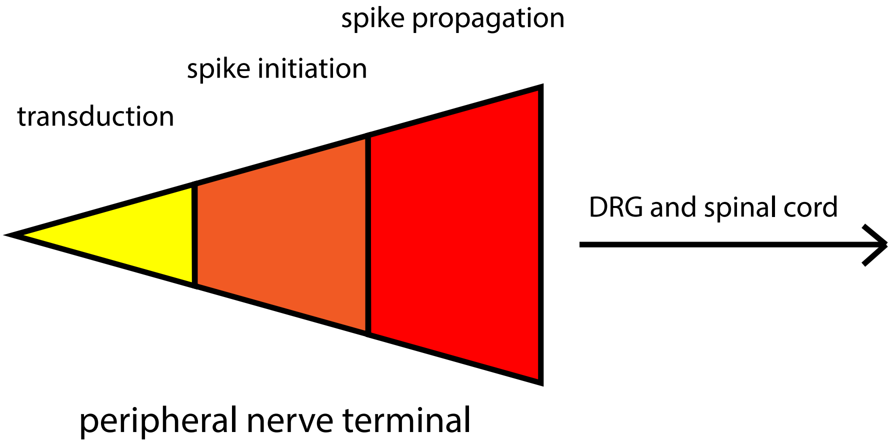


Figure 2.1: Transduction and propagation of spikes.

The nerve fibers of the mechanoreceptors differ from those of the other receptors of the somatosensory system in that they feature a thicker myelinated structure: This enables a higher speed of impulses (∼35-60 m/s against ∼5-35 m/s for thin myelinated and ∼0.4-2.0 m/s for unmyelinated receptors). Cutaneous mechanoreceptos differ from the others because of their sensitivity to low intensity stimuli (e.g. a light touch as opposed to a painful grip), therefore they are named low-threshold mechanoreceptors (LTMs).

The distribution of LTM endings varies between hairy skin and glabrous skin. In the hairy skin, nerve terminals spiral around the hair follicle base or run parallel to the hair shaft. In the glabrous skin, the endings are distributed at different skin depth depending on their type. The lips and the hands are the body parts that contain the highest concentration of mechanoreceptors, thus steering most of the research towards the behaviour of glabrous skin receptors.

In the present summary we will focus on the LTMs that are present in the human hand, being the hand the foremost limb for the interaction with real objects, and often the exclusive limb as concerns the interaction with digital devices. However, it must be underlined that the mechanisms underlying roughness perception may involve neurons that are not present in the hand as well [143].

2.2.2 LTM endings

Four classes of LTM endings are present in the glabrous skin, constituting, along with their afferent fibers, four different sensory systems. Sometimes such systems have been mapped to four sensory channels, which have been hypothesized to constitute four different streams of information reaching the brain [26].

Each LTM class possesses different encoding properties, namely different spatial resolution (or receptive field), frequency response and rate of adaptation. Such properties shape the spikes that result from the mechanotransduction of the stimuli. The rate of adaptation refers to the speed by which an LTM commutes from firing impulses at an elevated frequency in response to the beginning of a stimulus to subsiding back to a normal firing rate as the stimulus becomes constant. Such factor enables a distinction between "slowly adapting (SA)" and "rapidly adapting (RA)" mechanoreceptors.

The classes of LTM endings (along with their afferent fibers, noted in brackets) that are present in the glabrous skin are the following [124]:

- Merkel cells (SA I): slowly adapting, they can be found in groups of 25-75 thus forming a SA I unit named "Merkel cell–neurite complex", or "Merkel disk receptor". Each unit is innervated by a single afferent nerve. Merkel cells are sensitive to vibrations below 10 Hz;

- Ruffini endings (SA II): slowly adapting. They are sensitive to skin stretch, and have a smaller, localized receptive field. Interestingly, they cannot be found in monkey hands;

- Meissner corpuscles (RA, or FA I): rapidly adapting, they are grouped up to 15-25 per afferent axon, thus forming a FA I unit distributed over 2-4 dermal ridges. They are sensitive to vibrations approximately in the 8-70 Hz range;

- Pacinian corpuscles (PC, or FA II): rapidly adapting. Unlike the other LTMs, which reside in the dermis layer, PCs are located in the subcutaneous tissue. They are sensitive to vibrations approximately in the 60-300 Hz range and above.

The localization of LTM endings is depicted in Figure 2.2.

Innervation density and positioning vary among the LTMs as well, thus further characterizing the capabilities of locating a stimulus and detecting its temporal changes. In general, the innervation density is higher at the fingertips than at the palm (∼250 units/cm^2 versus ∼60 units/cm^2) [107].

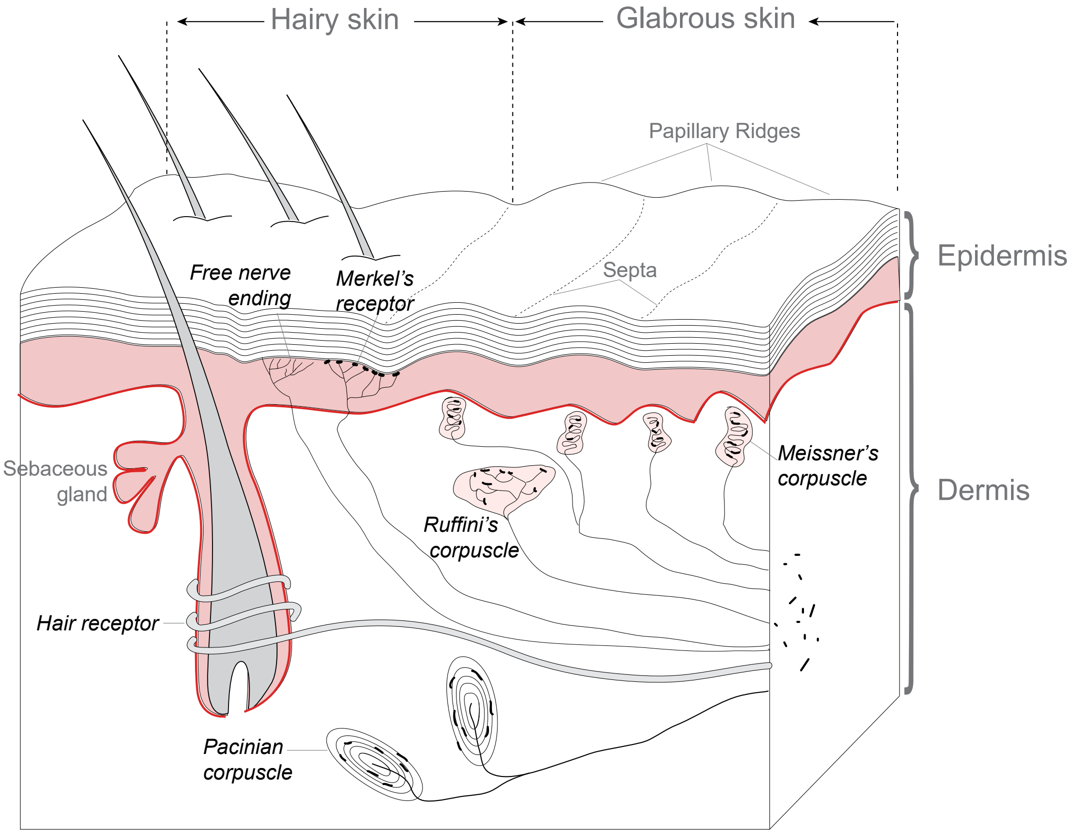


Figure 2.2: LTM endings in hairy and glabrous skin.

Theories on LTM specialization

The capabilities of the LTMs were tested in several experiments of surface probing, e.g. the scanning of Braille dots [168], or the sensing of textures presenting raised dot patterns with increasing diameter [24] or density [44]. The response of the different LTMs was isolated by means of the microneurographic technique [222], consisting in connecting an electrode to a single ANF at a time. A general correlation between increasing roughness of textures (i.e. sparser spatial pattern) and increasing firing and variability was observed for SA I, RA and PC alike [44]. Nonetheless, it is worth mentioning that such behaviour was not correlated to an increasing perception of roughness among the participants: Conversely, subjective roughness magnitude appeared to be an inverted U-shaped function of the dot spacing, while increasing dot diameter produced decreasing roughness sensations.

Later experiments on the encoding of natural surfaces [237] led to hypothesize that SA I are effective in detecting the spatial qualities of a surface, while RA and PC detect better temporal changes e.g. the slipping of an object. It is still unclear how spatial and temporal information are combined by the brain to produce a holistic tactile percept.

The following theories have been formulated concerning the encoding properties of the different LTMs [223]:

- SA I are specialized for spatial information, namely form and some types of texture;
- SA II are specialized for finger shape, possibly for static deformation of the finger pad;
- FA I are specialized for temporal information, namely tactile events, grip control (contact, slip), vibration, and fine texture;
- PC are specialized for vibration (also at high frequency), high sensitivity for distal events (i.e. tool-mediated interaction), very fine texture/roughness, poor localization capabilities.

In general, Merkel complexes are compared to low-pass filters, whereas both Meissner and Pacinian corpuscles are band-pass filters.

Tactile abilities and LTMs

The encoding properties of the LTMs have been mapped to the human tactile abilities as follows:

- Tactile acuity is related to spatial information, which is conveyed by SA I afferents which terminate as Merkel disks;
- PC, RA, SA I, and SA II mediate the detection of vibratory stimuli applied to glabrous skin, and the frequency of vibration determines which LTMs are going to fire;
- Texture perception relies on both a spatial mechanism (dominant for coarse textures) and a temporal mechanism (dominant for fine textures), and is mediated by SA I, RA and PC.

2.2.3 Muscles, tendons and joints

The neurons contributing to the kinaesthetic sense are located in skeletal striated muscles (muscle spindles), tendons (Golgi organs), and inside or around the fibrous capsules in joints (joint receptors):

- Muscle spindles detect changes in the length of muscles;
- Golgi tendon organs (GTO) sense changes in muscle tension;
- Joint receptors are found in the synovial junctions between bones, and detect mechanical deformation within the capsule and ligaments. Four types of endings constitute the joint receptors: Free nerve endings, Golgi type endings, Ruffini endings, and paciniform endings.

The transduction of stimuli into spikes takes place as follows: Stimuli cause pressure on the receptors, with consequent physical deformation of the joint receptor endings. The deformation induces a depolarization in the receptor (caused by the insertion of Na+ ions into its cells), which triggers an action potential.

2.2.4 Thermoreceptors

The thermoreceptors can be found in the dermal and epidermal layers of skin. They are categorized in warmth receptors and cold receptors: The firing rate increases with temperature for warmth receptors, while the opposite happens for cold receptors. The structure seemingly differs between the two classes, leading to a faster conduction velocity for cool receptors (which are thought to be thin myelinated, as opposed to the unmyelinated structure of the warmth receptors [48]), which causes different adaptation rates as well.

Warmth and cold receptors are organized in different populations, thus forming “warm spots” and “cold spots” that are distributed independently on the skin surface. Nonetheless, their receptive fields are only few millimeters wide, thus enabling the coexistence of warm and cold spots within a small area. Interestingly, the areas with the lowest detection threshold for coldness are found to have the lowest threshold for warmth as well [207]. It has been also found that thermal sensitivity is highly variable over the body, with its peak on the face, especially near the mouth.

Unlike the LTMs, thermoreceptors have non-specialized endings. Their firing frequency is related to the maintained skin temperature, and is altered by its changes, to contribute to the thermoregulation of the body.

Other receptors seem to be affected by temperature changes, especially the SA mechanoreceptors: Nonetheless, their possible role in either the perception or the processing of thermal information is yet to be ascertained [202].

2.3 Factors affecting tactile perception

2.3.1 Subjective factors in tactile perception

Several factors have been shown to affect the tactile sensitivity of a subject, such as age and gender. However, it has been demonstrated that such factors do not affect the tactile abilities coherently, but instead they have a deeper impact on certain sensations rather than others.

Differences in tissue conformation, skin thickness, skin hydration, temperature and mechanoreceptor density are among the possible causes that have been contemplated for the differences in tactile sensitivity [226].

The progressive neuronal loss in the brain and in the central nervous system due to ageing is hypothesized to affect tactile perception as well, although such behaviour has not been fully uncovered yet [239].

In presence of neural damage sensory dysfunction may take place. A flourishing literature in neuropsychology investigates what aspects of the tactile perception are lost in relation to what damage, and how the human brain and body cope with, and sometimes adapt to, such information loss.

Experiments on visually impaired individuals led to contrasting results, at times showing significantly lower detection thresholds in the blind compared to the sighted subjects [88]. Conversely, other studies did not highlight significant differences in performance either for the GOT or for vibrotactile frequency discrimination tasks [5]. A high inter-personal variability is thought to be a possible cause for such discordances, along with the impact of the previous experience with particular tasks, e.g. Braille reading, in the performance of the visually impaired [5].

In this section we will categorize the impact of subjective factors depending on the affected sensory ability, extending the search to proprioceptive and thermal sensations as well.

Tactile acuity

Women have higher sensitivity than men on spatial tactile acuity, namely they perceive finer surface details than men [166]. This effect has been hypothesized to be due to the density of SA I and FA I endings. Indeed, women and men have approximately the same number of Merkel disks and Meissner corpuscles: This leads to a higher density of such classes of LTMs in smaller fingers, such as those of women compared to those of men [58]. Therefore, women and men with the same finger size hypothetically possess the same level of tactile acuity.

Tactile acuity has also been proved to decline with age [206], and to depend on the applied contact force [88].

Vibrotaction

It has been shown that both age and gender affect the vibrotactile detection thresholds (VDTs) in passive touch [226]. However, gender has been proven to be a relevant factor only in older subjects (over 65-70 years of age), and for sensations that are conveyed to the hand rather than to the face: In such context, women's VDT is significantly lower than men's. Indeed, an age-dependent loss of sensitivity was detected at frequencies that mostly activate PC (40 to 600 Hz). Conversely, thresholds at lower frequencies tend to have a more differentiated decline pattern along the life span of an individual [227].

Texture roughness

The tactile roughness discrimination threshold (TRDT) and the tactile spatial resolution threshold (TSRT) at the index fingertip were investigated in parallel by using age and gender as factors [144]. Results showed that neither age nor gender affected the TRDT, while TSRT degraded significantly with age. The shown lack of correlation between the two thresholds confirmed the presence of two different neural mechanisms, one associated with PC afferents (TRDT) and one with SA I afferents (TSRT).

Proprioception

The effects of aging are relevant for proprioception, in that the progressive structural modifications within articular and cutaneous receptors cause a decrease in joint position sense and an increase in movement detection threshold. In particular, the decline in muscle spindles seems to be the most important cause at peripheral level, and it couples with the degeneration at central level, consisting in a progressive loss of the dendrite system in the motor cortex, losses in the number of neurons and receptors, and neurochemical changes in the brain. A review of such phenomena can be found in [176].

Thermal sensitivity

Aging causes a progressive decline in thermal sensitivity, especially at the body extremities, while the central regions have a slower decay rate, if any [207]. Women are shown to be more sensitive to heat sensations than men, and to have more pronounced differences across the body parts [82].

The impact of physical activity to thermal sensitivity was investigated, showing a reduction of thermal magnitude sensation on both genders during exercise [82, 83].

2.3.2 Sensory adaptation

Sensory adaptation is a widespread phenomenon that takes place in the context of different sensations, from heat to pain. Concerning touch sensations, one example is the accustomedness to the contact with one's own clothes.

Tactile adaptation

Following the aforementioned early studies on selective sensory adaptation [94], others focused on such phenomenon trying to map it to the presence of different sensory channels. Concerning the LTMs, the adaptation rate and the recovery rate of the different afferent neurons were investigated by varying the frequency and the amplitude of the adapting stimulus, as well as the frequency of the test stimulus [19, 141]. SA I, FA I, and PC were considered. For all three afferent types, the detection thresholds increased with amplitude and frequency of the adapting stimulus, although the desensitization effect was less pronounced for PC compared to SA I and FA I. Moreover, both the adaptation and the recovery followed an exponential time course. Thanks to a previous receptor model [78], such behaviour was explained by a change in the afferents' spiking threshold, namely it originates at the transduction site.

The data collected in [19] and [142] highlighted several contradictions arising from the behaviour of the LTMs in presence of a prolonged adapting stimulus, which have been summarized in [89]. First, SA I afferents are shown to adapt more rapidly to vibration than FA I and PC afferents (this does not apply to step indentation). The cause is hypothesized to lie in the aforementioned change in the spiking threshold at the transduction site, which is exponential in a shorter time constant for SA I than for FA I and PC. Second, the sensitivity in SA I and FA I seems to decrease with decreasing frequency, thus suggesting a symmetrical behaviour for high frequencies, namely a progressive increase. Conversely, the different frequency filtering properties of the different afferents revealed that the thresholds for SA I a FA I increased with increasing frequency.

Thermal adaptation

A continuous exposure to a thermal stimulus causes a decrease in neural responsiveness. Such phenomenon takes place both for cold and warm stimuli. The adaptation rate is in the order of 60 s for changes of $+/-1$ °C when the stimulus is within or close to the skin's neutral thermal region. Adaptation rates are much longer for more extreme temperatures and for stimuli that occur at less sensitive body areas (e.g. the forearm) [115].

2.3.3 Active and passive touch

A relevant factor is the context of perception, namely "active" or "passive" touch. Active touch refers to the situation where "the impression on the skin is brought about by the perceiver himself" [85], for instance when exploring a surface or manipulating an object. As such, it presents a more complicated scenario than simply trying to perceive a vibration that is unrelated to any user action (passive touch). First, the receptors located in joints and tendons are activated as well as the cutaneous mechanoreceptors (in this case, the Pacinian corpuscles [26]). Second, while an exploratory movement is purposive, passive touch is not: The cognitive implications are likely to affect sensitivity, in that active touch exploits the perception-action cycle. As a consequence, it has been shown that active touch, even when stationary, causes vibrotactile sensitivity thresholds to be lower than for passive touch [164].

2.3.4 Use of tools

Many everyday activities are carried out by means of a tool over an object: From handwriting and visual sketching to the tools for manufacture, haptic sensations are often conveyed to one's limb in a mediate form, as opposed to a direct manipulation scenario. At the level of percepts, direct touch differs from tool-mediated exploration in that the former gives a spatial, intensive measure of roughness, while the latter carries information about roughness, hardness and friction in the form of a multidimensional signal in the time variable.

At sensory level, while direct manipulation stimulates mostly SA and FA endings, PC endings are the main responsible for distal sensations due to their sensitivity to higher frequencies. SA II endings contribute as well, since skin stretch may be induced at the points of contact with the tool. Such skin stretch was shown to affect the perception of stiffness of a surface [172].

2.4 Haptic percepts and dimensions

Through the perception and the consequent brain processing of haptic stimuli we build ourselves an impression of determinate qualities of the external objects we are interacting with. Such impressions are a mental re-creation of such stimuli, and are generally named as “percepts”. Percepts can be altered in several ways, as will be shown in the next section.

Such qualities have been defined by Lederman and Klatzky as “dimensions” of the objects and surfaces as we perceive them [122]. Their percepts therefore represent the final result of the sensing process. The authors categorized the most important dimensions as follows:

- Roughness: It depends on two different sensing mechanisms according to the grain of the scanned texture: Vibrotactile (operated by PCs) for fine textures, and spatial (SA I) for medium to coarse-grained textures;

- Compliance: SA I units seem to be the main responsible for this dimension, especially at population level (that is, the combined response of a large group of receptors). Nonetheless, kinaesthetic cues need to be added to differentiate the stiffness of rigid objects;

- Thermal: The response of thermoreceptors may be responsible for the ability to discern different materials and, likewise, for the perceptual confusion between materials in case of similar response;

- Weight: The perception of weight is sharper when both cutaneous and kinaesthetic sensory systems are operating (i.e. when the hand is lifted). Expectations due to prior experience, sensory illusions, and a more direct estimate of moments of inertia are all considered to be factors affecting the weight dimension;

- Curvature, angle and orientation: Both cutaneous and kinaesthetic information contribute to the perception of curvature and angle, depending on the relative size of the object. Orientation has been shown to be affected by biasing factors such as the distance and, in general, the location of the object in relation to the perceiver’s body;

- Shape of two- and three-dimensional forms larger than a fingertip: Large objects seemingly necessitate a complex and expensive processing to integrate the local cues coming from multiple fingers into a single, coherent percept. As a consequence, humans perform rather poorly at processing large, single-material bi- and tri-dimensional objects;

- Size: The static perception of the length or width of an object is usually accurate in humans. However, illusions such as the vertical-horizontal illusion can distort such perception.

2.5 Haptic illusions

Haptic sensations can be deceived in several ways: For instance, by providing the perceiver with contrasting inputs from different senses, e.g. by adding a “wet” sound to the interaction with a dry object. Another possibility is to employ the sensory adaptation mechanisms, e.g. by submerging the hands into two bowls of water at different temperature, and then sense the temperature of a third bowl. At any rate, sensory illusions commonly refer to a “discrepancy between a physical stimulus and its corresponding percept” [136].

As pointed out in [132], sensory illusions are generated by the brain when it mis-interprets the signals incoming from the sensory channels. The brain aims to achieve sensory coherence as defined by constraints on mental estimations, and therefore tends to distort the percepts accordingly. Such goal can lead to illusions that apply to a single sensory channel, or to more than one. Inter-individual variability is a relevant factor, leading to different reactions in response to sensory conflicts.

A related topic is pseudohaptics, meaning the illusory generation of haptic sensations by means of stimuli coming from other sensory channels, usually the visual but the auditory as well. Such topic will be covered in the next Sections.

Lederman’s and Jones’s survey on haptic illusions [136] categorized them into those pertaining the properties of objects (further differentiated in material and geometric properties), and those related to haptic space (in turn differentiated between the observer’s body space and the external space). Such categorization is summarized below.

A notable repository of experiments concerning haptic illusions can be found in [97].

2.5.1 Object illusions

Material properties

The haptic illusions concerning the properties of a material were categorized as follows:

- Texture: They are further divided in the illusory enhancement of tactile sensitivity (e.g. by increasing the sensation of roughness of a surface) and in the virtual rendering of textures (e.g. by modulating the friction force during the scanning of a surface);

- Stiffness: The limitations of haptic devices in rendering features such as mechanical deformation and indentation have been approached through a multisensory strategy grounded in pseudohaptics. By adding audio and especially visual feedback to the interaction it is possible to alter one's perception of the stiffness of an object [135]. To this purpose, visual cues were shown to dominate over proprioceptive cues when presented simultaneously [203, 132];

- Temperature: Thermal illusions occur in the case of relevant and swift changes in temperature, generating the sensation of much more extreme values. In such situation, even if the absolute temperature is still innocuous (e.g. 20 °C to 40 °C), even pain sensations may arise [141];

- Weight: Material properties such as volume, shape, surface texture, temperature and density can alter the perception of the weight of an object [109]:
  - Size-weight illusion: When lifting two objects with identical mass, the smaller one is generally perceived as heavier. This illusion is stronger when the size of the objects can be assessed haptically and not only visually (which, conversely, generates a weaker effect);
  - Shape-weight illusion: The shape of an object influences its perceived volume, consequentially leading to a size-weight illusion even in presence of two objects of the same volume;
  - Density or material-weight illusion: The density of an object seems to alter its perceived weight, although concurring aspects such as the absolute mass value, the grip force and the way of lifting the object are likely to interfere with such effect;
  - Surface texture-weight illusion: By affecting the grip on an object, its surface texture affects the perceived weight in that a firmer grasp causes an object to feel lighter. Nonetheless, such effect takes place only when shear forces have to be contrasted by the grasp, that is when the object is held vertically: In such case, an increase in the required force for holding an object can either be attributed to an increase in weight or to an increase in slipperiness of the surface;

  - Temperature-weight illusion: Cold objects are perceived as heavier than objects at a neutral temperature. Warmer objects are perceived as heavier than neutral objects as well, albeit by a smaller difference.

Geometric properties

Size and shape illusions are among the most investigated when considering visual conditions. Nonetheless, such effects have been shown to take place when the objects are explored haptically as well.

Size illusions encompass:

- Müller-Lyer illusion: The perceived length of a line is altered by the shape of its end delimiters. Arrow-shaped delimiters cause a shorter perceived length as opposed to fin-shaped delimiters;
- Orientation-dependent illusions: The orientation of an object has been proven to affect the perception of the size of its composing segments. Examples of such phenomenon are the horizontal-vertical illusion (a vertically placed line is perceived as longer than a horizontally placed one with the same actual length), the bisection illusion (the length of the segments of T- and L-shaped figures are perceived differently according to their orientation), and the radial-tangent illusion (the length of a radial movement is overestimated compared to that of a tangential movement). Nonetheless, differences have been found between the effect of such illusions under visual and under haptic conditions.

The perceived shape of an object may be altered as an effect of different psychophysical phenomena, especially sensory adaptation to pressure. Many types of illusions have been categorized:

- Curvature illusion: The relative motion of the perceiver's hand and of the object can cause curved surfaces to feel flat, or vice versa;
- Rotation-induced illusions: A coin feels elongated under the fingertips of one hand when it is rotated by the other one, while a straight rod held at its middle while rotating feels thinner in the middle than at the extremities, like a hourglass;
- Ridge illusion: By letting a perforated surface roll while being lightly grasped between thumb and forefinger, a ridge rather than a series of holes is perceived, and the apparent thickness increases. The "computer paper illusion", the "bump illusion" and the "fishbone illusion" work similarly;

- Curvature aftereffect: After a prolonged exploration of a curved shape, a plain surface will feel curved, with an opposite concavity with respect to the actually curved surface;

- Simultaneous contrast effect for curvature: When a convex shape is sensed with the index finger while a flat shape is being sensed with the thumb, the flat shape will feel convex as well;

- Contour enhancement illusion: A single, threshold-level undulation over a surface is perceived more distinctly when a thin paper is manually passed across such surface. The cause may reside in the reduction of friction forces, which enables a better perception of normal forces;

- Tactile contact lens: A 2-dimensional array of pins, when superimposed to a irregular surface, enhances the perception of its irregularities brought to the fingertip. The cause resides in the added tangential stretch of the skin. By the same principle, by letting the array of pins move in a wave pattern, the illusion of a moving undulating surface is generated;

- Curved-plate illusion: The variation sensed by a fingertip when probing a curved surface can be mimicked by altering the displacement and/or inclination of said surface while the finger remains practically stationary. In general, by controlling the forces that are applied to the perceiver's fingertip, it is possible to generate shape illusions that may be even more robust than the geometric cues that are provided by the sensed object. By finding this, Robles de la Torre and Hayward suggested a concurrent contribution of both force and geometric cues [179], which are processed separately at neurophysiological level;

- Tactile diplopia: When a small object is in contact with two fingers that are crossed, it is perceived as two separate objects. This effect is due to a mis-perception of the spatial localization of the stimuli.

2.5.2 Haptic space illusions

Body space properties

Distortions in the spatial processing generate illusions concerning the location of tactile and thermal stimuli applied to the skin. Such distortions can be induced by manipulating the interaction between the spatial and temporal properties of the stimuli.

The tactile illusions on the skin have been categorized into those referring to distance, movement, localization, and thermal properties:

- Distance illusions: The perceived distance between stimuli depends on the temporal interval between their firing. In the "tau" effect, three stimuli are considered: The distance between the first and the second stimulus is twice the distance between the second and the third, yet the temporal interval between them is half. In this case, the second and the third stimuli can be perceived almost twice as far from each other as the first two. Related to such effect, it has been shown that the perceived distance covered by a moving stimulus across the skin depends on its speed: The faster the stimulus, the shorter the perceived distance. Weber's illusion relies on the differences in sensitivity across the body areas: Where the tactile acuity is higher, two nearby stimuli are perceived to be further apart compared to areas with lower acuity. Orientation affects the perceived distance as well: Transverse distances are sensed to be greater than the longitudinal of the same length. Nonetheless, such effect is variable across the body, possibly due to the influence of anatomic landmarks (e.g. joints) and asymmetrical receptive fields;

- Movement illusions: Discrete stimuli applied sequentially may generate the illusion of a single stimulus moving across the skin. Such effect is known as "phi phenomenon", or the "delta movement": The optimal inter-stimulus temporal interval for its arousal varies directly with the duration of the stimuli (such assessment was tested for a 25-400 ms duration range), and decreases with their number;

- Localization illusions: In the sensory funneling illusion, different simultaneous stimuli are perceived as a single stimulus localized at the central location among them. Such illusion takes place only for brief stimuli within a close spatial neighborhood. The "phantom" location is affected by the intensity of the stimuli as well. Sensory saltation refer to the impression of displacement of a single stimulus "hopping" across the skin rather than the actual firing of different, subsequent stimuli at different locations. As in the phi phenomenon, the number and the temporal interval of stimuli is relevant to the effect;

- Thermal illusions: The ability to localize thermal stimuli is normally poor in humans. As a consequence, spatiotemporal illusions are seemingly hard to obtain unless tactile cues are coupled with the thermal stimulation. Thermal sensations on the middle finger have been shown to be altered by the thermal sensations arising from the two adjacent fingers.

The perception of the relative position of limbs (the "body schema") can be distorted, resulting in proprioceptive illusions involving size, length, and position of such limbs:

- Rubber hand illusion: The tactile stimuli applied to an unseen hand may be perceived as applied to an artificial hand which, conversely, is visible. Moreover, the perceived position of the actual limb is slightly displaced towards the artificial one. Visual congruence (size, appearance), close position of the two hands and synchronicity between the stimuli applied to the two hands are necessary for the illusion. The conclusion is that visual, tactile, and proprioceptive information are combined to form the perception of a limb's position;

- Vibration illusion: By applying a vibration to a muscle tendon, the illusion of movement of the limb housing such tendon can be produced. Likewise, sensations of arm extension or flection can arise, even beyond the actual anatomical limits. Such effects are seemingly due to an increment of activity of the muscle spindles. Conversely, no effect takes place when the vibration is applied on joints.

External space properties

The position and orientation of external objects can be mis-interpreted. The reason is hypothesized to reside in the combination of the two frames of reference, the personal ("egocentric") and external ("allocentric") one, which leads to a non-euclidean haptic space within the horizontal plane:

- Parallelism: The task of placing two test bars in a parallel position on the horizontal plane caused large inter-individual differences. Nonetheless, the distance between the bars was relevant for all subjects, leading to poorer results as said distance increased;

- Oblique effect: Parallelization tasks like the one mentioned above are usually performed more poorly with oblique stimulus orientations than with either vertical or horizontal orientations. Nonetheless, the opposite effect has been reported to take place occasionally; The formulated hypothesis is that of the prevalence of the allocentric frame of reference in the regular oblique effect, as opposed of the prevalence of the egocentric one in the reverse oblique effect.

2.6 Technologies for haptic feedback

In this section we will overview the current state of the art of the available technologies for adding haptic feedback to the interaction with digital devices. A perceptual rather than technological standpoint will be adopted, and the categorization will comply with such perspective. Clearly, some technologies can be used to convey multiple haptic sensations: In such case, the most prominent use for each technology will be employed for the categorization.

As previously mentioned, vibrotactile feedback is the most commonly implemented, not only concerning the commercially available solutions, but concerning several innovative technologies as well. Vibrotactile devices were produced for being applied to various parts of the body, e.g. the back [113], or the feet [190]. Coherently with the rest of the chapter, in the present summary we will focus on the haptic devices that were designed for hand stimulation.

In experimental settings, many technologies were devised to provide haptic feedback in addition to those here reviewed. Examples of related prototypes are tactile displays employing air jets [10, 33] or shape memory alloys [96]. However, here we will focus to those technologies that either have been more recently adopted, or have a strong presence in academical or real-life practices.

A relevant aspect concerns whether the interaction is mediated by tools, e.g. a pen, or a stylus, or involves direct touch (clearly, this does not apply to thermal sensations): Not only the stimulated receptors are different, but the involved technologies enable the rendering of different sensations as well.

The co-location of display and input device is important for the performance in tool-mediated tasks as well [76].

2.6.1 Vibration

A common use of vibration has long been that of supporting alert functions within hand-held devices, such as mobile phones. In such context, the size and the cost of the apparatus are the most relevant factors, whereas the possibilities of modulating the output frequency and amplitude are secondary. As a consequence, the variety of vibratory cues that can be encoded is limited. Sensory adaptation to vibration poses further limits to the use of this type of feedback.

In the context of more complex haptic feeback, vibration is commonly used as a substitute for haptic sensations that are harder to achieve on a usual interactive interface, such as those generated by the scanning of the topographic or textural aspects of a surface. In such scenario, an inherent limitation of traditional vibrotactile feedback lies in the fact that the direction of movement is exclusively normal to the tool or limb used for the interaction. Consequently, lateral stimulation (e.g. shear forces) cannot be simulated effectively, as we will show in our experiments.

ERMs, LRAs, tactile transducers, piezoelectric actuators

The most traditional form of vibrotactile feedback involves the mechanical actuation, either of the manipulated object or of the tool that is used for the manipulation.

The common apparatus for providing vibrotactile feedback in mobile devices employs an eccentric rotating mass (ERM), which is an inexpensive solution and can be produced in small form factors. Beyond such context, common alternatives are linear resonant actuators (LRA), consisting of a voice coil driven to make a spring vibrate, and tactile transducers (or “shakers”), where the voice coil drives a small weight to induce vibrations to a resonant surface.

An alternative to such technologies are piezoelectric actuators, which consist of thin layers of piezoelectric materials, i.e. materials that accumulate electric charge in response to applied mechanical force, and vice versa can bend very quickly when a voltage is applied. Since no internal component is required to drive the vibrations, piezoelectric actuators can be produced in a thin form factor. Moreover, thanks to the properties of piezo materials, they can generate vibrations of a high intensity in relation to their size. However, due to the mechanical stimulation, piezo actuators have limited durability compared to the other aforementioned technologies.

Electricity-induced vibration

Several technologies were devised as an alternative to mechanical actuation. The following employ electricity to induce vibrations:

- Electrotactile (or electrocutaneous) stimulation [113]: The skin is electrically stimulated by means of surface electrodes. The reported effect (tingle, itch, pinch etc.) varies depending on various factors concerning the stimulation (voltage, current, waveform, electrode size and material), the interaction (contact force), and the skin (location, thickness, and hydration);

- Electrostatic vibration [245]: An electrostatic force is generated between a surface and thin conductive film slider, which users move with their fingers. The produced vibration is modulated according to patterns of voltage changes to simulate the vibration caused by the scanning of a real surface texture.

Pneumatic systems

Pneumatic systems rely on the manipulation of air to convey vibration or indentation to the skin. The advantages of such technologies are to enable subtler sensations than mechanical stimulation, and in some cases to enable mid-air interaction. The common drawback consist in the bulkiness of the necessary equipment (air pumps, motors etc.), and the difficulty of conveying sharp sensations (e.g. object edges).

Some examples of pneumatic systems are the following:

- Gloves containing fillable air-pockets to generate pressure, even static, to the hand [211];
- Ultrasonic transducers [102] or audio speakers [95] to generate force feedback by air movement;
- Surfaces with suction holes to generate the illusion of an attractive force [93];
- Rings of air emitted by one or more directional nozzles to generate the sensation of force and texture in mid air [200].

2.6.2 Friction

A solution for simulating the effect of lateral forces during the interaction with a flat surface is to modulate the contact friction between the probe and the surface itself. The currently available solutions are the electrovibration and the vibration generated by ultrasonic waves. Both focus on the use of the bare finger as a probe, and currently do not support multiple touch. Moreover, since friction takes place only in the case of movement, no sensation is generated when the probe is stationary.

Such technologies work as follows:

- Electrovibration: In the “TeslaTouch” [17] prototype and the following applications [119] the friction is increased by generating a voltage difference between the surface and the finger, which results in an attractive force that induces periodic skin deformation at the point of contact. The amplitude of the attractive force varies with the signal amplitude, thus generating periodic changes in friction, which can be modulated along with the frequency. The effect of amplitude variations depends on the signal frequency: At low frequencies (e.g. 80 Hz), higher amplitudes generate the sensation of stickiness or rubberiness, while at high frequencies (e.g. 400 Hz) a smoothness increase is sensed. The thickness of the fingertip skin and its moisture affect the force;

- Ultrasonic waves [236]: A surface acoustic wave (SAW) is generated that propagates between the surface and the finger. The result is a "squeeze air film effect", namely the creation of a thin air gap between finger and surface, causing intermittent contact and consequently the sensation of reduced friction. The generated gap can be modulated by means of the SAW's amplitude: By ranging from 0 to 2 μm sensations from "rough" to "very smooth" can be produced, provided that the frequency is sufficiently high (above 20 kHz). Nonetheless, the actual friction reduction is smaller than expected.

The two effects being independent, electrovibration and ultrasonic waves can be coupled to extend the range of achievable friction magnitude [87]. By shifting the wave amplitude and the voltage signal to avoid overlapping, the friction differences are enhanced, thus simulating a notched surface. Conversely, a total overlap can cancel such "step" feeling.

2.6.3 Skin stretch

Stretch at the fingertip's skin is sensed by SA II endings. It is relevant in the perception of coarse and medium textures, and complements the vibrotactile mechanism enabling the perception of finer grain textures. Moreover, skin stretch can be used to produce shape and movement illusions as well.

A typical solution to mechanically stretch the skin of the fingertip consists in pin displays. Other solutions can involve pneumatic systems, as mentioned above.

Pin displays

Pin displays consist of small-sized, two-dimensional arrays of round-tipped pins that are actuated independently to protrude from, and retreat into, a flat surface. They have been devised for conveying shapes, vibration, or static pressure to the user's fingertip. A common application consists in Braille displays, where pins are grouped in sets of 6 or 8, each to convey a single character.

Although complex sensations can be simulated at the fingertip by means of a pin display, the drawbacks of such device are its size and cost due to the presence of an actuator (either piezoelectric, or a solenoid) for each pin.

Laterally moving contactors

An alternative implementation of a pin display consists of enabling the inclination of the pins to induce lateral skin stretch. This has been employed to simulate the geometric features of an object by means of forces applied to the finger [179].

2.6.4 Force

The sensation of a force is among the least implemented in digital device: The generation and control of forces that affect not only the skin, but muscles and tendons as well, requires devices that are often not scalable in size. This undermines the possibility of embedding actual force feedback into portable digital interfaces.

Phantom and Falcon

Two common table-top devices for force feedback are the Sensable Phantom (now marketed as “Geomagic® Touch Haptic Device”[2]) and the Novint Falcon (no longer produced, as of 2017). They both consist of actuated arms connected to a handle, which may be a stylus (for the Phantom), a ball or a pistol grip (for the Falcon, see Figure 2.3). The arms are actuated by means of DC motors, and force feedback can be delivered in the three translation directions. The range of movement allowed by the arms’ extension is of several centimeters (160 mm width, 120 mm height, and 70 mm depth for the Phantom), thus targeting mainly the hand and the wrist. The two devices differ in technical specifications, price point and customer target: Nonetheless, the differences between the two devices in the effect of the force feedback on the performance in a pointing task were shown to be negligible [225].

Anyhow, since their first marketing (mid 1990’s for the Phantom, 2008 for the Falcon), no substantial upgrade of such technologies has been made available.

---

[2] http://www.geomagic.com/en/products/phantom-omni/overview/

Figure 2.3: Novint Falcon.

Electrorheological fluids

Electrorheological fluids (ERF) change their viscosity under electrical stimulation. As such, they can be investigated for prototypes of interfaces providing either force feedback or friction modulation to the user [167].

2.6.5 Temperature

Due to the subjectivity of thermal sensations, such type of feedback is hard to be used consistently. Moreover, the interpretation of thermal stimuli can be ambiguous. Nonetheless, some coherence among the users was found when comparing the interpretation of heat in four different real-life contexts, such as social media activity and online consumer reviews [243], leading to the conclusion that warm can be mapped to 1) more recent activity, 2) physical presence and busyness, 3) higher content use, and 4) positive reviews of a product or service.

Peltier cells

The most common devices that are used to provide thermal feedback in digital devices are Peltier cells. They consist of a thin semiconductor plate: When electricity runs through the plate, one side absorbs heat while the other emits it (see Figure 2.4), according to the direction of the current.

Their form factor enables experimentation in adding thermal feedback to mobile devices [242]. The main drawbacks of such technology encompass their cost, their low power efficiency, and their slow temperature switching speed.

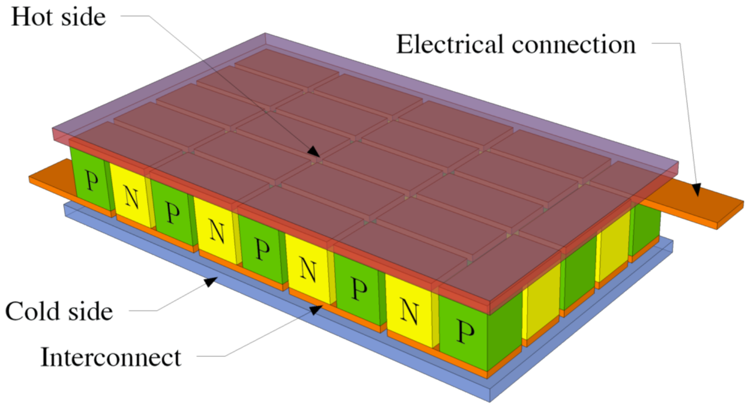


Figure 2.4: Peltier element. Electricity running through the plate enables heat absorption on one side and heat emission on the other.

Electro-resistive heating

The “ThermoTouch” haptic display [126] employs liquid cooling and electro-resistive heating to convey multi-point thermal sensations. A grid of copper resistors is printed over a PCB and placed over a container of liquid coolant: The copper coils heat up in response to the passage of electrical current, and cool down when no current is passing thanks to the coolant. The major drawback of such solution is the lack of portability due to the cooling system.

2.7 Haptic feedback beyond the hands

Several body parts have been investigated for conveying haptic sensations as alternatives to the hands. The reasons may reside in the necessity to leave the hands unencumbered by actuators, cables etc., to provide sensations located in the contact point with the interacting artifact, to enable a full-bodied, immersive haptic experience, or other.

Such alternative solutions for haptic feedback must mainly face the issue of sensitivity differences along the body surface: As mentioned before, the skin receptors for both tactile and thermal sensations have higher density on the hands (and lips) compared to the rest of the body. As a consequence, an accurate mapping of sensitivity across the selected body part is necessary before the implementation of the feedback.

Some examples of haptic feedback that is not conveyed to the hands are the following:

- Head: Vibrotactile headbands have been designed to provide guidance to visually impaired people, for instance to avoid the impact with moving objects [37]. More recently, the spread of VR headsets provided a hardware frame to incorporate actuators to enhance spatial awareness, for instance the localization of objects in the virtual space [49];
- Waist: Belts have been used to accommodate haptic actuators to provide guidance, especially as steering indications for the visually impaired [27]. Such type of tactile notification systems has other applications as well, for instance to provide feedback in live electronics music performance [194], or to inform about incoming messages from one's own mobile device [99];
- Feet: Experiments were conducted on providing feedback through vibrotactile-enhanced shoes to support musical activity such as foot-tapping [163];
- Whole body: Vibrotactile-enhanced suits were designed for the fruition of multisensory art installations [86] or haptic-only compositions [92].

Chapter 3

Sound for interaction

A relevant portion of the information we collect during our everyday activities, and during almost every interaction in general, comes from sound.

Sound enables us to foresee the effects of an action, to interpret a context, and to progressively adjust our actions in order to achieve the desired result. For instance, we often adjust the gear shift by listening to the RPMs of the engine, provided that we are sufficiently accustomed to the particular vehicle we are driving. Such example illustrates a closed-loop sonic interaction [180], in that auditory perception and action are tightly coupled.

Sound can have an affective power as well, which can be found in both natural and artificial sounds, and may lead to more engaging and proficient interactions [218].

Clearly, sound mixes with the incoming information from the other senses: A musician adjusts the position of her own fingers over the instrument's interface proprioceptive (the perceived position of the hands related to her body and the instrument's) information in addition to the visual (which may not be even available), yet the aural feedback (the intonation of the produced sounds) provides the final confirmation of the correctness of the movements.

Sound is the second most used medium in the interaction with digital devices. Every mobile device is provided with small and inexpensive sound reproduction systems which, while enabling the addition of sonic feedback, are usually incapable of providing a wide range of frequencies and of simulating the different localization of sound sources. Another omnipresent issue concerns the real-time feedback generation: Especially in time-critical tasks, delays in the sonic feedback should be avoided. Anyhow, low-quality hardware and non-optimized software usually introduce lags that exceed the perceptual tolerance of human hearing, thus breaking the perception-action cycle.

The present chapter briefly introduces the concepts and the issues concerning the use of sound in interactive scenarios. First, we introduce the notion of sound ecology, which is increasingly being taken into account when designing sounds. Then we introduce the discipline of Sonic Interaction Design, and some of the latest tools that have been developed to support it. Then, we depict the common procedures and practices in the sound design process, and their issues. Finally, we introduce the use of non-speech voice as a possible tool for sketching sounds.

3.1 Ecology of sound

An important aspect of the sounds we accompany to an interaction is their ecological suitability. Acoustic ecology is "a discipline studying the relationship, mediated through sound, between human beings and their environment"[1]. In this context, "ecological" means fitting to the behavior we expect from a natural, everyday interaction. For instance, the impact of two large objects is expected to produce a loud, booming sound. Materials and shapes play a role as well concerning our aural expectations: Indeed, we often adopt terms such as "glassy" or "wooden" to describe the qualities of a sound.

All of such predictions derive from our experience of real life events: When they are not fulfilled they may cause disorientation, incorrect cues, and an overall weaker interaction.

An approach that has been proposed to avoid such risks involves a consideration about what we actually perceive when we are immersed in a sonic environment: In fact, rather than the sounds themselves, we often perceive the attributes of the objects and events that generated them [81]. More specifically, the action properties are what we identify most promptly and accurately as opposed to, for instance, the material that the objects generating the sound are made of [138]. We perceive that a sound is generated by an object rolling, or bouncing, or being struck or scraped better than how we perceive that such object is made of metal, or wood. A possible explanation is that, while material perception is mainly based on the decay of the sound (at least for impact sounds) [11], the identification of an action relies on a wider set of cues, such a rhythm, duration, etc. [138], which varies across the different actions. Such finding confirms that a possible approach to produce ecological sounds is to model the physical processes that generated them, as we will see in a next paragraph.

Conversely, the interaction with digital devices or software has a tradition of relying on abstract sounds, such as beeps or sine waves, which do not resemble or recall any of our real life experiences, but instead were either

---

[1] https://en.wikipedia.org/wiki/Acoustic_ecology

generated due to technical limitations (especially back in the first decades of computing technology), or are the result of a poorly engineered sonic interaction.

Yet, the use of abstract, non-natural sounds can be intentional as well: The sonification of fictional interactions often requires the creation of ad-hoc, not-yet-existent sounds. Sounds from sci-fi movies and video games are a clear example of such scenario: An iconic example is the light saber's sound from the "Star Wars"® movie franchise. Moreover, sonic branding, which is a major topic in software design, industry and advertising, often require the creation of sounds that are somewhat familiar to the consumers, yet have innovative traits in order to be recognizable.

3.2 Sonic Interaction Design

In the last decade, to address the issues of creating appropriate sounds for interaction, the discipline of "Sonic Interaction Design" (SID) [186, 77] was outlined. The purpose of SID is to provide theoretical background, methods and guidelines to convey "information, meaning, aesthetic and emotional qualities in interactive contexts"[180]. Such purpose requires a multidisciplinary approach, encompassing sound and music computing, experimental psychology, computer science, engineering, and cultural studies. As such, some of the topics and issues that are investigated are the following:

- The sounds for interaction cannot be static, in that their features must progressively change to adjust to changes in the conveyed meaning or in the environment along with the the interaction (i.e., a continuous sonic interaction must be designed). As a consequence, the use of sound models whose synthesis controls can be manipulated in real time (also known as "procedural sound synthesis"), instead of the use of pre-recorded sound samples, is a widely adopted solution [31];

- Continuous sonic interaction is inherently hard to design, especially since designers are accustomed to the traditional discrete visual interactions. The resulting sonic interactive objects tend to be complex as well. A proposed approach is to adopt the methods of basic design, namely to start the investigation from well-defined actions or simple objects, and to model their acoustical behavior from a perceptual standpoint [185];

- The novelty of SID requires the creation of an appropriate pedagogy to train designers. "Problem Based Learning" as a discipline focuses on the definition of the object to be designed and on the motivations behind the design choices starting from the analysis of the problem space, domain and context, resulting in the definition of the problem and of its requirements. Team-working, multi-disciplinary approach and creative thinking are relevant to create a learning/design iterative process [160];

- Assigning a meaning to a sound is a subjective experience, due to personal differences but also to the phenomenological aspects of sounds (sound is omnipresent, temporally and physically dynamic, carries spatial information, and can generate emotional and shared experiences), to the typologies of listening (to listen to an isolated sound, or to a system of sounds in an environment - a "soundscape"), and to the domain of application (immersive experiences such as video games, or conversely self-contained interactive objects). In addition to a careful evaluation of such variables, a proposed approach is to try to suggest specific meanings without altering the auditory experience [105].

3.3 Tools for SID

To implement sound models for continuous sonic interaction, one approach is to reproduce specific perceptual effects by modulating the audio signal, regardless of the source that may have generated the sound in the first place. The other is to aim at reproducing the physics of such source, thus simulating the actual mechanism of sound production by means of virtual objects and interactions (impacts, frictions etc.). The latter approach is the most adopted, since it has been proven to be more effective in delivering ecological sounds, and more appropriate for a basic design investigation, since the connection between source and sound is more evident.

3.3.1 Programming frameworks and hardware

Several programming frameworks are available to create real-time sound generating modules, which enable the creation of convenient interfaces to control the synthesis parameters as well. Common frameworks are Max/MSP (currently known only as Max[2]) and PureData (commonly known as Pd[3]), which

---

[2] https://cycling74.com

[3] https://puredata.info

both enable modular architectures rather than line-code programming, as opposed to SuperCollider[4] and ChucK[5].

Real-time generation of sounds can be computationally intensive. To alleviate such problem, frameworks such as Max and Pd allow for the integration of external modules programmed in low-level languages such as C to improve the efficiency.

In the prototype phase of product design, sound synthesis is often run on compact hardware platforms, usually single-board microcontrollers, which can be embedded in the interactive objects and extended with sensors, actuators, and customized interfaces in general. Arduino[6] and Raspberry Pi[7] are general-purpose platforms that are commonly used for sound computing as well, while for instance Bela[8] is specifically optimized for audio and sensor processing.

3.3.2 Sound models and toolkits

Many researchers and designers investigated the physical modeling of sounds. The preliminary phase to such effort is an attempt to categorize the sounds of interest, which usually implies a categorization of their originating phenomena. This may produce large-scale classifications, although it is more common to focus on a well defined class of sounds, e.g. those produced by a precise action such as crumpling [39], which in turn may model different real actions, such as both crushing something and walking on a natural terrain [74]. In general, much effort is addressed to the synthesis of everyday sounds, in that they convey information about our surrounding environment: As such, human auditory perception is specialized at extracting such information [175].

In his book [69], Farnell implemented and collected a considerable catalogue of sound synthesis modules (in Pd), each of which can be used to reproduce one or a class of sounds, according to a categorization based on the perceptual aspects of the considered sounds. The sounds are decomposed in their constituting parts according to the physics of the object or event: For instance, an explosion can be decomposed in a blast, a shock wave, a subsequent pressure wave, and more, each causing a distinctive sound which overlaps to the others to form the final result (see Figure 3.1). An upper level of categorization divides the sounds in artificial sounds, idiophonics

---

[4] http://supercollider.github.io/

[5] http://chuck.cs.princeton.edu/

[6] https://www.arduino.cc/

[7] https://www.raspberrypi.org/

[8] http://bela.io/

(i.e. sounds produced by friction, scraping, rolling, impacts, crushing, and fragmentation), nature sounds, machines, and lifeforms. More advanced examples encompass explosive sounds, and examples of popular sci-fi sounds.

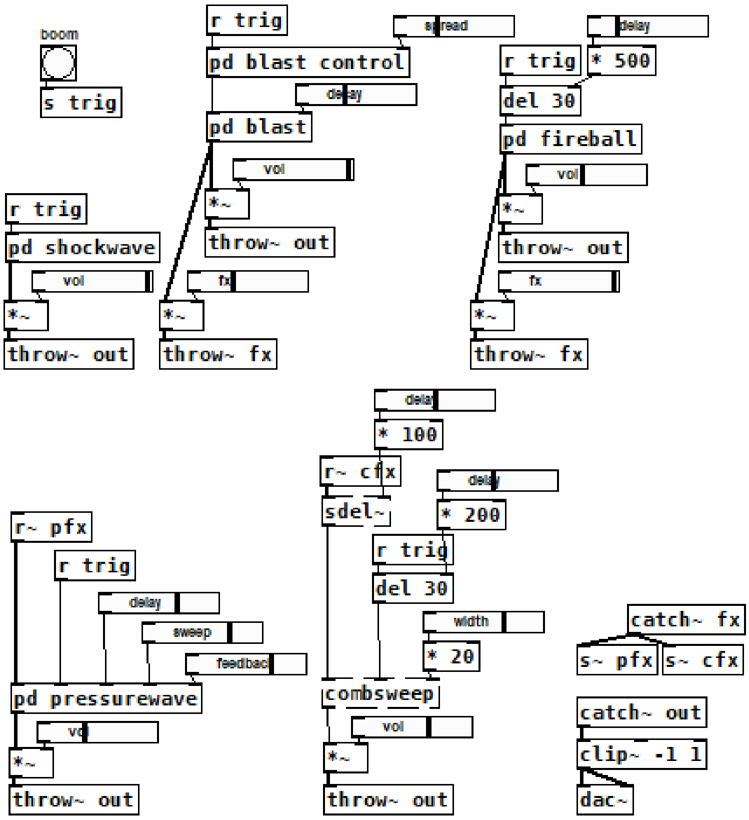


Figure 3.1: A code example by Farnell (in Pd) for the creation of different explosion sounds.

A dynamic impact model for synthesizing scratching, rubbing and rolling sound-actions was developed by Conan et al. [41, 42]. Impacts are distributed in time and controlled in amplitude according to stochastic models of such actions. The result is a single generic sound synthesis model allowing the reproduction of all the three classes of sounds: As a consequence, a single "action space" is provided to the user, who can morph seamlessly between different interactions by controlling the model parameters.

A previous effort by some of the same researchers focused on impact sounds, and to offer an intuitive control over the sound synthesis the resulting synthesizer enabled users to control material, size, and shape of the impacting objects [8].

Rath and Rocchesso [175] focused on modeling the sound of a rolling object, as it provides many cues about the ongoing natural interaction: In addition to material and size, a rolling sound can suggest the direction (e.g. straight or circular) and the velocity of the movement as well as the shape and the surfaces of the involved objects (e.g. a car tyre over a snowy alley), all of which variables can change in real time with an immediate consequence on the sound. Such research originated from the basic mathematical definition of the mechanics of a collision in terms of excitation produced by a point-like mass on a large surface and the resulting resonance [28, 12]. Starting from such formulas, sound models of low-level physical events such as impacts and frictions were formulated, and combined to simulate the acoustic outcome of higher level events, such as bouncing, dropping, breaking, crumpling, and the aforementioned rolling [174].

Starting from such models, and as an outcome of the EU-project "SOb - The Sounding Object"[9], a software package providing a set of physics-based models for interactive sound synthesis named "Sound Design Toolkit (SDT)"[10] was devised. The SDT is programmed in Max and Pd, with the addition of externals programmed in C.

The sound algorithms were designed according to the following criteria:

- Auditory perceptual relevance of the modeled events;
- Cartoonification: The most relevant aspects of the underlying physical events were exaggerated and simplified to increase both computational efficiency and perceptual clarity;
- Parametric temporal control, to achieve a natural and expressive control over the sonic processes.

The SDT has been progressively upgraded through following EU-projects such as "CLOSED"[11], "NIW"[12], and "SkAT-VG"[13], resulting in the modeling of events concerning liquids and gases as well as solids (see Figure 3.2). The refining process generated by-products such as efficient implementations of sound effects (e.g. reverb and pitch shifting) and sound signal analyzers (e.g. a pitch extractor and a spectral analyzer).

Moreover, particular categories of artificial sounds such as combustion motors and DC motors were modeled as well (see Figure 3.3), as part of an investigation in the process of product sound design.

---

[9] http://www.soundobject.org/

[10] http://soundobject.org/SDT/

[11] http://closed.ircam.fr

[12] http://niw.soundobject.org

[13] http://www.skatvg.eu

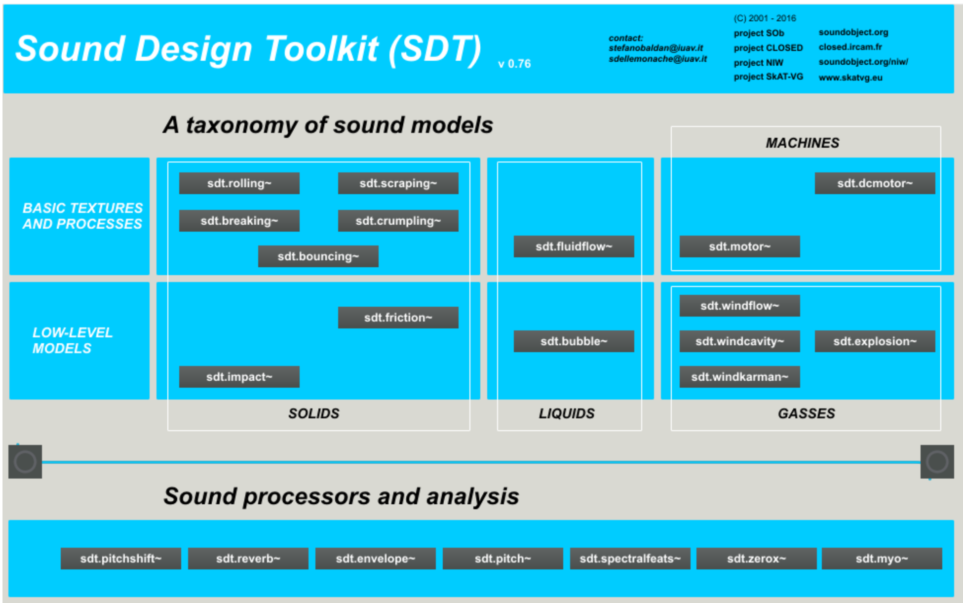


Figure 3.2: The overview of the sound models implemented by the SDT.

3.4 Sketching a sound

Creating sounds to accompany the interaction with objects while providing information and/or aesthetic qualities is a process that must take into account several aspects and, most often, the demands of different actors (the stakeholders, the customers, the developers). As a consequence, such process benefits from an engineered approach, which encompasses cycles of sketching, design and refinement. An industry product, as well as a visual piece of art, often starts from a sketch. Such sketch can be done by means of paper and pencil, or common sketching software. Such tools are flexible enough to enable designers to delete, redraw, adjust, refine sketches until a satisfactory result can be passed to the next phase, that is the actual design of the object as it will be produced in the future. Moreover, a visual sketch can be shared between designers, shown to stakeholders, thought upon during brainstorming sessions. At the present time, the design of sounds lacks of such tools and practices: How to represent a sound, to manipulate it in an effective and intuitive way, and to communicate it among individuals are tasks that are hard to envision. As a consequence, sound designers mostly operate alone, they do not share their results until late in the production phase (thus being prone to having taken the wrong direction of development, eventually leading to the disposal of the entire work), and they use a mixture of digital tools (digital audio workstations, synthesizers, sound editors, effects) of their

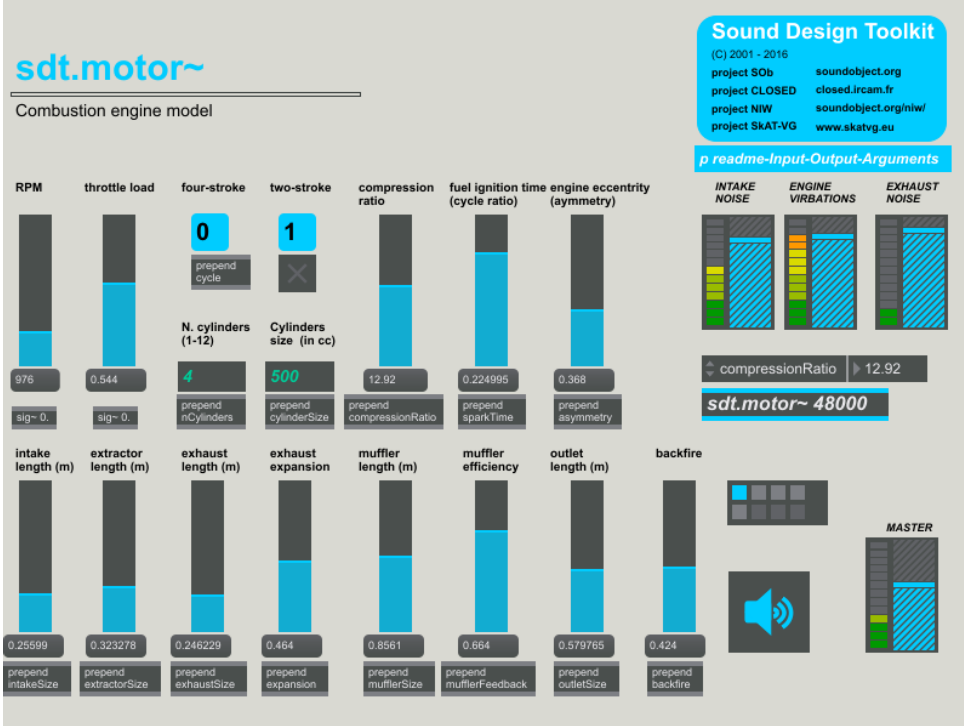


Figure 3.3: The interface of the motor sound model in the SDT.

own choice. Foley artistry, namely the production of sounds for movies other than the voices and the musical score, still relies largely on the use of physical props (see Figure 3.4). Moreover, the lack of sound sketching tools prevents designers to rapidly put into effect creative impulses, which are ephemeral in nature and might get lost or distorted if not promptly captured into a sketch.

3.5 Non-speech voice as a sound sketching tool

When we are lacking for words to describe something, we often use non-speech voice to mimic it, especially when it is linked to a dynamic phenomenon, and a concurrent sonic behavior. Non-speech voice represents a pre-speech, natural and immediate form of expression. As such, it overcomes the possible limitations in technical knowledge of design tools. Unlike onomatopoeias, it is culture-independent, therefore it allows for communication of sonic ideas across different cultures and languages, and theoretically it enables the implementation of an automatic, general-purpose recognizer. Lastly, the voice can be exploited simultaneously with manual interaction. In summary, non-

Figure 3.4: A “Foley pit” is traditionally used to reproduce the sound of an actor’s steps over different terrains.

speech voice represents a valid candidate for a sound sketching tool.

The degree of control of the voice apparatus is by no means inferior to the degree of control of the hands. Yet, the lack of suitable interfaces puts the voice in disadvantage when compared to hands in tasks that involve the fine control of many parameters. That is why non-speech voice appears to be more suitable for sketching rather than for sound refinement and prototyping. Moreover, the overall lack of an engineered approach to the discipline of sound design, with a clear definition of development phases missing, complicates the introduction of new practices in a designer’s workflow, such as the use of the voice to perform fast prototyping and to facilitate the communication of audio concepts.

Here follows a summary of past and current studies over the use of non-speech voice a sound sketching tool, as presented in [53]. Particular relevance is given to the EU-FP7 project “SkAT-VG - Sketching Audio Technologies using Vocalizations and Gestures” (2014-2016) [183], which aimed at the creation of sound sketching tools based on voice and gesture. An outline of the overall process of development of such a tool is provided.

3.5.1 Prior research on vocal sketching

"Vocal Sketching: a Prototype Tool for Designing Multimodal Interaction"

Vocal sketching in the context of multimodal interaction was investigated by Tahiroğlu and Ahmaniemi [217] in a series of experiments. Users were given the prototype of a graspable interactive object, and were asked to vocally imitate its expected auditory behavior while manipulating it. The shape, the affordance (i.e. the actions naturally suggested by the object), and the functionality of the object, as well as the specific gestures performed over it (i.e. moving, squeezing, and stroking) were meant to drive the imitations. The goal was to outline the sonic characteristics of the device and the information that would have been conveyed by the sounds before the actual realization of the accompanying sound synthesis. Likewise, the experiment intended to investigate the expectations concerning the auditory behavior of the object suggested to the users by the object's shape and affordance. Particular focus was given to the coupling between manual gestures and vocal sounds.

The vocal sketches produced by the participants could be mostly categorized under three types of sounds: real world (e.g. elevator sound), synthetic (e.g. sound effects), and abstract sounds. Changes in sonic features were triggered by manual gestures, e.g. a vertical movement of the device expressed a rise of the pitch, while continuous pitched sounds were conveyed by circular movements and horizontal rolling gestures.

Overall, the investigation highlighted the strong interconnection between sounds and interaction modalities. For instance, coupled sounds and gestures had always the same duration. Moreover, specific gestures prompted specific sounds, e.g. a shaking gesture was mostly accompanied by percussive vocal sounds.

"Using Vocal Sketching for Designing Sonic Interactions"

The use of vocal sketching as a methodology to approach sound design was investigated by Ekman and Rinott [66] by means of a workshop. A relevant aspect of sound design that was addressed consisted in the difficulties encountered by non-experts during the early stages of a sound design process. The proposed solution encompassed vocal sketching as a tool for helping designers to think and communicate about sonic ideas.

Participants were asked to use exclusively their voice to sketch the sonic behavior of objects. The resulting sounds were mostly complex, and often presented "organic" features: Clearly, such sounds would have required a high level of technical expertise to be produced by synthesis and to be used interactively in a design.

The investigation highlighted several limitations of vocalization which can affect a design process:

- The voice is essentially monophonic: As a consequence, only teamwork enables the production of harmonies or polyphonies;
- Specific complex sounds are inherently difficult to articulate;
- Single acoustic features are difficult to control separately;
- Long, continuous sounds are impossible to produce due to the limits of the breath cycle.

As a general remark, vocal sketching was attested to drive design particularly towards sounds that are hard to produce by means of current tools.

The VOGST project

The findings by Ekman and Rinott spurred a project carried by Franinović et al. named “VOGST - Voice-Gesture Sketching Tool”[14], whose goal was to develop a tool that enabled designers to use voice and gestures to sketch and improvise sonic interaction. Similar to such research, voice and gestures were proposed as tools to overcome the possible limitations in the technical knowledge of designers and artists while sketching interactive sound concepts. The process of designing the interactions between gesture and sound, and how to facilitate such task, was the main focus of the research.

Several problems were addressed in the process:

- “Ergo-audition”: The human voice is heard differently by the person producing the sound and the one hearing it. This can affect the communication of sonic ideas;
- Further refinement of vocal sketches: Visual sketches, such as pencil lines over a paper sheet, can be redrawn, corrected and changed at will. Conversely, investigations need to be made concerning how to achieve such elasticity by means of vocal sounds and gestures.

A workshop was held to test the resulting tool, namely a simple abstract object that could capture both voice an gestures. Interaction designers were asked to elicit possible design problems and to specify the iterative process of prototyping.

---

[14] http://blogs.iad.zhdk.ch/vogst/

VocalSketch: Vocally Imitating Audio Concepts

A large set of heterogeneous sounds, consisting of thousands of crowd-sourced vocal imitations, together with data representing the participants' ability to correctly identify such imitations, was collected by Cartwright and Pardo [36]. The resulting data set would "help the research community understand which audio concepts can be effectively communicated with this approach" [36].

Four categories of sounds were devised: "everyday", "acoustic instruments", "commercial synthesizers" and "single synthesizer". Users were asked to produce a vocal imitation starting from either a sound label or a reference sound. Users were discouraged to use onomatopoeias.

Participants were especially effective in conveying everyday sounds with vocal imitations. The familiarity of such sounds, but also the ease in reproducing them, were hypothesized to motivate such result. Indeed, the vocal imitations of sounds that are easily producible by the voice (e.g. yawning), or that present peculiar time-varied characteristics (e.g., police siren) were those which were recognized with the highest accuracy. Conversely, sounds consisting of many overlapping sonic events (e.g. glass shattering) led to the least accurate recognition.

As a general remark, inaccuracies mostly led to a description of similar or possibly more general concepts than the one that had been imitated. Authors argue that, as a consequence, more information might be needed for disambiguation. Such information may be provided verbally by users.

3.5.2 The SkAT-VG project

The "SkAT-VG" project aimed at enabling designers to use their voice and gestures to produce sketches of the auditory aspects of an object, whether an industrial product or an artistic effort. The final goal of the project was to devise an automatic system that would interpret the designers' intentions through their vocal sketches, and consequently select appropriate sound synthesis modules. The designers would then use such modules to perform the iterative refinement of a sketch and, potentially, collaborate in the sound design process. A block diagram depicting the phases of the creation of the vocal sketching tool and the involved research is shown in Figure 3.5.

The research that is behind the project, and constitutes the theoretical foundations to build such a tool, encompasses different disciplines such as phonetics, machine learning, and interaction design. Such research is structured in tasks such as:

- To identify the vocal sonic space that is specific of this context. In fact, such space exceeds in size that of spoken language, as vocal imitation employs also phonatory mechanisms that are rare or unused in language [98];

- To classify the sounds of interest, both on a perceptual and on a semantic basis. Product sound design was identified as the context for such sounds, and consequentially a set of 26 sound categories, organized in three main families ("abstract", "interaction", and "machine"), was experimentally defined [139];
- To implement an automatic classifier to associate a vocal sketch to a sound category: Specific audio descriptors that directly highlight the morphological aspects of sound were shown to produce acceptably accurate results in the classification of vocal imitations [148].

A prototypical tool named "miMic" (see Figure 3.6) was devised for enabling vocal sketching activities [184]. It consists of a microphone which has been augmented with two latching buttons and an inertial measurement unit. miMic is connected to a computer running the computations and displaying the visual information. miMic is meant to be used both as a tool for selecting one or more synthesis models ("select mode", activated by the first button) and as a controller for interacting with them ("play mode", activated by the second button). In the play mode, both voice and gestures affect sound synthesis: A control layer maps the envelopes of the voice and movement features into synthesis parameters. Keyboard-based interaction is therefore required only at a later stage, where such parameters may need further refinement. The result is a more natural, spontaneous interaction as opposed to the traditional mouse-and-keyboard operations.

An example of a sketching session is provided below, in the context of combustion motor sounds [14]. A possible application that is being investigated is the creation of sounds for not-yet-existent wheeled motor vehicles, which must address both safety issues (i.e. a virtually silent electric car represents a risk for pedestrians) and ecological constraints (e.g. a continuous alert signal may render cities unlivable):

0. **Selection** of a sound model, or of a mixture of sound models, by operating as follows:
    (a) The user presses "select" and performs a vocal sketch into the microphone;
    (b) The system analyzes the sketch and classifies it into a sound category, or a mixture of weighted categories (e.g., combustion engine and wind);

   (c) The system returns the sound synthesis modules that are relevant for the chosen sound categories.

1. **Mimicking** of the desired sound, which is made vocally after pressing "play". By means of the vocal signal, the designer drives the sound synthesis, validates the model selection, and familiarizes with the model's sonic space;

2. **Further exploration** of the possibilities of the synthesizer(s) by means of a creative use of the voice, that is beyond the simple imitation of the desired sound. Additional features of the sound can be manipulated by moving the microphone in various ways (rotating, tilting, swinging);

3. **Iterative refinement** of the sketch by manipulating the individual sound synthesis parameters on the computer, until a sound prototype is finally obtained.

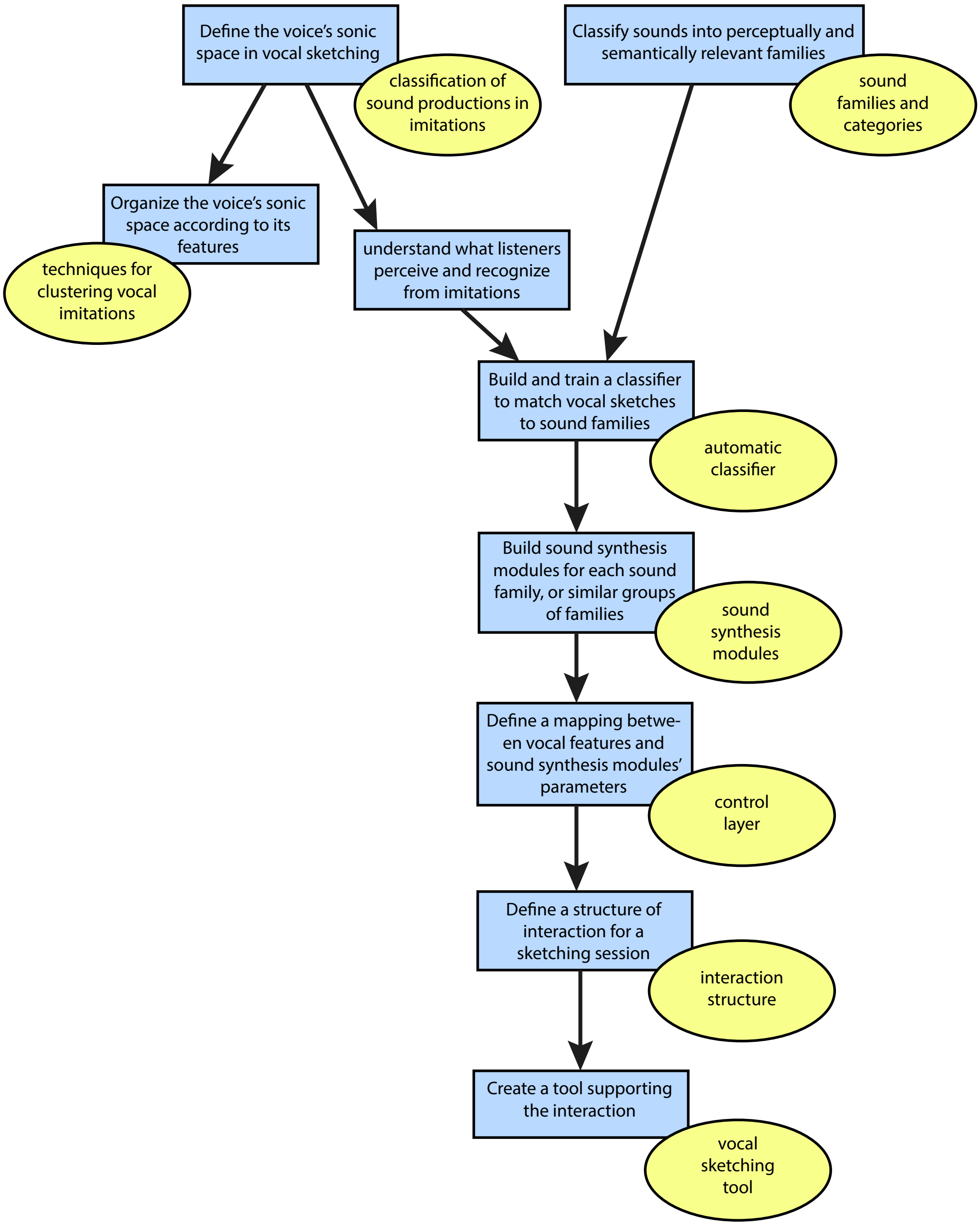


Figure 3.5: SkAT-VG project: phases of the creation of a vocal sketching tool (in blue), and their corresponding outcomes (in yellow).

Figure 3.6: miMic captures the designer's voice and gestures to enable the vocal sketching, the selection of a sound synthesis module and its exploration.

Chapter 4

Multimodality

We usually form ourselves the experience of external objects or events from the merging of information flows incoming from different senses. If coherent, such flows reinforce the perception, since each of them can convey different aspects of the same phenomenon.

Multisensory feedback has many applications. By presenting different information through different sensory channels, it can be used to reduce visual data in many data-intensive tasks, such as driving and remote operations (e.g. in medicine, or in industry). Nonetheless, the stimuli can refer to the same external properties as well, thus providing some form of redundancy: For instance, visual speech gesture information in addition to audio signal reinforces the perception of speech, especially when the audio is degraded [34].

Redundancy is fundamental in real life situation such as in mobile or ubiquitous contexts [221], and in general where sensory impairments, multitasking activities, environmental noise and/or occlusion phenomena take place. Another factor is personal preference over different sensory modalities, for which a multimodal approach to interface design may appeal to more users at once.

The sharing of experiences is another application of multisensory feedback: Haptic sensations, although omnipresent and pervasive of our perception, are highly subjective (see Section 2.3), and consequently hard to transmit among individuals. By substituting such sensations with those of other sensory channels, for instance the auditory one, it may be possible to achieve such sharing.

The nervous system must discriminate information generated by a single external event, or more than one: For such task, spatial and temporal information about the incoming stimuli are used to resolve ambiguity (e.g. whether two stimuli from different senses are generated by the same event or not). Apparently, temporal correlations dominate some cross-modal effects, while spatial correlations dominate others.

The terms "multimodal" and "multisensory" are often used interchangeably, although the former seems more adequate in describing multiple interaction styles, while the latter shall refer to the perceptual sphere. Focusing on the perceptual phase, in the present chapter we will adopt the term "multisensory" as a preference, except for the case of specific, well-established terminology.

This chapter is organized as follows: First, we consider the factors affecting the integration of multisensory stimuli; Then, we exemplify different classes of multisensory integration, depending on the involved senses. Then, we analyze cross-modality and pseudo-haptics (with an overview of the related phenomena), while mentioning the role of amodality and sensory substitution. Finally, we show the impact of multisensory feedback in several contexts.

4.1 Factors affecting multisensory integration

4.1.1 Temporal coincidence

Stimuli incoming from different sensory channels must be received approximately simultaneously to contribute to a coherent percept. However, the human perception and sensory integration process allows for a certain amount of time to occur between different stimuli (e.g. a beep sound and a flashing light) while still being interpreted as originated by a single event. Such mechanism presumably takes place to compensate for the differences in the transmission times for different sensory stimuli, both in the external environment (e.g. light and sound traveling at different speeds) and at neural level (e.g. different propagation speed in different nerve fibers, see Section 2.2.1) [234]. Moreover, such time span has been proven to be adaptable, depending on several factors, as will be shown below. The adaptation process is named temporal recalibration.

In neuropsychology, two similar metrics have been devised to evaluate such time span, which usually amounts to several hundred milliseconds. In particular, the time span within which stimuli are still perceived as synchronous is named "temporal binding window" (TBW) [235]; A complementary concept is that of the point of asynchrony at which separate stimuli are most likely perceived as synchronous, namely the "point of subjective simultaneity" (PSS).

In real life, the capability of rapidly adjusting perceptual representations constitutes a clear advantage, as it often leads to an improved representation of the sensory environment. Such process is named "perceptual learning" [68].

The capability of integrating not perfectly simultaneous stimuli into a single percept is convenient in terms of the implementation of multi-sensory interfaces, in that it provides some tolerance to technology-dependent signal delays. Conversely, there are clinical conditions which lead to an excessive size of TBW, such as autism spectrum disorder (ASD) [50]: In such cases the temporal acuity, namely the capability of segregating temporally different stimuli, is impaired. Moreover, it has been shown that a large TBW leads to a weaker overall multi-sensory integration [210]. As a consequence, there is apparently a limited range within which TBW is desirable in a practical setting.

Albeit considerably variable across individuals, and probabilistic in nature, TBW and PSS have been shown to depend on, and be altered by, several factors:

- Complexity of stimuli: Simple audiovisual stimuli (e.g. flashes and beeps) lead to narrower TBWs, while wider TBWs are reported for more complex stimuli (e.g. audiovisual speech) [209];

- Training: By training individuals to the perception of stimuli before the actual tests, TBW or, vice versa, temporal acuity can be enhanced. Apparently, the difficulty level of training stimuli affects such capabilities: While easily-discernible stimuli ultimately increase the TBW, hard-to-segregate stimuli lead to an enhancement of temporal acuity [50];

- Presence of feedback during training: Feedback information is known to integrate sensory experience in altering the rate of perceptual learning, or in enabling it when sensory experience is insufficient [196]. Distinctions were shown to exist between feedback that informs (i.e. negative feedback) and feedback that confirms (i.e. positive feedback), and the time scale by which the recalibration occurs. In fact, rapid recalibration (i.e. the recalibration due to the feedback to the immediately precedent action) shows a different behavior as opposed to cumulative recalibration (i.e. the long-term recalibration effect): While rapid recalibration tends to rely on positive feedback, cumulative recalibration relies more heavily on negative feedback [51]. In general, feedback signals also produce rapid improvements in multi-sensory temporal acuity;

- Involved senses: The TBW for tactile stimuli is seemingly less adaptable than the TBW for auditory and visual stimuli [158]. Moreover, recent findings seem to indicate that rapid recalibration occurs only with

> audiovisual stimuli, while audiotactile and visuotactile stimuli are prone to cumulative recalibration only [224]. This phenomenon is thought to be due to performance requirements in critical audiovisual tasks such as speech processing.

Temporal recalibration is important in the perception-action cycle, in which case the recognition of the causality between one's own action and the consequent sensory feedback supports the cycle itself. The perception of synchrony can be shifted after exposure to an induced delay, thus generating a temporal recalibration effect (TRE). Sugano et al. [212] tested the impact on TRE of visual and auditory delayed feedback in response to the participants' actions. While consistent asynchronies in both modalities induced large TREs, auditory feedback was shown to have a stronger impact than visual feedback in mixed conditions (i.e. asynchronous audio with synchronous visuals and vice versa). Such finding is presumed to be related to the higher sensory precision of audition compared to vision.

4.1.2 Spatial coincidence

The importance of the co-location of stimuli that are meant to contribute to a unified percept has been investigated concerning all combinations of visual, auditory, and haptic feedback. It has been hypothesized that the neural receptive fields form a map of space into which a visual, an auditory and a somatosensory map overlap [205].

Temporal and spatial localization affect each other. Simultaneous events are likely to originate from the same event, and consequently from the same point in the space. Likewise, it was shown how the temporal order of a series of stimuli is blurred as the originating sources are closer in space [121]. Moreover, a spatial bias was shown to exist in relation to the pitch of auditory stimuli: When presented in a left-to-right order, the central of three temporally contiguous brief tones with different pitch tends to be reported by listeners as the first or the last one of the series [182]. Conversely, such phenomenon does not occur when the spatial order is right-to-left.

An experiment attempting to measure the minimum angle needed for participants to discriminate an auditory stimulus from a tactile event showed how humans are very sensitive to spatial source differences [6]: In fact, such minimum angle was estimated in several degrees (5.3°), which is similar to the localization blur (i.e. the smallest change that can be perceived as a change of location) between two sounds.

It must be mentioned that such experiments consider horizontal displacement, while vertical localization of auditory stimuli is usually poor [188] and affected by the frequency content of the sound [70]. Moreover, in Chapter 2 we showed how overall tactile sensitivity is variable depending on body region and frequency of the stimulus. Tool-mediated interaction may further undermine the sensations of spatial coincidence of multiple stimuli.

4.1.3 Intensity matching

In real-life situations, a sound is often accompanied by a haptically perceivable vibration [154]. It has been hypothesized that vibratory stimuli are employed in animal kingdom as an extension of hearing for those frequencies which are below the audible range, thus forming a continuum of detectable frequencies [240].

In the attempt of reproducing multimodal percepts, especially in experimental settings, it is important to match the intensity of the stimuli to avoid bias towards one sensory modality over the other. The result of such matching process is intended to be a base-line level where visual, auditory and haptic intensity (or any combination of them) are judged as comparable by the users.

Experiments were run where an auditory signal was coupled to a whole-body vibration, for instance caused by a vibrating seat. Tests with sinusoidal signals at different frequencies were performed [73, 154], as well as with noise in the context of railways [104]. Later experiments considered time-varying stimuli as well [153], where participants were asked to match the intensities of a sound and a vibration produced by the same sinusoidal signal while the overall duration of such stimuli was modulated. Up to a certain time threshold, the perceived intensity of both vibration and sound was shown to increase with time. After several hundreds of milliseconds, adaptation effects presumably take place.

General results show large inter-individual differences, while single participants tend to be self-consistent within their judgments. One reason behind the inter-individual differences is the body-related transfer function (BRTF), which states that different human bodies transfer vibrations differently. Such outcome stresses the importance of subjective adjustments of feedback levels in perceptual experiments as well as in real life activities.

To date, no systematic technique for intensity matching has become a common practice in experimentation. Usually, participants are asked to perform a cross-modal matching before performing the experiment, which compensates for both inter-modal and inter-individual sensitivity differences. Pairs of stimuli are usually considered, such as visual-auditory, visual-tactile, and auditory-tactile. Nonetheless, variables such as the order of presentation of stimuli and the controls used to adjust intensities may bias such matching.

Moreover, one single match value may be not enough to match a pair, since the sensitivity curve is likely to be different for different senses.

An experiment asking participants to match the intensity of pairs of basic stimuli (a blink of light, a beep, and a tactile vibration) with two different matching techniques alternatively (either sliding scale and computer mouse or keyboard arrow keys) confirmed the significance of both the matching techniques and the matching pairs [169]. An interaction between the order of presentation and the intensity of the reference cue to be matched was demonstrated as well. In general, a high within-subject variability was detected, in addition to the usual between-subject variability: As a consequence, the authors express the need of averaging repeated matches for each stimuli pair and participant. Nevertheless, for auditory vs. tactile displays in the absence of visual feedback the cross-modal matching is not critical, as the aforementioned factors are much less relevant.

4.2 Dynamics of multisensory integration

The process that takes place when multiple unimodal information streams must be merged by the brain has been, and currently is, matter of investigation.

Driver and Spence [61] assumed that generally external events stimulate different senses simultaneously: Such stimuli are “convergent” meaning that they provide information about the same external property. Logically, the combination takes place to obtain the best estimate of such property, that is to reduce the uncertainty (or variance [67]) in its mental representation. As a consequence, it is hypothesized that the most acute sense with respect to the nature of the property is weighed more in the combination. For instance, in a localization task vision overcomes audition, while in a temporal ordering task audition is superior [67].

However, cross-modal interactions may take place even if the information provided by one modality is task-irrelevant, namely it is ”orthogonal”. For instance, it was shown that synchronous salient events (e.g. a high-pitched sound amongst low-pitched sounds) lead to the isolation of events of other modalities (e.g. a particular image shown amongst others) [233].

For the sake of simplicity, many studies focus on bi-modal interactions. Some are summarized in the following subsections.

4.2.1 Visuo-tactile integration

Vision and touch were shown to cooperate in enhancing the robustness of a percept. Viewpoint was found to be a relevant factor in that we usually experience hand-sized objects from a front view, while we touch and manipulate mostly their back and sides [159]. As a consequence, the two senses contribute to the perception of the whole three-dimensional object.

In addition to the evidence of the impact of visual stimuli over tactile stimuli, the opposite was shown to take place as well: A tactile cue can enhance the judgments for visual targets presented near the tactile cue with respect to those presented elsewhere [114].

4.2.2 Audio-visual integration

The spatial localization of differently positioned audio and visual stimuli was investigated under varied visual conditions, namely different levels of blurriness of vision. While vision dominated over sound in good visual conditions, the opposite happened when vision is severely blurred. In mixed situations, the position was evaluated to be in a mean position between the two sources. In the case of bi-modal co-localized stimuli, localization was more accurate than with the single unimodal feedbacks [4].

4.2.3 Audio-tactile integration

The perceptual integration of auditory and tactile stimuli has been investigated for decades. Whether with constructive or detrimental effects, an interaction between the two sensory modalities was often reported.

Former studies reported a masking effect, namely an increase in the detection thresholds, in case of simultaneous supra-threshold pulsed stimuli [84]: Such effect was stronger when an intense sound was used to mask a weak vibration rather than the opposite.

Nonetheless, more recent studies investigating sinusoidal stimuli highlighted a reinforcing effect between near-threshold auditory and vibrotactile stimuli, which significantly improved their detection rate as opposed to both unimodal cases (auditory only and vibrotactile only) [241]. The conditions under which such phenomenon took place were the simultaneity of the stimuli and their similarity in frequency content. The frequency range within which the reinforcement took place was roughly 50 to 300 Hz, namely the range of activity of Pacinian corpuscles: Conversely, little integration was found to take place below 50 Hz.

In a companion study, temporal synchrony was shown to be relevant, while the relative phase of the stimuli was not [240].

The performance in frequency-discrimination tasks was shown to worsen when incongruous auditory tones were provided [177]. Conversely, the intensity of auditory stimuli was shown to be affected by congruous vibrotactile stimuli, which increased the sensation of loudness [195].

4.3 Cross-modality

The influence of one sensory modality over another has been demonstrated to occur in frequent scenarios.

Such phenomenon was hypothesized to take place in the brain as follows: In multimodal neural structures, back projections to the primary modality in unimodal brain areas may be generated; As a consequence, the primary modality is modulated by stimulation in a second modality [147].

The role of vision is usually deemed as dominant over the other senses. For instance, it was shown that visual information can alter the perception of phonemes (the McGurk's effect [151]) and even musical notes [191]. Such phenomenon was called "visual capture" [187]. Moreover, due to the higher spatial resolution of vision compared to audition and touch, the localization of auditory and tactile events was shown to be affected by visual cues [64].

However, as shown above, there are contexts into which other senses may prevail in the construction of a percept.

In addition to the abovementioned TRE example, the predominance of auditory stimuli in multi-sensory perception was shown in other contexts as well:

- Auditory signals can modulate the sensations of tactile stiffness [57] and roughness [111];
- A single flash is perceived as multiple flashes when accompanied by multiple auditory beeps [197, 198];
- The number of taps delivered to a fingertip can be modulated by the number of simultaneous auditory signals, both in the case of single [103] and multiple taps [30];
- In a famous experiment, it was shown that most participants linked the sound of the imaginary word "takete" to a drawn shape containing sharp edges, while conversely the word "maluma" was linked to an object with round features [125, 199].

Such findings represent examples of cross-modal modifications and, in particular, lead the researchers to formulate the following hypotheses:

- Vision is not always the predominant sense in multi-sensory perception. Conversely, the percept concerning a one or more visual stimuli can be altered in duration, frequency, timing, and intensity e.g. by means of adequate auditory feedback;
- The percept of a continuous stimulus in one modality can be affected by a discontinuous stimulus in another modality, and such effect is stronger than the opposite.

Cross-modality is also intended as the practice of providing the same information to the user via different sensory channels. The main goal is the message reinforcement, where it is possible for users to employ an alternative source of information when one is impaired by environmental factors (e.g. noise, vibration, or occlusion): In such case, sensory substitution [113] takes place and ideally the same information is perceived.

Such behavior requires the formulation of attributes that are amodal, namely content that can be instantiated interchangeably in two or more modalities [221].

The use of cross-modality to enhance interaction was further validated by showing the possibility to learn cross-modal icons even by training with one modality and then executing a task with another [99]. In such experiment, “Earcons” (auditory cues) and “Tactons” (equivalent vibrotactile cues) were used to enhance a message application by conveying multi-modal information about type of message, urgency, and sender. Rhythm, roughness and spatial location of stimuli were used to convey the three pieces of information respectively, each implemented in two or three different levels, resulting in a total of 18 different cross-modal messages.

4.4 Pseudo-haptics

Pseudo-haptics is “a form of illusion exploiting the brain’s capabilities and limitations” [171]. Unlike the illusions listed in Chapter 2.5, it consists in causing the arising of haptic sensations by means of stimuli conveyed through other senses, rather than substituting haptic stimuli with others of the same nature, but more convenient e.g. concerning their implementation.

Ever since its definition, pseudo-haptics has been related to the use of vision to convey the stimuli causing the illusion [132]. This leverages the aforementioned demonstration that visual cues overcome haptic cues in the event of sensory discrepancy between the two information flows [203, 132]. Nevertheless, sound [111] and even thermal sensations [208] were proven to affect haptic perception as well.

The intensity of pseudo-haptic effects varies substantially among individuals. The reasons behind such strong variability, and the strategies concerning how to manage it, are still matter of research to date.

4.4.1 Principles of pseudo-haptics

In their theoretical work, Pusch and Lécuyer [171] classified pseudo-haptic phenomena into two categories: The illusions arising when attempting to achieve a unified percept of a haptic property starting from conflicting stimuli, and the cognitive ability of users to map visual stimuli into haptic feedback. While the first category is said to be inescapable, the second requires conscious decision-making and learning. At any rate, the first category is the one that most research focuses on, for both the study of its psychophysical implications and the application of its practical aspects.

To include pseudo-haptic phenomena into a general frame, the authors adopted the Interacting Cognitive Subsystem (ICS) model [16]. Such model structures the cycle of human perception, cognition and action into four levels, starting from the sensory subsystem (where the encoding of sensory inputs takes place) up to the effector subsystem (where action takes place in response to the interpretation of information).

At all levels, prior knowledge such as former experiences can integrate or substitute the processed sensory cues. Evidently, consistency is required for the integration, or blending, to take place. Such blending process, which is affected by the different weighting of the representations (e.g. vision is preferred over haptics), generates the pseudo-haptic effects. However, excessive gaps or uncertainties among representations increases the impact of subjective interpretations, thus increasing the variability among users. Conversely, biasing users by providing hints leaning towards the desired interpretation ("user priming") may reinforce the percept.

In addition to the incoming multisensory data and the user's memory and experience, the task to be performed is conjectured to affect the final (and, in this case, illusory) percept as well.

Finally, Pusch and Lécuyer proposed a set of general steps to the proper implementation of pseudo-haptic effects, which focus on the use of vision to generate them [132, 171]:

1. Observe the user during the experience with the real haptic property that is to be simulated;

2. Identify the law, or laws, controlling a haptic property;
3. Associate such law(s) with related spatial parameters;
4. Test the suitability of each parameter to its simulation, according to perceptual and technical constraints. Avoiding too many and too strong sensory conflicts is necessary for the effect;
5. Set up a visuo-haptic sensory conflict focusing on one or more parameters;
6. Envisage complementary stimuli to compensate for the characteristics of the haptic property that are not controlled by the simulation;
7. Modify the visual feedback of the parameters according to the actual execution of the perception-action loop;
8. Harmonize the user's prior knowledge by means of deliberate user priming.

4.4.2 Survey of pseudo-haptic feedback

Lécuyer assembled a survey of the pseudo-haptic experiments conducted by his colleagues and him [132], which is summarized in this Section.

It must be stressed that, in all of such experiments, the action point was displaced with respect to the feedback point: For instance, the participants were asked to operate a mouse, or a physical surface, while the effect of their actions was displayed on a separate screen.

The illusions are organized according the generated haptic sensation.

Friction

Changes of friction were visually simulated in several experiments, resulting in effective pseudo-haptic sensations.

In one experiment, users were asked to move a cube across a virtual environment by means of a mouse [135]: While the cube crossed virtual surfaces of different colors, the movement speed was either accelerated or slowed down. This corresponds to alter the "Control/Display ratio" (C/D ratio) [133], i.e. the ratio between the displacement of the actual interface and the visual displacement of the virtual object. The result was a reported sensation of "friction" turning to "sliding" or vice versa according to the speed gradient.

A second experiment aimed at simulating the sensation of resistance to movement due to strong wind [134]: The position of the visual feedback of the user's actual hand was manipulated, thus generating the pseudo-haptic feedback.

Stiffness

The degree of stiffness of an object, or "compliance" as we defined it in Section 2.1.1, was simulated by manipulating the visual deformation of a virtual object on screen in response to the pressure exerted on a physical interface [135]. Higher visual deformation caused the object to appear softer to users. Moreover, comparing the hardness between a virtual spring and an actual spring yielded similar results to comparing two actual springs.

Another effect that was observed during such experiments was the blurring of the perception of the position of the user's thumb acting on the physical interface due to the visual displacement. Such proprioceptive illusion was interpreted in the light of the importance of co-location for multisensory integration: In fact, it was found that spatial de-location of stimuli promotes the use of the dominant sense (in this case, vision), while co-location helps the integration of multisensory information into a coherent percept [43].

Weight

By altering the C/D ratio, the perception of weight of a virtual object was altered as well: A faster movement speed induced into the participants the impression of a lighter object, and vice versa a slower movement was related to a heavier object [59]. Such finding was confirmed in a follow-up experiment involving a haptic interface: The visual amplification of the interface's motion was shown even to reverse the weight judgment concerning two objects.

Texture

Lastly, acceleration and deceleration of the mouse pointer was used to simulate the slope of a virtual relief or macroscopic texture that was displayed on screen [133]: The deceleration referred to the climbing up movement, while the acceleration referred to the consequent sliding after reaching the top of a bump or ridge.

4.4.3 Sound in pseudo-haptics

Although sound usually takes on a secondary role in pseudo-haptic feedback, it has been shown how it can either alter the haptic perception or suggest haptic properties. The first reported example of alteration is the "parchment-skin illusion" [111], which affects the perceived roughness: In such experiment, when participants rubbed their palms together, capturing the produced sound and enhancing its high frequencies (above 2 KHz) generated the sensation of drier skin in most participants. Conversely, by dampening such frequencies a sensation of increased moisture was induced. The size of the TBW (see Section 3.1) was found to be relevant for the illusion: A delay of the audio feedback by more than 100 milliseconds with respect to the tactile feedback decreased the effectiveness of the pseudo-haptic effect.

Further experimentation confirmed the effect of auditory frequency manipulation on the sensation of tactile roughness, specifically in the case of abrasive surfaces [91]. Such effect was produced by replacing the original touch-produced sound with white noise as well, whereas pure sinusoids did not bring about the illusion [216]. Conversely, the perception of length was not affected by either type of sounds.

Once again, inter-individual differences are relevant to the effectiveness of the illusion, with the consequence that slight modifications in the experimental design lead to contrasting results as opposed to those mentioned above (for a review, see [201]).

Nonetheless, a growing application of sound design concerns the enhancement of the perceived quality of industry products, from home appliances [214] to food products [63, 228], by means of manipulating the sounds originated by the operation or the interaction with such objects. Such strategies leverage the emotional impact of sound to induce the sensation of pleasantness [161, 215], and may aim at evoking haptic sensations such as crispness of food.

4.5 Applications of multimodal feedback

4.5.1 Teleoperation

Operating remote devices such as vehicles (e.g. military drones, undersea pods) or tools (surgery devices, maintenance factory tools) suffers from the limited sensory feedback and from the delays in providing it [149]. This leads to navigation problems such as underestimating the magnitude of a required steering action. Moreover, status information is often provided via the visual channel only, which may cause sensory overload.

Since restoring the natural feedback may be complicated especially regarding the state of the remote device (e.g. position, orientation, acceleration), an alternative is to provide artificial task-related feedback. Visual and haptic guidance are common types of feedback in such contexts.

Visual guidance may take the form of augmented reality [157], such as overlaying trajectory lines on the visual display, or grid lines enhancing depth perception [120]. Alternatively, it can be replaced by virtual reality such as the 3D representation of human inner organs [193].

Haptic guidance can be implemented as attractive/repulsive forces [127, 130], respectively driving the user towards a suggested path or away from an obstacle. A common implementation is force feedback applied to the steering controls, e.g. an aircraft's control stick.

Passive guidance forces, or "virtual fixtures" [189], are common as well. They are designed to limit the range or classes of motion, and can be either "hard", such as virtual walls [189], or "soft" [23], such as reproducing the resistance experienced when navigating from a material with low density to another with high density.

4.5.2 Multitasking

Multimodal feedback may support task performance and quality in multiple concurrent tasks.

Everyday activities commonly involve the execution of more tasks at a time. Such contexts vary from harmless and undemanding, such as drinking a beverage while watching TV, to complex and dangerous, such as operating multiple pieces of heavy machinery at a time.

Especially in multitask scenarios, the workload is an important factor for performance. To reduce competition for visual resources, task-relevant data may be communicated via non-visual channels, such as warnings in the form of haptic vibrations e.g. in driving [204].

More information can be conveyed by structuring the feedback in the form of messages, i.e. by forming codes (e.g. the abovementioned tactons). Such codes may be based on spatial information e.g. stimuli presented at different locations, or non spatial, such as temporal e.g. stimuli varying in rhythm. Anyhow, decoding such messages may imply an effort resulting in competition for cognitive resources, namely the working memory. Moreover, redundancy may be introduced in that a single piece of information may be conveyed not only by means of different sensory channels, but by means of different codes as well (e.g. both as the location of a vibration, and as a vibration rate). In such case, the gain generated by a more efficient workload distribution may be canceled by the cost in decoding the messages: Redundancy gain and cost [238] must be evaluated.

Nonetheless, Ardoin et al. [9] showed potential avantages of multi-code redundancy at least for elementary tasks such as image selection and stimulus localization. Conversely, redundancy cost may prevail in more complex tasks.

Multitasking in general must be tuned to the humans' innate capabilities of parallel sensorimotor operation and information processing, and to their limitations: The difficulty and the number of tasks affect the ability of performing parallel tasks [29].

Hypothetically, each multitasking scenario places itself between two extreme situations: completely parallel multitasking and interleaved sequential multitasking [192]. On one extreme, the divided attention was identified as the cause for possible drop in productivity and job quality [146], when not a potential source for harm or damage. On the other, the cost of switching and resuming tasks was deemed to increase the mental workload [7].

Multimodality can be used to inform the user about the state of the overall operating scenario ("system/situation awareness", or SA): Multimodal cues were proven to reduce SA decrement due to unexpected state changes better than unimodal cues, e.g. icons or earcons [112]. Another benefit of the increased SA is to avoid cognitive tunneling, namely the tendency to focus excessively on one task while neglecting the others.

Resuming a task after switching from another requires recalling its context, and rehearsing the mental representation of the related problem. Such activity implies a cost in the form of cognitive effort, named "retention overhead", which can be mitigated by a multimodal representation of the problem and the context.

The impact of multimodal feedback in the execution of a progressively increasing number of concurrent tasks was investigated [117], showing that indeed multimodal feedback improves the performance with four or more concurrent tasks as opposed to unimodal feedback.

4.5.3 Entertainment

In entertainment contexts, multimodality enables an improvement of the immersiveness and of the perceived quality of an experience. Cinema theaters equipped with vibrating or even moving seats have a long story, while their employment in gaming scenarios is more recent [106]. Multiple vibrating actuators embedded in the seat enable the implementation of haptic illusions such as movement illusion, localization illusion, and sensory saltation (see Paragraph 2.5.2). The addition of haptic feedback to attractions in theme parks is the reason behind a part of the aforementioned research concerning innovative haptic technologies [17, 200].

Experiments in live concert enjoyment focus on providing the listeners with body vibrations that are coherent with the perceived sounds [155]. Such vibrations have been shown to improve engagement even when not openly acknowledged by the audience. Moreover, vibrations can be used as an independent channel in multisensory art installations where visuals, sound and haptics convey different messages [86]. Such implementations make use of haptically-enhanced suits for the audience, and stem from the pioneering research in composing concerts for the haptic sense in the form of musically structured patterns of vibration [92].

4.5.4 Music performance

Much research has been devoted to the application of multimodality in music production and performance. The experience of playing a traditional instrument involves a whole-body interaction, in which vision supports not only the execution of a music piece but the perceived instrument quality as well [79]. Likewise, the tactile vibrations transmitted to the player's hands and body were proven to be relevant for the performance experience [75].

The adoption of digital music instruments, either in the form of customized hardware or in the form of software used by means of general-purpose computer interfaces, typically impoverishes the multisensory experience of music performance. Such new instruments usually lack the tactile response of the materials a traditional instrument is made of (e.g. wood, brass etc., which are substituted by glass and plastic), as well as the vibrations generated by the excited parts (e.g. strings, skins etc.) and amplified by the resonant cavities of an instrument.

Ways to improve the interaction by means of visual and tactile feedback have been explored for decades, as corroborated by conferences such as New Interfaces for Music Expression[1] and Sound and Music Computing[2]. Moreover, multimodality has been investigated as a means to design innovative ways for music creation, for instance by Cadoz et al. [32], or to create virtual music instruments (see [140] for a recent overview).

---

[1] http://www.nime.org

[2] http://smcnetwork.org

Chapter 5

Experiment 1 - Multisensory texture exploration

A task that we perform every day, almost unknowingly, is to sense the surface of the objects around us. Such experience can be obtained by feeling the objects with our bare fingers as well as by using tools. Moreover, vision and sound contribute to form our percept as well. Nonetheless, such experience is generally unavailable in the interaction with virtual objects displayed on digital devices.

An experiment was conducted [52] to assess qualitatively the possibility of rendering the texture exploration experience in a tool-mediated context by simulating the expected multisensory (visual, auditory, and vibrotactile) feedback.

By using the "Sketch a Scratch" experimental platform [55], the experiment enabled participants to explore by means of a stylus different images displayed on a tablet screen, representing different surfaces and materials, and to experience them visually, aurally and haptically (see Figure 5.1).

This chapter is organized as follows: First, we overview the motivations for this experiment and the previous research concerning multisensory texture exploration. Then, we present the "Sketch a Scratch" framework and its uses so far. Then, we describe the experimental setup and setting. A summary of the participants' reactions and the following discussion precede the conclusions drawn from the experiment.

5.1 Motivation

A relevant part of our knowledge about the surrounding environment runs through our fingertips, in the form of a tactile experience of the surface of

Figure 5.1: Multisensory tool-mediated exploration of virtual textures.

the objects we interact with. The quality of a fabric, the finish of a piece of furniture, are qualities that we assess through such experience. Different actions, such as rubbing or scraping, can be carried out to achieve such knowledge, and tools can be employed as well. In fact, tool-mediated exploration of textures and materials is almost as common as direct touch, in that writing or different types or figurative art are deeply affected by the tactile feedback that we receive due to the qualities of the tool, of the surface, and of the ongoing interaction between them. To use a pencil on an overly rough sheet of paper may generate unpleasant sensations as well as complicate the task.

Indeed, the exploration of a surface is a multisensory experience: As we may anticipate, whether correctly or not, how a texture will feel by looking at it before actually touching it, the sound that is produced by the interaction - especially when using a tool - completes the percept as well. Product designers and advertisement firms employ such evidence to enhance the perceived quality of a product, or to convey the tactile features of an object remotely, e.g. through a television ad.

The interaction with digital devices is mostly deprived of such pervasive, deeply informative, and sometimes affective experience: Regardless of the effort put by engineers in creating interfaces with a pleasant feel, such feature is obviously permanent in their constituting material and its finish, thus it cannot be altered to simulate different textures. Yet, to reproduce such sensations is important in virtual reality and tele-operation settings.

As shown in Chapter 2, two separate mechanisms are responsible of texture perception, according to the grain of the surface. Yet, in tool-mediated exploration, the vibrations conveyed to the hand are primarily sensed by the Pacinian endings.

"Sketch a Scratch" is an experimental platform enabling the multisensory exploration of surface textures, consisting of vibrotactile-augmented tablet and stylus, an audio system, and a set of software modules for the generation and control of the feedback. For the sake of conciseness, the experimental tool presented here will be addressed as "Sketch a Scratch" as well, in that it represents one of the variants (see Section 5.3.2) of the same platform.

For a set of different textures displayed on a screen, auditory and vibrotactile feedback was recreated by modeling the response of the particular material and surface to different types of exploring actions. A physical modeling of the materials was employed. Aspects such as the interaction style and the qualities of real materials in terms of force dissipation were taken into account, as well as the peculiarities of the particular surface being displayed.

In addition to that, the possibility of designing the features of a surface starting from a sound was investigated. Participants were enabled to experiment vocal sketching (see Section 3.5) as a tool for rendering the acoustical outcome of a surface exploration, which is closely related to the surface's actual features.

Investigating the multisensory exploration of a virtual surface can lead to a better understanding of how vision, sound, and touch integrate in the forming of similar experiences in the real world. Different aspects of complex physical phenomena are rendered through such channels. While some of these aspects may be impossible or impractical to render accurately in a digital environment, their modeling helps investigating how they might possibly be replaced or imitated by means of other sorts of stimuli. For instance, some haptic sensations such as the lateral forces [181] are usually hard to convey through common devices: Thus, their rendition requires alternative solutions, often resorting to the fields of haptic illusions and/or pseudohaptics.

5.2 Background

5.2.1 Visual simulation

Haptic feedback is important to convey a plausible experience of a texture in virtual and augmented environments [178]. However, force-feedback devices are usually expensive or of impractical use in many contexts. As a consequence, pseudohaptics (see Section 4.4) is widely employed, in that it does not require additional hardware to be implemented. Besides, as we mentioned, the visual information is proven to effectively replace haptic information in many settings.

A common way to employ visual feedback to evoke haptic illusions is to manipulate the control/display ratio (see Section 4.4.2): The rate between the movement given to the hardware controller (e.g. a computer mouse) and the on-screen cursor, which is usually fixed, can be temporarily increased to convey a sensation of “stickiness”, that is an increase of friction between surface and cursor which obstacles the movement [116, 244].

An evolution of such approach is to enable the independent movement of the cursor, as if it was placed on a slope and therefore had the tendency to move in a certain direction. Mensvoort et al. [152] compared the visual simulation of macroscopic surface features such as bumps and holes with actual mechanical feedback. A mouse-controlled on-screen cursor was driven in specific directions by virtual force fields representing the descending slope of the simulated surface feature. They observed the effectiveness of the illusion, and the fact that coherent visual illusion and actual force feedback are additive in conveying the effect. In the case of contrasting behavior between visual and haptics, the evaluation varied considerably among the participants, as each weighted the two feedback modalities differently.

5.2.2 Haptic rendering

Direct touch and tool-mediated surface exploration convey different sensations, and consequently different sets of information: While touching a surface with the bare hand gives a spatial, intensive measure of roughness, the use of a tool carries information about roughness, hardness and friction in the form of a multidimensional signal in the time variable. In both settings, the co-location of stimulus (the user’s touch) and response (the feedback) is important to convey the sensation correctly (see Section 4.1.2), in that it recalls the context of many traditional manual activities that afford the development of skills through practice, such as painting or drawing.

Mechanical as well as alternative implementations of haptic feedback are summarized in Section 2.6. The haptic sensation of a surface can be provided by actuating either the interactive surface [80] or the probe [137]. Common probes are styluses for touch graphic tablets. Actuating the surface enables the use of direct touch, yet presents several technical drawbacks: Especially with large surfaces, the feedback is not homogeneous across the active area due to the flexibility of the interface. Moreover, since in traditional mechanically-actuated surfaces the motors are located on the sides of the surface, co-location is hard to simulate. Alternative technologies, such as ultrasonic waves [236], are in the prototypical phase to date: At the time of the experimentation, reliability issues undermined the possibility to employ such technologies in extensive experimentation, and prototype screens measured only a few centimeters in size. Conversely, in tool-mediated surface exploration, the possibility of actuating the probe presents some advantages: By embedding the actuator in the probe, the intensity of feedback is independent from the size and the geometry of the surface. Moreover, the round, plastic tip of a stylus exerts less friction than the finger tip on the glass surface of a touch screen: Starting from a more "neutral" actual friction condition, the implementations of haptic feedback have a wider range of levels of friction at their disposal to be simulated.

5.2.3 Auditory feedback and physical modeling

Exploring a surface by means of a tool often generates an audible signal. Such signal carries information about surface roughness, hardness, and friction [123] as well as about the type of action that is being carried on the surface (see Section 3.1). The inverse process is possible as well: A sound signal can be interpreted as a surface profile, and consequently be appreciated with other senses. Thus, the qualities of a surface can be evoked by means of synthesized or acquired sound signals.

Actions such as rubbing, scraping, or rolling can be described by microscopic contact events occurring between the probe and the surface. Such events can be simulated by the physical modeling of impact and friction phenomena. In sound synthesis, several models can be used to replicate such phenomena: One is friction, which is based on stick-slip commutation [13], namely the pattern of sudden jerks due to the force applied to the probe overcoming the surface friction. Other phenomena, such as rolling, are rendered by patterns of impacts [175]. In such types of modeling, surfaces are often specified as one-dimensional height profiles, either sampled (see Figure 5.2) or algorithmically generated.

Besides the type of action characterizing the exploration, the variables that must be taken into account in such models refer to both surface and bulk properties of a material (see Section 2.1.1), such as the local geometry of the surface around the point of contact between probe and surface on the one hand, and local energy dissipation on the other.

Several physical models were introduced in Section 3.3.2, such as those by Conan et al. [41], and the Sound Design Toolkit (SDT) [15].

Concerning the visual representation of a surface, its exploration has its salient points in the regions of maximal change, e.g. in brightness or color. Such landmarks may represent actual discontinuities such as ridges, as well as the evenly-colored parts of the image may represent flat regions. Therefore,

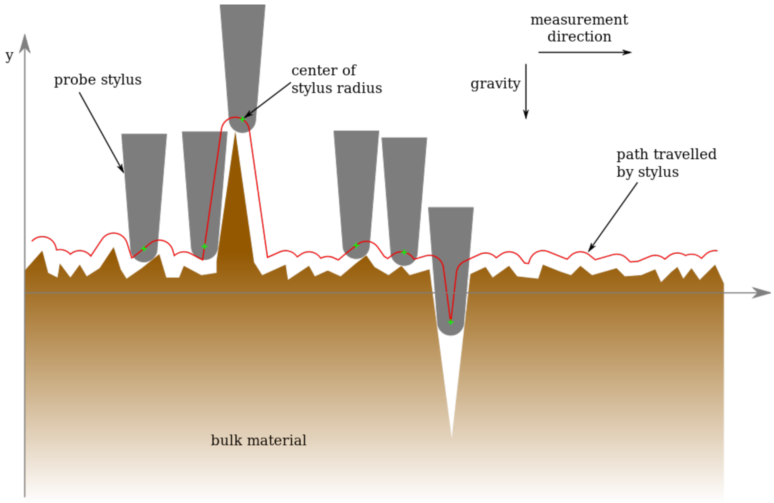


Figure 5.2: A one-dimensional height profile (in red), as a result of a linear scan by a probe.

by means of image analysis it is possible to extract visual cues that represent the salient regions of a surface, and by parametrizing their features it is possible to drive other feedback forms whose intensity is related to such features (e.g. the loudness of a scraping sound, or the intensity of a vibratory impulse). In summary, the representation of any kind of input as a visual surface enables its later multisensory exploration.

5.3 The “Sketch a Scratch” framework

5.3.1 Concept

“Sketch a Scratch” is an abstract experimental workbench conceived in a “research-through-design” perspective. Such research approach aims at gaining insight over human-computer interaction issues by designing loosely-constrained devices and interactions [247]. In this case, the topics that are to be explored are (1) the sonic sketching of surface qualities, (2) creative texture modeling and multimodal exploration, and (3) the exploration of auditory contents rendered by means of auditory, visual and tactile feedback.

In the present experiment, the focus was put on the contributions of visual, auditory, and haptic feedback in the probe-mediated exploration of surface textures: The goal is to reconstruct the surface exploration experience by exploiting the SDT as a basis for modeling the reaction of different materials to probe contact. The interaction takes place over an interactive surface, namely the touch screen of a graphic tablet, on which a digitized texture is displayed. The user runs the tip of a vibrotactile-augmented stylus on the screen. A physical sound model of a real material (e.g. wood, glass, dry soil) is driven by the exploration of the different features of the surface such as even areas, bumps, creases and ridges. A sound output is consequently generated in real-time, as well as vibrations for the stylus. A local visual deformation helps keeping track of the contact position between probe and screen, and completes the multisensory experience. An overview of the whole system is depicted in Figure 5.3.

The interaction with the system is structured in two phases: First, a two-dimensional image is generated starting from several alternative sources (see below). Such phase is defined as "Sketch" in that it enables the user to define arbitrarily the qualities of a surface by controlling the process of generation of the image. The second phase consists in the exploration of the surface ("Scratch"): An analysis of the image, in conjunction with the physical modeling of the virtual material, defines the response to different styles of interactions, e.g. rubbing or scraping. Such response is primarily an audio signal, which is then used to generate vibrotactile feedback as well.

The qualities of a surface can be sketched through several alternative representations of a texture:

- A digital image, whose regions of maximal change of grey-level are interpreted as depth shifts such as bumps and ridges;

- An audio signal, e.g. a vocal recording, whose features can be preliminarily converted into a visual mesh and finally interpreted as a map;

- A vibration, which can be generated by scanning a real surface with a probe to acquire its linear profile.

The system affords various types of contact (styles of interaction): scraping, rubbing and rolling, obtained by specific combinations of an impact and a friction model.

The sensory cues are meant to be synchronous and coherent: A vibrating motor attached close to the stylus tip co-locates the haptic feedback at the point of contact with the surface, and auditory feedback is presented as close as possible to the interactive surface. Besides, intensity and frequency of sonic and vibratory impulse are generated proportionally with the gradient of gray-levels in the area of the image that is being crossed by the stylus.

Indeed, the choice of vibration inherently poses limitations to the conveyable tactile sensations. The context of distal interaction and vibratory stimuli is likely to cause the activation of the Pacinian corpuscles exclusively (see Section 2.2.2). Being the vibration responsible for the sensation of fine textures, medium and coarse textures are presumably hard to reproduce. In general, lateral forces such as those experienced when crossing a bump or a crease on a surface are here substituted by vibration: The effectiveness of such sensory illusion is to be tested, and it is likely to vary among subjects nonetheless.

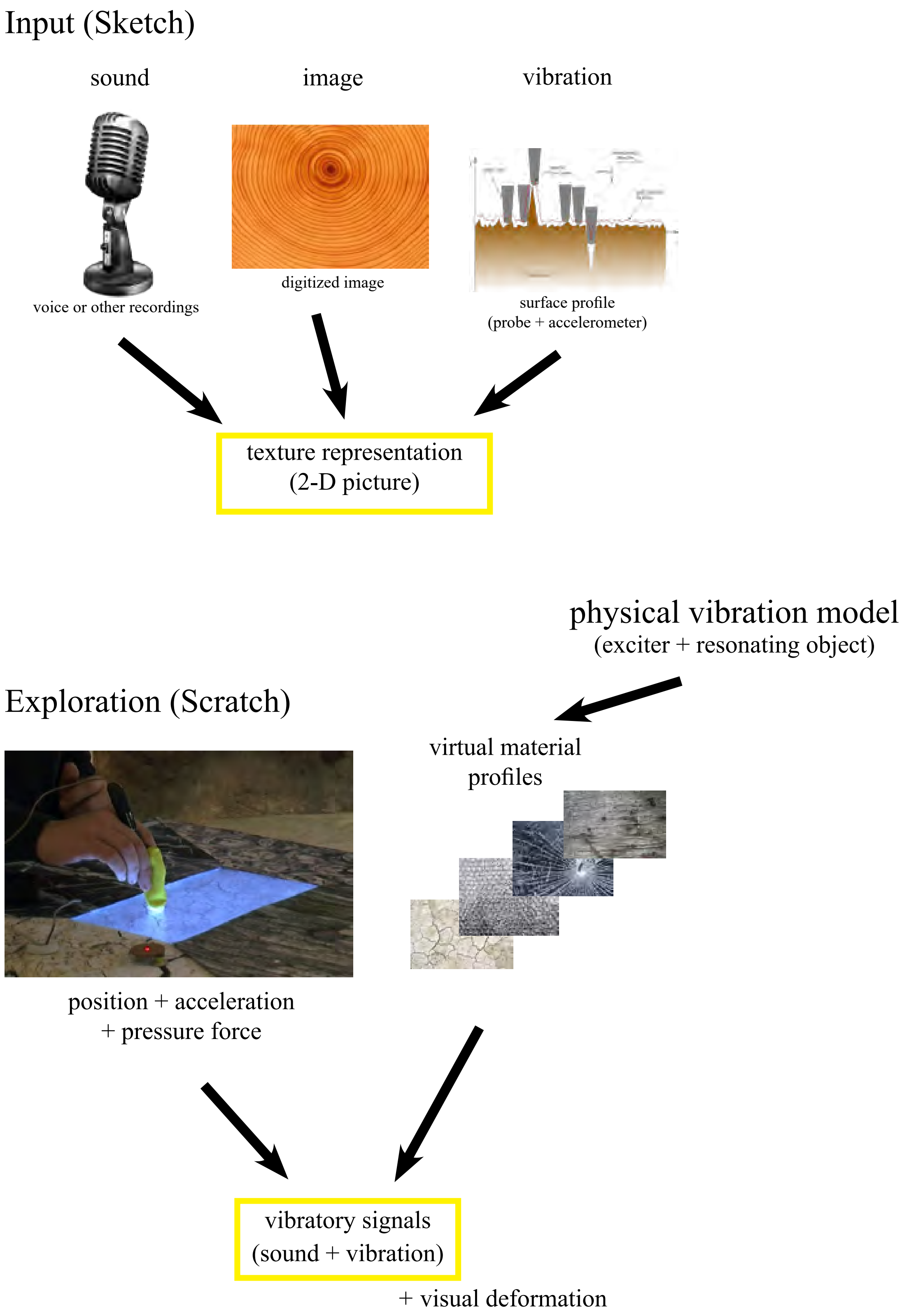


Figure 5.3: Sketch a Scratch concept. The interaction is divided in an input phase (“Sketch”) and an exploration phase (”Scratch”).

5.3.2 Previous demonstrations and experiments

The basic configuration served as workbench to investigate the potential of Sketch a Scratch in different contexts of use, and for a variety of purposes: demonstrations, experimental research, live performances and installations.

Demonstrations

Extensive demonstrations [55] enabled the collection of several comments from naïve users as well as from persons with specific drawing or vocal skills. The alteration of the feeling of the surface induced by the addition of sound and vibration was largely assessed. The perception of the probe shape was affected as well. An enrichment of the experience was generally reported, with an increase of engagement in activities such as drawing. In particular, some illustrators pointed out the lack of the rich sensory experience of drawing with different pens and pencils on various paper materials that they are used to experience when drawing on a tablet.

Performance

The potential of Sketch a Scratch as a tool for artistic performance was experimented in the occasion of the 2014 World Voice Day[1]. Two exemplars of Sketch a Scratch were played by a quartet: One vocalist provided vocal textures that were cyclically explored while the impact and friction parameters were dynamically manipulated by another performer. At the same time, one drawer acted with the stylus on the tablet to explore four different kinds of material textures, each corresponding to one movement of the piece, under direction of a fourth laptop performer.

Similarly to the context of musical performances with tangible user interfaces [110], the performers' body movements, the aural result, and the visual projection of the interface enable the transfer of the localized sensations (the vibrotactile feedback) to the audience.

Video footage can be found at the following link[2].

Experimentation

Sketch a Scratch was used as an experimental apparatus [181]. Participants were asked to follow a curvilinear path from one side of the tablet screen to the other, as shown in Figure 5.4, in what is called a “path following” task, or a “steering” task [1]. The path was represented by a set of bars (placed

---

[1] http://en.wikipedia.org/wiki/World_Voice_Day

[2] https://vimeo.com/93417532

in normal direction with respect to the path direction, like railway ties), and could be either visualized on screen or sensed by means of sound or vibration occurring when the stylus tip hit the bars, or by any combination of these three types of feedback. Additionally, a local visual distortion of the bars was implemented, taking place parallel to the surface to mimic the deformation of a membrane being pushed by the stylus.

The purpose was to find evidence of the effectiveness of image, sound and vibration as sensory substitutes of lateral forces, in this case represented by the friction of the probe against the path’s bars. More generally, the tool was used to investigate how different feedback modalities affect constrained exploratory gestures.

Such experiment inspired further experimentation concerning path following, which is reported in the next Chapter: Starting from a similar setting, the focus was set on the effectiveness of non-visual feedback for completing a path following task.

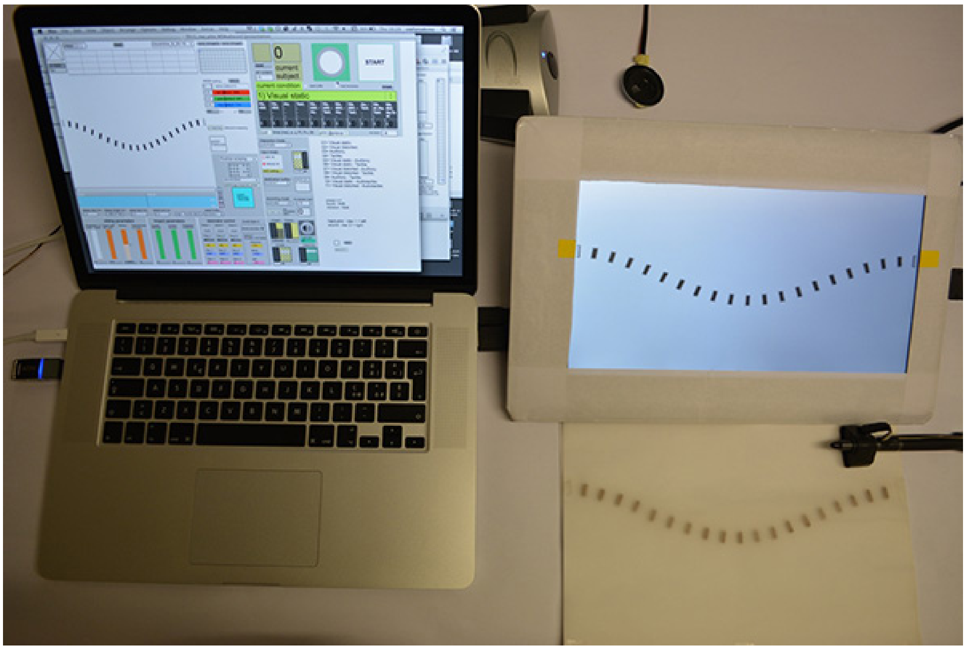


Figure 5.4: Experimentation in path following.

5.3.3 Physical modeling

The main model employed in the simulation describes the impact between two colliding bodies [12], a point-mass (exciter) and a resonating object.

The contact force f_i is a function of the generated object compression x and compression velocity $\dot{x}$ computed as follows:

$$f_i(x, \dot{x}) = \begin{cases} -kx^\alpha - \lambda x^\alpha \dot{x}\,, & x > 0 \\ 0\,, & x \leq 0 \end{cases} \tag{5.1}$$

where k accounts for the object stiffness, λ represents the force dissipation, and α describes the local geometry around the contact surface. $x \leq 0$ indicate the lack of contact.

In addition to the impact model, a friction model [13] describes the relationship between the relative tangential velocity v of two bodies in contact, and the produced friction force f_f. In this case, the exciter is the "rubbing" object, while the resonator is the "rubbed" object. The hypothesis behind the model is that friction results from a number of microscopic elastic bristles, accounting for stick-slip phenomena:

$$f_f(z, \dot{z}, v, w) = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v + \sigma_3 w \tag{5.2}$$

where z is the average bristle deflection, $\dot{z}$ the average bristle deflection velocity, the coefficient σ_0 is the bristle stiffness, σ_1 is the bristle damping, and the term $\sigma_2 v$ accounts for linear viscous friction. $\sigma_3 w$ is a noise component representing surface irregularities. In particular, the variable z describes the three regimes of friction that are simulated by the model:

- "Elastic": The rubbed object is fixed and does not vibrate, while the rubbing object moves tangentially;
- "Elasto-plastic": The rubbed object vibrates, while the rubbing object moves tangentially;
- "Plastic": The rubbed object does not vibrate and is dragged by the rubbing one.

The two models are used in conjunction to simulate complex vibratory phenomena. The combination of f_i and f_f is weighted according to the material to be simulated. A simulated surface profile is used in the impact model to modulate the relative displacement offset between the exciter and the resonating object (i.e. the stylus and the surface). The normal force applied to the stylus is also used to feed the impact model. In addition, when driven by the stylus' tangential motion and the normal reaction force f_i produced by the simulated micro-impacts, the friction model generates stick-slip phenomena.

In more detail, considering a single impact and a small portion of surface profile having slope δ around the contact point, the impact force f_i is returned along the direction normal to such slanted surface. Its horizontal component $f_i \sin \delta$ is derived and used to drive the sliding force parameter of the friction model [150] and, similarly, the vertical component $f_i \cos \delta$ can be used in the friction model to control the normal pressure on the surface.

5.4 Experimental design

The experiment here described aimed at evaluating the effectiveness of a combination of sensory illusions and pseudohaptics, provided via the Sketch a Scratch framework, in rendering the multisensory experience of the exploration of different textures and materials. The hypothesis was that such combination would have been able to induce convincing percepts in the participants, especially when compared to their everyday experience with interactive digital surfaces.

The impact and friction models produce vibratory signals, which can be used to generate sound (the sonic result of a texture exploration) as well as to drive a vibration transducer, or even a haptic device. In Sketch a Scratch, similarly to [150], a vibrotactile transducer attached to the stylus is driven by the low-frequency components of the synthesized sound. In addition, a local image deformation is applied at the point of interaction to mimic superficial vertical and lateral forces exerted by the stylus. Such deformation is meant both to enhance the impression of interacting with a physical material, and to compensate for the visual occlusion caused by the stylus tip.

In the experiment, the interaction is divided in an input phase ("Sketch"), where an image is either loaded or created from an audio signal - here a vocal excerpt -, and an exploration phase ("Scratch"), where the user explores the image displayed on the tablet screen by means of the stylus (see Figure 5.3). Between the two phases, the parameters of the physical model concerning the material properties and the interaction style can be manipulated.

5.4.1 Apparatus and functions

The main component of the system is a Max patch which is based on the models provided by the SDT; such patch implements the experiment management as well. It runs on a Apple laptop, to which a 13.3" Wacom Cintiq graphic tablet (1920 × 1080 pixels) and its stylus are connected. A Tactile-Labs Haptuator Mark II vibrotactile transducer is attached to the stylus, close to its tip. A Sonic Impact T-Amp amplifier drives both the transducer and a dynamic speaker, which is placed below the tablet for a localized emission of sound and vibration. A portable digital audio recorder enables the audio signal acquisition (in this case, the vocal sketch).

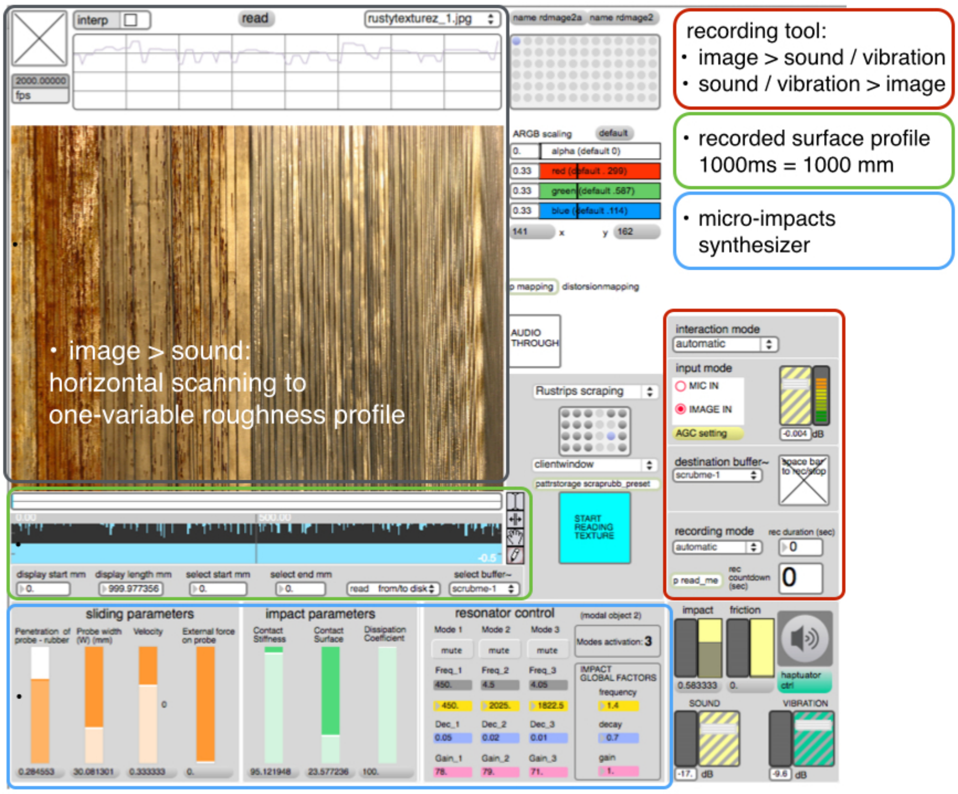


Figure 5.5: Sketch a Scratch GUI.

Through the graphical user interface (see Figure 5.5) the operator can load images of surface textures, record audio tracks and convert them into surface profiles, manipulate the parameters of the physical model, save and recall presets concerning different virtual materials and interactions.

Different kinds of virtual materials (e.g. glass-like, metallic, wooden) and interactions (e.g. bouncy, sticky) can be synthesized and saved as presets. Up to six different roughness profiles can be recorded as audio signals and recalled, to drive the synthesis engine. The “impact parameters” describe the quality of the single collision (stiffness, sharpness, and energy dissipation affecting the occurrence of bouncing phenomena). The “sliding parameter” layer is used to interpret the stored surface profile and drive the impact model accordingly. The vertical penetration of the probe sets the threshold level of the roughness profile above which the signal is detected, while the probe width parameter sets the size of the sliding window on the roughness profile (in mm, large = rubber, small = sharp object). The probe is advanced every Δt ms by a distance $\Delta x = v\Delta t$, where v is the sliding velocity in m/s.

Additional parameters are Δt in ms and the diameter of a single contact area in cm.

As a result, the profiles can be explored with virtual probes of different characteristics, to simulate scraping, and rubbing. In particular, the tilt of the stylus is used to virtually change the width of the probe, thus shifting the interaction style from scratching (stylus perpendicular to the screen) to rubbing (maximum tilt of the stylus). Theoretically, such feature may be used expressively in the context of artistic performative acts. Furthermore, the vertical force relative to the stylus' tip on the screen is used as a control of the vertical penetration of the probe on the virtual surface profile.

Single-variable audio or vibrotactile signals can be used to produce an image in different ways. One basic, yet effective transformation used in our tool is the stacking of luminance-translated audio signals to produce rows of pixels (see Figure 5.6). Such sound-to-image transformation affords different kinds of subsequent image-based exploration of the sound material (e.g. temporal expansion, inversion, interlacing, etc.).

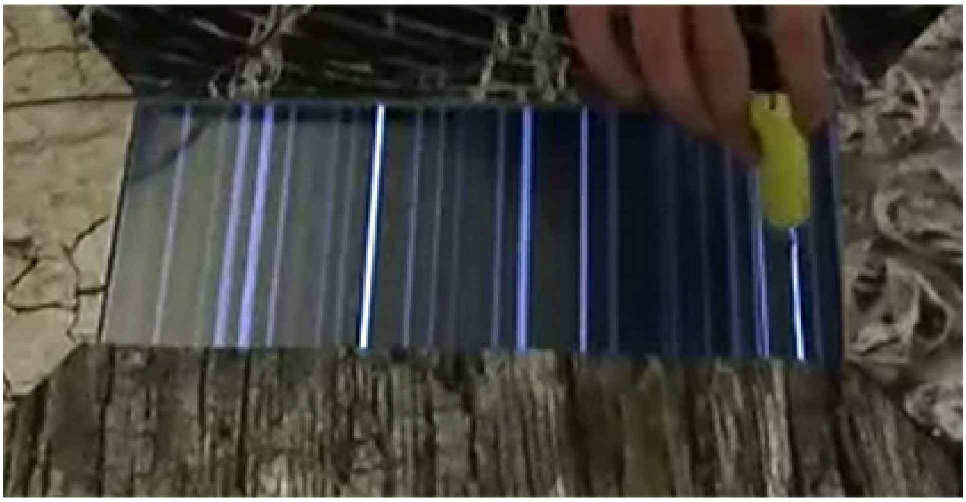
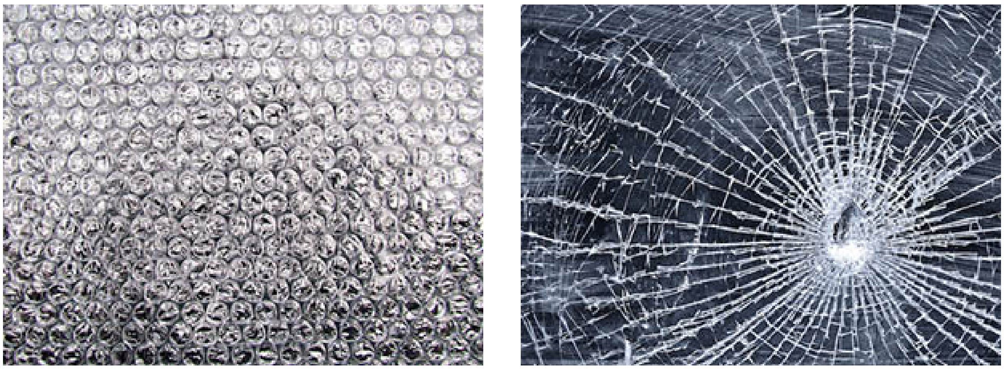


Figure 5.6: Exploration of an image generated from a vocal sketch.

5.4.2 Setup and procedure

The experimental setup took the form of a publicly-accessible installation. Participants were able to use the vibrotactile-augmented stylus to virtually scratch and scrape on four different surface textures (bubble wrap, broken glass, wooden board, and dry soil, see Figure 5.7) displayed on the screen of the tablet. They were prompted to explore the surface features such as bumps, ridges, and creases, while receiving a vibro-acoustic feedback coherent with the material characteristics of the 2D image displayed on the screen (e.g. plastic, wood, glass). In alternative to the exploration of pre-defined textures, participants were enabled to sketch a texture by means of vocal imitation: They could record short vocal excerpts to be converted to visual profiles, thus enabling their tactile exploration as shown in Figure 5.6.

Although the installation mainly focused on the multisensory rendition of actual physical textures, the possibility of vocal input aimed to illustrate the connection between the sound generated by the texture exploration and its tactile counterpart, and to show how the former might be used to design (or rather sketch) the latter. For such purpose, involving the voice was meant to provide the participants with an immediate tool to experiment such connection, with no former skills required.



(a) Bubble wrap. (b) Broken glass.

(c) Wood. (d) Dry soil.

Figure 5.7: Macros of the four textures available for exploration.

The main experimentation took place during an academic celebration: The occasional visitors served as participants, obviously naïve to the nature and purpose of the experiment.

Such setting required the apparatus to be appropriately encased for safety and aesthetic reasons: The final result was a self-contained interactive installation[3] in which the system was contained in a $0.5 \times 0.5 \times 1.15$ m multilayer cardboard parallelepiped (see Figure 5.8).

The tablet was embedded on the top side of the box. On a shelf just below it, two small loudspeakers were placed to achieve the co-location of visual, auditory and vibrotactile feedback. Lower shelves hosted the laptop carrying the computations, the amplifier, and the image switching system.

(a) External appearance.

(b) Internal slots.

Figure 5.8: Sketch a Scratch installation. On the right, details of the hardware embedded in the box.

Further adaptations due to the setting were an exaggeration of both auditory and vibrotactile feedback. The former was due to the background noise, and was operated by adding an active speaker placed on the bottom shelf of the box. As a drawback, the frequency response of such loudspeaker and the resonance of the box caused the friction to sound darker than what one would naturally expect from a real-world situation. The latter was due to the quite short time of stay per visitor, and aimed at a clear and prompt perception of the vibrotactile feedback by the visitors. For this reason, the intensity of vibration was increased in comparison to what was deemed as adequate in a lab environment.

---

[3] https://vimeo.com/111889017

As shown in Figure 5.8a, each side of the parallelepiped is covered with a print of a macro-image of a texture surface. The four textures were chosen in order to elicit diverse interactive experiences, and possibly prompt different responses and interaction styles. In addition, the impact and friction parameters were specifically adjusted for strengthening the expectations and interaction with the materials displayed.

The stylus was camouflaged to decrease its similarity to a pen, which might bias the interaction style. Moreover, the participants were prompted to hold the stylus between their index and middle finger, to avoid the metaphor of writing and facilitate the full experience of touch.

The participants could browse the available textures at will by positioning a token on one of the four switches located on the table top, each associated to one side of the shell. The switches were implemented as simple open circuits painted on paper with conductive ink[4].

Finally, users could record their voice by approaching a clearly-visible digital audio recorder, and engage in direct explorations of their sketches.

5.5 Experiment and discussion

Information was collected via different channels: Video recordings, direct observations, talking-aloud impressions and post hoc comments by the participants, especially regarding their own expectations. The goal was the evaluation of the effectiveness of the simulation, with the outlook of revising the system for improvement and envisioning new creative and functional scenarios.

We filmed the most engaged participants (12, 7 female, mean age 30 approximately), namely those who lingered enough to acquire a basic understanding of the system and of its features. Several remarks were extracted. For instance, a professor of modern art history suggested the application of the Sketch a Scratch framework to enhance the navigation experience in art galleries for the visually impaired.

Different interaction styles were employed (see Figure 5.9), from the regular, neat stroke of a painter to the irregular touch intensity shown by non-trained individuals.

In general, the installation was positively received. A variety of personal and creative exploration styles was elicited by the direct observation of the participants. For instance, the visual deformation localized at the contact point of the stylus with the screen prompted many participants to challenge

---

[4] The conductive ink and the magnetic led component are part of the Circuit Scribe system: http://www.123dapp.com/circuitscribe

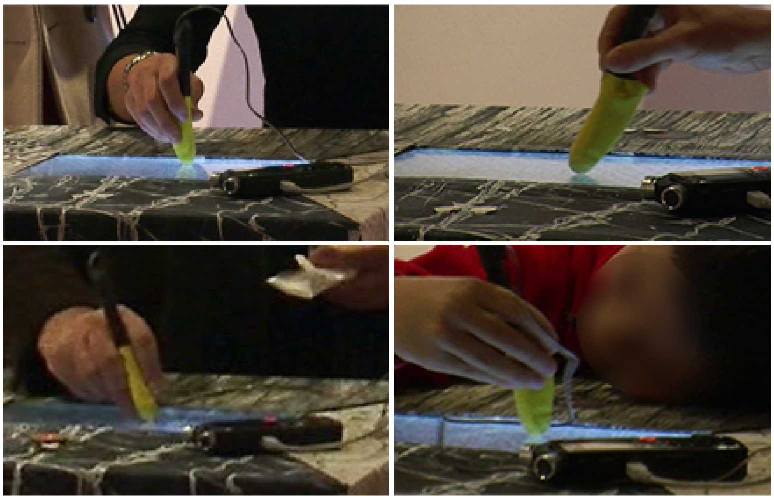


Figure 5.9: Different styles of interaction employed by visitors. From top left, clockwise: vertical popping, painter-like slanted stroking, quick scribbling, slow crossing of the texture's features.

its limits by means of a "popping" gesture (i.e., a vertical bouncing movement of the stylus resulting in short, discrete impacts). Such behavior was especially prompted by the bubble wrap texture: The participants expected to receive the auditory response of the actual material to such interaction (a popping noise, in fact), and eventually deemed the simulation as accurate, resulting in a fun experience.

Conversely, the visual deformation was barely noticed by other participants, for which the images were more appreciated as a navigation guidance than as a feedback source.

Other participants focused on evaluating the responsiveness and the fidelity of the feedback, e.g. by slowly crossing the texture features such as the cracks on the glass. Some minor latencies were reported, which can be attributed to the system as well as to the size of the stylus tip: Indeed, a large tip was chosen to decrease the friction on the display, at the cost decreasing the accuracy in the impact detection due to possible aliasing in the contact point.

Among the three sensory feedbacks, the vibrotactile feedback generated the strongest impressions: Although warned beforehand, participants had a generally surprised reaction at its appearance. Nonetheless, it was assessed as the most effective.

The auditory feedback was positively received as well. Residual inaccuracies in the auditory response were not deemed as important, thus suggesting that users were more focused on vision and touch than on hearing. As a consequence, sound represented more of a tool for augmenting the immersiveness of the experience rather than a guide for the exploration.

The audio sketching mode was received with milder interest. The idea behind it, admittedly not immediate, was hard to grasp for the participants until a practical demonstration was provided. Moreover, the participants showed a general shyness in attempting the creative use of their voice in public. Nevertheless, after a brief demonstration and explanation, they generally gained a clear understanding of the causality between the vocal gesture an its visual rendition. For instance, the visual impression of peculiar vocal inputs such as sustained sounds, rhythmic patterns, or trills, was easily spotted within the displayed image.

5.6 Conclusions

The present experiment investigated the possibility of rendering the experience of the exploration of a virtual texture by means of multisensory feedback based on physical modeling.

Although qualitative in its form, the experiment allowed for several considerations about the effectiveness and the limitations of such simulation. Other related topics were brought to the fore, such as the perceptual and cognitive aspects involved in such experience, and the potential of this kind of interaction for design purposes and performative uses, and require further investigation.

Participants generally evaluated the experience as positive and accurate: The physically-informed approach to sound synthesis and the pseudohaptic use of vision and sound resulted effective in conveying the salient aspects of contact phenomena such as scraping and rubbing. Moreover, coherent multisensory stimuli certainly increase the naturalness in the interactions with virtual surfaces, resulting in a higher expressiveness during creative efforts.

Nonetheless, it must be stressed that the presented vibrotactile feedback is inherently incapable of rendering the lateral forces. As a consequence, the sensory substitution (normal vibrations instead of lateral forces) is likely to show its limits in more critical applications, such as professional settings (e.g. industrial, or medical).

Another limitation consists in the general preference by the participants in employing the visual feedback as the main guidance for the navigation. Such behavior was confirmed by [181] as well.

On these premises, a possible investigation concerns the addition of actual lateral forces, in the form of 3D textures, to the current vibroacousticallyaugmented setup. In particular, by superimposing a thin 3D texture on the display, the two-dimensional information (speed and location of the stylus) extracted by the tablet can be integrated with the stylus information (tilt and force) deriving from the actual interaction with the real asperities of the overlay. Some preliminary explorations with 3D textures of few millimeters of thickness were carried out. For instance, Figure 5.10 shows a three-dimensional print (here representing four vocal sketches) that can be superimposed to the tablet while still not impairing its sensing capabilities.

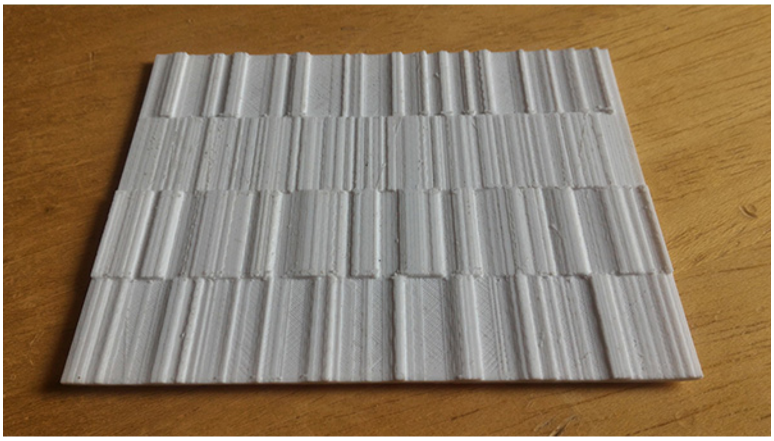


Figure 5.10: A 3D printed texture representing four different vocal sketches.

Sound and vibration can be exploited to enhance the experience of creative acts such as painting and drawing, when these activities are performed on interactive surfaces. In addition, the stylus could be used not only as a probe, but also as an active tool for texture manipulation. A designer might wish to flatten or curl a region of the virtual surface, or to displace it. Finally, the integration of vocalizations in the sketching process might lead to a scenario where voice and hands are in a continuous conversation, thus collaborating seamlessly in the molding of the creative result.

In this respect, Sketch a Scratch is a modulator of problem space, and serves as an open workbench for our design research in virtual texture modeling.

Chapter 6

Experiment 2: Path following in non-visual conditions

Many everyday activities involve the following of a path, a constrained route that we must traverse to reach a goal. This may happen when steering a vehicle or simply when moving our hands, or a tool, according to a certain trajectory connecting two points in space to perform a task. Such task is common in the interaction with digital devices as well: It is being currently investigated, commonly under the name of "steering task", to evaluate the factors affecting human performance in its execution. The related experiments usually take place under visual conditions, that is when the participants are able to see both the path and the tool, or limb, used for the navigation. Moreover, experimental settings usually consider only basic path shapes, such as straight lines or circles.

All of these assumptions may not be fulfilled in real-life situations, where vision may be occluded or engaged in other tasks, and the paths to be traversed can be intricate.

The experiment reported here [54] investigated the possibility of using audio and vibrotactile feedback as substitutes of visual information to complete a path-following task on an interactive surface. The participants navigated by using their index finger over such surface, while their speed, accuracy in terms of adherence to the path, and exerted force were recorded. Moreover, two different, irregular path shapes were provided. Feedback was strictly affirmative, i.e. participants were exposed to continuous feedback (sound or vibration, or both) only when their finger was on-track.

The chapter is organized as follows: First, we elaborate on the motivations behind this work. Then, we provide an overview of past research concerning steering tasks and the use of multisensory feedback in such context. Then, we describe the experiment as regards its technical setup, rationale, and procedure. Then we show the results of the experiment, followed by their discussion and the final conclusions. More material concerning the experiment, such as the interviews to the participants, can be found in the Appendix.

6.1 Motivation

Path-following tasks, or steering tasks, consist of navigating through a tunnel to reach a goal while remaining within the tunnel's boundaries. It is a common activity in everyday interaction with digital devices, where two-dimensional tunnels may be represented by a multi-level menu, being the goal a particular menu item. Another example is a virtual keyboard such as 'Swype'[1] et similia, which enables users to type a text by dragging a pointer from one letter to the next, thus forming a continuous track to form a word: In this case, the path is formed by the virtual segments connecting the characters of the word, and the goal is the final character.

Steering tasks are common in tele-operation as well, such as driving remote vehicles (e.g. space exploration vehicles, undersea pods etc.) or operating remote devices (e.g. machine-operated surgery tools, maintenance factory tools etc.).

Such tasks usually rely on visual feedback: When operating on a touch screen, the position of the finger, or stylus, in relation to the displayed interface, informs the user about the task progress and accuracy. The same information is conveyed by the cursor position in traditional WIMP ("Windows, Icons, Menus, Pointers") interfaces. As the task becomes more complicated, and possibly critical, additional information can be conveyed, thus resulting in a visually intensive interaction. Steering vehicles remotely, such as a harvester for instance, requires to be aware of the vehicle's position, speed and direction. Moreover, the task in its entirety may necessitate supplementary information to be monitored, such as the vehicle's mechanical status, the gas level, the level of storage of the crop etc., which may lead to an overly saturated visual interface. A common risk in such situations is the overload of the visual channel, with the consequent overlooking of important information leading to task failure.

Moreover, real life situations often encompass visual impairment, partial occlusion of the interface, and multitasking: Such factors may prevent users from completing the task by relying only on visual feedback. For instance, the interface for an In-Vehicle Information System (IVIS) is nowadays usually a touchscreen presenting a GUI for controlling the GPS navigator, the

---

[1] http://www.swype.com/

car stereo, the air conditioning, and for displaying maintenance information. Indeed, it would be safer and more effective to enable drivers to perform such tasks without drawing their visual attention from the road.

A relevant field of application concerns navigation aids for the visually impaired, which clearly rely on sound and touch. A brief survey of related implementations can be found in [246]. Smartphone-based solutions are common, such as [46] or the Google TalkBack app.

Lastly, specific steering tasks explicitly benefit from a form of non-visual feedback: For instance, the robot-assisted driving of a surgical tool requires haptic feedback conveying force and textural information to reduce the probability of unintentional injuries or traumas for the patient [162].

6.2 Background

6.2.1 Steering tasks

Pointing, crossing, steering

A fundamental study for the HCI discipline is Fitts' analysis of the human performance in pointing tasks [71]: In examining a repeated, fix-width movement such as tapping alternately with a stylus on two targets, Fitts focused on the information capacity of the whole human receptor-neural-effector system as limited by the statistical variability among trials. The result was a model of the task difficulty, in which the task completion time depends on an index of difficulty ID defined as follows:

$$ID = \log_2(\frac{D}{W} + 1) \tag{6.1}$$

where D is the distance between the targets (i.e. the width of the movement) and W is the target width in the movement direction. A formulation of the consequent estimation of the completion time, commonly known as "Fitts' law", is the following:

$$T = a + bID \tag{6.2}$$

where a and b are constants. The target size orthogonal to the movement (i.e. the height of the target) is not considered in such formulas, thus being assumed as unlimited. However, such dimension was shown to represent a constraint as well, leading to a more accurate formulation of ID [3]:

$$ID = \log_2(\sqrt{(\frac{D}{W})^2 + \eta(\frac{D}{H})^2} + 1) \tag{6.3}$$

where H is the target height (see Figure 6.1), and η is a constant < 1. Note that common interactions with digital interfaces, such as object selection and drag-and-drop, can be modeled after either law.

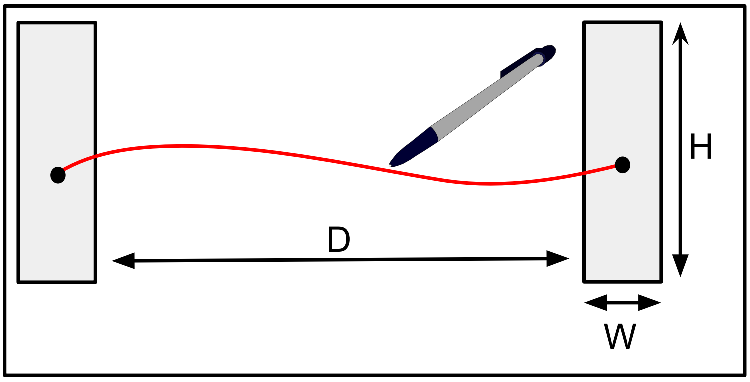


Figure 6.1: A pointing task with two-dimensional targets.

The additional dimension enables the modeling of another type of task, namely goal crossing.

Pointing and crossing tasks can be seen as "discrete", in that they presumably require a single movement and consequently a simple planning to be executed. Conversely, due to their lateral boundaries (which may be curvilinear), steering tasks require a continuous action-perception loop to be performed, which guides the user's actions towards a goal. Anyhow, paths can be split in sections that must be crossed, each representing a target per se: Therefore, a steering task can be seen as the integration of a series of crossing tasks (see Figure 6.2).

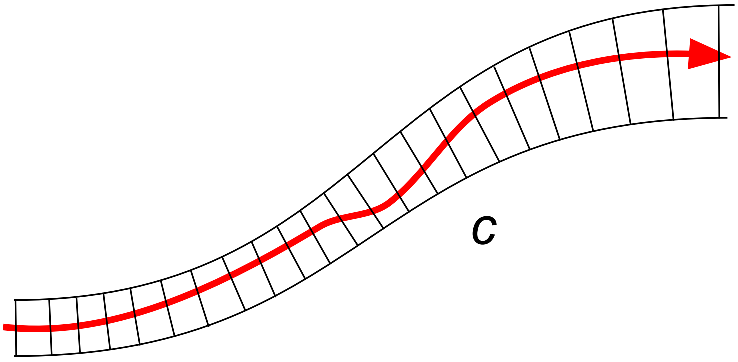


Figure 6.2: A steering task along a path c can be decomposed in a series of subsequent crossing tasks.

This intuition underlies the formulation of the "steering law" by Accot and Zhai [1], which relates the task completion time to the width of the tunnel to be navigated, as foreseen by Rashevsky [173] and Drury [62] concerning vehicle control.

Such law states that the ID in executing a steering task depends on the path width $W(s)$, which is integrated along the path c. The resulting formula of the estimated completion time T_c for path c is the following:

$$T_c = a + b \int_c \frac{ds}{W(s)} \tag{6.4}$$

where a and b are empirically determined constants. Such constants encompass the effect of the shape of the path, along with other context-dependent conditions. Indeed, the authors recommended to treat different path shapes separately.

The steering law is now a popular tool in HCI for predicting performance in steering tasks under defined circumstances (e.g., the use of a particular input device).

2/3 power law

The impact of path shape over the performance was investigated by Viviani et al.: The "isogony principle" [231] states that in handwriting and drawing movements the angular velocity tends to remain constant or, in terms of tangential velocity:

$$V = kr \tag{6.5}$$

where V is the tangential velocity, r is the radius of curvature of the trajectory, and k is a constant. In general, the authors found that trajectories could be split in curvilinear segments, each with a different tangential velocity. However, by observing that changes in such velocity occur also away from geometrical singularities, they theorized the impact of more global aspects of the trajectories as well as the radius of curvature. As a result, the relationship between velocity and radius was found to be less than linear. The consequent formula for the instantaneous velocity $V(t)$ is the following [129]:

$$V(t) = kR^{1/3}(t)) \tag{6.6}$$

where k is a "gain factor" which remains constant for the segment and $R(t)$ is the radius of curvature as a function of time. The more common formulation involves the angular velocity $v_\theta(t)$ in the so-called "${}^2/_3$ power law":

$$v_\theta(t) = kc^{2/3}(t)) \tag{6.7}$$

where $c(t)$ accounts for the curvature.

Such law holds true for both constrained and free movements, e.g. scribbles with a pen.

Other hypotheses concerning the drawing of two-dimensional trajectories are the "minimum jerk"[72], which states the human tendency to perform smooth movements to minimize the energy cost and consequently minimize the rate of change of acceleration, and the "isochrony principle", namely the "increase of the average movement velocity with the linear extent of the trajectory" [230], which causes movement size and movement duration to be nearly independent. An integration of such hypotheses ($^2/_3$ power law, minimum jerk, and isochrony) was attempted as well [229], although a complete merging was not achieved.

Later studies brought to light more kinematic invariants of human movement, e.g. during cyclical arm movements [60] or movements following geometrically affine shapes [18].

Limitations and integration of laws

The practical limitations in the validity of the steering law and of the $^2/_3$ power law, which are likely to arise given the complexity of the human motor control system and of the hand-arm system, were investigated. For instance, the validity of the steering law at different scales was tested [2], concluding that such law applies only to a 'middle range' scale of paths: The constraints of motor joints shift and human motor precision yield a U-shaped performance function.

Concerning the $^2/_3$ power law, handwriting was analyzed [170] resulting in limitations to the types of strokes that comply with such prediction, namely elliptical or hyperbolic trajectories.

An integration of the two laws was experimented [131] in the context of evaluating the tolerance in spatial selection tasks such as circumscribing an object with a pen stroke. Results showed that the steering law overcomes the $^2/_3$ power law as the path becomes narrower (see Figure 6.3), while the $^2/_3$ power law prevails when width constraints are looser.

Kulikov et al. [128] extended the Accot-Zhai formulation – which makes use of only error-free responses – by introducing the measurement of the out of path movement, namely the percentage of sample points outside the path boundaries. They also achieved a better prediction of the task execution time by analyzing the effective path width used by participants.

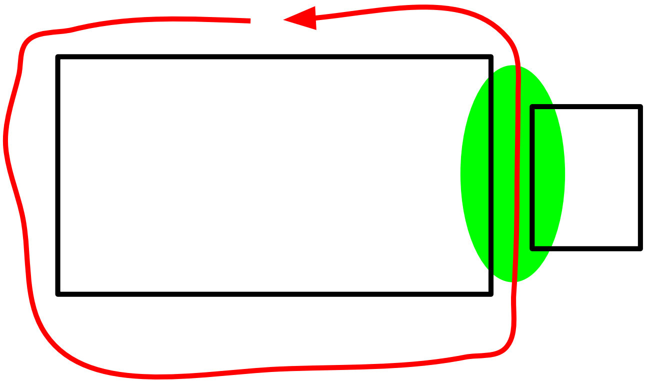


Figure 6.3: The steering law prediction is more accurate than the $^2/_3$ power law on narrow path regions (green zone), and vice versa the $^2/_3$ power law prevails on wider path regions.

6.2.2 Multisensory feedback in path following

In Section 4.1 we overviewed the factors affecting a multisensory percept, such as spatial and temporal coincidence, and intensity matching across the provided feedback modalities. Such factors are relevant in steering tasks as well: It has been stressed that, to be effective, both auditory and haptic feedback should be consistent with the visual information. Consistency should be sought with regard to synchronization [35] as well as to continuous adaptation in response to the user's movements. For instance, variations in pitch, loudness or rate of a 'vibrato' effect have been applied to continuous auditory feedback to inform the users about their progress along the edge of a graph [40]. Thoret et al. [220] investigated the ability to deduce the shape of simple trajectories from friction sounds generated by velocity profiles following the $^2/_3$ power law: Variations in the sound were shown to recall the steering movements that provoked them, and consequentially provide guidance in the task.

Several steering task experiments investigated the effect of haptic feedback at the user's hand in addition to visual information [35, 56]: Additional haptic feedback improved performance regardless of what tool was used for the navigation. Sun et al. [213] went further by comparing the performance in every combination of auditory, visual and tactile feedback in the navigation of a circular path by means of a stylus. The main result was that participants performed most accurately with tactile feedback, although they generally preferred the audio-visual modality. It is worth noticing that the path was always visible, and that the multisensory stimuli were related to error conditions, i.e. additional feedback was given when participants went off-track. Negative feedback aimed at avoiding possible fatigue due the continuous presence of audio or tactile cues, which may lead to concentration decrease and sensory adaptation. Conversely, [35] and [56] adopted an affirmative feedback strategy, i.e. stimuli were provided as long as participants stayed on the path.

The experiment described here is inspired to the work by Sun et al. [213], and represents a follow-up to a former study [181] (the Sketch a Scratch experiment summarized in Section 5.3.2). In both experiments all combinations of visual, auditory and vibrotactile feedback were randomly submitted to the participants, yet in [181] the path was invisible when in non-visual conditions. In such experiment, when the path was visible, complementary auditory or vibratory stimulation seemingly had no impact on the performance. When presented as alternatives, visual feedback greatly outperformed both auditory and vibratory feedback. Interestingly, non-visual feedback modes caused trajectories to be different than in visual mode.

6.3 Experimental design

In the present steering task, participants were asked to navigate through a path connecting the two sides of an interactive surface, left to right. They were required to use the index finger of their dominant hand and to perform the task as quickly and accurately as possible, based only on non-visual cues (i.e., the path was not visible). Different path shapes and non-visual feedback modes were provided.

Strictly affirmative continuous feedback was used: When on track, sound and/or vibration were provided, while no feedback was produced when off track. Time, finger position and normal force were recorded.

6.3.1 Apparatus

The surface used for the experiment was the active portion of the Madrona Labs Soundplane[2], a computer music controller. Such surface measures 560×140 mm and is capable of sensing position and normal force of up to ten fingers. It is usually employed as a 30×5 matrix of keys (gaps are actually carved between the keys): However, its sensor density enables its use as a continuous surface as well.

---

[2] http://madronalabs.com/soundplane/

Figure 6.4: The experimental setup. The Soundplane's surface is covered with a plastic foil, and participants wear a glove to minimize friction. Black velcro stripes serve as tactile landmarks indicating the starting position.

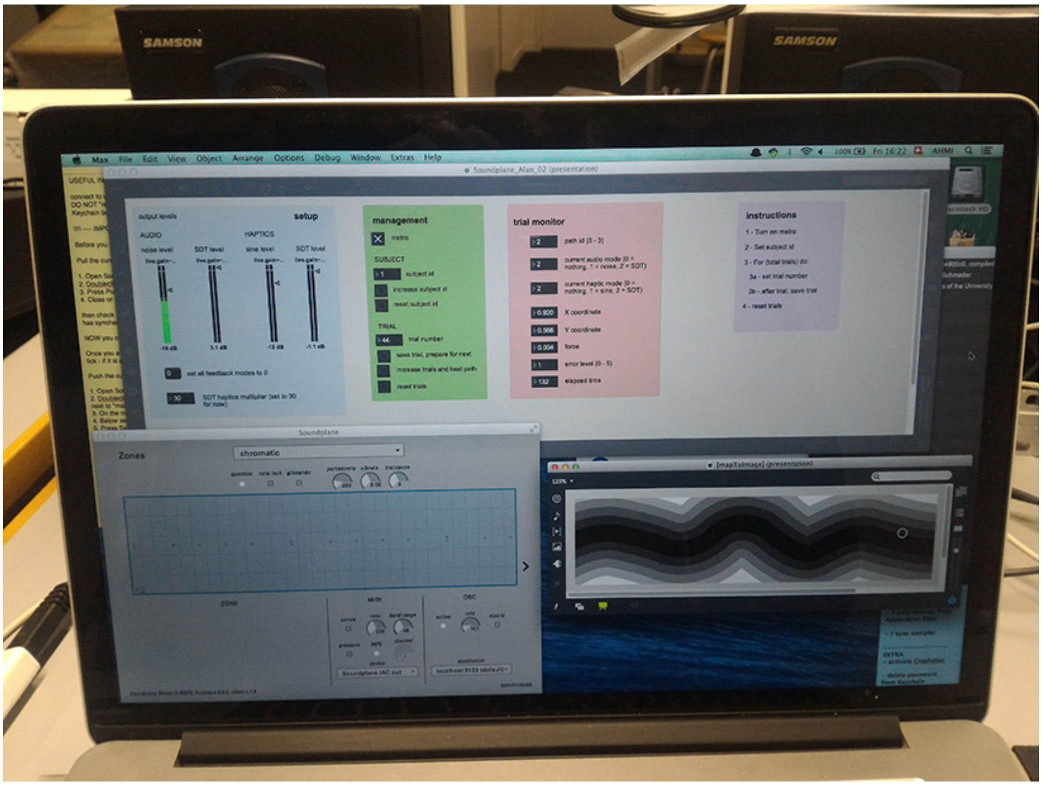
A vibration transducer (shaker) was fixed to the bottom of the Soundplane.

The participants' finger position and force were collected through the Soundplane's software client, which was interfaced to the experiment's management system developed in Max (see Figure 6.5). Data were sampled every 10 ms.

Feedback was generated interactively, according to finger position, by means of the SDT (see Section 3.3.2), and the same signal was used for rendering both auditory and vibrotactile cues. The model of a rolling sound was used to simulate an object rolling over an uneven terrain, such as a car tyre. Apart from the steering metaphor that such signal was intended to evoke, the wide frequency spectrum (similar to filtered noise) was meant to decrease the impact of selective adaptation under vibrotactile conditions (see Section 2.1.2).

The synthesized signal was sent to two output channels of an RME FireFace 800 audio interface, respectively leading to a pair of Beyerdynamic DT-770 Pro headphones and to a power amplifier connected to the shaker. The signal routed to the shaker was first band-pass filtered in the $80-250$ Hz range to maximize the transducer's efficiency while minimizing audible frequencies. The remaining sound spillage from the shaker was masked by a noise signal (pink noise, with a small amount of white noise added for covering high frequencies), sent to the headphones during the vibrotactile mode only.

The Soundplane's surface was covered with a thin, opaque plastic foil (see Figure 6.4) to make the surface uniform by covering the gaps between the keys, and to reduce finger friction. It had been previously verified that such foil would affect neither the precision in position detection nor the detected force. Velcro stripes were glued at the starting position of the paths,



output levels

setup

AUDIO

noise level -18 dB

SDT level 3.1 dB

HAPTICS

sine level -12 dB

SDT level -1.1 dB

0 set all feedback modes to 0

30 SDT haptics multiplier (set to 30 for now)

management

☒ metro

SUBJECT

1 subject id

○ increase subject id

○ reset subject id

TRIAL

44 trial number

○ save trial, prepare for next

○ increase trials and load path

○ reset trials

trial monitor

2 path id (0 - 3)

2 current audio mode (0 = nothing, 1 = noise, 2 = SDT)

2 current haptic mode (0 = nothing, 1 = sine, 2 = SDT)

0.820 X coordinate

0.566 Y coordinate

0.004 force

1 error level (0 - 5)

132 elapsed time

instructions

1 - Turn on metro

2 - Set subject id

3 - For (total trials) do:

3a - set trial number

3b - after trial, save trial

4 - reset trials

Figure 6.5: The experiment management system. From the top, clockwise: The Max interface for the control of the feedback levels and the task progress, the visualization of the current position with respect to the path, and the Soundplane's client.

thus allowing participants - who were blindfolded - to locate it by means of touch. The Soundplane rested over a keyboard stand, and rubber foam was interposed between the interface and the support to avoid unwanted resonances due to spurious standing waves, and to minimize vibration propagation through the floor at the same time.

6.3.2 Test conditions

A within-subjects design was used, i.e. all participants experimented all test conditions. Six conditions were available as the combinations of three feedback modes and two path shapes. The feedback modes were: audio ('A'), vibrotactile ('T'), and audio + vibrotactile ('A+T'). The two paths had a constant width of 28 mm, namely twice the average fingertip's contact area proposed in the literature [47]. After a pilot test, such width was chosen to accommodate possible changes due to the varying exerted force and inclination of the wrist, and to limit the overall difficulty of the task.

The paths were labeled 0 and 1: Path 0 (see Figure 6.6a) features a single curvilinear trajectory which mostly retains the same, mild curvature, while path 1 (see Figure 6.6b) features several changes in direction, with a pronounced slope. Since the paths shared the same left and right boundaries given by the surface's frame, the more curvilinear path 1 was 13.1% longer than path 0.

The conditions were presented in a guided random order - in the case of two subsequent occurrences of the same condition, they were manually distanced - to avoid learning and sensory adaptation effects. Each condition was repeated ten times, for a total of 60 trials, resulting in an average duration under 60 minutes.

(a) Path 0.

(b) Path 1.

Figure 6.6: The two path shapes to be followed.

6.3.3 Design rationale

The task relied on the presence of affirmative feedback, in agreement with previous studies [35, 56, 181]. Indeed, when exploring a space it feels more natural to look for a path (i.e., for clues signaling its location) rather than the opposite.

After a series of pilot tests involving seven participants and four path shapes, the following design decisions were taken:

- Number of paths: the number was reduced to two, to limit the duration of the experimental sessions to approximately 60 minutes while achieving statistical relevance (i.e. a sufficient number of trial repetitions per participant);

- Path shape: More complex shapes (e.g. the one displayed in the bottom-right corner of Figure 6.5) were excluded due to their excessive execution time and possible arising of frustration among the participants. The chosen paths were devised to minimize the learning effect: They start in different directions (one upwards, one downwards), they are not symmetrical with respect to the middle, they end at different locations of the right-hand side of the surface, they maximize the vertical range of movement by reaching both the higher and the lower border of the surface, and they do not encourage ergonomic assumptions (e.g. a trajectory forming an arc centered on the participant's body). Yet, the starting position is at mid height for both paths, to avoid bias in the first movements of the participants;

- Feedback differentiation: The experiment was designed to provide multiple feedback intensity levels according to the participant's position (e.g. decrease the feedback intensity as the distance from the correct track increases). However, the differences in vibrotactile sensitivity among individuals suggested a first version of the experiment with only two feedback levels, "on" and "off". Indeed, a preliminary tailoring of feedback intensity would be required for each participant, and such procedure would be complicated by the presence of multiple levels, since each must be clearly distinguishable from the others while none must cause discomfort due to excessive intensity.

6.3.4 Hypotheses

The following hypotheses were tested:

1. The average task completion time with non-visual feedback does not follow the Accot-Zhai prediction for steering tasks in visual mode, due to the different participant behavior according to the feedback modalities observed in [181];

2. The path shape affects speed and accuracy in the task. Specifically, path 1 yields worse performances than path 0 due to the number and steepness of changes in direction which, in non-visual conditions, cannot be foreseen;

3. Since auditory and vibrotactile feedback are synchronous and originate from the same signal, the combined auditory+tactile mode produces a better performance than a single sensory mode [241];

4. Manual skills, demographic factors, psychophysical factors, and preference over the feedback modality affect the task execution. For instance, manual skills presumably imply a higher level of tactile sensitivity, which might improve the performance in the presence of tactile feedback;

5. The exerted force generally accords to a coherent behavior, e.g. it is higher for a particular path shape and/or feedback mode.

6.3.5 Procedure

The experiment took place in a quiet room at the Zurich University of the Arts, involving 30 participants: 16 female and 14 male, mean age 28.7 years (female = 27.4, male = 30.1, SD = 9.0). All participants but one were right-handed. Twelve of the participants declared to possess a trained manual ability, e.g. as musicians, painters or sculptors. Before the experiment, each participant signed a customary consensus form following the rules of the local Ethics committee. Participants were rewarded a voucher valid at the local canteen and cafeterias.

During the briefing phase, the experimenter demonstrated the execution of one trial. Then participants had to wear headphones, and a light cotton glove that minimized friction with the plastic foil covering the surface. They were blindfolded to avoid the visual guidance of the Soundplane's frame borders as well as to prevent them to mentally project a path over the surface (a strategy observed in some pilot test participants). Participants were allowed to gain familiarity with the task by freely exploring the surface with one of the paths randomly loaded, while all feedback modes were provided in sequence. They were instructed to consider both speed and accuracy as important to the task, while feeling free to choose their navigation style. The experiment could be performed either sitting on an adjustable piano stool or standing. The level of fatigue and stress was constantly monitored. Mandatory breaks were given every ten minutes.

During the debriefing phase at the end of the experiment, participants had to report about physical and mental fatigue. Then, they were asked to express their preference over the three feedback modes. Lastly, they could express their opinion about the task and the overall experience.

6.4 Results

Completion time, distance, and average exerted force were the main variables recorded during the experiment. The analysis required the time and distance data to be manipulated to obtain metrics of performance in terms of speed and accuracy (derived variables).

The main variables showed a considerable variability among the participants, particularly with regard to force: The coefficient of variation (i.e. the ratio of the standard deviation to the mean) was 0.31 for completion time, 0.23 for distance and 0.46 for exerted force.

The mean values of the results are reported in Section 8.1.2.

By visualizing means, medians and standard deviations for all the main and derived variables, no obvious grouping could be detected among the participants. A further attempt at grouping was done according to subjective factors such as gender (age could not be used due to little variability), presence of trained manual abilities (e.g., as musicians), and physical or mental fatigue self-reported at the end of the session. Anyhow, non-parametric Mann-Whitney U-tests detected no significant differences in performance among such groups (see Section 8.1.1).

ANOVA tests were performed with path and feedback mode as factors. The hypotheses of normal distribution and sphericity of data were often not fulfilled, thus justifying the application of ϵ-corrections and the adoption of non-parametric tests whenever necessary. For such tests, Matlab[3] and a set of Java scripts developed by Ian Scott McKenzie[4] were used alternatively, according to the required functions.

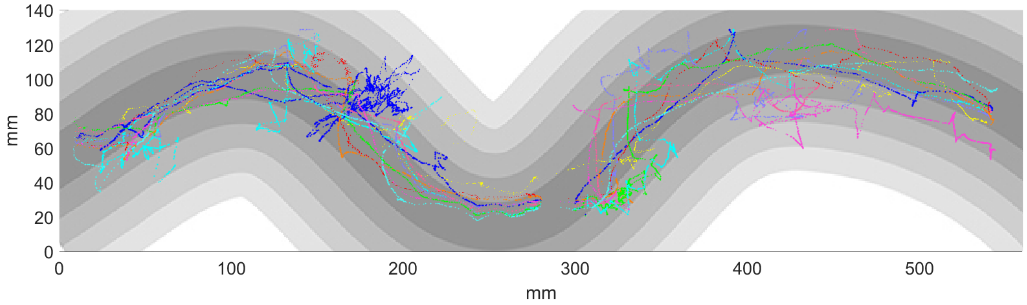


Figure 6.7: Tracks performed by a participant on path 1 in audio+vibrotactile mode. Different colours correspond to different repetitions, and the thickness is proportional to the exerted force. The shown 14 mm-wide gray stripes enable a quick evaluation of distances. After the second turning point, a "zig-zag" strategy is apparent.

---

[3] https://www.mathworks.com/

[4] http://www.yorku.ca/mack/HCIbook/

6.4.1 Task completion time

The mean time for completing a single task repetition was 41839 ms (SD = 13092 ms). Such time largely exceeds what would have been necessary to complete the task under visual conditions, namely far less than five seconds as informally tested.

The average execution time for path 1 was 39.6% longer than for path 0, and this difference was significant ($F_{1,29} = 94.435$, $p < .005$). Since there was only a 13.1% difference in length (594 mm for path 0 and 672 mm for path 1), this suggests the presence of other factors affecting performance.

On both paths, T required considerably longer times than A and A+T (see Table 8.4). Conversely, the differences between A and A+T were much smaller. Such differences were significant ($F_{2,58} = 23.363$, $p < .005$). The interaction between the two factors was marginally significant ($F_{2,58} = 3.624$, $p = .0329$). However, ϵ-corrections on the separate paths (Huynh-Feldt = 0.859 for path 0, Greenhouse-Gasser = 0.659 for path 1, $F_{1,24} = 9.049$, $p < .005$ and $F_{1,19} = 1.789$, $p = .189$ respectively) showed that the feedback mode was significant on path 0 but not on path 1.

6.4.2 Speed

Two measures of speed were considered: gesture speed and task completion speed. Gesture speed was computed by dividing the total distance run in a single trial by its completion time. Task completion speed was computed by dividing the original path length by the completion time.

Indeed, task completion speed is affected by trajectory accuracy, since the considered distance is fixed regardless of the actual trajectory drawn by the participant. However, such measure complies with the definition of "movement time", consisting in the "time taken to move from the start line to the end line" [128].

Gesture speed

Gesture speed was 4% slower on path 1 than on path 0, and such difference was significant ($F_{1,29} = 4.499$, $p = 0.043$). T was the slowest feedback mode (8.5% slower than A, and 9.7% slower than A+T), while A and A+T were almost equivalent. The difference between the feedback conditions was significant ($F_{2,58} = 25.788$, $p < .005$). Post-hoc comparisons after Friedman test showed A and A+T were not significantly different on both path 0 and path 1.

Task completion speed

Analysis of task completion speed confirmed the results related to gesture speed, with even larger differences. Task completion speed on path 1 was 19.9% slower than on path 0, and such difference was significant ($F_{1,29} = 30.977$, $p < .005$). Concerning feedback modes, T was 18.5% slower than A and 22.4% slower than A+T on path 0; On path 1, T was 13.0% slower than A and 8.0% slower than A+T. Such result was significant ($F_{2,58} = 41.827$, $p < .005$). Post-hoc comparisons after Friedman test showed no significant difference between A and A+T.

6.4.3 Accuracy

Two measures of accuracy were considered: Time-related and space-related accuracy.

Time-related accuracy was computed by dividing the total time spent on-track during a trial by the trial execution time. Such measure was inspired by the "out of path movement" metric as found in [128, 213], which consists in the percentage of sample points outside the constraint lines.

Space-related accuracy is a measure of trajectory error, and was computed as the mean of the Euclidean distances of each sampled position point from the nearest edge of the correct track. The sample resolution was 1 mm over the x-axis.

Time-related accuracy

The trials performed on path 0 were 16.6% more accurate than those performed on path 1, and such difference was significant ($F_{1,29} = 47.446$, $p < .005$).

By running the Friedman test on the two paths separately, post-hoc comparisons showed that the differences among feedback modes were not significant on path 0, while they were significant when comparing T to A and A+T on path 1.

Overall, time-related accuracy shows an interesting result: Participants spent on average 67.1% of their trial time on-track. Considering a space of 14 mm around the track - comparable to the average diameter of a fingertip [47] - the above value rises from 67.1% to 93.0%. Such result indicates that participants spent most of the time on the correct track or in its immediate vicinity, meaning that the participants responded promptly to the variations in feedback. Moreover this result, in conjunction with the low movement speed, also predicts small trajectory errors.

Trajectory error

The feedback mode had no significant effect on trajectory accuracy ($F_{2,29} = .690,\ p = .506$). Such result supports the observation that the participants adopted the same cautious navigation style with all feedback modes. Moreover, this determined a generally small average trajectory error (see Table 8.10).

Path 1 resulted in an error 7.5% higher than path 0, and this result was significant ($F_{1,29} = 19.103,\ p < .005$).

6.4.4 Exerted force

The average exerted force was of 2.218 N, with a 1.014 N standard deviation. Upon inspection of the inter-subject differences (see Figure 6.8), it is apparent that the values of the exerted force do not induce any grouping among the participants. Indeed, 25 out of 30 participants are roughly comprised within the range 1–3 N. Clearly, the variances are not related to the absolute values of the means.

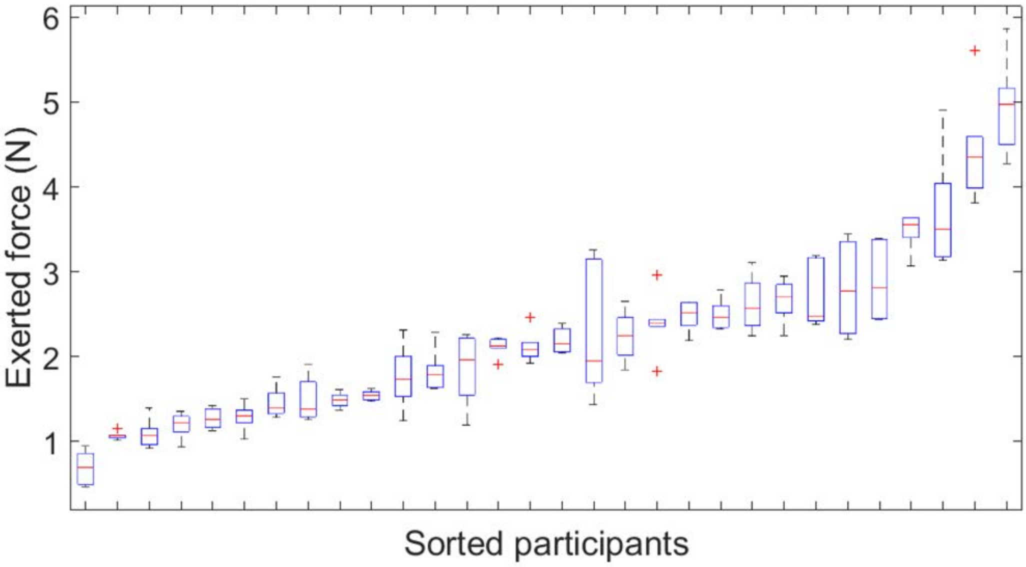


Figure 6.8: Box plot of the exerted forces for each participant.

The path factor was not significant ($F_{1,29} = 0.121,\ p = .730$). Conversely, the feedback mode significantly affected the exerted force ($F_{2,58} = 22.312,\ p < .005$): On average, participants exerted a stronger force when executing the task with vibrotactile feedback than with the other two modalities (see Table 8.6).

Post-hoc comparisons after Friedman test, run on the two paths separately, showed that there was no significant difference between A and A+T on both paths.

On average, the force exerted by participants when off-track was slightly (+4.5%) higher than when on-track, and such difference was significant ($F_1 = 28.409,\ p < .005$). This result is confirmed by a Wilcoxon Signed-Rank test ($z = -4.494,\ p = 0.0000$).

Lastly, correlations among all variables were computed to highlight possible redundancies in the analysis. No unpredicted correlations were found (see Section 8.1.3). In particular, the force does not seem correlated neither with speed ($r(28) = -.011,\ p = .954$ with gesture speed, $r(28) = -.037,\ p = .846$ with task completion speed) nor with trajectory errors ($r(28) = -.016,\ p = .933$).

6.4.5 Analysis on normalized data

One approach to tackle the relevant variability in the recorded values among the participants is to normalize the data by one of the three feedback modes. For instance, the trajectory errors in A and A+T can be divided by those in T. Thus, by analyzing error ratios instead of absolute values, the impact of inter-subject variability is likely to be reduced.

Two normalization were attempted: the division by the recorded values in T (thus enabling a comparison between A and A+T), and the division by the recorded values in A+T (thus enabling a comparison between A and T). Task completion times, trajectory errors and exerted forces were considered.

With the first normalization, the differences between A and A+T were shown not to be significant on either time ($F_{1,29} = 0.092,\ p = .763$), trajectory errors ($F_{1,29} = 1.435,\ p = .240$), and force ($F_{1,29} = 1.402,\ p = .246$). This complies with the former analysis. With the second normalization, the differences between A and T were significant for time ($F_{1,29} = 43.545,\ p < .005$) and force ($F_{1,29} = 18.972,\ p < .005$), but not for the trajectory errors ($F_{1,29} = 1.037,\ p = .317$). This also complies with the former analysis.

6.4.6 Trend analysis

The participants' performance was investigated to detect possible trends, and possibly hypothesize their causes. Trajectory error and task completion speed were considered as measures of performance. Additionally, the exerted force was evaluated.

Trend analysis was computed 1) at repetition level, that is monitoring the evolution along a single trial, and 2) at trial level, that is considering the trend along the 10 repetitions for each test condition.

Trend at repetition level

This analysis would help discern two factors that might affect performance: path length and shape. It was hypothesized that if, after covering the same distance over the two paths, performance was comparable, then the path shape would not be relevant. If, otherwise, performance on path 1 was worse than on path 0, then the path shape would be proven a difficulty factor.

Task completion speed and trajectory error were computed on the portions of the repetitions that corresponded to the first 600 mm of space run by each participant, sliced in twelve 50 mm-long sections. Speed was averaged over each section, while errors were cumulated. The size of such portion was chosen to maximize the length that had been run on each trial repetition, to enable a proper analysis regardless of the different lengths run according to participants and conditions. Average values were computed over all participants and all trials on the six available conditions.

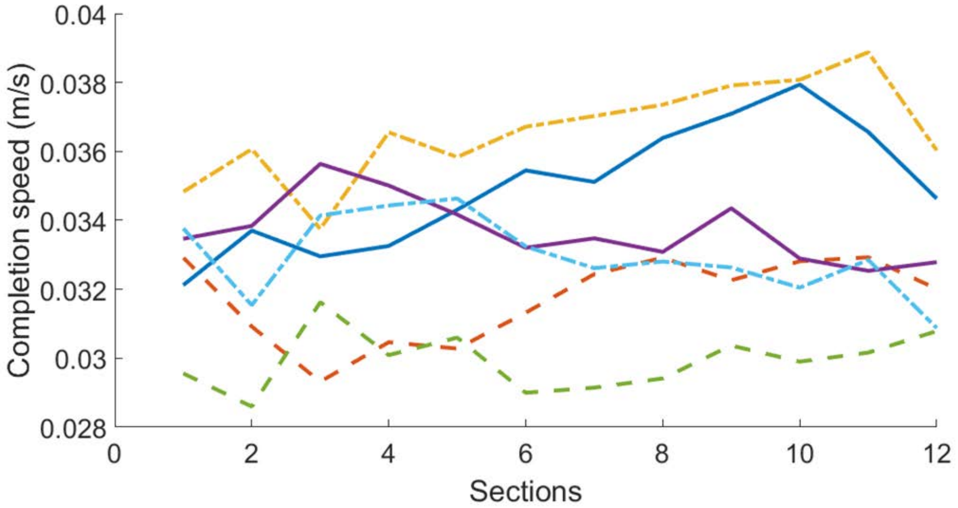


Figure 6.9: Completion speed along the first 600 mm of the paths, according to conditions (path 0: blue solid line = A, orange dashed line = T, yellow dash-dotted line = A+T. Path 1: purple solid line = A, green dashed line = T, cyan dash-dotted line = A+T).

As shown in Figure 6.9, the speeds do not show any particular trend, except for the lower speed of trials with T (shown in dashed lines) compared

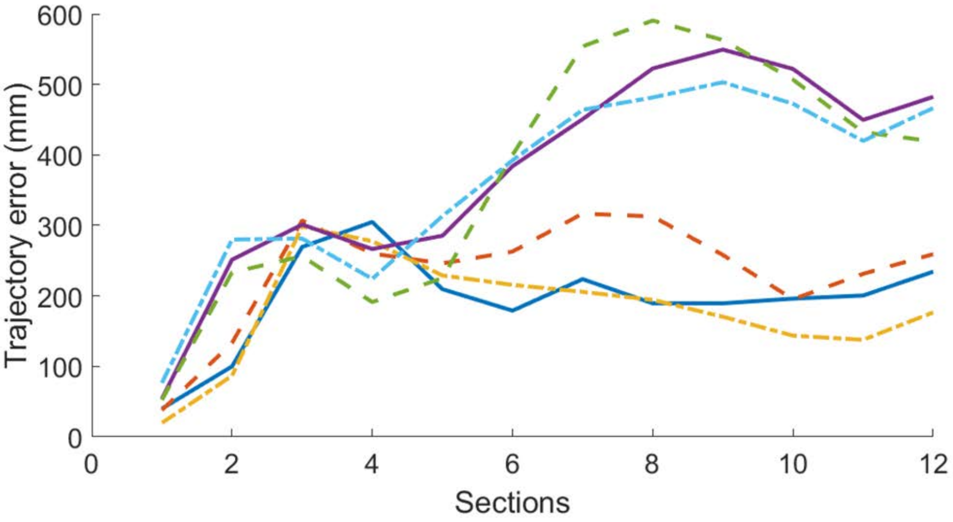


Figure 6.10: Cumulative trajectory error along the first 600 mm of the paths, according to conditions (path 0: blue solid line = A, orange dashed line = T, yellow dash-dotted line = A+T. Path 1: purple solid line = A, green dashed line = T, cyan dash-dotted line = A+T).

to those with A or A+T on the same path, whose speeds are instead similar.

Conversely, Figure 6.10 shows that trajectory errors are comparable in the first sections, but soon diverge depending on the path: The repetitions on path 0 caused lower trajectory errors compared to those on path 1.

Such visualizations suggest what follows:

1. Trajectory errors are affected by the path shape, which consequently is a difficulty factor for the task;
2. Due to the lack of clear trends in speed, the presence of a linear speed/accuracy trade-off cannot be evaluated (see Figure 6.11).

To investigate further the relationship between path shape and trajectory errors, the magnitude of the estimated curvature of path 1 was visualized against the average trajectory errors (see Figure 6.12). Although a rigorous interpretation cannot be given, it is apparent how the curvature of path 1 seems to modulate the trajectory errors.

Trend at trial level

The values of trajectory errors, task completion speeds and exerted forces were averaged across the participants for each of the 6 conditions, thus form-

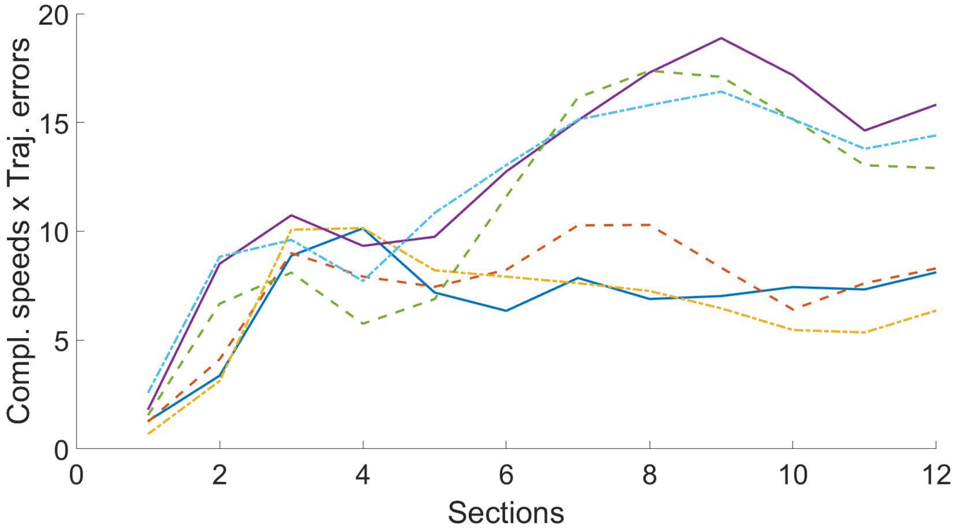


Figure 6.11: Trajectory errors / accuracy trade-off (path 0: blue = A, orange = T, yellow = A+T. Path 1: purple = A, green = T, cyan = A+T).

ing piecewise linear curves formed by 10 points (one per repetition), as shown in Figures 6.13, 6.14, and 6.15.

The trends hardly show any regularity, except for the completion speeds, which increase almost linearly on all conditions (Figure 6.13).

In absence of a generalized improvement in performance, we may exclude the influence of learning effects along with the repetitions. At most, the increasing speed may be related to an increasing familiarity of the participants with the task.

Figure 6.14 clearly shows the offset in average trajectory errors between the two paths.

As shown in Figure 6.15, force data were even more mixed, and no trend can be spotted over the ten repetitions. However, it is evident how participants used more force in the vibrotactile feedback condition.

6.4.7 Participants' behavior

Due to the invisibility of the paths, participants generally adopted a cautious navigation style. Such behavior complied with [181], and took place regardless of the feedback mode. A peculiar navigation strategy was occasionally adopted by some participants, consisting in zig-zagging across the path (see Figure 6.7). In their intention, this would have enabled a constant awareness of the path boundaries: Conversely, such strategy did not improve their ac-

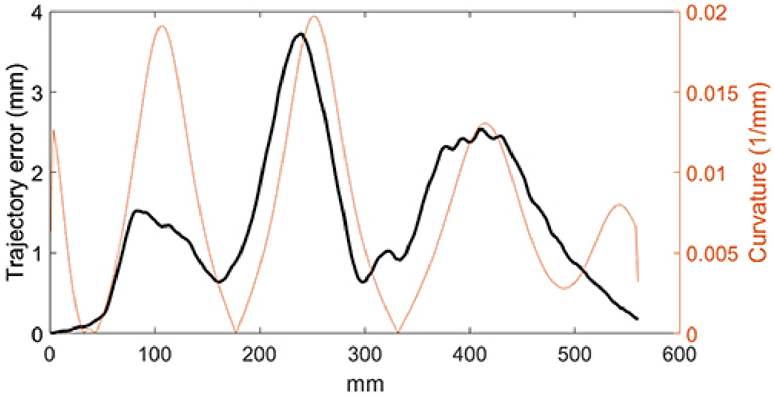


Figure 6.12: Average trajectory error (in black) and absolute value of curvature of path 1 (in orange).

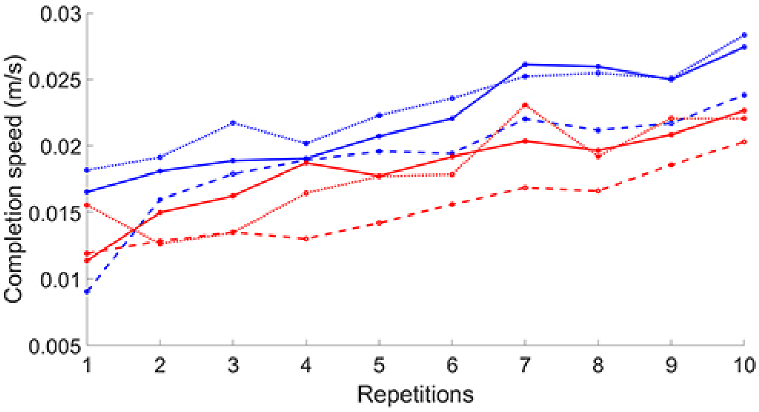


Figure 6.13: Completion speed over the 10 repetitions (solid = A, dashed = T, dotted = A+T). The blue lines refer to path 0.

curacy significantly, while it did cause considerably longer task completion times (up to +20% in some cases).

Even though, in case of lost track, participants were instructed to explore the surface vertically to intercept the correct track again, such behavior was rarely adopted: Instead, a common behavior consisted in back-tracking to the last known correct finger position, and resuming the navigation from that spot.

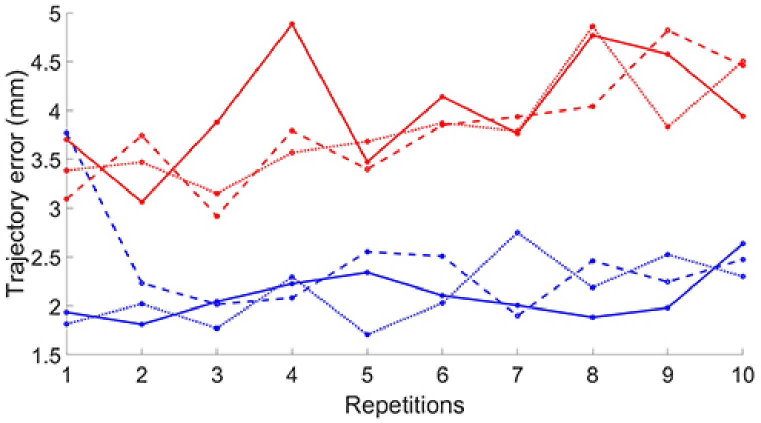


Figure 6.14: Average trajectory error over the 10 repetitions (solid = A, dashed = T, dotted = A+T). The blue lines refer to path 0.

6.4.8 Debriefing

In general, the task was deemed as demanding by the participants: 10 participants reported mental fatigue due to the concentration. On the whole, 5 participants considered the task fun or challenging, while 4 considered it boring or frustrating.

Sixteen participants reported physical strain, namely a slight stiffness to the wrist and/or shoulder or, more frequently, to the finger.

The participants were deliberately not informed about the available number of paths, which were just two in random succession. However, probably due to the randomization of conditions, such number was supposed to be much higher than two by many of the participants who tried to guess.

When asked about their preference concerning the feedback mode, 22 participants said to have preferred A, 3 preferred T, 3 A+T, and 2 had no preference. Indeed, the vibrotactile condition was the most difficult for many participants.

Ten participants reported difficulties in perceiving vibratory cues. Six reported a progressive desensitization towards the end of each trial with vibrotactile feedback.

The main comments that were collected are provided in Section 8.2.

6.5 Discussion

A few issues affected the experimental setup and the data collection during the experiment, namely:

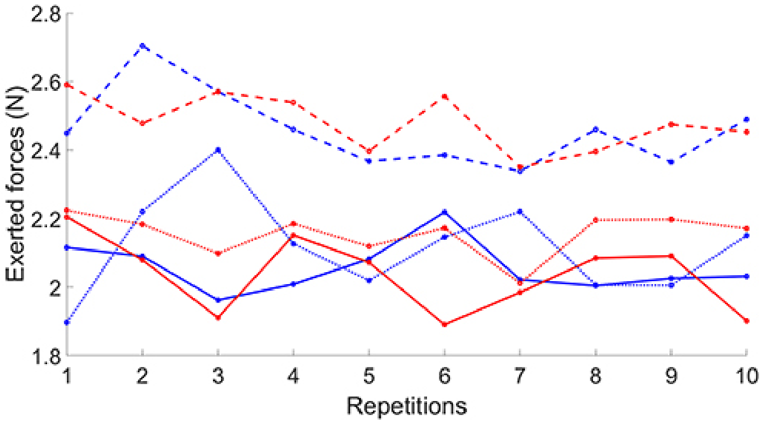


Figure 6.15: Exerted force over the 10 repetitions. The blue lines refer to path 0. The two dashed lines refer to the vibrotactile feedback conditions.

- The intensity of vibration was upper-bound by the concurrent acoustic spillage originating from the shaker. Also, the loudness of masking noise was limited by comfort requirements;

- The finger is a rather inaccurate pointing tool for touch interfaces: The softness of the user's skin causes touch position to be sensed anywhere within the contact area between the user's fingertip and the device, and changes in finger inclination may affect the detected contact point. This is an aspect of the so-called "fat finger problem" [232];

- Force measurement showed varying accuracy across the surface: A characterization of force sensing was performed by means of test weights placed across the surface, showing a mean coefficient of variation of 0.183. The measured data were fitted resulting in a calibration curve, which was then used in the analysis phase to compensate for varying force sensing accuracy across the surface.

Two crucial factors guided, and somewhat limited, the interpretation of the experimental data: the noisiness in trajectories and the generally high variability of data among the participants. Noisiness was likely mainly due to the use of non-visual feedback, which prompted the participants to adopt an exploratory behavior. As a consequence, trajectories often presented U-turns and discontinuities. The position sensing issues mentioned above may also have contributed to additional trajectory noise. Overall, noise in trajectories invalidated the estimate of navigation progress: The direction in a particular segment of movement might have been opposite to the path direction.

Therefore, instant velocity data was meaningless in such scenario. As a final consequence, comparisons with the predictions of the steering law or the $^2/_3$ power law were intrinsically not viable.

As previously mentioned, variability of data among the participants prevented their grouping based on any of the considered parameters. In particular, the exerted forces varied greatly among the participants.

In contrast with the findings of [213], accuracy was only partially affected by the considered feedback modalities. As a point of contact, in both studies subjective evaluation highlighted a general preference for auditory feedback over vibrotactile feedback.

Using audio and vibrotactile feedback simultaneously did not improve performance: The results were comparable to those of audio feedback alone. Indeed, 5 participants declared that, in the combined audio-tactile condition, they had focused on the audio feedback alone. Moreover, while 8 participants found the combined condition helpful, 5 found it confusing.

An hypothesis that was confirmed in [181] is that the presence of visual feedback practically nullifies the benefits of multisensory feedback: In such experiment, when visual feedback was available, the performances were very similar regardless of simultaneous presence of other feedback modes, and far superior to the "non-visual only" feedback modes. Since the present experiment showed the prevalence of auditory feedback over vibrotactile, we may hypothesize a hierarchy for steering task scenarios, in which visual information prevails over audio, which in turn overcomes tactile cues.

Although vibration feedback was designed to prevent selective adaptation, nonetheless desensitization took place to some extent, possibly due to the relatively high intensity and continuous character of the vibratory cues. Indeed, since several participants reported either desensitization or difficulties in sensing vibrations, it is possible that individual level adjustments before each experimental session may result in better acceptance of vibratory feedback, and even improved performance. Such intensity matching, however, is impractical in user interfaces, where each user may be expected to set the auditory and vibratory intensity levels in a comfort zone, with no further joint calibration.

Concerning the exerted force, no significant correlation was found with performance or experimental conditions. Nevertheless, participants exerted more force under vibrotactile feedback conditions. Since vibrotactile perception is enhanced for increasing finger pressing forces [164], we suppose that participants instinctively increased their pressing force aiming at maximizing tactile sensation. Consistently with this interpretation, a slight increase in force was recorded when participants were off-track and searching for the correct path.

Overall, as previously reported in [181], the present experiment confirms that the use of strictly affirmative vibrotactile cues as guidance in steering tasks is not trivial, even when targeting participants within a limited age range, meaning without considering the progressive deterioration in sensitivity and neural elaboration of stimuli that is associated with aging [226]. However, the following considerations must be taken into account:

- It is important to stress once again that non-visual conditions are common in real life scenarios. Moreover, the audio channel, which is the second most used sensory channel, can be impaired as well, due to noisy environments or concurrent auditory tasks;

- The paths proposed in this experiment were much longer and more complex than what is commonly required by human-machine interfaces (e.g., the navigation of a menu requires movements that are short, discrete, and rectilinear);

- The accuracy in the task was similar among all the (non-visual) conditions.

As a consequence, there seems to be the possibility to devise optimized vibrotactile interfaces that would enable users to perform simple steering tasks in a sufficiently effective manner.

Such hypothesis must confront the participants' opinions that were collected. The general feeling was that of a demanding task, either physically or mentally, or both. It is likely that the physical fatigue might be addressed with the above mentioned adoption of new, more effective haptic technologies, and with a more accurate customization of the intensity levels. Indeed, it would be advisable to avoid the latter, since a real life application is likely not to provide time for performing such procedure before every interaction.

Regarding the concentration demand, it is to be ascertained to what degree it would be decreased by the adoption of better interfaces. In particular, would path following tasks in non-visual conditions become suitable for secondary tasks, such as in the IVIS scenario depicted in Section 6.1. To date, complex auditory tasks were shown to degrade driving performance critically [25]. Yet, studies concerning such performance when the secondary task is based on visual feedback confirmed its safety implications [118], thus prompting further investigation for such scenario.

6.6 Conclusions and future work

Path following tasks are being investigated for their multiple application scenarios. The impact of non-visual feedback modalities and different path shapes, however, is more rarely addressed. The present experiment represented a preliminary assessment of the relevance of each factor with respect to the user's performance. Its natural continuation would aim at understanding what features of such factors are the most relevant and to what quantifiable extent, and how factors interact.

The main conclusion drawn from this experiment is that different non-visual feedback modalities did not affect accuracy, yet they affected speed. In particular, vibrotactile feedback caused slower gestures than audio feedback. Moreover, when in presence of audio cues, vibrotactile feedback became irrelevant to the performance. Combining these results with the post-hoc reports, which highlighted a strong preference for audio feedback, it can be argued that vibrotactile cues suffered from the inter-individual sensitivity differences. However, despite the taken precautions, we cannot exclude the impact of selective sensory adaptation, possibly resulting in desensitization.

The shape of the path was relevant to the task accuracy: Pronounced curves seemingly affected trajectory errors. Although noise in trajectories prevented accurate comparisons with the predictions of the steering law and the $^2/_3$ power law, such models seem not to hold for steering tasks under non-visual conditions. The cause resides in the exploratory nature of the gestures when a path is invisible, leading to irregular trajectories.

The forces exerted by the participants with their finger over the surface seem not to be related to performance. However, they were considerably higher under vibrotactile conditions, and slightly higher when participants were outside of the track.

In the light of such findings, further research may gain from employing more advanced technology for vibrotactile feedback [17, 87], aiming at decreasing the impact of the aforementioned perceptual issues. With minimal modifications to the current experimental setup, it is possible to repeat the experiment using a negative feedback strategy, and compare the performance to the present results.

The effect of feedback differentiation (e.g. the use of different intensity levels, or frequency content) may be assessed as well. For instance, more intense feedback might be provided when the user is closer to the tunnel edges.

Chapter 7

Conclusions

In this thesis I assembled an overview of the most common forms of multisensory feedback, namely those involving haptics and audition. The purpose was to provide a unifying framework for the numerous aspects, issues, and possibilities that must be taken into account when designing a multisensory interaction. Such framework represents both the motivation and a tool for analyzing the results of my research, consisting of two experiments concerning multisensory feedback.

Albeit similar at a first glance, the two experiments presented several relevant differences: the use of tools versus free-hand gesture, free exploration versus goal-oriented task. The setting as well, a public exhibit versus a controlled laboratory setup, clearly affected the interaction. All of such factors stress on the one hand the amount of variables an interaction designer must pay attention to, on the other the possibilities provided by using additional sensory channels to overcome such variables to achieve an effective, sometimes even enjoyable, interaction.

The introduction to haptic feedback was particularly ample for several reasons: First, the term “haptic” encloses different senses, which are separate as regard physiology, processing and function. Unlike audition, the involved sensory systems are not fully understood yet, nor is the possibility of interaction among them. For instance, we mentioned how a weight sensation may depend on both tactile and proprioceptive mechanisms, and it is clear how a cutaneous sensation of heat or force can turn into a pain sensation, which apparently is elaborated differently at cortical level. Second, the complex sensations that we receive when interacting with real objects are difficult to reproduce in a digital setting. As a result, concurrently with a constant research for technological improvement, alternative approaches are investigated to deceive the senses and achieve a plausible percept nonetheless.

Concerning audition, I focused on the discipline of Sonic Interaction Design, in that it arguably represents the most effective and holistic approach to the design of auditory feedback in interactive settings. Notwithstanding the theoretical foundations for such discipline, grounded in multiple fields such as physiology, psychology, cultural studies, sound and music computing etc., the practice is usually fragmented and heterogeneous among designers. Indeed, the invisible nature of sound is the main cause for the lateness in the creation and the general adoption of standardized procedures and tools supporting such activity. However, an improved awareness of the designer's needs – achieved through workshops and led by a "research through design" approach – and the analysis of the possible use of the voice as sound sketching tool are promising steps forward the strengthening of SID practices.

The integration of different sensory modalities encompasses several issues as well, carrying the risk of detrimental effects to the interaction: First, an effective merging requires the calibration of several factors. Second, the use cases need to be evaluated, in that critical tasks may suffer from attention division among the sensory channels.

The experimentation highlighted the limitations inherent in the use of commonly available technologies. Nonetheless, especially in the texture exploration scenario, it became apparent that an acceptable impression can be achieved with a proper use of pseudo-haptics and tactile illusions, even when a faithful reproduction of real stimuli is not achievable. In particular, the novelty of the vibrotactile augmentation was welcomed by the participants as a feature that increased realism. The path following task, however, is more critical: When performance is relevant, the perceptual issues (e.g. difficulty in perceiving a stimulus, or conversely desensitization due to excessive stimulation) become more important to the participants, thus possibly leading to concentration strain or even frustration. In particular, the use of vibrotactile feedback requires careful design and evaluation. Indeed, in the path following task the vibrotactile feedback caused longer task completion times than the auditory feedback. Besides, the coexistence of the two non-visual feedback modalities did not yield any improvement in the performance.

The auditory section of the simulations was indeed the most successful among the participants, in that sounds were more clearly perceivable and recognizable than vibrotactile stimuli. Yet, interactions often take place in noisy environments, and auditory attention is usually as much divided among different stimuli as the visual attention is. Moreover, the omnipresent mobile devices do not provide adequate audio systems to bypass such issues by means of sound synthesis and ad-hoc modulation.

Indeed, innovative haptic technologies are not ready for use yet. Most of the available interfaces to date consist of a flat surface, with consequent ergonomic limitations. Prototypes of interactive surfaces based on electrovibration or ultrasonic waves in substitution of mechanical vibration are already available, yet relevant issues need to be addressed (e.g. the impact of finger moisture and sonic spillage respectively), and multi-touch support is not achievable to date. The use of pneumatic systems beyond theme park augmentation is still to be assessed, in that the required mechanical parts pose a limit to their miniaturization. Yet, even when using more traditional technologies, a better knowledge about the human sensorimotor system and brain processing is likely to lead to improvements in the interaction. In such perspective, there are several open questions: For instance, the inter-individual differences in tactile sensitivity are usually not addressed in a systematic manner. Moreover, the reasons behind the variability of pseudo-haptic effects are still to be clarified.

All of these open questions, concerning different disciplines, directly affect the design and the evaluation of digital interfaces. Human-computer interaction researchers have the task of applying the discipline's general principles to a continuously changing situation concerning both the insight we have achieved about how we sense, and the technologies that are available to instantiate such knowledge. Likewise, general models of interaction may require to be corrected or expanded in the light of new interactive scenarios made possible by innovative technologies, or generated by new necessities from the users.

Chapter 8

Appendix A

8.1 Experiment 2: statistical analysis

8.1.1 Mann-Whitney test on subjective factors

Table 8.1: Task completion speed

	z value	**p value**
Gender	-0.956	0.3390
Manual skills	-1.312	0.1894
Physical fatigue	-1.039	0.2987
Mental fatigue	-0.748	0.4545

Table 8.2: Exerted force

	z value	**p value**
Gender	-0.665	0.5060
Manual skills	-1.312	0.1894
Physical fatigue	-1.455	0.1457
Mental fatigue	-0.836	0.4032

Table 8.3: Trajectory errors

	z value	**p value**
Gender	-1.580	0.1142
Manual skills	-0.042	0.9662
Physical fatigue	-1.039	0.2987
Mental fatigue	-0.792	0.4284

8.1.2 Mean values for main and derived variables

Table 8.4: Average task completion times (in ms).

	Path 0	**Path 1**	*mean*
A	33072	44278	38675
T	40448	53920	47184
A+T	31240	48077	39658.5
mean	34920	48758.33	41839.17

Table 8.5: Average distances (in mm).

	Path 0	**Path 1**	*mean*
A	905.62	1218.8	1062.21
T	998.68	1347.8	1173.24
A+T	901.92	1266.6	1084.26
mean	935.41	1277.73	1106.57

Table 8.6: Average exerted force (in N).

	Path 0	**Path 1**	*mean*
A	2.056	2.037	2.046
T	2.459	2.480	2.469
A+T	2.119	2.156	2.137
mean	2.211	2.224	2.218

Table 8.7: Average gesture speed (in m/s).

	Path 0	**Path 1**	*mean*
A	0.031	0.030	0.030
T	0.028	0.027	0.028
A+T	0.032	0.030	0.031
mean	0.030	0.029	0.030

Table 8.8: Average task completion speed (in m/s).

	Path 0	**Path 1**	*mean*
A	0.020	0.016	0.018
T	0.016	0.014	0.015
A+T	0.021	0.015	0.018
mean	0.019	0.015	0.017

Table 8.9: Average time-related accuracy (in % of time spent on-track).

	Path 0	**Path 1**	*mean*
A	72.6	61.0	66.8
T	71.7	63.2	67.4
A+T	72.5	61.6	67.1
mean	72.3	62.0	67.1

Table 8.10: Average trajectory error (in mm).

	Path 0	**Path 1**	*mean*
A	2.097	4.020	3.059
T	2.424	3.806	3.115
A+T	2.140	3.811	2.976
mean	2.220	3.879	3.050

8.1.3 Correlations among the variables

Note: The correlation coefficients range from -1 to 1, where values close to 1 indicate a positive linear relationship between the data columns, values close to -1 indicate a negative linear relationship (anticorrelation), and values close to or equal to 0 suggest there is no linear relationship between the data.

Table 8.11: Correlations

	T	F	D	G S	C S	T E
Completion Time	1	0.0275	0.6063	-0.6150	-0.8615	-0.3741
Force	0.0275	1	0.0373	-0.0109	-0.0371	-0.0160
Distance	0.6063	0.0373	1	0.1317	-0.4768	-0.3194
Gesture Speed	-0.6150	-0.0109	0.1317	1	0.7874	0.2779
Completion Speed	-0.8615	-0.0371	-0.4768	0.7874	1	0.5201
Trajectory Error	-0.3741	-0.0160	-0.3194	0.2779	0.5201	1

8.2 Experiment 2: participants' comments in the debriefing phase

- “There is need for fade in/out for the feedback, vibration at first is not very noticeable, but then the intensity was just right because too much makes the finger tickle. The task was not boring, rather it was challenging”;
- “I got lost at times, couldn't find the path again”;
- “At first I didn't feel the vibration at all. Only sometimes. Nonetheless I feel better with T than with A, generally. Sometimes I thought that the path was always the same”;
- “T is hard. On the whole it was exhausting, it demanded concentration. Also, I grew impatient because the paths seemed of the same kind more or less”;
- “The task was easy because of the limited directions you can take, otherwise it would have been harder”;
- “I had more fun with T, maybe it was easier”;
- “It was fun but exhausting. The T was subtle, sometimes I couldn't tell if it was there or not”;

- "I think there were about 8 different paths. I felt like the path was made of steps";
- "On A I was expecting changes in sound. A+T is disturbing because you don't know which to follow. Concentration wears off. It would have been better to have fade in/out when going on/off track";
- "It was hard to switch modes from A to T. The task required concentration. I had to swap fingers sometimes";
- "I had the impression that the track was moving, so I had to wait for a moment for it to settle. It was frustrating when I was losing the track, so I pressed more when in doubt. I had to change finger after a while";
- "I tried to draw the track in my mind, but that led to wrong assumptions. The attention is usually higher towards the end of the track";
- "It's hard to get back when losing the track. T requires more concentration. With A it looks easier to foresee the track. Are there any repeating paths? In A+T I focused on A";
- "Fun, but it requires concentration";
- "I feel like it's made of steps";
- "Difficult. It demanded concentration. When pressing more it feels like it doesn't work well, so I have to relieve it a bit. Sometimes I had the impression to feel the correct line underneath the finger";
- "I was trying to memorize the paths. Are some the same? I think there were about 4 different paths";
- "I tend to press harder when searching for the path. When in T, the finger keeps vibrating after task completion. I felt like the path was made of steps";
- "It was boring, sometimes frustrating. I'm usually impatient";
- "T is feeble. Relaxing. It is not boring, but tiring: concentration wears off in the end. Are there applications for the blind? I pressed more for feeling more and having more stability";
- "T is very feeble, requires more concentration. There are some failures that make the feedback twitch, which is irritating";

- “T is tricky, requires much concentration”;
- “T gave the impression of having more surface to touch, so it gave more security. Going fast made it harder to find the path”;
- “Sometimes I was in a flow, while sometimes I lost concentration. It worked better when I didn’t overthink. Sometimes it’s a bit frustrating not to see any improvement. Sometimes I feel I’m pushing the sound/vibration, instead of being guided by it. Sometimes I felt like a string I had to follow, but then it was cut and I had to find the other part”;
- “It is unusual to me. Sometimes I felt like “is the vibration there or am I imagining it?”. It was exhausting for the concentration”;
- “It was very annoying, frustrating. It required a lot of concentration. After a while I lost sense of position. The stuttering is annoying”;
- “Difficult. Using the volume to denote position could be useful. You need to get accustomed to the interface and to the posture”;
- “It was much better with A. With sound, the space feels ten times bigger. It is like drawing with one stroke”;
- “It was easier with A. T is more problematic towards the end of the task: it becomes hard to understand whether the vibration is there or not”;
- “A is clearer, but with T the information can be better transferred to the movement of the finger. Fun. Blindfolded: sometimes am I going in the right direction? T required to pay more attention, especially in the beginning. Sometimes T seems uneven.”

Bibliography

[1] Accot, J., and Zhai, S. Beyond fitts' law: Models for trajectory-based hci tasks. In *Proc. of the ACM SIGCHI Conf. on Human Factors in Computing Systems* (New York, NY, USA, 1997), CHI '97, ACM, pp. 295–302.

[2] Accot, J., and Zhai, S. Scale effects in steering law tasks. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (New York, NY, USA, 2001), CHI '01, ACM, pp. 1–8.

[3] Accot, J., and Zhai, S. More than dotting the i's—foundations for crossing-based interfaces. In *Proceedings of the SIGCHI Conf. on Human Factors in Computing Systems* (Minneapolis, MN, USA, 2002), ACM, pp. 73–80.

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

[5] Alary, F., Duquette, M., Goldstein, R., Chapman, C. E., Voss, P., Buissonnière-Ariza, V. L., and Lepore, F. Tactile acuity in the blind: A closer look reveals superiority over the sighted in some but not all cutaneous tasks. *Neuropsychologia 47*, 10 (2009), 2037 – 2043.

[6] Altinsoy, M. E. The effect of spatial disparity on the integration of auditory and tactile information. In *International Workshop on Haptic and Audio Interaction Design* (2010), Springer, pp. 20–25.

[7] Altmann, E. M., and Gray, W. D. An integrated model of cognitive control in task switching. *Psychological review 115*, 3 (2008), 602.

[8] Aramaki, M., Gondre, C., Kronland-Martinet, R., Voinier, T., and Ystad, S. Thinking the sounds: an intuitive control of an

impact sound synthesizer. In *International Conference on Auditory Display (ICAD'09)* (Copenhague, Denmark, May 2009), pp. 119–124.

[9] Ardoin, W.-J. V., and Ferris, T. K. Investigating redundant encoding methods for tactile messaging in multitask scenarios. *IEEE Transactions on Human-Machine Systems 46*, 3 (2016), 451–459.

[10] Asamura, N., Yokoyama, N., and Shinoda, H. Selectively stimulating skin receptors for tactile display. *IEEE Computer Graphics and Applications 18*, 6 (1998), 32–37.

[11] Avanzini, F., and Rocchesso, D. Controlling material properties in physical models of sounding objects. In *Proceedings of International Computer Music Conference (ICMC)* (La Habana, Cuba, 2001).

[12] Avanzini, F., and Rocchesso, D. Modeling collision sounds: Nonlinear contact force. In *Proc. COST-G6 Conf. Digital Audio Effects (DAFx-01)* (2001), pp. 61–66.

[13] Avanzini, F., Serafin, S., and Rocchesso, D. Interactive simulation of rigid body interaction with friction-induced sound generation. *IEEE Trans. Speech and Audio Proc. 13*, 5 (Sept 2005), 1073–1081.

[14] Baldan, S., and Delle Monache, S. His engine's voice: towards a vocal sketching tool for synthetic engine sounds. In *Proceedings of the XX Colloquium on Musical Informatics (CIM)* (Rome, Italy, 2014), Università IUAV di Venezia.

[15] Baldan, S., Delle Monache, S., and Rocchesso, D. The sound design toolkit. *Software X in press* (2017).

[16] Barnard, P. J., and Teasdale, J. D. Interacting cognitive subsystems: A systemic approach to cognitive-affective interaction and change. *Cognition and Emotion 5*, 1 (1991), 1–39.

[17] Bau, O., Poupyrev, I., Israr, A., and Harrison, C. Teslatouch: Electrovibration for touch surfaces. In *UIST'2010* (2010), ACM.

[18] Bennequin, D., Fuchs, R., Berthoz, A., and Flash, T. Movement timing and invariance arise from several geometries. *PLoS computational biology 5*, 7 (2009).

[19] Bensmaïa, S. J., Leung, Y. Y., Hsiao, S. S., and Johnson, K. O. Vibratory adaptation of cutaneous mechanoreceptive afferents. *Journal of Neurophysiology 94*, 5 (Nov. 2005), 3023–3036.

[20] Bergmann Tiest, W. M. Tactual perception of material properties. *Vision research 50*, 24 (2010), 2775–2782.

[21] Bergmann Tiest, W. M., and Kappers, A. M. Haptic perception of force. *Scholarpedia 10*, 4 (2015), 32732. revision #151740.

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

[23] Bettini, A., Marayong, P., Lang, S., Okamura, A. M., and Hager, G. D. Vision-assisted control for manipulation using virtual fixtures. *IEEE Transactions on Robotics 20*, 6 (dec 2004), 953 – 966.

[24] Blake, D. T., Hsiao, S. S., and Johnson, K. O. Neural coding mechanisms in tactile pattern recognition: The relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness. *Journal of Neuroscience 17* (1997), 7480–7489.

[25] Blanco, M., Biever, W. J., Gallagher, J. P., and Dingus, T. A. The impact of secondary task cognitive processing demand on driving performance. *Accident Analysis & Prevention 38*, 5 (2006), 895 – 906.

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

[27] Borenstein, J. The navbelt-a computerized multi-sensor travel aid for active guidance of the blind. *Proceedings of the CSUN's Fifth Annual Conference on Technology and Persons with Disabilities* (mar 1990), 107 – 116.

[28] Borin, G., De Poli, G., and Rocchesso, D. Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems. *IEEE Transactions on Speech and Audio Processing 8*, 5 (2000), 597–605.

[29] Borst, J. P., Taatgen, N. A., and van Rijn, H. The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, memory, and cognition 36*, 2 (2010), 363.

[30] Bresciani, J.-P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., and Kheddar, A. Feeling what you hear: auditory signals can modulate tactile tap perception. *Experimental Brain Research 162*, 2 (2005), 172–180.

[31] Bresin, R., Delle Monache, S., Fontana, F., Papetti, S., Polotti, P., and Visell, Y. Auditory feedback through continuous control of crumpling sound synthesis. In *Proc. CHI workshop on sonic interaction design* (2008), pp. 23–28.

[32] Cadoz, C., Lisowski, L., and Florens, J.-L. A modular feedback keyboard design. *Computer music journal 14*, 2 (1990), 47–51.

[33] Caldwell, D. G., Tsagarakis, N., and Giesler, C. An integrated tactile/shear feedback array for stimulation of finger mechanoreceptor. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on* (1999), vol. 1, IEEE, pp. 287–292.

[34] Callan, D. E., Jones, J. A., Munhall, K., Kroos, C., Callan, A. M., and Vatikiotis-Bateson, E. Multisensory integration sites identified by perception of spatial wavelet filtered visual speech gesture information. *Journal of Cognitive Neuroscience 16*, 5 (2004), 805–816.

[35] Campbell, C. S., Zhai, S., May, K. W., and Maglio, P. P. What you feel must be what you see: Adding tactile feedback to the trackpoint. In *Proceedings of INTERACT'99: 7th IFIP Conference on Human Computer Interaction* (1999), IOS Press, pp. 383–390.

[36] Cartwright, M., and Pardo, B. Vocalsketch: Vocally imitating audio concepts. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2015), CHI '15, ACM, pp. 43–46.

[37] Cassinelli, A., Reynolds, C., and Ishikawa, M. Augmenting spatial awareness with haptic radar. In *2006 10th IEEE International Symposium on Wearable Computers* (Oct 2006), pp. 61–64.

[38] Chambers, M. R., Andres, K. H., Duering, M. v., and Iggo, A. The structure and function of the slowly adapting type ii mechanoreceptor in hairy skin. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences 57*, 4 (1972), 417–445.

[39] Cirio, G., Li, D., Grinspun, E., Otaduy, M. A., and Zheng, C. Crumpling sound synthesis. *ACM Transactions on Graphics (TOG) 35*, 6 (2016), 181.

[40] Cohen, R. F., Haven, V., Lanzoni, J. A., Meacham, A., Skaff, J., and Wissell, M. Using an audio interface to assist users who are visually impaired with steering tasks. In *Proc. of the 8th Int. ACM SIGACCESS Conf. on Computers and Accessibility* (New York, NY, USA, 2006), Assets '06, ACM, pp. 119–124.

[41] Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Kronland-Martinet, R., and Ystad, S. Navigating in a space of synthesized interaction-sounds: Rubbing, scratching and rolling sounds. In *16th International Conference on Digital Audio Effects (DAFx)* (2013), pp. 202–209.

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

[43] Congedo, M., Lécuyer, A., and Gentaz, E. The influence of spatial delocation on perceptual integration of vision and touch. *Presence: Teleoperators and Virtual Environments 15*, 3 (2006), 353–357.

[44] Connor, C. E., Hsiao, S. S., Phillips, J. R., and Johnson, K. O. Tactile roughness: neural codes that account for psychophysical magnitude estimates. *Journal of Neuroscience 10*, 12 (12 1990), 3823–2836.

[45] Craig, J. C. Grating orientation as a measure of tactile spatial acuity. *Somatosensory & Motor Research 16*, 3 (1999), 197–206.

[46] Croce, D., Giarré, L., Rosa, F. G. L., Montana, E., and Tinnirello, I. Enhancing tracking performance in a smartphone-based navigation system for visually impaired people. In *2016 24th Mediterranean Conference on Control and Automation (MED)* (June 2016), pp. 1355–1360.

[47] Dandekar, K., Raju, B. I., and Srinivasan, M. A. 3-D finite-element models of human and monkey fingertips to investigate the mechanics of tactile sense. *J. of Biomechanical Engineering 125*, 5 (Oct. 2003), 682–691.

[48] Darian-Smith, I., Johnson, K. O., LaMotte, C., Shigenaga, Y., Kenins, P., and Champness, P. Warm fibers innervating palmar and digital skin of the monkey: responses to thermal stimuli. *Journal of Neurophysiology 42*, 5 (1979), 1297–1315.

[49] de Jesus Oliveira, V. A., Brayda, L., Nedel, L., and Maciel, A. Experiencing guidance in 3d spaces with a vibrotactile head-mounted display. In *2017 IEEE Virtual Reality (VR)* (March 2017), pp. 453–454.

[50] De Niear, M. A., Koo, B., and Wallace, M. T. Multisensory perceptual learning is dependent upon task difficulty. *Experimental Brain Research 234*, 11 (2016), 3269–3277.

[51] De Niear, M. A., Noel, J.-P., and Wallace, M. T. The impact of feedback on the different time courses of multisensory temporal recalibration. *Neural Plasticity 2017* (2017), 12.

[52] Del Piccolo, A., Delle Monache, S., Rocchesso, D., Papetti, S., and Mauro, D. A. To "sketch-a-scratch". In *Proceedings of the 12th Sound and Music Computing Conference (SMC15)* (2015).

[53] Del Piccolo, A., and Rocchesso, D. Non-speech voice for sonic interaction: a catalogue. *Journal on Multimodal User Interfaces 11*, 1 (2017), 39–55.

[54] Del Piccolo, A., Rocchesso, D., and Papetti, S. Path following in non-visual conditions. *submitted for review* (2017).

[55] Delle Monache, S., Rocchesso, D., and Papetti, S. Sketch a scratch. In *Eight International Conference on Tangible, Embedded and Embodied Interaction* (2014), TEI '14.

[56] Dennerlein, J. T., Martin, D. B., and Hasser, C. Force-feedback improves performance for steering and combined steering-targeting tasks. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (New York, NY, USA, 2000), CHI '00, ACM, pp. 423–429.

[57] DiFranco, D. E., Beauregard, G. L., and Srinivasan, M. A. The effect of auditory cues on the haptic perception of stiffness in virtual environments. In *Proceedings of the ASME Dynamic Systems and Control Division* (1997), vol. 61, American Society of Mechanical Engineers, pp. 17–22.

[58] Dillon, Y. K., Haynes, J., and Henneberg, M. The relationship of the number of meissner's corpuscles to dermatoglyphic characters and finger size. *Journal of Anatomy 199*, 5 (2001), 577–584.

[59] Dominjon, L., Lécuyer, A., Burkhardt, J.-M., Richard, P., and Richir, S. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *Virtual Reality, 2005. Proceedings. VR 2005. IEEE* (2005), IEEE, pp. 19–25.

[60] Dounskaia, N. Kinematic invariants during cyclical arm movements. *Biological cybernetics 96*, 2 (2007), 147–163.

[61] Driver, J., and Spence, C. Multisensory perception: beyond modularity and convergence. *Current biology 10*, 20 (2000), R731–R735.

[62] Drury, C. G. Movements with lateral constraint. *Ergonomics 14*, 2 (1971), 293–305. PMID: 5093722.

[63] Duizer, L. A review of acoustic research for studying the sensory perception of crisp, crunchy and crackly textures. *Trends in Food Science & Technology 12*, 1 (2001), 17 – 24.

[64] Eimer, M. Multisensory integration: How visual experience shapes spatial perception. *Current Biology 14*, 3 (2004), R115 – R117.

[65] Ekman, G., and Hosman, J. amd Lindstrom, B. Roughness, smoothness, and preference: A study of quantitative relations in individual subjects. *Journal of Experimental Psychology 70*, 1 (July 1965), 18–26.

[66] Ekman, I., and Rinott, M. Using vocal sketching for designing sonic interactions. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (New York, NY, USA, 2010), DIS '10, ACM, pp. 123–131.

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

[68] Fahle, M. Perceptual learning: specificity versus generalization. *Current Opinion in Neurobiology 15*, 2 (2005), 154 – 160. Cognitive neuroscience.

[69] Farnell, A. *Designing sound.* Mit Press, 2010.

[70] Ferguson, S., and Cabrera, D. Vertical localization of sound from multiway loudspeakers. *J. Audio Eng. Soc 53*, 3 (2005), 163–173.

[71] Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J. of experimental psychology 47*, 6 (June 1954), 381–391.

[72] Flash, T., and Hogan, N. The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience 5*, 7 (1985), 1688–1703.

[73] Fleming, D., and Griffin, M. A study of the subjective equivalence of noise and whole-body vibration. *Journal of Sound and Vibration 42*, 4 (1975), 453 – 461.

[74] Fontana, F., and Bresin, R. Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. In *Proc. Colloquium on Musical Informatics* (2003), pp. 109–114.

[75] Fontana, F., Papetti, S., Civolani, M., del Bello, V., and Bank, B. An exploration on the influence of vibrotactile cues during digital piano playing. In *Proceedings of the International Conference on Sound and Music Computing (SMC), Padua, Italy* (2011).

[76] Forlines, C., and Balakrishnan, R. Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2008), CHI '08, ACM, pp. 1563–1572.

[77] Franinovic, K., and Serafin, S., Eds. *Sonic Interaction Design*. The MIT Press, 2013.

[78] Freeman, A. W., and Johnson, K. O. Cutaneous mechanoreceptors in macaque monkey: temporal discharge patterns evoked by vibration, and a receptor model. *The Journal of Physiology 323*, 1 (1982), 21–41.

[79] Fritz, C., Curtin, J., Poitevineau, J., Morrel-Samuels, P., and Tao, F.-C. Player preferences among new and old violins. *Proceedings of the National Academy of Sciences 109*, 3 (2012), 760–763.

[80] Fukumoto, M., and Sugimura, T. Active click: Tactile feedback for touch panels. In *Proceedings of CHI 2001* (2001), ACM, pp. 121–122.

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

[82] Gerrett, N., Ouzzahra, Y., Coleby, S., Hobbs, S., Redortier, B., Voelcker, T., and Havenith, G. Thermal sensitivity to warmth during rest and exercise: a sex comparison. *European Journal of Applied Physiology 114*, 7 (2014), 1451–1462.

[83] Gerrett, N., Ouzzahra, Y., Redortier, B., Voelcker, T., and Havenith, G. Female thermal sensitivity to hot and cold during rest and exercise. *Physiology & Behavior 152, Part A* (2015), 11 – 19.

[84] Gescheider, G. A. Some comparisons between touch and hearing. *IEEE Transactions on Man-Machine Systems 11*, 1 (March 1970), 28–35.

[85] Gibson, J. J. Observations on active touch. *Psychological review 69* (1962), 477–491.

[86] Giordano, M., Hattwick, I., Franco, I., Egloff, D., Frid, E., Lamontagne, V., TeZ, C., Salter, C., and Wanderley, M. Design and implementation of a whole-body haptic suit for "ilinx", a multisensory art installation. In *12th International Conference on Sound and Music Computing (SMC-15), Ireland, July 30, 31 & August 1, 2015* (2015), vol. 1, Maynooth University, pp. 169–175.

[87] Giraud, F., Amberg, M., and Lemaire-Semail, B. Merging two tactile stimulation principles: electrovibration and squeeze film effect. In *World Haptics* (2013), IEEE, pp. 199–203.

[88] Goldreich, D., and Kanics, I. M. Tactile acuity is enhanced in blindness. *Journal of Neuroscience 23*, 8 (2003), 3439–3445.

[89] Goodwin, A. Paradoxes in tactile adaptation. focus on "vibratory adaptation in cutaneous mechanoreceptive afferents" and "time-course of vibratory adaptation and recovery in cutaneous mechanoreceptive afferents". *Journal of neurophysiology 94*, 5 (2005), 2995–2996.

[90] Graziano, M. S. A., and Botvinick, M. M. How the brain represents the body: insights from neurophysiology and psychology. *Common Mechanisms in Perception and Action: Attention and Performance 19* (2002), 136–157.

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

[92] Gunther, E., and O'Modhrain, S. Cutaneous grooves: Composing for the sense of touch. *Journal of New Music Research 32*, 4 (2003), 369–381.

[93] Hachisu, T., and Fukumoto, M. Vacuumtouch: Attractive force feedback interface for haptic interactive surface using air suction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2014), CHI '14, ACM, pp. 411–420.

[94] Hahn, J. F. Low-frequency vibrotactile adaptation. *Journal of Experimental Psychology 78*, 4 (pt.1) (Dec. 1968), 655–659.

[95] Hashimoto, Y., Nakata, S., and Kajimoto, H. Novel tactile display for emotional tactile experience. In *Proceedings of the International Conference on Advances in Computer Enterntainment Technology* (New York, NY, USA, 2009), ACE '09, ACM, pp. 124–131.

[96] Hasser, C. J., and Weisenberger, J. M. Preliminary evaluation of a shape-memory alloy tactile feedback display. In *Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems* (1993), ASME Winter Annual Meeting, pp. 73–80.

[97] Hayward, V. A brief taxonomy of tactile illusions and demonstrations that can be done in a hardware store. *Brain Research Bulletin 75*, 6 (2008), 742 – 752. Special Issue: Robotics and Neuroscience.

[98] Helgason, P. Sound initiation and source types in human imitations of sounds. In *FONETIK 2014* (Stockholm, Sweden, 2014), Stockholms Universitet.

[99] Hoggan, E., and Brewster, S. Designing audio and tactile crossmodal icons for mobile devices. In *Proceedings of the 9th international conference on Multimodal interfaces* (2007), ACM, pp. 162–169.

[100] Hollins, M. Somesthetic senses. *Annual Review of Psychology 61* (1 2010), 243–271.

[101] Hollins, M., Bensmaïa, S., Karlof, K., and Young, F. Individual differences in perceptual space for tactile textures: Evidence from multidimensional scaling. *Perception & Psychophysics 62*, 8 (2000), 1534–1544.

[102] Hoshi, T., Takahashi, M., Iwamoto, T., and Shinoda, H. Non-contact tactile display based on radiation pressure of airborne ultrasound. *IEEE Transactions on Haptics 3*, 3 (July 2010), 155–165.

[103] Hötting, K., and Röder, B. Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological Science 15*, 1 (2004), 60–64.

[104] Howarth, H., and Griffin, M. The relative importance of noise and vibration from railways. *Applied Ergonomics 21*, 2 (1990), 129 – 134.

[105] Hug, D. Towards a hermeneutics and typology of sound for interactive commodities. In *Proc. CHI Workshop on Sonic Interaction Design* (2008), pp. 11–16.

[106] Israr, A., Kim, S.-C., Stec, J., and Poupyrev, I. Surround haptics: tactile feedback for immersive gaming experiences. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (2012), ACM, pp. 1087–1090.

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

[108] Jones, L. *Thermal Touch.* Atlantis Press, Paris, 2016, pp. 257–262.

[109] Jones, L. A. Perception of force and weight: Theory and research. *Psychological bulletin 100*, 1 (1986), 29.

[110] Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (New York, NY, USA, 2007), TEI '07, ACM, pp. 139–146.

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

[112] Kaber, D. B., Wright, M. C., and Sheik-Nainar, M. A. Investigation of multi-modal interface features for adaptive automation of a human–robot system. *International Journal of Human-Computer Studies 64*, 6 (2006), 527 – 540.

[113] Kaczmarek, K., Webster, J., Bach-y Rita, P., and Tompkins, W. Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering 38(1)* (1991), 1–16.

[114] Kennett, S., Eimer, M., Spence, C., and Driver, J. Tactile-visual links in exogenous spatial attention under different postures: Convergent evidence from psychophysics and erps. *Journal of Cognitive Neuroscience 13*, 4 (2001), 462–478.

[115] Kenshalo, D. R., and Scott, H. A. Temporal course of thermal adaptation. *Science 151*, 3714 (1966), 1095–1096.

[116] Keyson, D. V. Dynamic cursor gain and tactual feedback in the capture of cursor movements. *Ergonomics 40*, 12 (1997), 1287–1298. PMID: 9416013.

[117] Kim, G., and Kim, H. C. Designing of multimodal feedback for enhanced multitasking performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), ACM, pp. 3113–3122.

[118] Kim, H., Kwon, S., Heo, J., Lee, H., and Chung, M. K. The effect of touch-key size on the usability of in-vehicle information systems and driving safety during simulated driving. *Applied Ergonomics 45*, 3 (2014), 379 – 388.

[119] Kim, S.-C., Israr, A., and Poupyrev, I. Tactile rendering of 3d features on touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology* (2013), ACM, pp. 531–538.

[120] Kim, W., Tendick, F., and Stark, L. Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator. *IEEE Journal on Robotics and Automation 3*, 5 (October 1987), 418–425.

[121] Kitagawa, N., Zampini, M., and Spence, C. Audiotactile interactions in near and far space. *Experimental Brain Research 166*, 3 (2005), 528–537.

[122] Klatzky, R., and Lederman, S. Human haptics. In *New encyclopedia of neuroscience*, vol. 5. Elsevier Amsterdam, 2009, pp. 11–18.

[123] Klatzky, R. L., and Lederman, S. J. Multisensory texture perception. In *Multisensory object perception in the primate brain*, M. J. Naumer and J. Kaiser, Eds. Springer Verlag, 2010, pp. 211–230.

[124] Knibestöl, M., and Vallbo, A. B. Single unit analysis of mechanoreceptor activity from the human glabrous skin. *Acta physiol. scand.*, 80 (1970), 178–195.

[125] Köhler, W. Gestalt psychology (1929). *New York, NY: Liveright* (1947).

[126] Kratz, S., and Dunnigan, A. Thermotouch: Design of a high dynamic temperature range thermal haptic display. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2016), CHI EA '16, ACM, pp. 1577–1582.

[127] Kuiper, R. J., Heck, D. J., Kuling, I. A., and Abbink, D. A. Evaluation of haptic and visual cues for repulsive or attractive guidance in nonholonomic steering tasks. *IEEE Transactions on Human-Machine Systems 46*, 5 (2016), 672–683.

[128] Kulikov, S., MacKenzie, I. S., and Stuerzlinger, W. Measuring the effective parameters of steering motions. In *CHI '05 Ext. Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2005), CHI EA '05, ACM, pp. 1569–1572.

[129] Lacquaniti, F., Terzuolo, C., and Viviani, P. The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica 54*, 1 (1983), 115 – 130.

[130] Lam, T. M., Boschloo, H. W., Mulder, M., and van Paassen, M. M. Artificial force field for haptic feedback in uav teleoperation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 39*, 6 (Nov 2009), 1316–1330.

[131] Lank, E., and Saund, E. Sloppy selection: Providing an accurate interpretation of imprecise selection gestures. *Comput. Graph. 29*, 4 (Aug. 2005), 490–500.

[132] Lécuyer, A. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments 18*, 1 (Jan. 2009), 39–53.

[133] Lécuyer, A., Burkhardt, J.-M., and Etienne, L. Feeling bumps and holes without a haptic interface: the perception of pseudo-haptic textures. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2004), ACM, pp. 239–246.

[134] Lécuyer, A., Coquillart, S., and Coiffet, P. Simulating haptic information with haptic illusions in virtual environments. Tech. rep., DTIC Document, 2001.

[135] Lécuyer, A., Coquillart, S., Kheddar, A., Richard, P., and Coiffet, P. Pseudo-haptic feedback: can isometric input devices simulate force feedback? In *Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048)* (2000), pp. 83–90.

[136] Lederman, S. J., and Jones, L. A. Tactile and haptic illusions. *IEEE Transactions on Haptics 4*, 4 (October 2011), 273–294.

[137] Lee, J., Dietz, P., Leigh, D., Yerazunis, W., and Hudson, S. Haptic pen: A tactile feedback stylus for touch screens. In *UIST'2004* (2004), ACM, pp. 291–294.

[138] Lemaitre, G., and Heller, L. M. Auditory perception of material is fragile while action is strikingly robust. *The Journal of the Acoustical Society of America 131*, 2 (2012), 1337–1348.

[139] Lemaitre, G., Voisin, F., Scurto, H., Houix, O., Susini, P., Misdariis, N., and Bevilacqua, F. A large set of vocal and gestural imitations. Tech. rep., Deliverable 4.4.1 of the Project SkAT-VG, FP7-ICT-FET-618067, 2015.

[140] Leonard, J., and Cadoz, C. Physical modelling concepts for a collection of multisensory virtual musical instruments. In *New Interfaces for Musical Expression (NIME 2015)* (2015), pp. 150–155.

[141] Leung, A. Y., Wallace, M. S., Schulteis, G., and Yaksh, T. L. Qualitative and quantitative characterization of the thermal grill. *Pain 116*, 1 (2005), 26–32.

[142] Leung, Y. Y., Bensmaïa, S. J., Hsiao, S. S., and Johnson, K. O. Time-course of vibratory adaptation and recovery in cutaneous mechanoreceptive afferents. *Journal of neurophysiology 94*, 5 (2005), 3037–3045.

[143] Libouton, X., Barbier, O., Berger, Y., Plaghki, L., and Thonnard, J.-L. Tactile roughness discrimination of the finger pad relies primarily on vibration sensitive afferents not necessarily located in the hand. *Behavioural Brain Research 229*, 1 (2012), 273 – 279.

[144] Libouton, X., Barbier, O., Plaghki, L., and Thonnard, J.-L. Tactile roughness discrimination threshold is unrelated to tactile spatial acuity. *Behavioural Brain Research 208*, 2 (2010), 473 – 478.

[145] Loomis, J. M., and Lederman, S. *Tactual perception, in Handbook of Perception and Human Performances*, vol. 2. Wiley, New York, 1986.

[146] Loukopoulos, L. D., Dismukes, R., and Barshi, I. The perils of multitasking. *AeroSafety World 4*, 8 (2009), 18–23.

[147] Macaluso, E., Frith, C. D., and Driver, J. Modulation of human visual cortex by crossmodal spatial attention. *Science 289*, 5482 (2000), 1206–1208.

[148] Marchetto, E., and Peeters, G. A set of audio features for the morphological description of vocal imitations. In *Proceedings of the 18th Int. Conference on Digital Audio Effects (DAFX'15)* (Trondheim, Norway, 2015).

[149] Massimino, M. J., and Sheridan, T. B. Variable force and visual feedback effects on teleoperator man/machine performance. In *Proceedings of the NASA Conference on Space Telerobotics* (jan 1989), vol. 1, pp. 89 – 98.

[150] McDonald, C. G., and Kuchenbecker, K. J. Dynamic simulation of tool-mediated texture interaction. In *World Haptics Conference (WHC), 2013* (Apr. 2013), IEEE, pp. 307–312.

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

[152] Mensvoort, K. V., Vos, P., Hermes, D. J., and Liere, R. V. Perception of mechanically and optically simulated bumps and holes. *ACM Transactions on Applied Perception (TAP) 7*, 2 (2010), 10.

[153] Merchel, S., and Altinsoy, E. Comparison of the perceived intensity of time-varying signals in the tactile and auditory domain. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (2016), vol. 253, Institute of Noise Control Engineering, pp. 6797–6802.

[154] Merchel, S., and Altinsoy, M. Cross-modality matching of loudness and perceived intensity of whole-body vibrations. In *International Conference on Haptic and Audio Interaction Design* (2010), Springer, pp. 1–9.

[155] Merchel, S., and Altinsoy, M. E. Vibratory and acoustical factors in multimodal reproduction of concert dvds. In *International Conference on Haptic and Audio Interaction Design* (2009), Springer, pp. 119–127.

[156] Michail, G., Dresel, C., Witkovskỳ, V., Stankewitz, A., and Schulz, E. Neuronal oscillations in various frequency bands differ between pain and touch. *Frontiers in human neuroscience 10*, 182 (2016).

[157] Milgram, P., Rastogi, A., and Grodski, J. J. Telerobotic control using augmented reality. In *Proceedings 4th IEEE International Workshop on Robot and Human Communication* (jul 1995), pp. 21 – 29.

[158] Navarra, J., Soto-Faraco, S., and Spence, C. Adaptation to audiotactile asynchrony. *Neuroscience letters 413 1* (2007), 72–76.

[159] Newell, F. N., Ernst, M. O., Tjan, B. S., and Bülthoff, H. H. Viewpoint dependence in visual and haptic object recognition. *Psychological science 12*, 1 (2001), 37–42.

[160] Nordahl, R., Serafin, S., and Timcenko, O. Contextualisation and evaluation of novel sonic interfaces using problem based learning. In *CHI 2008 Workshop on Sonic Interaction Design.* (2008).

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

[162] Okamura, A. M. Haptic feedback in robot-assisted minimally invasive surgery. *Current opinion in urology 19*, 1 (2009), 102.

[163] Papetti, S., Civolani, M., and Fontana, F. Rhythm'n'shoes: a wearable foot tapping interface with audio-tactile feedback. In *New Interfaces for Musical Expression (NIME 2011)* (2011), pp. 473–476.

[164] Papetti, S., Järveläinen, H., Giordano, B. L., Schiesser, S., and Fröhlich, M. Vibrotactile sensitivity in active touch: Effect of pressing force. *IEEE Transactions on Haptics 10*, 1 (Jan 2017), 113–122.

[165] Parsons, K. *Human thermal environments*, 2nd ed. Taylor & Francis, London, 2003.

[166] Peters, R. M., Hackeman, E., and Goldreich, D. Diminutive digits discern delicate details: Fingertip size and the sex difference in tactile spatial acuity. *The Journal of Neuroscience 29*, 50 (2009), 15756–15761.

[167] Pfeiffer, C., Mavroidis, C., Bar-Cohen, Y., and Dolgin, B. P. Electrorheological-fluid-based force feedback device. In *Proceedings of SPIE* (1999), vol. 3840, pp. 88–99.

[168] Phillips, J. R., Johansson, R. S., and Johnson, K. O. Representation of braille characters in human nerve fibres. *Experimental Brain Research 81*, 3 (1990), 589–592.

[169] Pitts, B., Riggs, S. L., and Sarter, N. Crossmodal matching: A critical but neglected step in multimodal research. *IEEE Transactions on Human-Machine Systems 46*, 3 (2016), 445–450.

[170] Plamondon, R., and Guerfali, W. The 2/3 power law: When and why? *Acta Psychologica 100*, 1 (1998), 85 – 96.

[171] Pusch, A., and Lécuyer, A. Pseudo-haptics: From the theoretical foundations to practical system design guidelines. In *Proceedings of the 13th International Conference on Multimodal Interfaces* (New York, NY, USA, 2011), ICMI '11, ACM, pp. 57–64.

[172] Quek, Z. F., Schorr, S. B., Nisky, I., Okamura, A. M., and Provancher, W. R. Augmentation of stiffness perception with a 1-degree-of-freedom skin stretch device. *IEEE Transactions on Human-Machine Systems 44*, 6 (2014), 731–742.

[173] Rashevsky, N. Mathematical biology of automobile driving iv. *The bulletin of mathematical biophysics 32*, 1 (Mar 1970), 71–78.

[174] Rath, M., Avanzini, F., Bernardini, N., Borin, G., Fontana, F., Ottaviani, L., and Rocchesso, D. An introductory catalog of computer-synthesized contact sounds, in real-time. In *Proc. Colloquium of Musical Informatics* (2003), pp. 103–108.

[175] Rath, M., and Rocchesso, D. Continuous sonic feedback from a rolling ball. *IEEE MultiMedia 12*, 2 (2005), 60–69.

[176] Ribeiro, F., and Oliveira, J. Aging effects on joint proprioception: the role of physical activity in proprioception preservation. *European Review of Aging and Physical Activity 4*, 2 (2007), 71–76.

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

[178] Robles-De-La-Torre, G. The importance of the sense of touch in virtual and real environments. *IEEE Multimedia 13*, 3 (2006), 24–30.

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

[180] Rocchesso, D. *Explorations in sonic interaction design.* Logos Berlin, 2011.

[181] Rocchesso, D., Delle Monache, S., and Papetti, S. Multisensory texture exploration at the tip of the pen. *Special Issue on Data sonification and sound design in interactive systems - International Journal of Human-Computer Studies (IJHCS)* (2015).

[182] Rocchesso, D., Lemaitre, G., and Grassi, M. Evidence for a spatial bias in the perception of sequences of brief tones. *The Journal of the Acoustical Society of America 133*, 5 (2013), EL346–EL350.

[183] Rocchesso, D., Lemaitre, G., Susini, P., Ternström, S., and Boussard, P. Sketching sound with voice and gesture. *ACM Interactions 22*, 1 (2015), 38–41.

[184] Rocchesso, D., Mauro, D. A., and Delle Monache, S. miMic: The microphone as a pencil. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction* (Eindhoven, The Netherlands, 2016), TEI 2016, ACM.

[185] Rocchesso, D., Polotti, P., and Delle Monache, S. Designing continuous sonic interaction. *International Journal of Design 3*, 3 (2009).

[186] Rocchesso, D., Serafin, S., Behrendt, F., Bernardini, N., Bresin, R., Eckel, G., Franinovic, K., Hermann, T., Pauletto, S., Susini, P., et al. Sonic interaction design: sound, information and experience. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems* (2008), ACM, pp. 3969–3972.

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

[188] Roffler, S. K., and Butler, R. A. Factors that influence the localization of sound in the vertical plane. *The Journal of the Acoustical Society of America 43*, 6 (1968), 1255–1259.

[189] Rosenberg, L. B. The use of virtual fixtures to enhance operator performance in time delayed teleoperation. Tech. rep., Armstrong Lab Wright-Patterson AFB OH Crew Systems Directorate, 1993.

[190] Rovan, J., and Hayward, V. Typology of tactile sounds and their synthesis in gesture-driven computer music performance. *Trends in Gestural Control of Music* (2000), 297 – 320.

[191] Saldaña, H. M., and Rosenblum, L. D. Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics 54*, 3 (1993), 406–416.

[192] Salvucci, D. D., Taatgen, N. A., and Borst, J. P. Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2009), ACM, pp. 1819–1828.

[193] Satava, R. M. Virtual reality, telesurgery, and the new world order of medicine. *Journal of Image Guided Surgery 1*, 1 (1995), 12–16. PMID: 9079422.

[194] Schumacher, M., Giordano, M., Wanderley, M. M., and Ferguson, S. Vibrotactile notification for live electronics performance: A prototype system. In *Proceedings of the International Symposium on Computer Music Multidisciplinary Research (CMMR)* (2013), pp. 516–525.

[195] Schürmann, M., Caetano, G., Jousmäki, V., and Hari, R. Hands help hearing: facilitatory audiotactile interaction at low sound-intensity levels. *The Journal of the Acoustical Society of America 115*, 2 (2004), 830–832.

[196] Seitz, A. R., and Dinse, H. R. A common framework for perceptual learning. *Current Opinion in Neurobiology 17*, 2 (2007), 148 – 153. Cognitive neuroscience.

[197] Shams, L., Kamitani, Y., and Shimojo, S. Illusions: What you see is what you hear. *Nature 408*, 6814 (2000), 788.

[198] Shams, L., Kamitani, Y., and Shimojo, S. Visual illusion induced by sound. *Cognitive Brain Research 14*, 1 (2002), 147–152.

[199] Shinohara, K., Yamauchi, N., Kawahara, S., and Tanaka, H. Takete and maluma in action: A cross-modal relationship between gestures and sounds. *PLoS ONE 11*, 9 (aug 2016).

[200] Sodhi, R., Poupyrev, I., Glisson, M., and Israr, A. Aireal: interactive tactile experiences in free air. *ACM Transactions on Graphics (TOG) 32*, 4 (2013), 134.

[201] Spence, C., and Zampini, M. Auditory contributions to multisensory product perception. *Acta Acustica united with Acustica 92*, 6 (2006), 1009–1025.

[202] Spray, D. C. *Thermoreceptors.* Birkhäuser Boston, Boston, MA, 1988, pp. 132–133.

[203] Srinivasan, M. A., Beauregard, G. L., and Brock, D. L. The impact of visual information on the haptic perception of stiffness in virtual environments. In *ASME Winter Annual Meeting* (1996), vol. 58, pp. 555–559.

[204] Stanley, L. M. Haptic and auditory cues for lane departure warnings. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2006), vol. 50, Sage Publications Sage CA: Los Angeles, CA, pp. 2405–2408.

[205] Stein, B. E., and Meredith, M. A. *The merging of the senses.*, vol. XV. MIT Press, Cambridge, MA, U.S.A., 1993.

[206] Stevens, J. C. Aging and spatial acuity of touch. *Journal of Gerontology 47*, 1 (1992), P35.

[207] Stevens, J. C., and Choo, K. K. Temperature sensitivity of the body surface over the life span. *Somatosensory & motor research 15*, 1 (1998), 13–28.

[208] Stevens, J. C., and Green, B. G. Temperature–touch interaction: Weber's phenomenon revisited. *Sensory processes* (1978).

[209] Stevenson, R. A., and Wallace, M. T. Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research 227*, 2 (2013), 249–261.

[210] Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance 38*, 6 (2012), 1517.

[211] Stone, R. Virtual reality and telepresence. *Robotica 10*, 5 (1992), 461–467.

[212] Sugano, Y., Keetels, M., and Vroomen, J. Auditory dominance in motor-sensory temporal recalibration. *Experimental Brain Research 234*, 5 (2016), 1249–1262.

[213] Sun, M., Ren, X., and Cao, X. Effects of multimodal error feedback on human performance in steering tasks. *J. of Information Processing 18* (2010), 284–292.

[214] Susini, P., McAdams, S., Winsberg, S., Perry, I., Vieillard, S., and Rodet, X. Characterizing the sound quality of air-conditioning noise. *Applied Acoustics 65*, 8 (2004), 763–790.

[215] Susini, P., Misdariis, N., Lemaitre, G., and Houix, O. Naturalness influences the perceived usability and pleasantness of an interface's sonic feedback. *Journal on Multimodal User Interfaces 5*, 3 (May 2012), 175–186.

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

[217] Tahiroğlu, K., and Ahmaniemi, T. Vocal sketching: A prototype tool for designing multimodal interaction. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (New York, NY, USA, 2010), no. 42 in ICMI-MLMI '10, ACM, pp. 42:1–42:4.

[218] Tajadura-Jiménez, A., and Västfjäll, D. *Auditory-Induced Emotion: A Neglected Channel for Communication in Human-Computer Interaction.* Springer, Berlin, Heidelberg, 2008, pp. 63–74.

[219] Taylor, M. M., and Lederman, S. J. Tactile roughness of grooved surfaces: A model and the effect of friction. *Perception & Psychophysics 17* (1975), 23–26.

[220] Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J. L., and Ystad, S. From sound to shape: Auditory perception of drawing movements. *J. of Experimental Psychology: Human Perception and Performance 40:3* (2014), 983–994.

[221] Tuuri, K., Eerola, T., and Pirhonen, A. Leaping across modalities: Speed regulation messages in audio and tactile domains. In *International Workshop on Haptic and Audio Interaction Design* (2010), Springer, pp. 10–19.

[222] Vallbo, Å. B., and Hagbarth, K. Impulses recorded with microelectrodes in human muscle nerves during stimulation of mechanoreceptors and voluntary contractions. *Electroencephalography and Clinical Neurophysiology 23*, 4 (1967), 392.

[223] Vallbo, Å. B., Johansson, R. S., et al. Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Human Neurobiology 3*, 1 (1984), 3–14.

[224] Van der Burg, E., Orchard-Mills, E., and Alais, D. Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental brain research 233*, 1 (2015), 53–59.

[225] Vanacken, L., De Boeck, J., and Coninx, K. *The Phantom versus the Falcon: Force Feedback Magnitude Effects on User's Performance during Target Acquisition.* Springer, Berlin, Heidelberg, 2010, pp. 179–188.

[226] Venkatesan, L., Barlow, S. M., and Kieweg, D. Age- and sex-related changes in vibrotactile sensitivity of hand and face in neurotypical adults. *Somatosensory & Motor Research 32*, 1 (2015), 44–50.

[227] Verrillo, R. T. *Sensory Research: Multimodal Perspectives.* Psychology Press, 1993, ch. 17. The effects of aging on the sense of touch, pp. 285–298.

[228] Vickers, Z. Sound perception and food quality. *Journal of Food Quality 14*, 1 (1991), 87–96.

[229] Viviani, P., and Flash, T. Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance 21*, 1 (1995), 32.

[230] Viviani, P., and Schneider, R. A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology: Human Perception and Performance 17*, 1 (1991), 198–218.

[231] Viviani, P., and Terzuolo, C. Trajectory determines movement dynamics. *Neuroscience 7*, 2 (1982), 431–437.

[232] Vogel, D., and Baudisch, P. Shift: A technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2007), CHI '07, ACM, pp. 657–666.

[233] Vroomen, J., and de Gelder, B. Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of experimental psychology: Human perception and performance 26*, 5 (2000), 1583.

[234] Vroomen, J., and Keetels, M. Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics 72*, 4 (2010), 871–884.

[235] Wallace, M. T., and Stevenson, R. A. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia 64* (2014), 105 – 123.

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

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

[238] Wickens, C., Prinet, J., Hutchins, S., Sarter, N., and Sebok, A. Auditory-visual redundancy in vehicle control interruptions:

Two meta-analyses. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2011), vol. 55, Sage Publications Sage CA: Los Angeles, CA, pp. 1155–1159.

[239] Wickremaratchi, M. M., and Llewelyn, J. G. Effects of ageing on touch. *Postgraduate Medical Journal 82*, 967 (2006), 301–304.

[240] Wilson, E. C., Reed, C. M., and Braida, L. D. Integration of auditory and vibrotactile stimuli: effects of phase and stimulus-onset asynchrony. *The Journal of the Acoustical Society of America 126*, 4 (2009), 1960–1974.

[241] Wilson, E. C., Reed, C. M., and Braida, L. D. Integration of auditory and vibrotactile stimuli: Effects of frequency. *The Journal of the Acoustical Society of America 127*, 5 (2010), 3044–3059.

[242] Wilson, G., and Brewster, S. A. Multi-moji: Combining thermal, vibrotactile & visual stimuli to expand the affective range of feedback. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI '17, ACM, pp. 1743–1755.

[243] Wilson, G., Davidson, G., and Brewster, S. A. In the heat of the moment: Subjective interpretations of thermal feedback during interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2015), CHI '15, ACM, pp. 2063–2072.

[244] Worden, A., Walker, N., Bharat, K., and Hudson, S. Making computers easier for older adults to use: Area cursors and sticky icons. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1997), CHI '97, ACM, pp. 266–271.

[245] Yamamoto, A., Nagasawa, S., Yamamoto, H., and Higuchi, T. Electrostatic tactile display with thin film slider and its application to tactile telepresentation systems. *IEEE Transactions on Visualization and Computer Graphics 12(2)* (2006), 168–177.

[246] Zhang, J., Ong, S. K., and Nee, A. Y. C. Navigation systems for individuals with visual impairment: A survey. In *Proceedings of the 2Nd International Convention on Rehabilitation Engineering & Assistive Technology* (Kaki Bukit TechPark II,, Singapore, 2008), iCREATe

'08, Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre, pp. 159–162.

[247] Zimmerman, J., Forlizzi, J., and Evenson, S. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2007), CHI '07, ACM, pp. 493–502.