

COLLEGE	MATHS, TELECOMS
DOCTORAL	INFORMATIQUE, SIGNAL
BRETAGNE	SYSTEMES, ELECTRONIQUE

# THÈSE DE DOCTORAT DE

L'INSTITUT NATIONAL DES  
SCIENCES APPLIQUÉES DE RENNES

ÉCOLE DOCTORALE N° 601

*Mathématiques, Télécommunications, Informatique, Signal, Systèmes,*

*Électronique*

Spécialité : *Informatique*

Par

**Lendy MULOT**

**Advancing Ultrasound Mid-Air Haptics: Perception Studies, Rendering Methods, and Design of Virtual Reality Interactions**

Thèse présentée et soutenue à Rennes, le 25 Novembre 2025

Unité de recherche : IRISA (UMR 6074)

## Rapporteurs avant soutenance :

Sriram SUBRAMANIAN Full Professor, UCL, Royaume-Uni  
Lynette JONES Senior Research Scientist, MIT, États-Unis

## Composition du Jury :

Présidente :	Marie BABEL	Professeure des Universités, INSA Rennes, France
Examinateurs :	Sriram SUBRAMANIAN	Full Professor, UCL, Royaume-Uni
	Lynette JONES	Senior Research Scientist, MIT, États-Unis
	Sarah MCINTYRE	Assistant Professor, Linköping University, Suède
	Orestis GEORGIOU	R&D Funding Lead, Acoustofab, Royaume-Uni
Directrice de thèse :	Maud MARCHAL	Professeure des Universités, INSA Rennes/IUF, France

## Invités :

Co-directeur de thèse :	William FRIER	Haptic Perception Team Lead, Ultraleap, Royaume-Uni
	Claudio PACCHIEROTTI	Chargé de Recherche HDR, CNRS, France



# Table of Contents

<b>Introduction</b>	<b>9</b>
Context . . . . .	9
Scientific Challenges . . . . .	12
Contributions . . . . .	15
<b>List of Publications</b>	<b>17</b>
<b>1 Related Work</b>	<b>19</b>
1.1 The Tactile System in the Hand . . . . .	19
1.1.1 Biomechanics of the Skin . . . . .	20
1.1.2 Tactile Afferents and Mechanoreceptors . . . . .	22
1.2 Haptic Rendering . . . . .	27
1.2.1 Haptic Interfaces . . . . .	27
1.2.2 Haptic Rendering Schemes . . . . .	28
1.3 Mid-Air Haptics . . . . .	30
1.4 Ultrasound Mid-Air haptics . . . . .	32
1.4.1 Phased Array and Focal Point . . . . .	33
1.4.2 Modulation Techniques . . . . .	40
1.4.3 Shape Rendering . . . . .	47
1.4.4 Stimulus Design and Validation . . . . .	51
1.5 Designing Interactions with UMH . . . . .	54

---

## TABLE OF CONTENTS

---

1.5.1	Coupling Schemes . . . . .	55
1.5.2	UMH for Interactions . . . . .	56
1.6	Conclusion . . . . .	61
<b>I</b>	<b>Understanding Human Perception of UMH Stimuli</b>	<b>63</b>
<b>2</b>	<b>Characterization of Primary Tactile Afferent Response</b>	<b>65</b>
2.1	Introduction . . . . .	65
2.2	Neurophysiological Recording Study: Tactile Afferent Response in Glabrous Skin	66
2.2.1	Materials and Methods . . . . .	67
2.2.2	Preliminary Results and Discussion . . . . .	70
2.3	In-Depth Skin Deformation Measurements . . . . .	71
2.3.1	Background on OCT Recording . . . . .	71
2.3.2	Apparatus and Setup . . . . .	72
2.3.3	Preliminary Results and Discussion . . . . .	74
2.4	Neurophysiological Recording Study: Tactile Afferent Response in Hairy Skin	76
2.4.1	Materials and Methods . . . . .	76
2.4.2	Expected Results . . . . .	77
2.5	Conclusion . . . . .	77
<b>3</b>	<b>Perception of Dynamically-Rendered 2D shapes</b>	<b>79</b>
3.1	Introduction . . . . .	79
3.2	Spatio-Temporally-Modulated Tactile Pointers . . . . .	80
3.2.1	Rendering Algorithm . . . . .	80
3.2.2	Stimuli Parameterization for Experimental Validation . . . . .	82
3.3	User Study: Evaluation of Stimulus Intensity . . . . .	83
3.3.1	Hypotheses . . . . .	83
3.3.2	Materials and Methods . . . . .	84
3.3.3	Results . . . . .	86
3.3.4	Discussion . . . . .	87
3.3.5	Limitations and perspectives . . . . .	87
3.4	Naive Shape Perception . . . . .	88
3.4.1	Materials and Methods . . . . .	88
3.4.2	Results and Discussion . . . . .	89
3.4.3	Limitations and Perspectives . . . . .	91

---

TABLE OF CONTENTS

3.5	User Study: Clarity of Rendered Shapes . . . . .	92
3.5.1	Hypotheses . . . . .	92
3.5.2	Materials and Methods . . . . .	92
3.5.3	Results . . . . .	94
3.5.4	Discussion . . . . .	95
3.6	Discussion and Perspectives . . . . .	97
3.6.1	Limitations and Future Work . . . . .	97
3.6.2	General Discussions . . . . .	97
3.7	Conclusion . . . . .	98
<b>II</b>	<b>Optimized Primitives for Improved Rendering Algorithms</b>	<b>99</b>
<b>4</b>	<b>Harnessing Shock Waves in the Skin: Spiral-Based Rendering Primitives</b>	<b>101</b>
4.1	Introduction . . . . .	101
4.2	Harnessing Shear Shock Formation: Spiral Scanning . . . . .	103
4.2.1	Physics-Based Derivation . . . . .	103
4.2.2	Numerical Experiments . . . . .	105
4.3	Spiral-Based Shape Rendering . . . . .	108
4.3.1	2D Shape Rendering . . . . .	109
4.3.2	Haptic Shape Shading . . . . .	110
4.3.3	3D Shape Rendering . . . . .	111
4.3.4	Generalization and Discussion . . . . .	111
4.4	User Study: 3D Object Perception . . . . .	112
4.4.1	Materials and Methods . . . . .	112
4.4.2	Results . . . . .	114
4.4.3	Discussion . . . . .	115
4.5	User Study: Haptic Shading . . . . .	116
4.5.1	Materials and Methods . . . . .	117
4.5.2	Results . . . . .	118
4.5.3	Discussion . . . . .	118
4.6	Conclusion . . . . .	119
<b>III</b>	<b>UMH Rendering Pipeline: Formalization</b>	<b>121</b>
<b>5</b>	<b>2D Rendering: Sampling Strategies</b>	<b>123</b>

---

**TABLE OF CONTENTS**

---

5.1	Introduction . . . . .	123
5.2	Stimulus Design Formalization . . . . .	124
5.3	The DOLPHIN Framework . . . . .	125
5.3.1	Design Tool . . . . .	126
5.3.2	Emitter . . . . .	131
5.3.3	Interfacing for Perception Evaluation . . . . .	132
5.4	Proof of Concept . . . . .	132
5.4.1	Materials and Methods . . . . .	133
5.4.2	Results and Discussion . . . . .	135
5.5	Interoperability . . . . .	138
5.6	Conclusion . . . . .	139
<b>6</b>	<b>3D Rendering: Intersection Strategies</b>	<b>141</b>
6.1	Introduction . . . . .	141
6.2	Intersection Strategies: Definition . . . . .	143
6.2.1	Component Representations . . . . .	143
6.2.2	Definition of Intersection Strategies . . . . .	144
6.3	Using Intersection Strategies to Generate Haptic Rendering Schemes . . . . .	147
6.3.1	Intersection Strategies . . . . .	147
6.3.2	Chosen Component Representations . . . . .	148
6.3.3	Effects on Non-Haptic Factors . . . . .	148
6.4	Pilot Study: Impact of Haptic Rendering Schemes on Perceived Shape Properties	151
6.4.1	Materials and Methods . . . . .	152
6.4.2	Results and Discussion . . . . .	154
6.5	User Study: Modifying the Perception of Infill Density . . . . .	155
6.5.1	Materials and Methods . . . . .	155
6.5.2	Results . . . . .	156
6.5.3	Discussion . . . . .	158
6.6	Conclusion . . . . .	159
<b>IV</b>	<b>Enhancing Virtual Reality Interactions Using UMH</b>	<b>161</b>
<b>7</b>	<b>Hand Guidance</b>	<b>163</b>
7.1	Introduction . . . . .	163
7.2	UMH Strategies for Hand Guidance . . . . .	164

---

TABLE OF CONTENTS

7.3	User Study: Evaluation in a Static VR Environment . . . . .	167
7.3.1	Materials and Methods . . . . .	167
7.3.2	Results . . . . .	170
7.4	User Study: Evaluation in a Dynamic VR Environment . . . . .	173
7.4.1	Materials and Methods . . . . .	173
7.4.2	Results . . . . .	174
7.5	Conclusion . . . . .	176
<b>8</b>	<b>Visuo-Tactile Cues for Contact Rendering</b>	<b>179</b>
8.1	Introduction . . . . .	179
8.2	Contact Rendering Techniques . . . . .	180
8.2.1	Interpenetration Computation . . . . .	180
8.2.2	Visual Rendering . . . . .	181
8.2.3	Haptic Rendering . . . . .	183
8.2.4	Design Discussion . . . . .	185
8.3	Pilot Experiment: Haptic Trigger Calibration . . . . .	186
8.3.1	Materials and Methods . . . . .	186
8.3.2	Results and Discussion . . . . .	189
8.4	User Study: Generating Coupling Schemes . . . . .	191
8.4.1	Hypotheses . . . . .	191
8.4.2	Materials and Methods . . . . .	191
8.4.3	Results . . . . .	192
8.4.4	Discussion . . . . .	195
8.5	User Study: Comparing Coupling Schemes . . . . .	196
8.5.1	Hypotheses . . . . .	197
8.5.2	Materials and Methods . . . . .	197
8.5.3	Results . . . . .	199
8.5.4	Discussion . . . . .	204
8.6	General Discussion . . . . .	205
8.6.1	Limitations . . . . .	205
8.6.2	Discussion . . . . .	206
8.7	Conclusion . . . . .	206
<b>9</b>	<b>Bimanual Interactions</b>	<b>209</b>
9.1	Introduction . . . . .	209
9.2	Bimanual UHM Manipulation System . . . . .	210

## TABLE OF CONTENTS

---

9.2.1	System Hardware . . . . .	211
9.2.2	Coupling Schemes . . . . .	212
9.3	User Study: Object Grasping and Holding with Bimanual Haptics . . . . .	214
9.3.1	Hypotheses . . . . .	215
9.3.2	Materials and Methods . . . . .	216
9.3.3	Results . . . . .	220
9.3.4	Discussion . . . . .	222
9.4	User Study: 3D Object Shape Exploration with Bimanual Haptics . . . . .	223
9.4.1	Hypotheses . . . . .	224
9.4.2	Materials and Methods . . . . .	224
9.4.3	Results . . . . .	226
9.4.4	Discussion . . . . .	229
9.5	Global Discussion . . . . .	230
9.6	Use Cases . . . . .	231
9.7	Conclusion . . . . .	233
<b>10</b>	<b>Conclusion</b> . . . . .	<b>235</b>
10.1	Summary and Short-Term Future Work . . . . .	235
10.2	Perspectives . . . . .	239
<b>Bibliography</b>		<b>243</b>
<b>A Résumé Long en Français</b>		<b>275</b>

# Introduction

Ultrasound mid-air haptics (UMH) aims to improve natural gestures and interactions by rendering tactile sensations directly in mid-air through the use of focused ultrasound waves. This thesis advances the field by investigating the perception and rendering of UMH stimuli, as well as their conjoint use with virtual reality interactions. Below, we go into more details about the context of this thesis, as well as the underlying scientific challenges tied to UMH. Finally, we introduce the associated contributions we propose to address these challenges.

## Context

The **sense of touch** is a fundamental part of the human somatosensory system, as it contributes to our interaction with the environment in both physical and emotional contexts. The somatosensory system encompasses various sub-senses, including:

- Proprioception [[Tuthill et al. 2018](#)] refers to one's perception of their own body position, movement, and efforts;
- Thermoception [[Jones et al. 2002](#)] is the perception of heat and changes in temperature;
- Nociception [[Sneddon 2018](#)] is responsible for the perception of harmful stimuli;
- Pruriception [[Andersen et al. 2017](#)] uses similar mechanisms to provide itch sensations;
- **Tactile** sensations [[Jones 2018](#)] refer to the **perception** of mechanical stimuli on the skin (e.g. vibration, texture).

Among those, this thesis focuses on tactile sensations, which is currently the most studied sub-sense, providing a wide range of extremely rich sensations in response to mechanical stimulation on the skin. While often taken for granted, tactile feedback plays a crucial role in how we perceive our environment, or even communicate socially. For instance, we rely on a specific set of hand motions to probe and explore our environment using touch (e.g. applying pressure to gauge the hardness of an object, or contour following to identify its shape) [Leiderman et al. 1987]. Tactile stimulation has also been used to compensate for other sensory deficiencies (e.g. visual impairments) [Lacôte et al. 2024] by lightening the load on impaired channels, and taking advantage of the under-used tactile channel. Tactile sensations is also an important part of social and affective communication, from friendly taps to affective caresses[Hauser et al. 2019; McIntyre et al. 2019; McIntyre et al. 2022].

To properly design efficient sensations, it is crucial to understand how tactile stimuli are perceived. Tactile perception, and perception in general, is a very complex pipeline [Goldstein et al. 2022a]: when a stimulus, referred to as **distal stimulus**, is present or generated in the environment, it then propagates through a medium until it reaches sensory receptors, at which point it is considered a **proximal stimulus**. For tactile stimulation, this corresponds to the deformation of the skin towards **mechanoreceptors**. Then, the associated **afferent neurons** may generate an electric signal in response to the stimulus during a process called **mechanotransduction**. The individual signals from many afferent neurons then travel towards the brain, at which point they are combined in a process called **neural processing**. Finally, the processed signals are turned into behavioural responses, including a conscious awareness of the stimulus, which is then recognized and understood by the recipient based on their knowledge and prior experience. Finally, this recognized sensation may lead the user to take motor actions (e.g. move their hand to explore another part of the environment), which allows them to experience new stimuli.

To efficiently study and leverage the sense of touch, it becomes very useful to design technologies that can artificially provide controlled and consistent tactile sensations. Such devices, referred to as **haptic interfaces** or **haptic displays**, have been designed to provide specific types of stimuli (e.g. vibrations, pressure) [Jones 2018] through a process called **haptic rendering**. Not only do these interfaces enable the consistent evaluation of tactile stimuli, and provide artificial touch sensations to users, but they can also often act as a user interface for 3D interactions. Such interactions usually follow the interaction loop, illustrated in [Figure 1](#). Actions by the users are detected by input devices (e.g. hand tracker, force sensors), and are then translated in a physical or virtual environment according to the implemented interac-

tion technique. Feedback is then computed in response to the new state of the environment, and is sent to the output devices (e.g. screen, vibration motor). The feedback is then rendered, generating a stimulus that can be perceived by the user. This pipeline is handling most types of 3D interactions. Yet most of these interactions could benefit from the addition of tactile feedback for improved performances, better user experience, or even reduced cognitive load. More specifically, the interactions designed in this thesis focus on **virtual reality** (VR) uses. VR is an interactive and often immersive technology that lets users interact with a fully virtual environment through computer-generated sensory feedback. Immersive VR often relies on equipment like head-mounted displays or haptic interfaces to provide the best experience possible.

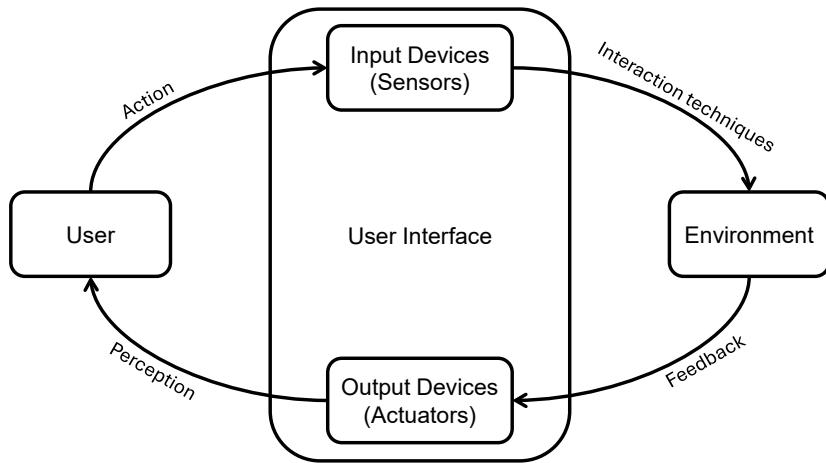


Figure 1 – The interaction loop governing 3D interactions.

Among the many types of haptic interfaces, this thesis focuses on ultrasound mid-air haptic (UMH) stimulation [[Rakkolainen et al. 2021](#)]. UMH interfaces (shown in [Figure 2\(A\)](#)) rely on a collection of ultrasound transducers, which are small speakers turning electricity into an ultrasonic vibration. By adjusting the phase of each transducer independently, it is then possible to generate constructive interference at a given location in space, called **focal point**, at which the local pressure becomes very high (see [Figure 2\(B\)](#)). Then, by quickly and repeatedly moving this focal point along a path, it is possible to induce a perceivable vibration on nearby skin (see [Figure 2\(C\)](#)) [[Kappus et al. 2018](#); [Frier et al. 2018](#)]. While this technology can generate tactile sensations in mid-air, it remains a relatively recent technology, introduced in 2008 [[Iwamoto et al. 2008a](#)], and much about the perception, rendering, and application of UMH stimuli is still not fully understood.

This is why this thesis focuses on *advancing ultrasound mid-air haptics: perception studies, ren-*

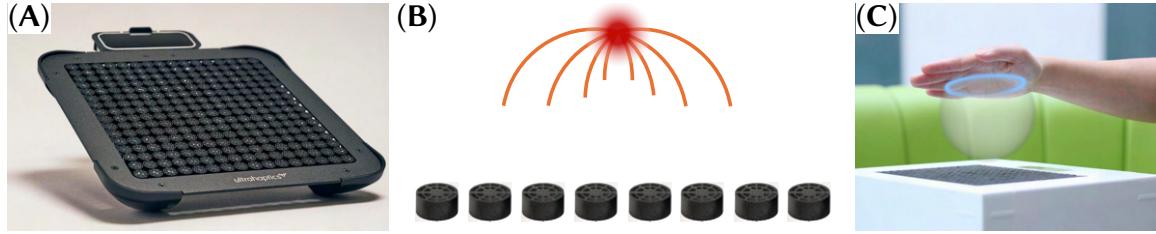


Figure 2 – (A) UMH interface, composed of a collection of ultrasound transducers, (B) are focusing ultrasound waves to target positions in space. (C) By moving this focal point, they can generate tactile stimulation along a shape. Sources: (A) and (C) are from [Ultraleap](#) while (B) is adapted from [\[Rakkolainen et al. 2021\]](#).

*dering methods, and design of virtual reality interactions.* We propose advances to the field by investigating the perception and rendering of UMH stimuli, as well as their conjoint use with VR interactions. We study how UMH stimuli are **perceived** through physiological measurements and psychophysics studies, to expand our understanding of UMH stimuli and allow sensation designers to optimize their rendering techniques accordingly. We also propose and discuss new **rendering** primitives to expand the UMH design toolkit and provide more freedom to UMH designers, while pushing the limits of what can be rendered with UMH. We then formalize, discuss, and evaluate the rendering pipeline for UMH to provide a standardized discussion of the rendering pipeline, pushing for more reproducibility in the field, and a better understanding of the low-level rendering mechanisms behind the design of UMH sensations. Finally, we contribute to the design and evaluation of UMH feedback and associated **interaction** techniques in VR. We chose VR as the scene for such interactions as we believe that both technologies synergize very well as they both promote natural immersive gestures and interactions in a seamless environment.

## Scientific Challenges

While the innovative aspect of UMH is undeniable, it remains too understudied to be optimally used. More specifically, in the context of this PhD, we identified four main challenges related to perception (**I**), rendering (**II, III**), and interaction (**IV**).

**I. Understanding Human Perception of UMH Stimuli.** In order to provide any kind of effective sensory stimulation, it is very important to understand how the stimulus will be perceived by the user. Yet, the human perceptual process is extremely complex, composed of many steps [\[Goldstein et al. 2022a\]](#), as previously discussed. But as sensation designers, we

can only control the distal stimulus.

In a way, haptic (and more generally any sensory) rendering can be viewed as information *encoding*, and perception as the decoding *process*. Indeed, the stimulus is a representation of some information that we want the participant to perceive or decode. And similarly to information encoding, the decoded information is generally more important than the encoded one, since the main purpose of any sensory rendering is to elicit target sensations.

By studying this complex stimulus-sensation relationship, we would obtain a better understanding of this *decoding* process, which would enable sensation designers to provide highly specific and precisely controlled stimuli that elicit precise sensations, thus providing a better user experience. And while this is true for any sensory stimulation, the relative novelty of UMH implies that only few specific perception studies are available, which often leads designers to rely on perception results from contact vibrotactile studies or other technologies without necessarily accounting for the specifics of UMH (e.g. [Howard et al. 2023a]). This is why understanding this *decoding* process is a core goal of this thesis, allowing for more effective stimulus design guidelines.

**II. Improving Rendering Algorithms and Providing Optimized Primitives.** The quality of a rendering algorithm is highly dependent on the context of the task and interaction, meaning that there is not a universally best UMH rendering algorithm. For instance, depending on context, we may want to render the same piece of information, optimized for high perceived intensity, low operating noise, or feedback sharpness. And different rendering algorithms would be used for each of these specific objectives.

Also, UMH rendering algorithms can usually be thought as a sequence of steps that achieve general purposes [Frier et al. 2019; Mulot et al. 2021b]. So having a toolkit full of options, that we call *primitives*, for each step is hugely beneficial to sensation designers, as it helps create a larger range of sensations. At the time of writing this thesis, a few UMH primitives exists (e.g. [Iwamoto et al. 2008b; Kappus et al. 2018; Takahashi et al. 2018]), but the UMH rendering toolkit lacks diversity.

**III. Providing Unified Rendering Baselines.** UMH has gained the attention of a few research groups around the world, with different backgrounds and focus. While this diversity in expertise and scientific approach is very beneficial to the field, it also means that similar concepts are being reported in different ways. This makes it hard, especially for new researchers and designers, to find parallels, or to understand subtle differences between render-

ing techniques developed by different groups. For instance, even the name of the technology is not a consensus: while some works, including ours, call it “ultrasound mid-air haptic” interfaces [Mulot et al. 2021b; Seifi et al. 2023], others refer to it as an “airborne ultrasound tactile display” [Suzuki et al. 2021; Iwamoto et al. 2008a], or even “holographic haptic display” [Reardon et al. 2023]. Other examples include the design of similar rendering techniques: spatiotemporal modulation (STM) [Kappus et al. 2018; Frier et al. 2018] and lateral modulation (LM) [Takahashi et al. 2018], which sometimes leads to confusions.

This discrepancy extends to most aspects of the reporting of our collective works, with similar concepts being discussed in different ways. This unfortunately makes it hard to compare existing works from different groups, or to reproduce existing studies. To address this issue, we argue for a unified formalization of UMH rendering, that can represent any rendering algorithm in a very precise and reproducible way. By adopting such formalization, the community would report their results in a more coherent way, thus improving reproducibility in the field.

**IV. Enhancing Natural Interactions in Virtual Reality.** While VR technologies have received tremendous technological improvements in the past few years, now allowing for the visual rendering of very realistic environments, the lack of multisensory feedback when interacting with the environment reduces immersion. This is why researchers groups have worked on developing and integrating tactile interfaces in VR.

If our goal is to provide a natural and realistic VR interactions, we need to have our hands free from any equipment, and our motions completely unrestricted to more closely resemble real-life interactions. This makes UMH a great candidate as it can provide tactile feedback, while its mid-air nature fits these constraints.

Yet UMH also brings its own set of limitations, such as its low feedback intensity, or its medium-sized workspace not suited for all types of interactions. For instance, bimanual interactions with two hands next to each other would force them at the limit of the workspace for traditional UMH interfaces. It is therefore important to study how UMH can be integrated and coupled to virtual environment to provide rich unobtrusive tactile cues in VR. This is where the study and design of coupling schemes becomes important to represent the relations between the simulated virtual environment being updated in response to the user’s virtual interactions, and the resulting tactile feedback.

## Contributions

While the previous section highlights four scientific challenges related to the study of UMH, this thesis contributes to addressing them, expanding our understanding of these technologies and improving their capabilities. This thesis is organized in four parts, providing contributions respectively addressing the four challenges discussed previously.

[Chapter 1](#) first discusses related work on UMH. More specifically, we delve into the behaviour of the skin and its tactile afferents, before discussing hardware and software solutions to stimulate this tactile system. Then, we go deeper into the specifics of mid-air haptics and specifically, UMH, discussing the mechanisms behind it, the rendering methods, and the tools that have been designed for it. Finally, we discuss user interactions involving UMH, and other applications.

The following two chapters, included in [Part I](#) address parts of **Challenge I**, focusing on the perception of UMH stimuli.

[Chapter 2](#) aims at measuring the individual response of first-order tactile afferents in the glabrous skin to a wide range of UMH stimuli. We use microneurography techniques to record the activation of single tactile units responding to UMH stimuli rendered at different modulation frequencies. We also provide in-depth measurements of skin deformation in response to UMH stimulation, using optical coherence tomography (OCT) methods. This provides insights about how the acoustic waves propagate in the skin, and how the nervous system initially encodes the unique UMH signals.

Then, [Chapter 3](#) presents the design of a new UMH rendering technique, specifically designed to provide dynamic sensations of shapes. This technique takes inspiration from a previous method [[Hajas et al. 2020](#)], but aims at providing a higher perceived intensity with a similar discrimination accuracy. We conduct perceptual studies to evaluate our method, and compare it to the state-of-the-art.

Then, [Part II](#) focuses on **Challenge II** to provide physics-based rendering primitives for UMH. Specifically, [Chapter 4](#) presents a spiral-shaped rendering primitives that can be used to focus shear-shock waves resulting from the fast motion of the focal point. We validate this primitive through numerical simulations, and discuss the effect of its parameters. Finally, we use this primitive to create a full shape-rendering algorithm, and we evaluate its high-level potential through a set of user studies.

[Part III](#) aims at providing a general formalization of the UMH rendering pipeline, thus address-

ing **Challenge III**.

In [Chapter 5](#) we formalize *sampling strategies*, which represent general rendering algorithms for 2D shapes. We also present *DOLPHIN*, a new framework for the design of UMH stimuli and their perceptual evaluation. We also conduct a psychophysics experiment to show how we can use Dolphin in practice to generate sensations, and how we can interface it to external perception evaluation tools such as PsychoPy [[Peirce et al. 2019](#)].

[Chapter 6](#) extends this formalization to 3D shape rendering by introducing *intersection strategies*. We design different intersection strategies to illustrate the pipeline, and conduct user studies to show how we can use intersection strategies to alter the perception of object properties.

Finally, [Part IV](#) explores ways to improve VR interactions using UMH feedback, thus contributing to **Challenge IV**.

[Chapter 7](#) uses touch as a complementary channel we can use to provide additional information to the user without clogging the visual feedback. Specifically, we design UMH feedback methods to encode 2-DoF position and orientation information, that can be used to guide the user's hand towards a given pose. We conduct two user studies to evaluate these feedbacks, both in a static and dynamic environment.

Conversely, [Chapter 8](#) uses the touch channel to provide redundant information, coherent with the visual stimulation. Precisely, we proposed a set of visual and haptic feedbacks to represent the interpenetration with a virtual object. As tactile stimulation cannot prevent interpenetration by itself, and since VR is often lacking haptic feedback during manipulation, combining visual and UMH stimulation can be mutually beneficial. We thus conduct a user study to pair each visual feedback with the best associated UMH feedback, thus generating visuo-haptic feedback techniques. Finally, we conduct a second user study to evaluate each of these visuo-haptic pairs, in order to find the best candidate.

Then, [Chapter 9](#) tackles the workspace limitations of UMH by proposing a hardware system dedicated to bimanual interactions. We also propose a rendering scheme dedicated to bimanual haptic manipulations with UMH. We then conduct two user studies to measure the effect of UMH during exploration and grasping VR tasks.

Finally, [Chapter 10](#) concludes this thesis and discusses perspectives of future work to further explore these scientific challenges.

# List of Publications

## Journals

**Lendy Mulot**, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. “Ultrasound Mid-Air Haptics for Hand Guidance in Virtual Reality.” In: *IEEE Transactions on Haptics*, vol. 16, no. 4, pp. 497-503, 2023. Discussed in [Chapter 7](#).

**Lendy Mulot**, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. “Improving the Perception of Mid-Air Tactile Shapes with Spatio-Temporally-Modulated Tactile Pointers.” In: *ACM Transactions on Applied Perception*, vol. 20, no. 4, pp. 1–16, 2023. Discussed in [Chapter 3](#).

**Lendy Mulot**, Thomas Howard, Guillaume Gicquel, Claudio Pacchierotti, and Maud Marchal. “Bimanual Ultrasound Mid-Air Haptics for Virtual Reality Manipulation.” In: *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 9, pp. 4821-4832, 2025. Discussed in [Chapter 9](#).

## International Conferences

**Lendy Mulot**, Guillaume Gicquel, Quentin Zanini, William Frier, Maud Marchal, Claudio Pacchierotti, and Thomas Howard. “DOLPHIN: A Framework for the Design and Perceptual Evaluation of Ultrasound Mid-Air Haptic Stimuli.” In: *ACM Symposium on Applied Perception (SAP)*,

2021. Discussed in [Chapter 5](#).

**Lendy Mulot**, Thomas Howard, Sarah Emery, Claudio Pacchierotti, and Maud Marchal. “Designing 3D Object Rendering Techniques for Ultrasound Mid-Air Haptics using Intersection Strategies.” In: *ACM Symposium on Applied Perception (SAP)*, 2024. Discussed in [Chapter 6](#).

## Work in Progress

**Lendy Mulot**, Guillaume Gicquel, William Frier, Maud Marchal, Claudio Pacchierotti, and Thomas Howard. “Curvature Discrimination for Dynamic Ultrasound Mid-Air Haptic Stimuli.” In: *IEEE World Haptics Conference (WHC)*, 2021.

**Lendy Mulot**, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. “Can We Increase the Perceived Intensity of Mid-Air Haptic Shapes Rendered With Dynamic Tactile Pointers?” In: *IEEE World Haptics Conference (WHC)*, 2023.

# Chapter 1

## Related Work

### Contents

---

1.1	The Tactile System in the Hand	19
1.2	Haptic Rendering	27
1.3	Mid-Air Haptics	30
1.4	Ultrasound Mid-Air haptics	32
1.5	Designing Interactions with UMH	54
1.6	Conclusion	61

---

In this chapter, we review the scientific literature which served as the foundations for our own contributions. Namely, we start by presenting the skin and its role in tactile perception ([Section 1.1](#)), before discussing the software and hardware solutions that have to be implemented for controlled haptic stimulation ([Section 1.2](#)). Then, we converge towards mid-air haptics ([Section 1.3](#)) and specifically ultrasound mid-air haptics (UMH, [Section 1.4](#)), before discussing the design of UMH feedback for human-computer interactions ([Section 1.5](#)). We also use this opportunity to position our different contributions in relation to the existing literature.

### 1.1 The Tactile System in the Hand

As mentioned in the [Introduction Chapter](#), this thesis focuses on mid-air tactile stimulation, which is mediated by the skin. The skin is the largest and heaviest human organ, covering the

whole body. This is a fundamental difference between touch and the other traditional senses that rely on localized organs (e.g. eyes for sight). With its large nature, the skin properties change depending on the location on the body [Corniani et al. 2020]. Mostly, two types of skin are usually considered: **glabrous skin** (e.g. on the palm) and **hairy** skin (e.g. on the arm). Most of the discussion below are tailored for glabrous skin in the hand, which is the main target of UMH, especially in the context of this thesis.

To design efficient stimuli, it is crucial to understand how one perceives tactile sensations, which requires an understanding of the behaviour of the skin (Section 1.1.1), and tactile afferents (Section 1.1.2), which we present in this section.

### 1.1.1 Biomechanics of the Skin

**Skin Properties** As tactile stimulation is mostly mediated by skin deformation caused by exterior forces, it is important to understand how skin behaves in order to predict its response to designed stimuli. Specifically, the skin is a physical medium with notable properties (see [Kalra et al. 2016] for review):

- the skin is **viscoelastic** [Everett et al. 2013], meaning it combines viscous (fluid-like) and elastic (spring-like) response components to stress. This also implies that the skin response is dependent on the frequency of the deformation. To characterize this effect, it has been shown that the elastic component predominates at frequencies below 100 Hz, while the viscosity predominates for higher frequencies [Moore 2007; Wiertlewski et al. 2012; Nickell et al. 2000].
- the skin is **anisotropic**: its properties vary depending on the considered direction. For instance, it has been shown that the skin’s light scattering coefficient widely varies depending on the direction considered [Nickell et al. 2000]. Anisotropy can also be observed in response to mechanical stimulation, where specific directions exhibit a higher stiffness [Meijer et al. 1999; Eloune et al. 2023].
- the skin’s structure is actually **multi-layered**, composed of (from exterior to interior) the *epidermis* (protective layer in contact with the environment), the *dermis* (home of most tactile-involved components), and the *hypodermis* (interface with underlying structures), as shown in Figure 1.1. These layers can also be decomposed into sub-layers for more accurate representation. Systems have been developed, leveraging optical coherence tomography to visualize these layers in a completely non-intrusive way [Duvernay et al. 2023]. Importantly, the size and composition of each layer can vary depending on

the position on the body, which explains the location-specific underlying perceptual properties of the skin.

- all these properties contribute to the **non-linear** behaviour of the skin, that has been observed for electrical [Rahimi et al. 2019] and mechanical [Meijer et al. 1999] stimulations. Non-linearity implies that the response to multiple stimulation sources cannot be calculated as a combination of the individual responses. This property highlights the very complex behaviour of the skin, which results in computationally expensive models.

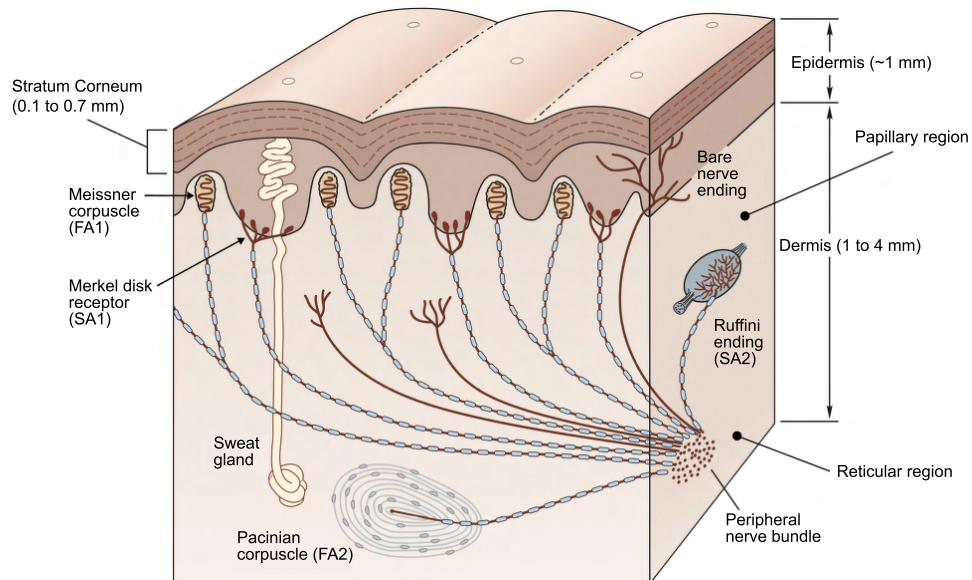


Figure 1.1 – Multi-layer structure of the glabrous skin. Some of the skin’s internal microstructures are also depicted, including mechanoreceptors, nerve endings, and sweat glands. Adapted from [Kandel et al. 2012a; Deflorio et al. 2022].

**Skin Modelling** When a mechanical stimulation is applied to a location of the skin, several waves are generated:

- **Compression waves** are longitudinal waves travelling in the normal direction to the surface.
- **Surface waves** include both a longitudinal and transverse, and travel at the surface of the skin.
- **Shear waves** are transverse and travel inside the skin, parallel to its surface.

These waves convey energy, with shear waves being the primary reason for this propaga-

tion [Sofia et al. 2013]. For intense stimulation, these waves can reach the bone, which can also propagate with less damping [Dobrev et al. 2017]. In the case of UMH, the mechanical deformation of the skin results in the generation of shear (e.g. [Reardon et al. 2019; Reardon et al. 2023]) and surface (e.g. [Frier et al. 2018; Cataldo et al. 2023]) waves that have been measured and discussed, as they are assumed to drive the perception, while compression waves have been mostly disregarded for being significantly weaker. Taking advantage of these waves may then be possible for perceptual purposes, such as to increase perceived intensity of the feedback [Frier et al. 2018].

The complex properties of the skin, combined with the added complexity of surface waves [Manfredi et al. 2012] has led researchers to attempt to model the skin behaviour using numerical rather than analytical techniques. These techniques often rely on variations of finite element models (e.g. [Larrabee Jr et al. 1986; Wu et al. 2004]). More recently, frameworks have been developed to specifically model acoustic waves propagation in the skin using pseudo-spectral methods (extensions of finite difference techniques) [Treeby et al. 2010; Treeby et al. 2014]. This has proven to be very useful for UMH simulation, to measure the effect of the focal point’s fast motion [Reardon et al. 2023].

In Chapter 4 we leverage the latter framework, building upon work by Reardon et al. [2023] to design physics-based rendering primitives, using a custom focal point motion to focus generated shear shock-waves in the skin to target locations.

**Skin Measurements** To assess the validity of these models, measurements have to be conducted on real skin, providing a baseline for comparisons. While many measurement tools have been developed, laser Doppler vibrometry [Rothberg et al. 2017] has been extensively used (e.g. [Reardon et al. 2019; Reardon et al. 2023; Cataldo et al. 2023]) to measure surface deformations, due to its high sensitivity, non-intrusive and non-contact nature. Optical coherence tomography (OCT) [Huang et al. 1991] methods are also getting popular, as they offer similar non-intrusive measurements with very precise results, based on interference measurements of light scattered by the skin. But while laser vibrometry is mostly limited to measurements at the surface of the skin, OCT methods allow for an in-depth visualization of the skin deformation [Duvernoy et al. 2023].

### 1.1.2 Tactile Afferents and Mechanoreceptors

On top of the complex biomechanical properties of the skin, it is also composed of varied structures responsible for the perception of tactile stimuli, as presented below.

**The Hand's Tactile Afferents** The glabrous skin of the hand is home to four types of **low-threshold mechanoreceptors** (LTMRs), that initiate the perceptual process [Johansson et al. 1979; Bolanowski Jr et al. 1988; Jones et al. 2006]. These mechanoreceptors, illustrated in [Figure 1.1](#), encapsulate the peripheral nerve endings of **tactile afferent** neurons (sometimes also referred to as tactile afferent units, or simply tactile afferents). A mechanoreceptor deforms when mechanical strain is applied, thus filtering the detected deformation by the nerve ending. The shape and localization within the skin layers of each receptor type make them suitable for detecting different types of stimulation. A mapping between afferent type, their corresponding mechanoreceptors, and their properties is presented in [Table 1.1](#). Namely, afferent units can be distinguished by their **adaptation rate** [Johansson et al. 1979]. Slow-adapting (SA) units continuously respond to a stimulus, while fast-adapting (FA) units mostly respond to the changes in stimulation.

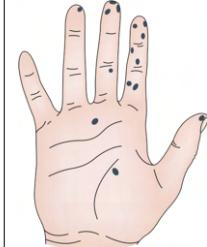
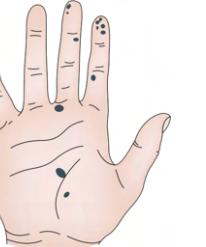
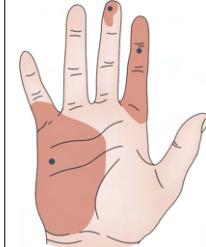
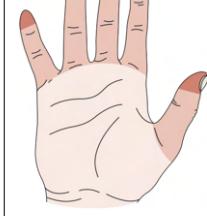
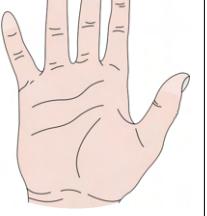
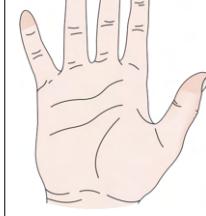
Both categories are also divided into type 1 and type 2, with type 1 units having a small **receptive field** (region that will activate the unit if stimulated) due to a small mechanoreceptor localized near the epidermis, while type 2 units have a large receptive field due to a large mechanoreceptor localized deeper in the skin tissues. Specifically, in the glabrous part of the hand [Vallbo et al. 1978; Vallbo et al. 1984; Johansson et al. 2009]:

- Merkel disks are associated to type 1 SA (SA1) units, and are sensitive to the perception of low-frequency variations and static force.
- Ruffini endings encapsulate SA2 afferents, responsible for static force and skin stretch perception.
- Meissner corpuscles encapsulate the FA1 units, which is the most common unit in the hand [Johansson et al. 2009], involved in the perception of mid-frequency vibrations and spatial discontinuities.
- Pacinian corpuscles are the largest, deepest, and rarest mechanoreceptors, encapsulating FA2 afferents. They are responsible for the perception of high-frequency vibrations.

Interestingly, recent work also observed the very scarce presence of C-tactile (CT) afferents innervating the glabrous part of the hand [Watkins et al. 2021]. These receptors are usually present in the hairy skin regions, and are mostly involved in the perception of social and affective touch [Löken et al. 2009], and are deeply linked to the perception of hair deflection [Moore et al. 2025].

It is worth noting that the characterization presented in [Table 1.1](#) come from studies using specific stimulation hardware, including von Frey hairs (e.g. [Vallbo et al. 1978]) for the receptive field mapping and activation threshold measurement, mechanical probes for assessing

Table 1.1 – Properties of the four LTMRs in the glabrous skin of the hand. Receptive field sizes were extracted from [Vallbo et al. 1984]. Innervation density estimates were aggregated from [Vallbo et al. 1978; Vallbo et al. 1984; Johansson et al. 2009]. Sensitivity information, and hand images were adapted from [Johansson et al. 2009].

Channel	SA1	SA2	FA1	FA2
End organ	Merkel disks	Ruffini ending	Meissner corpuscle	Pacinian corpuscle
Adaptation rate	Slow	Slow	Fast	Fast
Sensitivity	<5 Hz vibrations static force	Static force Skin stretch	5-50 Hz vibrations spatial discontinuities (e.g. edges)	40-400 Hz vibrations
Receptive field size (mm <sup>2</sup> )	11 	59 	12.6 	101 
Innervation density (cm <sup>-2</sup> )	Fingertip: 70 Palm: 8 	Fingertip: 9 Palm: 16 	Fingertip: 140 Palm: 25 	Fingertip: 21 Palm: 9 

constant-pressure response (e.g. [Johansson et al. 1982; Vallbo et al. 1984]), or even vibrators (e.g. [Bolanowski Jr et al. 1988]) or moving contactors (e.g. [Verrillo 1966]) or for characterizing the frequency response. Yet no measurements have been performed for UMH stimulation. This thesis addresses this issue by attempting to characterize the frequency-response of LTMRs to different UMH patterns, as presented in [Chapter 2](#).

**From the Hand to the Central Nervous System** The deformation of the mechanoreceptor leads to variations of the electrical charge distribution around the nerve fiber, which, after reaching a certain threshold, lead to the opening of ion channels. This generates an **action potential** (also often referred to as a **spike**), corresponding to an electrical charge propagating along the axon until it reaches the terminals [Kandel et al. 2012a]. This process is called **mechanotransduction**. For these LTMR, the axon extends all the way from the hand to the spinal cord, where they synapse to other neurons, as an entry gate to the central nervous

system (CNS).

It is worth noting that different types of afferents are composed of different types of nerve fibers [Kandel et al. 2012b; Varga et al. 2015]:

- Type A fibers are large myelinated fibers, grouped into  $A\alpha$  (proprioception),  $A\beta$  (touch, pressure),  $A\gamma$  (for muscle spindles), and  $A\delta$  fibers (pain, coldness).
- Type B fibers convey autonomic responses from the autonomic nervous system to ganglia. Those are not involved in touch.
- C fibers are small unmyelinated fibers, whose functions include the perception of pain and temperature.

The four LTMR of the glabrous skin of the hand all use  $A\beta$  fibers, which is a fast, large fiber. These fibers have also been shown to encode social touch information [Xu et al. 2024]. On the other hand, CT afferents use type C fibers. Nerve fibers also differ in spike duration ( $A < B < C$ ), conductance velocity ( $A > B > C$ ), and size ( $A > B > C$ ) [Varga et al. 2015].

It is also worth noting that the skin also contains receptors for other types of somatosensory cues. Mostly, thermoceptors (sensitive to temperature) and nociceptors (sensitive to pain) are present, and mostly innervated with  $A\delta$  and C fibers [Jones et al. 2008], but this is exceeding the scope of this thesis.

Overall, many nerve fibers are grouped into a single *fascicle*, and groups of fascicles are combined to form a peripheral nerve, transferring the signal from the receptor group to the CNS. Three nerves innervate the hand [Kandel et al. 2012b]:

- The median nerve innervates the thenar region of the palm, along with the middle, index, and thumb fingers.
- The ulnar nerve innervates the rest of the palmar side of the hand, including the pinky, ring finger, and hypothenar part of the hand, along with the back of the pinky, and half of the back of the ring finger.
- The radial nerve innervates the rest of the back side of the hand.

**Measurements** It is important to measure the activation of the tactile afferents to better understand how they react to specific stimuli, and therefore optimize the design of sensations accordingly. One of the main techniques involved is microneurography [Vallbo 2018]. It implies inserting a tiny metal needle inside the nerve fiber, and comparing the recorded potential

with that of a reference electrode. While this technique provides extremely valuable information about single-afferent activation, it is also complicated to apply, requiring highly trained experimenters to position the needle and classify the measured afferent. It also often requires very long experiments due to the complicated setup and manipulation, and requires the participant to remain very still, as micro-movements of the skin could move the needle out of the nerve fiber.

In [Chapter 2](#), we apply microneurography techniques to characterize the specific response of single afferents to UMH stimulation. Doing so, we provide the very first microneurography measurements for UMH.

**Cognitive Integration** After a spike emitted from the hand reaches the spinal cord, nerve fibers convey it to the brain through the lemniscal pathway, synapsing in the thalamus, before being relayed to the somatosensory cortex [[Goldstein et al. 2022b](#)]. Once there, combined signals from different segments will trigger different regions of the cortex, depending on the origin location of the signal [[Schott 1993](#)]. The combined signals from all nearby receptors are then processed together in the cortex, being interpreted and integrated together to generate the resulting sensation [[Goldstein et al. 2022b](#)].

On a higher-level, it is important to note that many factors influence this cognitive integration of tactile stimuli. For instance, added visual stimulation (and multimodal stimulation in general) influences the cognitive integration of tactile stimuli [[Longo et al. 2012](#)], especially depending on the congruence between the different modalities [[Evangelou et al. 2023; Evangelou et al. 2024](#)].

The type of interaction can also significantly alter cognitive integration. For instance, [Talvas et al. \[2014\]](#) review bimanual haptic interactions, explaining that the inputs from two hands working together on the same task can be integrated together. This can provide improved performances or avoid splitting the user’s attention. In [Chapter 9](#) we explore bimanual UMH interactions in VR and compare them to their unimanual equivalents.

Since we now presented how an external stimulation can be turned into a haptic sensation, it becomes tempting to use this knowledge to artificially induce desired haptic sensations, whether to further study haptic perception, or use the sense of touch to convey specific information.

## 1.2 Haptic Rendering

**Haptic interfaces** (also called haptic displays) [Jones 2018; Culbertson et al. 2018] are designed specifically to provide artificial haptic stimulation. This process of encoding a piece of information as a haptic stimulus is called haptic **rendering**. These technologies involve both hardware (Section 1.2.1) and algorithmic (Section 1.2.2) components, working in harmony with the user to provide haptic stimulation, or even a human-computer interface.

### 1.2.1 Haptic Interfaces

A very wide variety of hardware technologies has been developed to provide different kinds of stimulations in different scenarios [Jones 2018; Culbertson et al. 2018; Otaduy et al. 2016]. Culbertson et al. [2018] classified haptic interfaces into three categories:

- **Graspable** interfaces require the user to grasp part of the interface, often providing force feedback (e.g. [Silva et al. 2009; Quek et al. 2014]).
- **Wearable** interfaces include tactile systems providing localized stimulation to the user, and body-grounded (e.g. exoskeleton) force-feedback devices (e.g. [Cempini et al. 2015; Pacchierotti et al. 2016]).
- **Touchable** provide stimulation when the user actively comes in contact with the interface to explore it and receive the stimulation (e.g. [Bau et al. 2010; Bianchi et al. 2015]). The delivered stimulation is usually cutaneous.

While this is a great starting point to discuss haptic interfaces, there are non-empty intersections between these categories (e.g. the WeATaViX [Tinguy et al. 2020] is a wearable interface involving grasping). On the contrary, other interfaces, such as UMH interfaces [Rakkolainen et al. 2021], do not fit very well to those definitions (being closer to touchable) due to the contactless nature. Later, Adilkhanov et al. [2022] proposed another taxonomy, classifying interfaces first based on their level of wearability, and then with a second level designed specifically for each wearability type. But once again, UMH does not fit easily, being closest to the grounded-touchable category.

Haptic interfaces could instead be classified based on their different properties, such as:

- Its position relative to the user: grounded (e.g. [Silva et al. 2009; Howard et al. 2020]), or wearable (e.g. [Pacchierotti et al. 2016; Tinguy et al. 2020]).
- The type of contact delivered: permanent or semi-permanent (e.g. [Silva et al. 2009; Pacchierotti et al. 2016]), encounter-type (e.g. [Tinguy et al. 2020; Bouzbib et al. 2024]), or mid-air (e.g. [Adel et al. 2018; Cha et al. 2017]).

- The type of cues provided: cutaneous (e.g. [Bau et al. 2010; Pacchierotti et al. 2016]), or kinaesthetic (e.g. [Silva et al. 2009; Cempini et al. 2015]).

Combinations of these criteria allow for a decent classification of haptic interfaces, and allow a classification of UMH as a grounded mid-air cutaneous interface. Yet, it is worth noting that a given interface can potentially fit into several categories at once if they provide many features (e.g. WEART’s TouchDIVER Pro gloves<sup>1</sup> provide both kinaesthetic and cutaneous stimulation at the same time).

Apart from the intuitively necessary processing unit and actuators, most haptic interfaces also require sensors to be able to provide dynamic reactive feedback that follows the user’s interaction. These can include force or position sensors for servo control of the device (e.g. [Silva et al. 2009]), or hand trackers for position-sensitive stimulation (e.g. [Zhu et al. 2025]).

### 1.2.2 Haptic Rendering Schemes

In order to provide dynamic stimulation that is responsive to the user’s interaction, the haptic interface’s behaviour has to be algorithmically coupled to the user’s body tracking, and/or virtual environment. The algorithm is also tasked with the synchronization between components. We refer to these algorithms as **coupling schemes**. These coupling schemes aim to compute the information that needs to be conveyed to the user, and encode it as a haptic stimulus that will be delivered by the interface [Lin et al. 2008; Otaduy et al. 2013; Otaduy et al. 2016; Xia 2018].

Such coupling often involves interactions with an environment, in which *collision detection* can be performed to gain information about how the user is interacting with the environment. If the environment is virtual (e.g. in VR scenarios), the collision detection is often done using some sort of *physical simulation* which may be able to provide interpenetration or forces applied between components. These collisions can then result in *constraints*, which are mathematical conditions that restrict the behaviour of some objects (e.g. preventing interpenetration).

Interacting with a virtual environment allows for some controlled discrepancies between the virtual and real world. For instance, the position or speed of the virtual hand does not have to exactly match that of the real hand. Instead, a **proxy** hand can be used to provide smoother vir-

---

1. <https://weart.it/touchdiver-pro-haptic-gloves/>

tual experience or artificially increase the interaction workspace [Lobo et al. 2017; Matthews et al. 2019; Clarence et al. 2022]. Otaduy et al. [2013] formalizes this concept for force-feedback interfaces, presenting a coupling algorithm as an intermediary between the real haptic interface, the constrained virtual equivalent proxy, and the rest of the virtual environment, as illustrated in [Figure 1.2](#). The visual representation of the virtual hand can also be altered for computational, immersion, or task improvement purposes [Otaduy et al. 2016; Normand et al. 2024b].

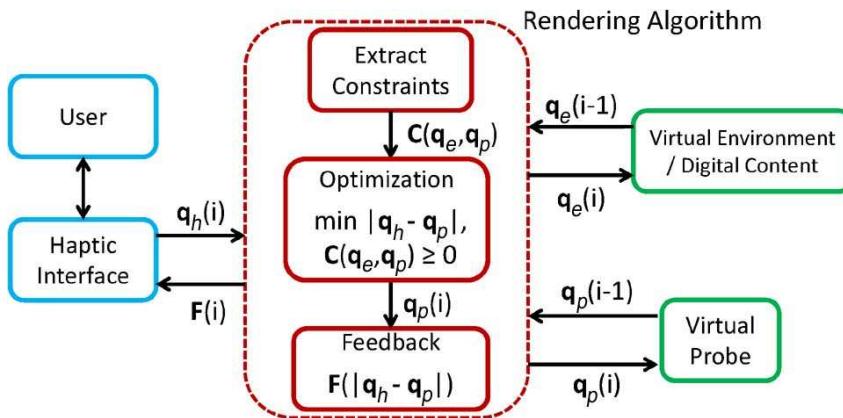


Figure 1.2 – General representation of the coupling algorithm (red) for a constraint-based force-feedback system.  $q_h$ ,  $q_e$ ,  $q_p$  represent the configurations of the haptic interface, virtual environment, and virtual probe (proxy) respectively.  $C$  represents constraints (e.g. preventing the virtual proxy from interpenetrating with a virtual object), and  $F$  represents the force feedback to generate. Image from [Otaduy et al. 2013].

Texture rendering also involves complex algorithmic coupling schemes. Indeed, haptic rendering of texture sensations aim at encoding micro-element exploration, that provides very different stimulations depending on the exploration speed and force applied on the surface [Ahmaniemi et al. 2010; Normand et al. 2024a]. The generated feedback thus has to be precisely coupled with the hand motion to provide realistic stimulation.

Talvas et al. [2014] reviewed bimanual haptic interactions, for which the coupling pipeline is illustrated in [Figure 1.3](#). These interactions contain an added layer of difficulty, as they not only need to couple the hands with the environment, but also need to couple the hands together [Dupin et al. 2015]. Yet despite the added difficulty, bimanual interactions can allow for a combined task integration of both hands, which can lead to improved performances or lower cognitive load [Talvas et al. 2014]. This is why in [Chapter 9](#) we propose a coupling scheme for bimanual UMH rendering in VR, and assess the benefits of bimanual rendering and bimanual interactions over their unimanual equivalents.

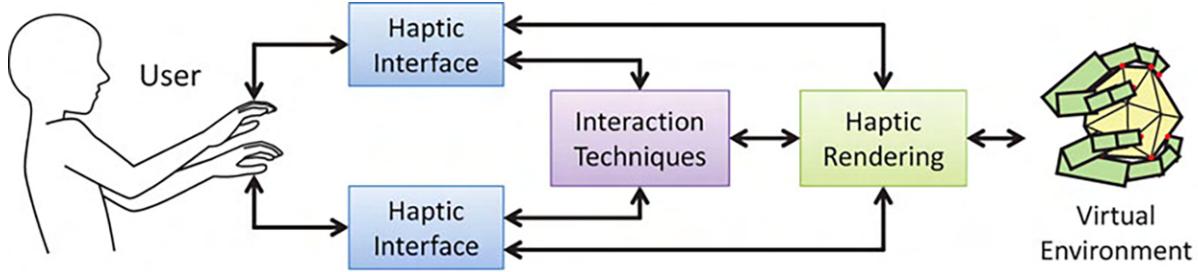


Figure 1.3 – Overall depiction of a bimanual coupling scheme for interactions with a virtual environment. The user interacts with a haptic interface that acts both as an input (sensor) and output (actuator) layer. The haptic rendering algorithm translates these actions into changes of the virtual environment, and translates these changes into new haptic feedback. The interaction techniques act as an intermediate to all these components, and dictates how interactions with the virtual environment must be conducted. Image adapted from [Talvas et al. 2014].

### 1.3 Mid-Air Haptics

**Mid-air or non-contact** haptic interfaces use an effector that is not directly in contact with the user's skin [Arafsha et al. 2015]. The stimulus thus has to somehow travel in the air until it reaches one's skin. This usually allows for very high interaction freedom, as the user is not constrained to touching or grasping a grounded interface. Many sources have been studied to provide mid-air stimulation, as illustrated in Figure 1.4. The main types are discussed below.

Historically, the first mid-air haptic interfaces leveraged **air jets**. For instance, Suzuki et al. [2005] proposed an interface composed of 100 air jet nozzles that could expel air at different speeds, and therefore different distances. A nearby hand would then feel air being blown onto the skin, generating a tactile sensation. Overall, this interface creates a sort of height map that could be interpreted as the shape of an object, or a surface to be explored. Air jets interfaces were later refined to provide richer sensations by adjusting the air flow (e.g. continuous, impulses, or modulated), and be scaled appropriately to the given application [Shen et al. 2024a].

**Air vortex** interfaces also as a source to convey the tactile sensation. By generating a precise flow of air swirling around a ring, one can create an air vortex that can move through the air, and apply a force when colliding with the hand, generating a tactile sensation [Gupta et al. 2013; Sodhi et al. 2013; Hein et al. 2022]. While air vortex technologies allow for a relatively long distance stimulation, they suffer from a high latency, due to the vortex' travel time. They also cannot provide sustained stimulation, but would instead use a sequence of vortices.

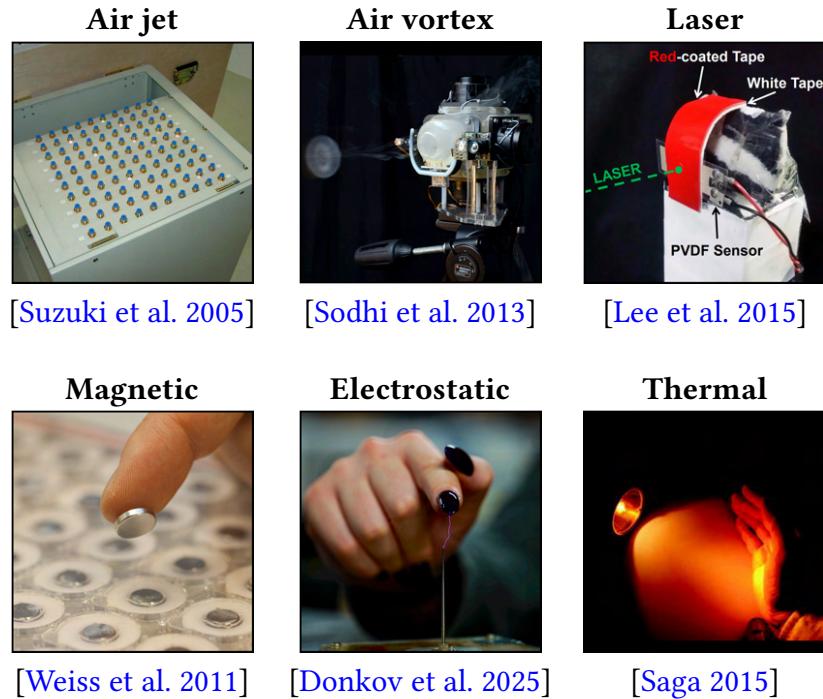


Figure 1.4 – Many types of non-contact interfaces have been developed, leveraging widely varied media.

For even longer-range stimulation, one can rely on **lasers**. More specifically, it is possible to point a high-intensity laser towards a thermoelastic medium in contact with the skin. The heat generated by the laser will thus deform the medium, which can in turn transmit an elastic wave to the skin, providing a tactile sensation [Lee et al. 2015; Cha et al. 2017]. Jun et al. [2015] also proved that it is possible to stimulate the skin directly, without the intermediate medium, thus allowing more natural interactions. Using femtosecond lasers, Ochiai et al. [2016] were able to focus light into single points in mid-air, resulting in the generation of light-emitting plasma. By quickly changing the position of the point, they successfully created holograms in a  $1\text{ cm}^3$  workspace, that could also provide tactile sensations on contact. These laser-based stimuli are usually perceived by users as a mostly mechanical or electrical sensation, similar to pricking or stinging [Jun et al. 2015].

By using an array of controllable electromagnets, and attaching a magnet to the user's finger, it is then possible to generate both repulsive and attractive forces to the finger through a controlled **magnetic** field [Weiss et al. 2011; Adel et al. 2018]. These interfaces can then also generate vibrating sensations by quickly alternating between repulsion and attraction.

By using a small Tesla coil, it is possible to generate electric arcs between a high-voltage pin

and nearby skin, providing **electrostatic** stimulation [Spelmezan et al. 2017]. To create more detailed spatial patterns, Spelmezan et al. [2017] combined many such pins in an array, allowing the user to move freely above the array while still being able to receive feedback. Donkov et al. [2025] followed a completely different approach: they showed that one can control the spark’s contact point, allowing for spatiotemporal stimulation. For instance, they measure the identification rate between line directions and simple shapes, obtaining an average of 40-50% correct answers across 14 possibilities.

One can also provide mid-air **thermal** cues by radiating a light onto one’s skin [Saga 2015; Iwai et al. 2019], or even by directly blowing air at the user [Xu et al. 2019]. While they can provide interesting cues, thermal interfaces also suffer from important limitations. Namely, the spread of heat in the skin makes it ill-suited to provide high spatial resolutions [Jones et al. 2008]. Thermal conduction is also a naturally slow process, which may result in latency, and difficulties in quickly alternating between cold and hot stimulation.

**Ultrasound-based** mid-air interfaces [Iwamoto et al. 2008b; Hoshi et al. 2009; Carter et al. 2013] leverage interference of acoustic waves to provide localized mid-air tactile sensations. This technology, referred to throughout this thesis as ultrasound mid-air haptics (UMH) but also known by several other names in the literature, has become the most mature form of mid-air haptic feedback. These interfaces are a very interesting candidate to provide mid-air stimulation, as they do not require any equipment on the user, while providing decently intense sensations in a relatively large workspace. They are also highly reconfigurable, giving large freedom to haptic designers, which allows for the rendering of rich cues. These interfaces are the main focus of this thesis, and are described in details in [Section 1.4](#).

Interestingly, UMH interfaces can also sometimes be combined with other types of mid-air stimulation, to provide multimodal stimulation. For instance, Hashizume et al. [2017] proposed a system that can simultaneously provide UMH cues and air vortices at the same location. Alternatively, Hashizume et al. [2017] combined UMH with thermal stimulation. They used the generated pressure field from the UMH interface to guide the heat towards the hand, and provide simultaneous thermal and UMH tactile cues.

## 1.4 Ultrasound Mid-Air haptics

UMH interfaces are currently the most mature form of mid-air haptic technologies. They do not require any equipment on the user, while still enabling interactions in a decently large

workspace. The large amount of rendering parameters also provides a high reconfigurability and enables the design of rich sensations. All of this explains the high amount of attention that UMH interfaces have received in the literature. Yet, they were first introduced in 2008 [Iwamoto et al. 2008b], making it a rather recent technology compared to most types of haptic interfaces. This is why we still have a lot to learn about how their stimuli are perceived, or how we can efficiently use these interface to efficiently render tactile patterns.

In this section, we start by presenting the hardware and physical mechanisms behind UMH ([Section 1.4.1](#)), before discussing the state-of-the rendering primitives ([Section 1.4.2](#)) and how they are used to render and study the perception of tactile shapes ([Section 1.4.3](#)). Finally, we discuss helpful tools that have been designed to help UMH designers create, model, or visualize UMH stimuli ([Section 1.4.4](#)).

### 1.4.1 Phased Array and Focal Point

**Ultrasound Transducers** To enable tactile stimulation in mid-air, UMH interfaces rely on a collection of ultrasound transducers, which roughly act as small speakers emitting an ultrasonic wave [Rakkolainen et al. 2021]. These transducers are also often referred to as **emitters**.

The transducers are often leveraging piezoelectric effects: applying an electrical signal generates a mechanical deformation of a small crystal. When alternated, this deformation can be thought as a vibration, which is transmitted to the air in its vicinity, creating a sound wave. These are usually operated at their resonant frequencies, in the ultrasonic range, above 20 kHz. This is the frequency of the generated wave, referred to as **carrier frequency**. Most current devices use a carrier frequency of 40 kHz [Iwamoto et al. 2008b; Carter et al. 2013; Inoue et al. 2016; Suzuki et al. 2021], with rarer examples at 70 kHz [Ito et al. 2016]. It is also worth noting that some prototype investigate the use of different transducer technologies for improved efficiency [Kamigaki et al. 2017; Van Neer et al. 2019; Marzo 2022].

**Generating a Focal Point** Then, many transducers are put together to create a complete UMH interface, as shown in [Figure 1.5](#) (B). Initially, the first UMH interface [Iwamoto et al. 2008b] used transducers that were not all independent of each other. Therefore, they positioned each one at the perfect position, so that the waves emanating from all transducers interfered constructively at a given position in mid-air. This constructive interference meant that the local pressure was increased at this so-called **focal point**. The focal point generation process is illustrated in [Figure 1.5](#) (A). Yet the non-independence of these emitters implied that the

focal point could only be moved along a line perpendicular to the interface, which is very restricting.

Therefore, later models were designed so that each unique transducer can be independently adjusted, and algorithms were developed to compute the correct set of phases and intensities to generate a focal point at a target position [Hoshi et al. 2009; Hoshi et al. 2010]. This enabled a whole new range of possibilities by allowing 3D control of the focal point.

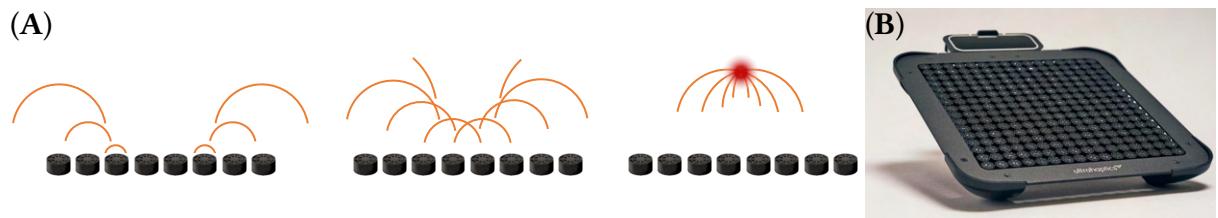


Figure 1.5 – (A) Generation of a focal point. Wave fronts emitted from all transducers are shown in orange, at three successive time steps. By correctly setting the initial phase offset for each transducer, all waves can constructively interfere at the focal point (shown in red). Image adapted from [Rakkolainen et al. 2021]. (B) Lots of transducers are then combined to create a full UMH interface. The model shown is a STRATOS Explore from Ultraleap, which is the device we used during most of the contributions of this thesis. Image from [Ultraleap](#).

**Underlying Physics Theory** When a solid medium is blocking the sound waves, it undergoes a **radiation force**, due to the **acoustic radiation pressure** [Awatani 1955]. This force acts in the direction the wave is moving, thus pushing on the blocking medium. This effect was first described in the case of UMH by Iwamoto et al. [2008]:

$$P = \alpha \frac{p^2}{\rho c^2}$$

where  $P$  [Pa] is the acoustic radiation pressure,  $p$  [Pa] is the root mean square (RMS) value of the sound pressure,  $c$  [m/s] is the sound speed,  $\rho$  [kg/m<sup>3</sup>] is the propagation medium's density, and  $\alpha \in [1, 2]$  is a constant related to the quality of the reflection on the blocking medium. In our case we are mostly interested in mid-air interactions on the skin, so  $\rho$  characterizes the ambient air, and  $\alpha \approx 2$  indicates that the skin reflects almost all the ultrasound energy. By considering the ultrasound waves follow a beam, the radiation force can be computed as:

$$F = PS \cos \theta \quad (1.1)$$

where  $F$  [N] is the radiation force,  $S$  [ $\text{m}^2$ ] is the surface area, and  $\theta$  [rad] is the incident angle between the acoustic beam and the surface that receives it [Awatani 1955; Hoshi 2022a]. This highlights the non-linearity of the acoustic effects, where the force received is proportional to the square of the sound pressure, and non-linearly decreases as the incidence angle increases.

Assuming that all transducers from a  $N \times N$  square matrix interface provide the same pressure  $p_r$  [Pa] at a distance  $r$  [m] above the center of the interface, we can approximate the pressure field at the focal plane during the time of maximum amplitude [Hoshi et al. 2010]:

$$p(x, y) = \sqrt{2}p_r N^2 \left| \frac{\operatorname{sinc}\left(\frac{Ndv_x}{2}, \frac{Ndv_y}{2}\right)}{\operatorname{sinc}\left(\frac{dv_x}{2}, \frac{dv_y}{2}\right)} \right| \quad (1.2)$$

with  $N$  the number of transducers on each line,  $d$  [m] the distance between transducers,  $v_x = \frac{k}{r}(x - x_c)$ ,  $v_y = \frac{k}{r}(y - y_c)$ , with  $k$  [rad/m] the wave number,  $(x_c, y_c)$  the position of the focal point, and  $\operatorname{sinc}(a, b) \equiv \frac{\sin a \sin b}{ab}$ . This pressure field is illustrated in Figure 1.6. Note that the shape of this pressure field is highly dependent on how the transducers are positioned relative to each other.

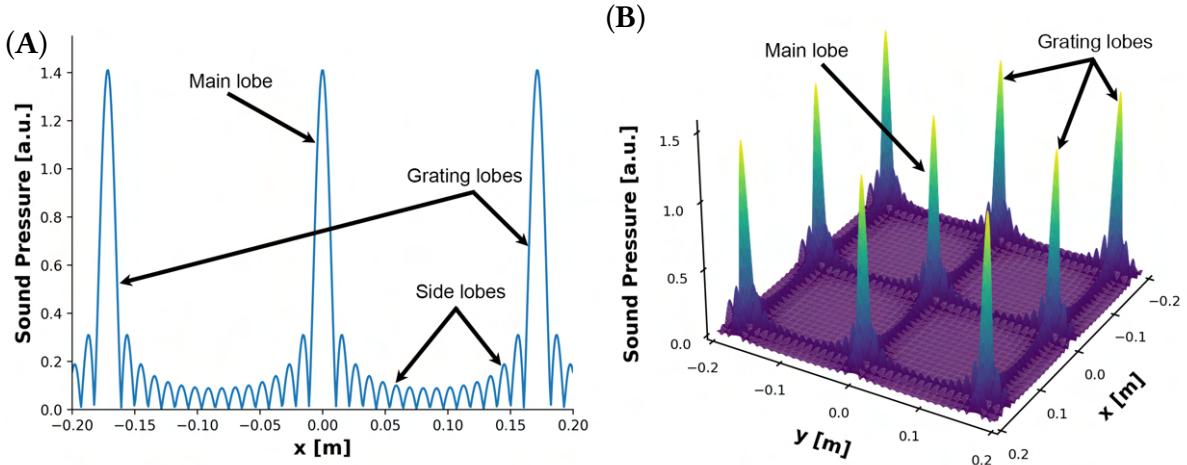


Figure 1.6 – Pressure field simulation (see Equation 1.2) generated from the creation of a focal point 20 cm above the center of a STRATOS Explore interface.  $x, y$  coordinates are relative to the center of the focal point, while the sound pressure values are reported in an arbitrary unit. (A) Pressure distribution in the  $y = 0$  cross-section. (B) Pressure distribution along a plane parallel to the interface, 20 cm away. In both illustrations, we can observe a main lobe of high pressure at the focal point, surrounded by smaller side lobes. High-pressure grating lobes can also be observed periodically. Figures reproduced from [Hoshi 2022a].

This pressure field highlights two interesting features.

First, we can clearly see on [Figure 1.6](#) (A) the main lobe at the focal point position, surrounded by small side lobes. While the focal point is not a solid object, it can be represented by its pressure field distribution, which has no objective size but is an indicator of the spread of acoustic energy. [Hoshi et al. \[2010\]](#) chose to use the width between the closest two zero-pressure points around the center of the focal point. Namely, they derived that:

$$w = 2\lambda \frac{r}{Nd} \quad (1.3)$$

with  $w$  [m] the defined focal point diameter, and  $\lambda$  [m] the ultrasound wavelength. Using similar methods, they were able to compute the focal “depth”  $\delta$  [m] of the focal point:

$$\delta = \frac{4\lambda r^2}{N^2 d^2} \quad (1.4)$$

For devices similar to that shown in [Figure 1.5](#) (B), this leads to  $w \approx 2$  cm,  $\delta \approx 5$  cm at  $r = 20$  cm. It is worth noting that these indicators still are not an objective size representation, as the focal point can only really be described as a field, but it gives a good idea of its behaviour. Yet, one could have chosen other indicators, such as the full width at half maximum (FWHM), which would lead to a smaller diameter measurement equivalent to  $0.61w$  [[Ito et al. 2016](#)]. In any case, these measurements underline the relatively large size of the focal point, which is the base unit of rendering. It is therefore very complicated, if not impossible, to render perceivable detailed cues.

The second interesting thing that we can see on the pressure distributions ([Figure 1.6\(B\)](#)) are these other tall lobes that appear periodically. These are referred to as **grating lobes**. They are side lobes, where fully constructive interference happen, leading to a high local pressure comparable to that of the focal point. The positions of these grating lobes are directly tied to the pressure distribution, and therefore to the interface’s shape. Yet, they are usually unwanted, and can provide confusing cues to the user [[Price et al. 2018](#); [Rakkolainen et al. 2021](#)]. From [Equation 1.2](#), it can also be derived that these grating lobes are periodically repeating, with a distance of  $\lambda \frac{r}{d}$  between them. Yet, this overall model suffers from a few limitations:

- The ultrasound wave’s energy fades out with the travelled distance. For instance, this means that the four grating lobes directly left, right, back, and in front of the main lobe can be perceived, while the others do not actually provide enough energy.

- Derivations for [Equation 1.2](#) assume that the wave emitted from a transducer is perfectly spherical. Yet this is not entirely true, as transducers are known to have some directivity [[Drinkwater 2022; Marzo 2022](#)].
- These equations model the field generated by an interface that uses a square grid of ultrasound transducers, but the result could be very different if other configurations were used [[Price et al. 2018](#)].

While this model is still used nowadays due to its relative simplicity and accuracy, other work has been conducted to provide more precise predictions of the acoustic field [[Bruus 2012; Drinkwater 2022; Marzo 2022](#)]. These attempt to overcome the previously mentioned limitations to improve the accuracy of the model, thus increasing the predictability of designed stimuli.

**Impact of the Carrier Frequency** In this section, we introduced a first rendering parameter: the carrier frequency  $f = \frac{c}{\lambda}$ . While this parameter is usually tied to the choice of the transducers, and therefore cannot usually be changed afterwards, it has critical implications.

First, [Equation 1.3](#) and [Equation 1.4](#) show that the size of the focal point is inherently proportional to the wavelength, which is inversely proportional to the carrier frequency. Increasing the carrier frequency would thus lead to a smaller, more precise focal point.

On the other hand, sound waves get attenuated as they travel through the air. Acoustics theory states that higher frequency waves are attenuated faster [[Bass et al. 1995; Hoshi et al. 2021; Rakkolainen et al. 2021; Battista 2022](#)]. This suggests that the usable interaction zone, often called **workspace**, gets smaller as the carrier frequency is increased. That said, results from [Ito et al. \[2016\]](#) seem to indicate that this difference in attenuation may be negligible between 40 kHz and 70 kHz carrier frequencies.

Transducers at 40 kHz are also quite easier to get, compared to 70 kHz ones, as they are currently mass-produced, resulting in cheaper products. Similarly, standard piezoelectric 40 kHz transducers provide a good balance between their size and required power [[Rakkolainen et al. 2021](#)].

**UMH Interface Shapes** Several UMH interfaces have been designed in many configurations and sizes to optimize the generated stimulus intensity, reduce the grating lobes, or provide a more comfortable user experience through a better workspace. Some of them are illustrated in [Figure 1.7](#). The first UMH interface [[Iwamoto et al. 2008b](#)] was designed using transducers aligned on concentric circles. This implied that all transducers on the same circle

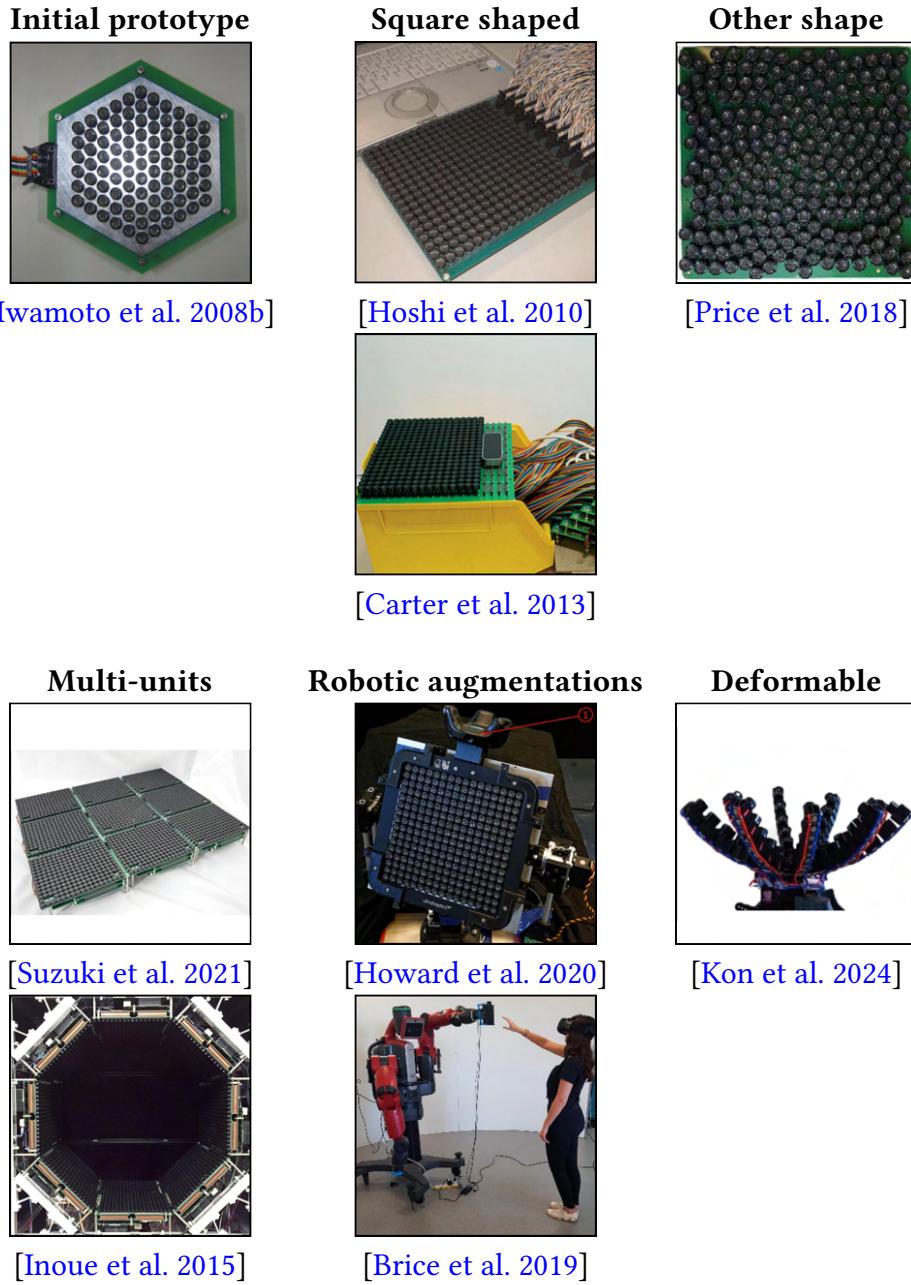


Figure 1.7 – Illustration of prototype UMH interfaces.

were at the same distance from the center, which meant they should share the same phase to create a focal point along the normal axis to the interface. While this allowed for a fairly simple control algorithm, it also imposes that the focal point could only be moved along a line perpendicular to the interface, which is very limiting. It is worth noting that this interface was introduced in 2008, making UMH a rather recent technology compare to most fields of haptics.

Although UMH has gained a lot of attention since then, its novelty implies that still currently have limited knowledge about how to best use these technologies.

The next interfaces [Hoshi et al. 2009] instead aligned transducers on a grid, and the internal phase solver was improved to allow for a full 3D control of the focal point. This design was very convenient to make, and is still used by the main interfaces today [Carter et al. 2013; Inoue et al. 2016; Suzuki et al. 2021]. In particular, rectangular interfaces could easily be combined to create larger ones [Suzuki et al. 2021; Inoue et al. 2022]. By adding more transducers, one can therefore increase the sound pressure at the focal point, leading to stronger sensations. As shown in [Equation 1.3](#), and measured by Hasegawa et al. [2018], creating an overall larger interface reduces the size of the focal point, which is desirable to render more precise cues.

Yet, matrix-shaped UMH interfaces offer one significant drawback: the generation of grating lobes ([Figure 1.6](#)). Therefore, other emitter arrangements were considered to mitigate the apparition of these unwanted stimulation spots. First, Gavrilov [2008] relied on a random arrangement of the transducers to break the symmetry and periodicity that was causing the grating lobes with the square layout. Then, spiral-based designs were proposed to reduce grating lobes, while also providing more control over the generated pressure field based on the sparsity of the spiral arrangement [Price et al. 2018; Allevato et al. 2022; Allevato et al. 2023].

Usable workspace is critical when interacting with a UMH interface, as it defines the area where the user can experience a stimulus. Due to the high attenuation of ultrasound waves, the usable workspace is usually limited to a few tens of centimeters above the interface. Thus, two main approaches were developed to increase the usable workspace. First, rectangular interfaces were put together to increase the planar surface covered by the transducers [Inoue et al. 2016; Suzuki et al. 2021]. Users could therefore move their hand above any of the units without losing the sensation. But it is also possible to combine UMH units for a different objective: allowing stimulation from different angles [Inoue et al. 2014; Inoue et al. 2015; Furumoto et al. 2019; Furumoto et al. 2020]. To do so, multiple UMH interfaces are positioned on a large structure, all pointing towards its center. This allows both the stimulation of the hand from any direction [Inoue et al. 2014; Inoue et al. 2015], and the manipulation of light objects by applying a force on them [Furumoto et al. 2019; Furumoto et al. 2020]. Similarly, we proposed a two-units system to enable bimanual interactions where each hand can be independently stimulated by its dedicated unit [Mulot et al. 2025] (detailed in [Chapter 9](#)).

The other approach consists in using robotic solutions to move the interface to follow the

user. In this regard, [Howard et al. \[2020\]](#) attached an UMH interface to a pan-tilt robot, which allowed the user to move their hand anywhere around it. Doing so, they reported a 14-fold increase in workspace volume. Similarly, [Brice et al. \[2019\]](#), and [Villa et al. \[2022\]](#) both proposed systems where a STRATOS Explore interface is attached to a robotic arm, which moves the interface to follow the user’s hand. [Park et al. \[2025\]](#) went a step further and attached a small UMH interface to the user so that it always lies under their hand. This turns the interface into a wearable haptic interface, bringing the wearable’s benefits (nearly infinite workspace), but also drawbacks (weight and shape leading to unnatural interactions).

As resolving the phase for each transducer is a costly process, [Kon et al. \[2024\]](#) recently attempted to provide deformable UMH interfaces. Instead of adjusting the phase of a transducer, one could change their position to adjust the distance between them and the focal point, thus reducing computations. The prototype uses transducers laid on flexible arms that can be curved in by pulling on a string. Doing so lowers the focal point, allowing for some spatial control. While very promising, this does not yet allow for a full 3D control of the focal point.

Despite these hardware advancements, a static focal point is limited in the sensations it can produce: a static pressure on the hand, and a vibration at the carrier frequency. Yet the carrier frequency is too high to be perceived by tactile afferents, and the very light static pressure sensation would be quickly adapted [[Hoshi 2022a](#)]. This means that a static focal point is not really perceivable, and modulation will be required to render perceivable sensations.

### 1.4.2 Modulation Techniques

Different algorithms were therefore developed to encapsulate the high-frequency vibration in a lower-frequency perceivable envelope. Doing so aims to elicit sensations by activating the skin’s tactile afferents through this lower-frequency stimulation. To achieve this goal, several methods have been developed [[Hasegawa et al. 2022](#)], as illustrated in [Figure 1.8](#).

**Amplitude Modulation (AM)** Initial studies in UMH took inspiration from contact vibrotactile methods, to vary the amplitude of the generated ultrasound signal with respect to time. This variation is then repeated periodically at a perceivable frequency, thus encapsulating the 40-70 kHz vibration inside a lower-frequency envelope [[Iwamoto et al. 2008b; Hoshi et al. 2010](#)], as illustrated in [Figure 1.8](#). Doing so allows the creation of a vibration that is closer to that of the tactile afferents’ detection range, creating a perceivable signal.

The choice of the modulation waveform also highly contributes to the generated sensation.

	Amplitude Modulation	Lateral Modulation	SpatioTemporal Modulation
Focal Trajectory	Immobilized	Relatively Short Line	Variously Scaled 2D Pattern
Radiation Force	Intensity 	Intensity 	Intensity 

Figure 1.8 – Main modulation techniques for UMH. The intensity of an AM point is varied at a perceivable frequency. LM and STM instead rely on the repetitive motion of the focal point along a trajectory to stimulate each position at perceivable frequencies while providing a constant intensity. Adapted from [Hasegawa et al. 2022].

While modulating the sound pressure or radiation force to follow a sine wave provides smooth sensation, and enables frequency-characterization, other waveforms could be considered. For instance, Hasegawa et al. [2013] proved that we can alter the perceived texture of the sensation by changing the modulation waveform, similarly to what can be achieved with contact vibrotactile stimulation [Nai et al. 2021]. While investigating the detection threshold for single AM points [Hasegawa et al. 2018; Howard et al. 2019], it was also discovered that square waveforms lead to a lower threshold at low modulation frequencies, while being most easily detected at 200 Hz. This seem to indicate that UMH stimuli are mostly detected by FA2 afferents, that have a peak sensitivity near 200 Hz, while the added lower frequency harmonics of the square wave may activate other receptors. Yet no concrete afferent response measurements have been conducted yet to confirm this hypothesis.

The level of details we can provide with this technology is highly related to the size of the focal point, and the minimum perceivable distance between a pair of them. The lower this distance is, the more details can be fit in a given area. To measure this, two-points discriminations (2PD) procedures were employed [Howard et al. 2023a; Cataldo et al. 2023], comparing thresholds at different modulations frequencies. While the measured thresholds do not perfectly match between the two studies, they reveal an average 2PD threshold between 15-35 mm on the palm at 200 Hz modulation, and slightly higher at 50 Hz, partially due to the lower perceived

intensity.

Overall, AM mainly introduced two new rendering parameters: the **modulation frequency**, and the **modulation waveform**.

**Spatiotemporal Modulation (STM)** Yet, AM suffers from a critical flaw: the intensity variation implies that we do not leverage the whole power output that the interface could deliver, resulting in very weak sensations. To fully utilize the output power of the device, we would need to render the focal point at 100% intensity all the time. But as discussed in [Section 1.4.1](#), a static constant-intensity focal point cannot really be perceived.

Therefore, a new modulation technique was proposed, STM, relying on a time-variation of the focal point’s position [[Kappus et al. 2018](#); [Frier et al. 2018](#)]. By quickly and repeatedly scanning the focal point along a fixed trajectory at a perceivable frequency, all positions along the trajectory can be interpreted as being amplitude-modulated at the same time, as shown in [Figure 1.8](#).

Not only does this generate largely stronger sensations, but it also makes sampled positions perceivable at the same time. So, instead of perceiving a sequence of stimulated positions moving through time, the user perceives constant stimulation across the whole scanned trajectory. This notably allows the stimulation of larger stimuli, including shapes, as discussed in [Section 1.4.3](#).

Overall, STM introduced several new rendering parameters, including:

- the **draw frequency**, meaning the frequency at which the trajectory is scanned, also corresponding to the modulation frequency at each sampled location.
- the **average focal speed**, which corresponds the scanning speed of the focal point if its motion were continuous. It is worth noting that this value does not directly represent a physical phenomenon, as the focal point is in fact jumping from one sample position to the next in a discrete manner, not as a continuous motion. One could also consider the **instantaneous focal speed**, corresponding to the equivalent speed of going from one sample position to the next during one time step if the motion were continuous.
- the **sample rate**, corresponding to the number of sampled positions on the trajectory.

All of these parameters are crucial when rendering anything with UMH, as they can widely alter the produced sensations. Notably, the perceived intensity is very tied to the tactile afferent’s peak sensitivity range, and is therefore very impacted by the draw frequency [[Frier et al. 2018](#); [Frier et al. 2019](#)].

Another notable effect is due to the average focal. Moving a pressure source on the skin elicits shear waves that can travel up to tens of centimeters from away from the source. But if the source is moved faster than the shear waves are travelling, shear shock waves are generated, and a shock front forms [Reardon et al. 2019; Reardon et al. 2023]. The shape of this shock front changes how the waves travel across the skin, and thus likely impact the perception of the underlying stimulus. These effects also correlate with results from Frier et al. [2018], who showed that matching the focal speed with the surface wave speed can generate a larger skin displacement, and increase the perceived intensity. This indicates that shear waves and surface waves may both play an important role in UMH stimulus perception.

**Lateral Modulation (LM)** LM follows the same approach as STM, using the focal point motion to modulate the signal instead of using its intensity [Takahashi et al. 2018; Takahashi et al. 2020]. The large similarities stem from the fact they were both independently designed and proposed at almost the same time. Initially, LM was introduced as a way to modulate a point through a back-and-forth motion of the focal point on a line around it [Takahashi et al. 2018], as shown in [Figure 1.8](#). The authors later expanded this method to a 2D motion of the focal point around a small circle to provide a high-intensity point-like sensation.

While LM and STM are very similar and use the same time-motion principle to create stronger signals, we now usually consider that STM aims at rendering large stimuli, while LM aims at generating point-like sensations [Hasegawa et al. 2022].

It is worth noting that the range of perceivable frequency for touch is mostly included in the range of audible frequencies. This implies that working UMH interface often generate a noise as it renders a stimulus [Morzyński et al. 2024]. Discontinuities in the focal point motion are also known to generate high-frequency audible harmonics, that can potentially worsen the user experience.

**Other Methods** While, at the time of writing, STM, LM, and AM are the most used methods, it is worth noting that some completely different rendering techniques exist, mostly based on pressure field reconstruction [Matsubayashi et al. 2022].

While the main rendering primitive is usually a focal point, UMH interfaces are able to generate more complex structures. Instead of attempting to create a very high pressure at a single position, the transducer phase computation algorithm may also be adapted to instead repro-

duce any arbitrary pressure field

Inoue et al. [2015] initially used a collection of UMH interfaces to generate standing waves corresponding to the sides and edges of 3D shapes. Matsubayashi et al. [2020] developed another pressure field reconstruction algorithm specifically designed to account for the wave scattering due to the hand.

While these previous methods did not rely on the use of any focal point, other techniques relied on multiple focal points at the same time. [Carter et al. 2013] initially presented the possibility of generating simultaneous focal points. An improved version was then proposed, relying on the computation of eigenvectors to render slices of 3D shapes [Long et al. 2014]. This algorithm was then further refined by adding correction factors to the transducer amplitudes [Plasencia et al. 2020]. Barreiro et al. [2020] used these promising multi-focal point methods in combination with STM to compute trajectories for each focal point, resulting in better approximation of the target pressure field [Barreiro et al. 2020]. Finally, these multi-focal points methods to reconstruct target pressure fields were used in several physical-based interactions. These interactions rely on the computations of forces applied during a hand-object interaction, and multiple focal points are generated to approximate this force field [Jang et al. 2020; Barreiro et al. 2021].

**Sensations** While haptic sensations are very difficult to describe with words, UMH stimuli provide three main types of tactile cues:

- The main component is a **vibrotactile** sensation. Modulation techniques generate mechanical waves in the skin within the frequency range of tactile afferents that can be interpreted as vibrations. This is similar to what contact vibrotactile devices provide [Verrillo 1966].
- The continuous radiation force applied on the skin also generates a **static pressure** sensation. It has also been shown that by using a strong enough interface with a low modulation frequency, it is possible to remove the vibratory sensation, to leave a sensation that is more similar to a constant static pressure being applied on the skin [Morisaki et al. 2022; Morisaki et al. 2024].
- The very high pressure at the focal point can sometimes reach more than 160 dB SPL [Hoshi 2014], which is very high. At such pressures, the behaviour of sound waves in the air turns into a turbulent flow, which can create a sensation of “wind” at the surface of the skin [Hoshi 2022a; Drinkwater 2022]. This phenomenon is called **acoustic streaming**.

**Using Modulation Techniques** Modulation techniques can be considered as **rendering primitives**, and can then be extended or combined to create a proper rendering algorithm. To that end, [Frier et al. \[2019\]](#) introduced the concept of **sampling strategies**, encompassing the sampling of the focal trajectory, and the applied modulation. We then further formalized this process, both for 2D [[Mulot et al. 2021b](#)] (detailed in [Chapter 5](#)) and 3D rendering [[Mulot et al. 2024](#)] (detailed in [Chapter 6](#)).

For instance, it is possible to change the position of an AM focal point after a few modulation periods to create a **sensation of motion** [[Hoshi et al. 2009; Hoshi et al. 2010](#)]. The continuity and overall quality of the sensation was later evaluated with respect to some rendering parameters such as the number of sampled positions [[Wilson et al. 2014](#)].

These modulation techniques can also be used to encode **physical properties**. For instance, [Marchal et al. \[2020\]](#) showed that it was possible to render stiffness sensations by adjusting the maximum intensity of an AM point when a user presses on a virtual piston.

**Textures** have also received some interest, as they provide additional challenges. Indeed, real-life textures are usually composed of really small elements that are smaller than a focal point (see [Section 1.4.1](#)). While they are still large enough to be perceived by tactile afferents [[Biswas et al. 2019](#)], they cannot be directly rendered using UMH. To address this texture rendering problem, [Freeman et al. \[2017\]](#) chose to approximate textured surfaces as a regular tessellation of pyramids or tetrahedrons, to render an AM focal point at each intersection location. They can then vary the intensity, modulation frequency and waveform to alter the sensation. This method was then refined to automatically infer the roughness associated to an image, and account for the usually irregular aspect of textures, and dynamic stimulation to different hand speeds [[Beattie et al. 2020](#)]. After the development of STM, it was shown that the draw frequency can alter the perceived roughness of the stimulus [[Ablart et al. 2019](#)]. Finally, [Wojna et al. \[2023\]](#) studied the impact of visual and haptic rendering parameters on the perception of visuo-haptic particle-based texture.

In order to design stimuli that can easily be detected and distinguished, it is also important to well understand the perception of UMH stimuli, and the effect the different rendering parameters. To this end, several studies have attempted to measure just-noticeable differences (JND) to assess the effect of different rendering parameters, including the stimulus intensity, modulation frequency [[Lim et al. 2024a; Wojna et al. 2023a](#)], or the focal speed [[Rutten et al. 2020](#)]. These results both provide a better understanding of UMH perception, and help UMH designers create better experiences.

**Rendering on Different Locations** Interestingly, all locations on the skin do not respond identically to UMH stimulation, due to variations in innervation and physical properties (e.g. skin thickness).

Most of the research and applications for UMH have focused on the **glabrous skin of the hand**, because its dense innervation allows for the perception of rich and intense cues. In this skin area, it has been proven that a single AM focal point is perceived stronger for modulation frequencies around 200 Hz [Hasegawa et al. 2018]. This is close to the peak sensitivity of FA2 afferents [Johansson et al. 1982], hinting towards their major role in UMH perception, along with the predominance of the vibrotactile component on the palmar side of the hand. Yet no physiological measurements have yet been conducted to verify this hypothesis and characterize afferent response to UMH stimuli. This is why, in [Chapter 2](#), we present the first microneurography measurements for UMH, aiming to provide a better understanding of UMH perception.

When changing from a single AM point to a full STM pattern, it is worth noting that the draw frequency is not the only rendering parameter controlling the sensation. For instance, [Frier et al. \[2018\]](#) showed that matching the focal speed with that of surface waves on the skin could increase the perceived intensity of the stimulus. This truly captures the complexity of UMH perception, combining the effects of many rendering parameters, as well as neurophysiological, and biomechanical effects.

The **hairy skin of the arm** is innervated quite differently, with FA2 units being much rarer [[Corniani et al. 2020](#)]. This translates to a way weaker, or even unperceivable vibrotactile sensation. To address this issue, two main strategies have been employed. [Takahashi et al. \[2020\]](#) used a very strong multi-unit UMH interface (as shown in [Figure 1.9\(A\)](#)) to supply strong sensations, combined with lateral modulation, to provide perceivable vibrotactile feedback. A completely different approach consists in using a very slow motion of the focal point, at only a few cm/s [[Tsumoto et al. 2021; Pittera et al. 2022](#)], in a similar fashion to what is usually done to convey affective touch sensations [[Hauser et al. 2019](#)], as it corresponds to the peak sensitivity of CT afferents in hairy skin [[Löken et al. 2009](#)]. Interestingly, such a slow motion would not be perceivable on the glabrous skin, similarly to a static focal point. Overall, the first study aims at boosting the signal to continue using vibrotactile methods designed for the glabrous skin, while the second one provided methods tailored for hairy skin, exploiting the acoustic streaming rather than the vibrotactile component of the sensations.

Some studies have also focused on stimulating around the **knuckles**, on the dorsal side of the

hand to provide notifications [Montano et al. 2024] (see Figure 1.9(B)) and guidance [Zhu et al. 2025] cues. They found that classical STM stimulation is perceivable, and that the stimulus is unexpectedly strong in between two fingers. This is assumed to be caused by the sound waves repeatedly hitting back and forth between the two fingers.

Finally, the **face** has also been an area of interest for UMH stimulation, despite the complete absence of FA2 units in the skin [Corniani et al. 2020]. Yet, the face is very densely innervated, with a very thin skin, which still allows for stimulation with classical techniques. Yet we observe peak sensitivities closer to 40 Hz modulation, contrary to the 200 Hz for on the palm [Gil et al. 2018; Lan et al. 2024] (see Figure 1.9(C)). This is hypothesized to be caused by the FA1 afferents, which have peak sensitivities around this range. Overall, face stimulation has been conducted on the lips, teeth, cheek, nose, and around the eyebrows [Shen et al. 2022; Gil et al. 2018; Lan et al. 2024], with especially high perceived intensity reported on the lips, probably for similar reasons to the knuckles.



[Takahashi et al. 2020]

[Montano et al. 2024]

[Gil et al. 2018]

Figure 1.9 – UMH setups for stimulation the arm (A), knuckles (B), and the face (C).

Overall, a few rendering primitives have already been developed for UMH, and creative uses have been proposed to render a wide variety of stimuli. Yet these rendering techniques vary a lot depending on the stimulated location, and the type of property to render. Providing more rendering primitives will contribute to the UMH rendering toolkit, providing more possibilities to UMH designers.

### 1.4.3 Shape Rendering

While UMH is assumed to be mostly perceived by FA2 afferents, their large receptive field and low innervation density does not make them great candidates to perceive precise geometrical cues. UMH designers therefore need to rely on optimized rendering algorithms to provide convincing and distinguishable stimuli.

A UMH stimulus will be considered **static** if its associated sensation does not significantly change over time. For instance, a small circle rendered using STM at 50 Hz draw frequency will provide a static sensation, where the whole shape is perceived at once, despite the motion of the focal point. On the other hand, **dynamic** stimuli lead to changing sensations. This could be due to changes in the focal trajectory (e.g. a circle with varying size), slow focal point motion, or changes to the rendering technique (e.g. slow variation of the intensity over time). Note that the rendering can also be considered static or dynamic, depending on whether the generated stimulus depends on the user’s actions. For instance, a simple STM-rendered circle following the user’s hand consists in a static stimulus with dynamic rendering, while a shrinking STM circle at a fixed position corresponds to a dynamic stimulus with static rendering.

**2D Static Stimulus** Initial shape rendering methods relied on either the simultaneous rendering of several AM focal points, or its motion. Yet, the focal point has to spend at least one modulation period around the same area for it to be decently stimulated, which highly restrict the range of possible focal speeds. Despite this, [Korres et al. \[2016\]](#) provided an algorithm to turn shapes into point-clouds. An AM focal point then successively moved through these positions to create a shape sensation. They used their technique to measure shape identification accuracy across four 2D shapes. Similarly, [Rutten et al. \[2019\]](#) measured identification accuracies across four static and four dynamic patterns rendered using four simultaneous AM focal points. They showed accuracies between 26 and 60% depending on the pattern.

The introduction of LM and STM enabled the straightforward generation of static stimuli, due to the fast motion of the focal point. Initial studies focused on the rendering of circles, trying to find the optimal rendering parameters to maximize the perceived intensity [[Frier et al. 2018](#); [Frier et al. 2019](#)]. Specifically, they first demonstrated that the circle’s perceived intensity could be maximized if the focal speed matched that of surface waves on the skin [[Frier et al. 2018](#)]. Their work was then extended to measure the effect of the update rate and circle size on perceived intensity [[Frier et al. 2019](#)]. This work revealed that it is usually not optimal to use the maximum available update rate for rendering circles at low frequencies. More recent work seem to indicate that the sample rate does not play such an important role for higher draw frequencies [[Reardon et al. 2023](#)].

Lines have also received attention. [Howard et al. \[2019\]](#) first measured the detection threshold for STM lines, before measuring direction discrimination performances at different intensities. They also showed that we can achieve sensations of bumps and holes by respectively increasing or decreasing the focal point’s intensity around the center of the line. They obtained con-

sistent identification rates above 90% for a four-direction discrimination. [Morisaki et al. \[2024\]](#) also showed that small LM circles can be interpreted as flat or bumpy surfaces.

As detailed in [Chapter 5](#), we also conducted a user study to measure the 75% JND in curvature for STM arcs, revealing no effect of the sample rate in the tested range [[Milot et al. 2021b](#)].

**2D Dynamic Stimulus** To further improve identification rates for UMH shapes, it can be helpful to generate a dynamic stimulus, relying on time-variation to sharpen the feedback. This concept has been extensively applied to the rendering of lines, especially to generate apparent motion sensations. [Hoshi et al. \[2009\]](#) initially showed perfect identification rate for the 2-ways slow motion of an AM focal point, as a proof of concept for the UMH interface they designed. They then extended their work to also measure position detection accuracy, and trajectory discrimination across four patterns, including lines, circles, spirals, and a random point cloud [[Hoshi et al. 2010](#)]. Later, [Perquin et al. \[2021\]](#) rendered six simultaneous AM focal points and moved them together along a line that could take any of three directions: along the distal or radial axis of the hand, or diagonally. For each overall direction, participants were asked to identify in which way the focal points were moving (e.g. towards the fingers or towards the wrist if the direction was distal). Doing so, they reveal a direction bias, where participants struggled more to identify the motion along the oblique direction. [Wilson et al. \[2014\]](#) and [Pittera et al. \[2023\]](#) used an AM point moving along a line to measure whether they could induce an apparent motion sensation from the start to the end point, even if the focal point did not actually render the center part of the line. They also both evaluated direction identification rates across four possible directions.

[Hajas et al. \[2020\]](#) formalized this technique relying on a slow motion of an AM point, naming it dynamic tactile pointer (DTP), and used it to render a wide variety of 2D shapes. This slow motion of the focal point helps the user following and understanding the trajectory. They also proposed an even better variant for polygonal shapes, where pauses at the vertices further improves identification accuracy. Indeed, this pause is assumed to break the percept into smaller parts, leveraging a cognitive process called chunking [[Zhang et al. 2012](#)], which eases the integration into the complete sensation. While it provided sharp sensations with high accuracy, DTP relied on AM, which meant that sensations were fairly weak. We therefore proposed a variant, detailed in [Chapter 3](#), that leverages STM instead, to obtain stronger sensations while keeping similar identification rates [[Milot et al. 2023b](#)].

While dynamic stimuli definitely help provide sharper sensations, their dynamic nature im-

plies that it takes some time for the full stimulus to be rendered. This duration can be a large drawback during interactions, as the user would need to pause to understand the feedback before resuming their interaction. This is why static stimuli are still extensively used in practice.

**3D Shape Rendering** When creating a static sensation, the available acoustic energy is spread across the trajectory. This is why we usually render the contour of a 2D shape instead of rasterizing it: the longer the trajectory, the lower the energy available at each location. This means that applying naive rendering algorithms to higher-dimension shapes could tremendously reduce its perceive intensity, to the point where 3D shapes would likely not be perceivable. Specific algorithms thus have to be designed to allow such a feat, and this limitation often means we have to rely on dynamic rendering methods, that adapt to the position of the user’s hands. Some of these existing methods are illustrated in Figure 1.10.

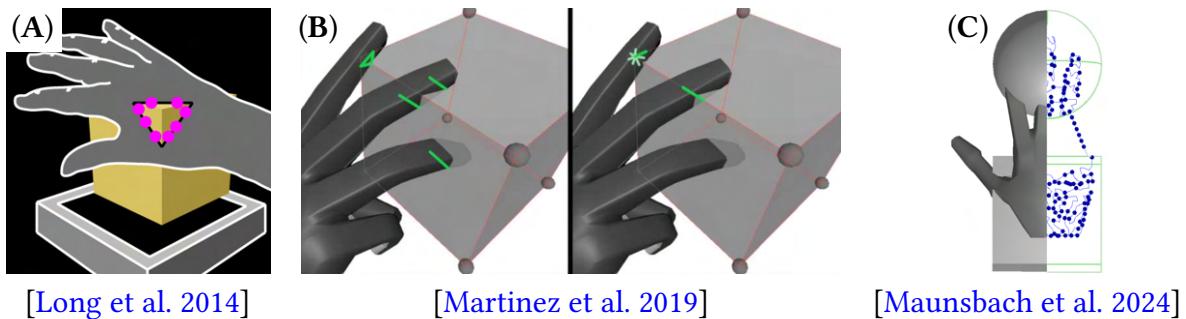


Figure 1.10 – Rendering methods for interactions with a 3D object. (A) AM points are scattered along the intersection contour. (B) Different rendering algorithms are proposed, highlighting the intersection contour or some of the shape’s salient features (e.g. vertices, edges). (C) Many collision points are detected, and then linked together to generate a long STM pattern.

[Long et al. \[2014\]](#) proposed the first 3D object dynamic rendering technique. They approximate the user’s hand as a collection of planes. By tracking the configuration of the hand, they then measure segment intersections between these planes and the object to render. Finally, they use multiple AM focal points distributed across these segments to represent the hand-object intersection with the object. By moving their hand, the user can therefore experience “slices” of the object, that can then be integrated into a full 3D shape. [Matsubayashi et al. \[2019\]](#) refined this method specifically for stimulation at the fingertips. Instead of considering the user’s hand as a small set of planes, they modelled it as a mesh, and rendered its intersection with the virtual object using AM.

A collection of algorithms were then proposed for STM rendering [[Martinez et al. 2019](#)]. The

authors proposed many variations that highlight different elements, such as edges, vertices, or hand-object intersection. Maunsbach et al. [2024] also proposed MAMMOTH, an open-source 3D object rendering algorithm by modelling the user's hand as a set of particles. By identifying which of these particles collide with the objects to render, they then solve a travelling salesman problem, allowing them to find the shortest path that links all these points. The result is then considered as the focal trajectory, representing the intersection surface. We used their insights to design a 3D object rendering method suited for bimanual interactions, as described in Chapter 9 [Mulot et al. 2025]. We also worked on a formalization of the 3D-object rendering pipeline, described in Chapter 6, and proposed a proof-of-concept study showing how the intersection strategy can alter the perceived properties of the object [Mulot et al. 2024]. Finally, in Chapter 4 we introduce a new spiral-based rendering primitives and use it to render 3D objects while measuring identification performances.

While we mentioned that rendering the whole shape at once is often useless as it provides low intensity at each point and provides energy to areas that are not in contact with the hand, it can still be possible to render the whole shape with the right tools. For instance, Inoue et al. [2015] used a massive octagonal-shaped interface composed of 3984 ultrasound transducers (while common interfaces usually use 200-300 transducers) to generate standing waves on each of the object's faces, thus successfully rendering the whole object at once.

#### 1.4.4 Stimulus Design and Validation

Designing UMH stimuli for experimentations or applications often involves a hefty amount of programming, which can be a bottleneck to one's creativity. Furthermore, reusing stimuli from one application to the other is generally not an easy task. It can therefore be very time-saving to use general design tools to quickly prototype stimuli that can be used in varied scenarios. It is also very handy to be able to simulate or visualise in real-time the generated stimuli in order to debug or validate one's approach.

**Design Tools** While very good design tools exist for contact vibrotactile stimulation [Pezent et al. 2021], they have been lacking for UMH for a long time. While the Ultraleap company proposed their Sensation Editor<sup>2</sup>, allowing the use of a few pre-defined sensations with some rendering parameterization, it was mostly designed as a demo and prototyping tool. This means it lacks in-depth control over the proposed sensation. Its closed-source nature also makes it

2. <https://support.ultraleap.com/hc/en-us/articles/360004432437-Sensation-Editor-tutorial-1-Using-the-tool>

ill-suited for compatibility with third-party hardware (e.g. other UMH interfaces) or software (e.g. experiment software).

To tackle this issue, we proposed DOLPHIN<sup>3</sup>, an open-source framework designed to provide full-control over the low-level stimulus’ rendering parameters [Mulot et al. 2021b]. It was designed to enable importation and exportation of stimuli to be used in other applications. A custom server was also provider to make integration with evaluation tools such as PsychoPy [Peirce et al. 2019] easier. Dolphin is presented in more details in [Chapter 5](#).

While DOLPHIN was a great tool for those who want full control on every detail of their sensation, especially for perceptual evaluation, it is less suited to more complex high-level sensations. Seifi et al. [2023] filled this need by introducing Feellustrator. This new design tool was specifically designed for high-level rendering, favouring creativity to high perceptual control. While the Sensation Editor was restricted to a fixed set of preset sensations, Feellustrator also allowed the designers to combine primitives into more complex and customizable sensations, thus providing more control.

While both DOLPHIN and Feellustrator allowed the export of the design stimuli to be used in other applications, the generated stimuli only used static rendering, or at best simply followed the user’s hand. AdapTics<sup>4</sup> was therefore introduced as another open-source toolkit that enabled the design of fully dynamically-rendered stimuli, linking to stimuli to metrics of the application [John et al. 2024]. For instance, they could design a beating sensation that could dynamically adapt to a virtual heart rate in a given application.

**Simulation Tools** While DOLPHIN [Mulot et al. 2021b], Feellustrator [Seifi et al. 2023], and Adaptics [John et al. 2024] all provided an illustration of the focal trajectory, and DOLPHIN also provides a 1D illustration of the generated pressure over time at a given location, none of these tools well illustrates the resulting pressure field.

Utraino [Marzo et al. 2018] and HandyBeam<sup>5</sup> were introduced as acoustic simulation tools. They simulate the behaviour of ultrasound transducers to approximate the resulting pressure field, which is very useful to evaluate a device’s performances.

Lim et al. [2024] then presented a simulation tool specifically designed to calculate the radiation forces applied at a given location on the skin. The authors validated their tool through

---

3. <https://gitlab.com/h-reality/dolphin>

4. <https://github.com/AdaptiveHaptics>

5. <https://github.com/ultraleap/HandyBeam>

comparative measurements between simulation and recording data. [Shen et al. \[2024\]](#) went a step further, combining a physical model to simulate the generated acoustic field, with a perceptual model to predict the perceived intensity of a given stimulus.

While acoustic field data is valuable, the previous tools do not model the complex biomechanics of the skin deformation in response to this field. Notable exceptions include a skin-like simulation of acoustic wave propagation, based on pseudo-spectral methods [Reardon et al. \[2019\]](#); [Reardon et al. \[2023\]](#). As discussed in [Section 1.1.1](#), their model simulates skin reformation resulting to STM stimulation. We used the same simulation setup in [Chapter 4](#) to evaluate physics-based rendering primitives.

**Visualisation Tools** While UMH provides a concrete tactile sensation, its acoustic origin means the stimulus is not directly visible. Yet being able to see the stimulus could be very helpful for debugging, investigating, or demonstration purposes. Some of the techniques discussed below are illustrated in [Figure 1.11](#).

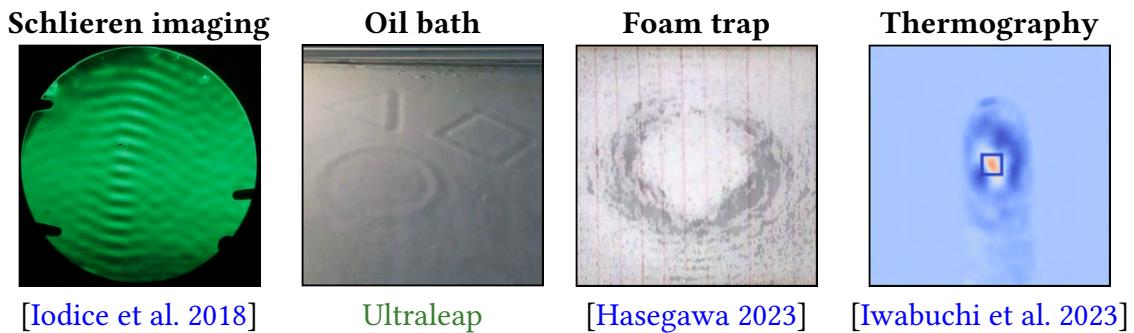


Figure 1.11 – Many types of non-contact interfaces have been developed, leveraging widely varied media.

First, it is worth noting that ultrasound transducers convert energy bidirectionally: we often use it to transform an electric signal into a mechanical vibration, but it can also generate electricity if a mechanical vibration is applied. By using an array of ultrasound transducers all attached to an LED, the transducers near the stimulated area will light up their LED, allowing for a simple visualization of the stimulated area.

[Iodice et al. \[2018\]](#) also proposed a system based on Schlieren imaging. The changes in pressure in the air lead to changes in refractive properties of the medium, which can be detected by a camera pointed at a parabolic mirror. This method shows the node and antinode regions of the pressure field across the measured region.

Later, [Abdouni et al. \[2019\]](#) proposed a simpler system, where the UMH interface generates a stimulus at the surface of an oil bath, creating small oil displacements. By shining light at the surface of the oil, the light is then reflected onto a sheet of paper. The differences in contrast relate to the oil displacements, and thus allow the user to visualize the shape of the generated stimuli.

To tackle the bulkiness and complexity of these previous systems, [Hasegawa \[2023\]](#) proposed a different approach, relying on the radiation force to push very light polystyrene beads. To do so, they trapped these beads between two acoustically-transparent layers [[Howard et al. 2023b](#)], meaning that the acoustic energy is able to go through them and push the beads. This creates an area with no beads near the stimulation area, visually representing the stimulus.

Finally, thermography techniques have also been used to visualise the UMH stimuli [[Iwabuchi et al. 2023](#); [Morgan et al. 2025](#); [Onishi et al. 2025](#)]. Indeed, the focal points generate a temperature gradient around it, that can be detected by a thermography camera.

Overall, the current section presented and discussed the bases of UMH rendering and perception, highlighting the challenges of conveying efficient stimuli, and the importance of understanding UMH perception to improve rendering algorithms. This is why [Part I](#) delves into UMH perception, with [Chapter 2](#) studying the response of single afferent units, and [Chapter 3](#) proposes and evaluates a new rendering algorithm for efficient UMH shape rendering. Then, [Chapter 4](#) proposes a new rendering primitive based on physical models to add more controlled techniques to the UMH rendering toolkit. [Part III](#) then focuses on formalizing the UMH rendering pipeline, with [Chapter 5](#) focusing on 2D rendering while [Chapter 6](#) extends it to the rendering of 3D objects. The following section then focuses on the use of UMH in interactive scenarios, to which we also contribute in [Part IV](#).

## 1.5 Designing Interactions with UMH

While [Section 1.4](#) mostly focused on static rendering methods for perception evaluation and optimization of rendering techniques, UMH can also be used in dynamic scenarios. Indeed, its mid-air nature allows for the stimulation of the tactile channel without interfering with the more commonly studied visual or audio channels. The non-wearable nature of UMH also allows for very natural interactions, and a “plug-and-play” experience for the user.

In this section, we discuss the design of coupling schemes tailored for UMH ([Section 1.5.1](#)),

before studying how UMH can be used to complement and improve human interactions ([Section 1.5.2](#)), with a specific focus in the context of virtual reality.

### 1.5.1 Coupling Schemes

Using UMH for interactions adds two challenges. First, rendering algorithms now have to be dynamic, adapting the provided sensation to the user's motion. Secondly, they require added hardware to be able to measure this hand motion.

Many different techniques have been used to track user hands, including:

- Using a head-mounted display (HMD)'s integrated hand tracking capabilities (e.g. [[Fardan et al. 2022](#)]);
- Using a wearable 6-DoF tracker attached to the hand (e.g. [[Mulot et al. 2023c](#)]);
- Using optical markers (e.g. [[Park et al. 2025](#)]);
- Using a dedicated external infrared stereo hand-tracker (e.g. [[Mulot et al. 2024](#)]) or depth camera (e.g. [[Matsabayashi et al. 2019](#)]).

Yet, combining all these components requires synchronization and coherency between them. This is the role of a **coupling scheme**: representing the joint evolution of two or more systems (e.g. how UMH systems varies when a hand gets closer to an object), as described in [Section 1.2.2](#). Well-designed coupling schemes are crucial to provide realistic or efficient stimulation, yet have been mostly overlooked for UMH rendering. Notable exceptions include the work of [Villa et al. \[2022\]](#), who designed a complex coupling algorithm to jointly control a robotic arm, to which is attached the UMH interface and a hand tracker. The coupling has to be carefully thought and implemented to ensure a smooth motion of the haptic stimulation while the robotic arm is moving.

[Matsabayashi et al. \[2019\]](#) also implemented a coupling method to control the provided haptic feedback depending on hand-object interpenetration detected using a depth sensor. They measured the finger's interpenetration with a virtual object, and use a proxy method to compute the associated UMH feedback.

To complement this lack of study on coupling schemes for UMH, [Chapter 9](#) contributes by proposing a tailored scheme for bimanual UMH interactions in VR, suitable for exploration and physical interactions with a virtual object.

### 1.5.2 UMH for Interactions

In the current subsection, we discuss different uses of UMH in interactive scenarios, including in **virtual** [Howard et al. 2022] and real-world (e.g. **automotive** [Brown et al. 2022a]) contexts. The contributions to interactive UMH in this thesis focus on VR interactions (**Part IV**).

#### Automotive

UMH stimulation has also been extensively used for automotive purposes [Brown et al. 2022a]. This stems from the current need to look at the car’s control screen or buttons. During this interaction time, the driver is therefore not looking at the road, which constitutes a safety risk. Using gesture control with haptic feedback would allow the use of another sensory channel, allowing the driver to fully focus on the road, without adding much workload. Georgiou et al. [2017] proposed one of the initial setups, with gesture control and UMH feedback, but while still keeping a screen to visualise the current state of the car when not directly interacting with it. It was then shown that UMH feedback effectively reduces the eye-off-the-road time, with even stronger effects if it is combined with audio or visual peripheral cues [Shakeri et al. 2018].

It is then worth investigating the kind of information we can convey through UMH, and how we can encode it as a haptic sensation. To do so, many haptic icons were proposed and evaluated for replacing a car’s screen or buttons [Brown et al. 2020; Brown et al. 2022b]. For each parameter to control, they evaluated the saliency of different UMH patterns so as to find the one that would generate the least added workload to the driver. Montano et al. [2024] also proposed to stimulate the knuckle region of the hand while it is on the wheel in order to provide notifications to the driver.

#### Mixed Reality

While UMH has been used in conjunction with technologies from the whole mixed reality spectrum, VR has particularly received a lot of attention. Note that in this subsection, we refer to VR in the broad sense, including both immersive and non-immersive (on-screen) scenarios. VR interactions are quite renowned for their good visual and audio feedback, but they often lack in haptic feedback. Many tactile modalities have been used in combination with VR to improve performances, immersion, or even user experience [Marchal et al. 2025]. UMH is a particularly interesting candidate for this due to its unobtrusive nature, allowing for very natural interactions [Howard et al. 2022].

**Improving User Experience** Agency and immersion are two key concepts in VR, although their exact definitions are still debated. Overall, agency can be thought of as the user's feeling of control over their actions and the environment [Kilteni et al. 2012], while immersion represents the level of user engagement in the virtual world, mediated by the engagement of each sensory channel [Berkman et al. 2019]. Higher levels of perceived agency and immersion are usually sought after, as they are associated with better user experience and performances. Since tactile channels are not usually stimulated in VR, UMH has been seen as a promising candidate to improve these two concepts [Evangelou 2025; Cornelio Martinez et al. 2017]. For instance, UMH has been shown to counter the negative effect of action-reaction latency on agency [Evangelou et al. 2021]. Yet, UMH cues have to be designed carefully, as non-congruent UMH cues can be worse than no UMH at all [Evangelou et al. 2024; Evangelou et al. 2023]. On another note, UMH feedback has also been used in an attempt to increase the level of immersion, such as improve the efficacy of spider exposure therapy in immersive VR [Brice et al. 2021].

A few studies have also leveraged the interesting and sci-fi-esque nature of UMH to complement and improve immersive experiences [Georgiou et al. 2022b]. For instance, Martinez et al. [2018] designed a supernatural demo, where the user could cast spells and experience them through abstract UMH sensations. Hwang et al. [2017] also used UMH to support the use of a virtual piano, reporting improved user experience.

Many day-to-day interactions require hand motions in a large workspace, which is not suited to the static nature of UMH interface. The physics of ultrasound focusing also provides weaker sensations if the hand's orientation does not well match that of the interface (Equation 1.1). This is why it is crucial to design UMH systems tailored for interactive scenarios. As discussed in Section 1.4.1, different robotic solutions have been proposed, involving attaching the interface to a robotic arm [Brice et al. 2019; Villa et al. 2022], or to a custom pan-tilt rotating device [Howard et al. 2020]. This allowed hand stimulations in a larger workspace, and with more hand orientation freedom.

Specifically for VR interactions, Sand et al. [2015] proposed to directly mount an UMH array onto the user's HMD, for hand stimulation. This makes the system fully wearable. A similar approach was followed by [Shen et al. 2022] for lips and teeth stimulation. Park et al. [2025] also proposed to turn an UMH interface wearable, but by attaching it to the user's wrist, stimulating the palm.

**Augmented Reality Interactions** UMH pairs very well with augmented reality since both technologies provide holographic stimulation, but for different sensory modalities. Despite this, only a few demos and studies have been conducted, indicating that augmented reality has currently received a lot less attention than virtual reality. A few notable exceptions include the heart bio-hologram by [Romanus et al. \[2019\]](#). They use cardiac measurements from a smart-watch to animate a synchronized virtual beating heart. The heartbeat can also be haptically experienced, using a STM circle of which the radius or intensity is synchronously pulsating. [Kervegant et al. \[2017\]](#) also provided a demo where the user can interact with a virtual globe. They can grasp it to rotate it, enabling visualization of the whole globe, while receiving UMH stimulation at contact points.

**Virtual Reality Interactions** In this thesis, we focus on VR interactions, to which we discuss related works below.

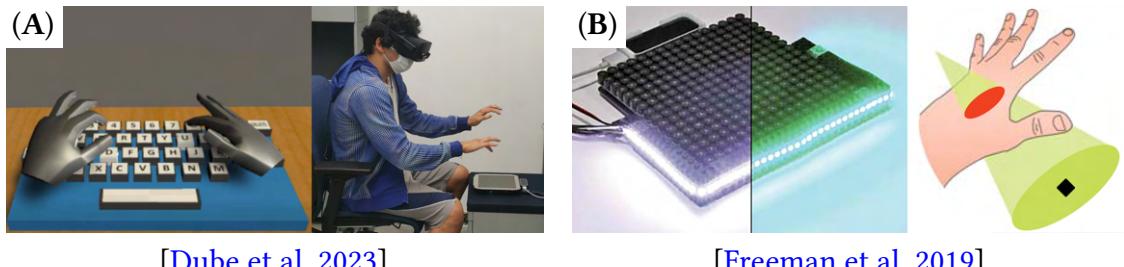
*Virtual Typing.* In VR, UMH really shines if no controller or handheld is used, to fully benefit from the mid-air nature and natural interaction. Yet, this begs the question: how can we replace contact-based input devices? To address this, a few studies dug into the design of mid-air keyboards for VR.

For instance, [Villa et al. \[2022\]](#) proposed a variety of use cases for their arm-mounted system, including the use of a virtual keyboard, user interface and virtual buttons, that could be accessed anywhere thanks to the system’s large workspace. To do so, they designed one-finger clicking sensations for button presses, or dial sensations for knob turning. UltraBoard [[Park et al. 2025](#)] also proposed to use their custom setup to provide a virtual keyboard. They had to design a custom coupling scheme between the UMH interface, VR environment, wearable device, and optical tracking system. Overall, they achieved typing speeds of around 12 words per minute, and observed improved performances and reduced frustration when UMH and visual feedback were combined. [Dube et al. \[2023\]](#) also proposed a mid-air keyboard coupling scheme, with different rendering alternatives, including a “touch and press”, “press”, and “gradual” feedback when hitting the virtual key, as shown in [Figure 1.12](#).

[Hwang et al. \[2017\]](#) designed another type of keyboard, rendering key presses on a piano. They also designed a coupling algorithm to render the key presses. To do so, they vary the maximum intensity of an AM point according to a custom law that they designed, allowing an intensity spike when contact with the key starts, followed by a slow decay, and sudden increase when the key is fully pressed.

*Guidance.* Many interactions require precise manipulation or motions of the hand, thus ben-

efiting from additional cues to help guide the user's hand to the best position or orientation. For instance, [Suzuki et al. \[2019\]](#) proposed to render a handrail, representing a path that should be followed by the user. This lifeline was rendered using a sequence of AM points. They measured average deviations of less than 50 mm, indicating good following performances. On the other hand, [Freeman et al. \[2019\]](#) provided a fixed position guidance system, combining UMH and visual cues, as shown in [Figure 1.12\(B\)](#). Specifically, they render the STM cross-section between the user's hand, and a virtual cone, coaxed with the UMH interface and pointing up. This means that the location of the cross-section on their hand would help them decide in which direction they should move to align with the interface. The interface was also covered in LEDs, of which the color would indicate the vertical proximity to the optimal position. Overall, their system would help the user reach the point of optimal stimulation in the UMH interface's workspace. Yet none of these systems has been evaluated for VR uses. Thus, in [Chapter 7](#), we propose tailored hand guidance feedback, evaluated in both static and dynamic scenarios in immersive VR. We also extend the previous work by proposing hand orientation guidance as well [[Mulot et al. 2023c](#)].



[Dube et al. 2023]

[Freeman et al. 2019]

Figure 1.12 – UMH rendering for improved interactions. (A) Tactile feedback is applied on the fingertips when pressing a virtual keyboard. (B) Visuo-haptic cues are provided to guide the user's hand toward the area that can receive the best UMH feedback.

*Object Manipulation.* Object manipulation is also a very common interaction in VR, that crucially needs haptic support so as not to break immersion when the virtual hand starts colliding with the object. It is therefore of no surprise that UMH has also been used to support such manipulations. Unfortunately, UMH suffers from a critical limitation in this case: they only provide a weak force, that is in no way sufficient to prevent the user from interpenetrating with the virtual object. This is why visual feedback is often used in addition to the UMH feedback to constrain the visual hand or increase the saliency of the feedback. For instance, [Matsubayashi et al. \[2019\]](#) proposed a system to manipulate a virtual object using one's fingers. The object was visually rendered using a stereoscopic display, and UMH feedback was provided at the areas of contact with the hand, coupled through a depth sensor. They also used

a god-object [Zilles et al. 1995] method to compute the interpenetration. In immersive VR, this god-object could be used to visually constrain the hand at the surface of the object. Frutos-Pascual et al. [2019] also proposed a single visuo-haptic combination to evaluate grasping and lifting of a single virtual object, while being stimulated on the palm with a single AM focal point with fixed intensity and modulation frequency.

In VR, contact is an important concept to render. Contact is often considered as a combination of several events and states, including the initial impact, friction or interpenetration resulting from hand motion, and detachment [Sreng et al. 2007]. Among those, Chapter 8 addresses the rendering of interpenetration by proposing and evaluating different visuo-haptic feedback combinations. We evaluate performances and user preferences in different occlusion configurations and hand orientations. In Chapter 9 we also attempt to formalize UMH coupling schemes in VR, and evaluate UMH feedback for object exploration and precise manipulation purposes.

*Bimanual Interactions.* While the previously discussed work on UMH focused on stimulating a single hand, bimanual interactions actually leverage different cognitive processes [Talvas et al. 2014]. In addition to enabling a whole new range of interactions (e.g. manipulation of larger objects), they can leverage a better integration of complex tasks, thus often improving the speed or accuracy of the interaction. Yet bimanual interactions have mostly been disregarded for UMH, as its workspace makes it impractical. Only a few demos have been developed, where the two hands are stimulated in the same plane, and either:

- they are both stimulated simultaneously, but each one only receives half of the available energy [Martinez et al. 2018];
- the interaction alternates between the two hands, meaning only one is stimulated at a time [Georgiou et al. 2018].

In Chapter 9, we address this issue by proposing a system that enables bimanual interactions where the two hands are facing each other and can receive simultaneous stimulation. We also evaluate the effect of UMH stimulation on manipulation performances in VR.

*Medical Applications.* The combination of UMH and VR has also been used for more concrete scenarios. For instance, Marchal et al. [2020] proposed to render the stiffness of an object by modifying the maximum intensity of an AM point. This was used for medical palpation scenarios. Balint et al. [2018] also studied VR palpation with UMH, while proposing a similar system aiming at training medical students.

Interestingly, it was found that UMH could be used to distract someone from pain [Karafotias et al. 2018]. Although the perceived pain is not reduced, it allows users to handle the pain for

a longer period of time.

## 1.6 Conclusion

This chapter proposed an overview of the current literature relating to UMH rendering, perception, and interaction, with context about general tactile perception and haptic rendering. We highlighted both the large potential and complexity of UMH rendering for holographic stimulation. Specifically, despite the wide interest it generated, UMH remains a very young technology that is not fully mature yet. There remain a lot of fundamental unknowns relating to UMH perception and rendering, that would dictate the quality of resulting interactions.

In this thesis, we therefore attempt to consolidate our common knowledge of UMH, by studying the specific mechanisms and perceptual response to UMH stimulation ([Part I](#)), providing new optimized rendering primitives specifically tailored to leverage the biomechanical effects of UMH on the skin ([Part II](#)), providing a general formalized pipeline for UMH rendering ([Part III](#)), as well as improving UMH interactions in VR ([Part IV](#)).

### Chapter Summary

- Tactile perception is an incredibly complex system, relying on a dense network of tactile afferents and mechanoreceptors in the skin.
- UMH stands out as the most mature mid-air technology, due to its decently large workspace, precision, and reconfigurability.
- UMH interfaces usually rely on the motion of a focal point, obtained through constructive interference of ultrasound waves.
- Further understanding of the physics ([Part II](#)) and perception ([Part I](#)) is crucial for providing optimized rendering algorithms ([Part III](#)), and tailored feedback to support dynamic interactions ([Part IV](#)).



---

## **Part I**

# **Understanding Human Perception of UMH Stimuli**

---



# **Chapter 2**

## **Characterization of Primary Tactile Afferent Response**

### **Contents**

---

2.1	Introduction . . . . .	65
2.2	Neurophysiological Recording Study: Tactile Afferent Response in Glabrous Skin . . . . .	66
2.3	In-Depth Skin Deformation Measurements . . . . .	71
2.4	Neurophysiological Recording Study: Tactile Afferent Response in Hairy Skin . . . . .	76
2.5	Conclusion . . . . .	77

---

### **2.1 Introduction**

Part I aims at providing a deeper understanding of human perception of UMH stimuli. To do so, the current chapter is positioned at the physiological level, attempting to measure the impact of UMH stimuli on individual tactile afferent responses. While afferent response characterization has been the cornerstone of tactile stimulus design for more popular technologies (e.g. contact vibrotactile, or tangible-based pressure [Vallbo et al. 1978; Johansson et al. 1982]), guiding the design of sensations, no such studies have yet been conducted for UMH. This has led

UMH designers to rely on characterizations that were made for other technologies, and that do not account for the specificities of UMH stimulation, such as its spatiotemporal aspect, large stimulation area, or acoustic streaming.

This chapter focuses on characterizing the afferent response to tactile stimuli, in order to provide solid foundations for future perception studies and rendering optimizations. To achieve this, we rely on microneurography techniques for afferent response recording, and optical coherence tomography (OCT) to measure the associated in-depth skin deformation.

In this chapter, we present the following preliminary contributions:

- The first neural recordings for UMH stimulation on glabrous skin and its associated experimental setup ([Section 2.2](#)).
- OCT measurements to visualize in-depth skin deformation in response to the acoustic radiation force ([Section 2.3](#)).
- An experimental protocol for recording the neural response to UMH stimulation on hairy skin ([Section 2.4](#)).

Contrary to the following chapters, this chapter presents preliminary contributions of an ongoing project. Yet this is an important direction in the context of this thesis, contributing to **Challenge I** and providing a better understanding of UMH perception, which is why we chose to still include this chapter in the thesis. We discuss experiments that have been designed, but whose analysis remains ongoing. These contributions have not yet been published, but correspond to a collaboration between the following researchers (in alphabetical order): Basil Duvernoy, William Frier, Ewa Jarocka, Maud Marchal, Sarah McIntyre, Lendy Mulot, and Claudio Pacchierotti.

## 2.2 Neurophysiological Recording Study: Tactile Afferent Response in Glabrous Skin

As discussed in [Section 1.4.2](#), there is evidence of different skin locations responding extremely differently to different types of UMH stimuli. For instance, the glabrous skin of the arm responds well to classical modulation, with a peak response around 200 Hz [[Hasegawa et al. 2018](#)]. On the other hand, while it remains possible (but harder) to stimulate the hairy skin of the arm [[Takahashi et al. 2020](#)], it is easier to elicit strong sensations using a slow motion of

the focal point, around 1-10 cm/s [Tsumoto et al. 2021; Pittera et al. 2022].

This first section thus focuses on the characterization of afferent response in the glabrous skin of the hand. Given the 200 Hz peak sensitivity, we assume that the tactile perception is mostly driven by fast-adapting type 2 (FA2) afferent response. Specifically, we address the following research question: how do the different afferent classes respond to various UMH stimuli on the glabrous skin of the hand, as a function of the modulation frequency?

### **2.2.1 Materials and Methods**

The proposed experimental setup aims at enabling microneurography measurements during UMH stimulation, with synchronized signals.

#### **Apparatus and Setup**

The setup for this study is illustrated in [Figure 2.1](#). We positioned participants in a Fowler's position on a treatment chair, with cushions for comfort. Their right arm was positioned on a vacuum cushion to provide support. Their index, middle, and ring fingertips were lightly glued to a custom hand support system, ensuring that they keep the same pose during the whole experiment, while requiring as little force as possible to do so. A custom UMH interface, composed of a  $16 \times 16$  array of transducers, was attached to a moveable support arm, ensuring easy manipulation of the interface to target different stimulation locations. This UMH interface was linked to the experimenter's computer, running both the experiment program (generating the successive stimuli) and experiment UI (managing the experiment).

A reference electrode was also positioned on the surface of the participant's arm, near their shoulder. A tungsten recording needle electrode was inserted percutaneously into the median nerve. Weak electrical stimuli were used to help localize the nerve. The differential signal between the two electrodes was then sent to a neurophysiological amplifier, and then observed on another computer using LabChart<sup>1</sup>. Once the recording electrode was inserted inside the nerve, small movements were manually made until a signal from a single afferent was observed.

To synchronize the microneurography recording with the UMH stimulation, the experimenter's computer was linked to an Arduino board that sent a TTL signal to the amplifier connected to LabChart during UMH stimulation.

---

1. <https://www.adinstruments.com/products/labchart>

Experiment sessions usually involved at least 3 people: one for the manipulation of the microneurography electrodes and classification of the measured afferent, one handling the recording computer (starting the recording, adding live comments, etc.), and one for running the haptics part (positioning the UMH interface and handling the experimenter computer).

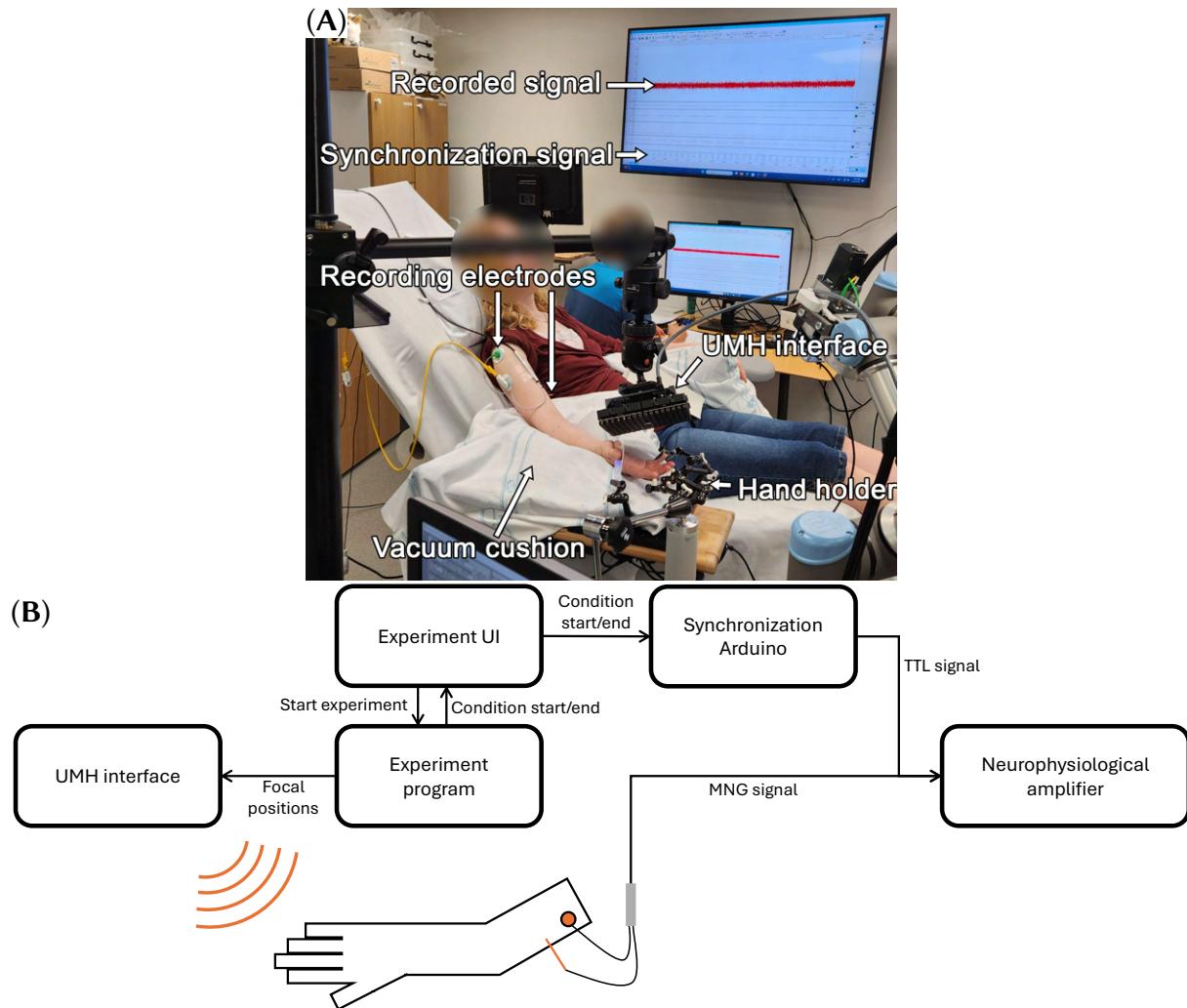


Figure 2.1 – (A) Physical setup for the microneurography recording experiments. (B) Diagram of all components of the setup.

## Conditions

We implemented a total of 100 conditions, corresponding to the combination of 10 modulation frequencies, and 10 UMH patterns. Specifically, we used the following patterns:

- A single point, with a sinusoidal waveform AM;

## *2.2. Neurophysiological Recording Study: Tactile Afferent Response in Glabrous Skin*

---

- A single point, with a square waveform AM;
- A small STM circle (diameter = 2 cm);
- A small STM line, aligned along the interface's x axis (length = 2 cm);
- A small STM line, aligned along the interface's y axis (length = 2 cm);
- A small square (side length = 2 cm);
- The same four previous shapes, with a doubled size.

These patterns were chosen as they represent the most common shape primitives for UMH, and because we wanted to experience with different sizes.

The 10 modulation frequencies were as follows: 5, 10, 30, 50, 100, 200, 250, 300, 400, and 500 Hz to cover a wide range of frequencies, while having more frequencies around the peak vibrotactile sensitivity of FA2 afferents. All stimuli were rendered 8 cm away from the UMH interface, which corresponds to this device's optimal distance for peak intensity.

### **Procedure**

Once a unit was found, its mechanical threshold was measured using Von Frey filaments. The unit type was then determined based on its response to Von Frey filaments, and the target stimulation point was chosen as the point on the receptive field that provided the most spikes during the recording. The UMH interface was then positioned with its center 8 cm away from the target point, and its normal aligned with the local skin normal to the best of our capabilities.

Once the interface was well positioned, all 100 conditions were successively performed, in order, with a stimulation duration of 200 ms each and a pause of 50 ms between conditions. After the 100 conditions, we lightly pressed on the receptive field while looking at our recordings to ensure that we did not lose the unit due to involuntary movement of the recording electrode. If we did not lose it, we then performed another full set of recordings, and repeated this procedure again as needed. We usually managed to perform 3 sets of recordings per unit.

During the recording, participants were asked to remain as still as possible to avoid moving the recording electrode. Each session lasted between 5 and 6 hours, including explanations, preparation of the participant, measurements, and breaks. Participants were compensated 200 SEK/h.

## Participants and Collected Data

We recruited 6 participants (5 women, 1 man, aged  $m \pm s.d. = 37 \pm 11.4$ ) for this experiment, and 2 others for a preliminary pilot study aimed at testing and refining our setup.

We collected information about each unit found (localization, type, mechanical threshold), and its electrical response during stimulation.

### 2.2.2 Preliminary Results and Discussion

Overall, we obtained recordings from 8 SA1, 6 SA2, 16 FA1, and 1 FA2 units. Examples of recorded signals are shown in [Figure 2.2](#). We observed consistent activation of the FA2 unit, but no activation from all other units, apart from a few individual spikes that could not confidently be attributed to the UMH stimulation. This result provides the first direct physiological evidence that FA2 are the primary drivers of UMH perception in the glabrous skin of the hand, corroborating with psychophysical findings that report peak perceptual sensitivity around 200 Hz [[Hasegawa et al. 2018](#)].

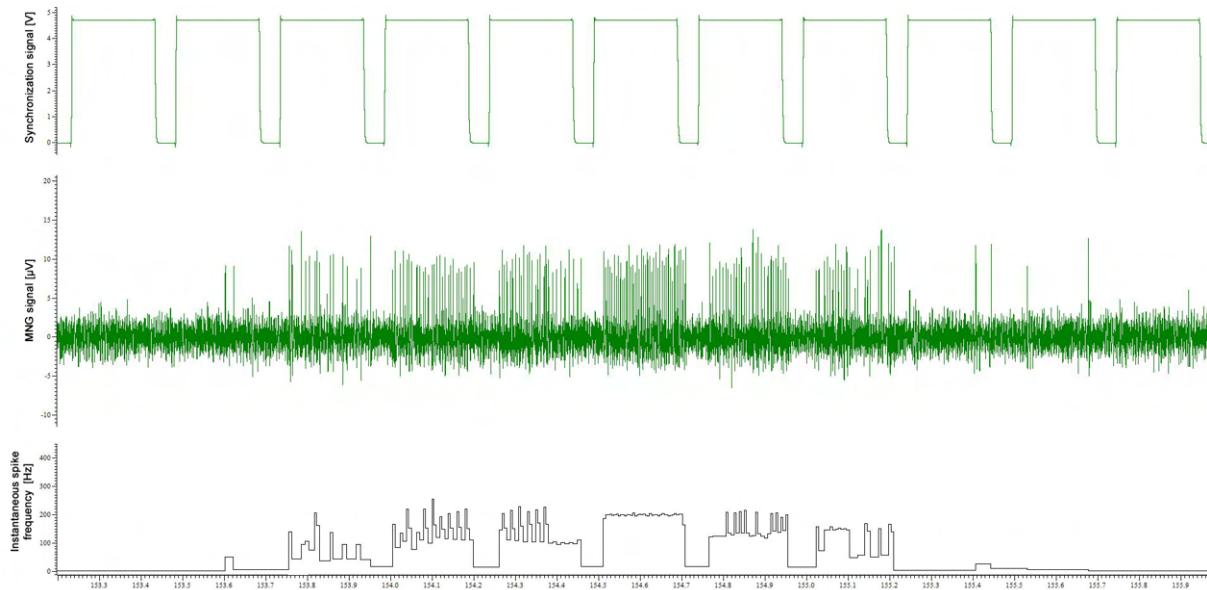


Figure 2.2 – Example recordings from the FA2 unit. The top plot shows the synchronization signal, representing when the UMH interface was emitting (high value), while the second plot represents the measured microneurography signal. The bottom plot represents the instantaneous spike frequency, defined as the inverse of the inter-spike interval.

That being said, [Morisaki et al. \[2024\]](#) showed that it was possible to generate static contact sensations rather than vibratory stimuli, by using a very low LM modulation frequency. Such

frequency should not (or barely) stimulate FA2 afferents [Johansson et al. 2009], which is not on par with our preliminary results. This can be explained by the higher output power of the device used in the aforementioned study. Indeed, their setup uses a total of 996 ultrasound transducers, while our interface only has 256 transducers. The significantly higher amount of transducers leads to a stronger output, which may be enough to then exceed the mechanical threshold of other types of afferents.

As previously mentioned, this project is currently still ongoing, implying that further analyses will be conducted. For instance, we will precisely locate and count all spikes from the FA2 measurements to provide a more quantitative characterization of UMH-based mechanotransduction. We also measured the force generated by our interface by positioning a small cylinder on top of a precision scale, and rendering an unmodulated focal point on the cylinder's surface. This resulted in an equivalent weight of 0.63 g ( $F = 6.2 \text{ mN}$ ) when the UMH interface was positioned 8 cm away from the cylinder. Such measurements should be compared with the obtained mechanical thresholds for all units, which should correlate with the activation or non-activation of each afferent unit.

## 2.3 In-Depth Skin Deformation Measurements

In this section, we aim to correlate the microneurography results with a visualization of skin deformation in response to UMH stimulation, using OCT techniques.

### 2.3.1 Background on OCT Recording

When light reaches an interface between two media, some of it is back-scattered towards its source of origin, while the rest can be refracted, absorbed, or forward-scattered. When reaching back to the source, the returning light has a different phase from the emitted light due to the distance travelled, and a different amplitude due to the absorption and scattering. This difference can create an interference pattern between the two types of light. Such interference can be analyzed using an interferometer, providing information about the interface between the two media, such as its distance to the light source.

Spectral-domain OCT (SD-OCT) techniques treat tissues as a continuous sequence of interfaces. By shining a beam of light onto the medium and analyzing the spectrum of the interference pattern caused by the back-scattered light, it is possible to extract information about

the medium at different depth levels [Huang et al. 1991]. More precisely, a light source emits a beam of light in a certain bandwidth. The beam is then split into two parts: a reference beam, and a sample beam. The reference beam then hits a mirror at a controllable distance before heading back towards a photodetector. The sample beam hits another controllable mirror, which redirects it towards the skin (or any other sample tissue to analyse). The back-scattered light follows the inverse path, reaching the same photodetector as the reference beam, allowing for interference to form. Finally, just before reaching the sensor, a diffractor separates the resulting combined beams by their wavelength, allowing each pixel of the sensor to receive a single wavelength. This process is illustrated in [Figure 2.3\(A\)](#).

By default, OCT provide data about a 1D in-depth line in the medium, called A-scan. By adjusting the two mirrors, it is then possible to change the position of the recorded A-scan. Then, combining multiple A-scans together can then generate a 2D cross-sectional image of the tissue, called B-scan (see [Figure 2.3\(B\)](#)). To measure skin movement instead, we can acquire A-scans at a constant position but at different timesteps. Such techniques have been used to visualize skin deformation in response to mechanical deformation (e.g. for contact vibrotactile stimulation [[Duvernay et al. 2023](#)]). These types of visualization can reach a few millimeters in depth.

### 2.3.2 Apparatus and Setup

The overall setup for the OCT measurements is illustrated in [Figure 2.4](#). Measurements were conducted on a single participant, who was laying down on a treatment chair. Their right hand was maintained still using a vacuum cushion and a custom hand holder. The same custom UMH with  $16 \times 16$  transducers was attached to a moveable support arm, and positioned 8 cm away from the target positions. It was used to render the same conditions as for the Microneurography experiment ([Section 2.2.1](#)) On the other hand, a SD-OCT system (Thorlabs' TEL320C1 with an OCT-LK3 lens <sup>2</sup>) was attached to a robotic arm to allow for precise positioning. The OCT's light source emitted light with a roughly gaussian enveloppe around 1300 nm wavelength, with a usable bandwidth of 200 nm. The OCT's sensor contained 2048 pixels, allowing for A-scan images of  $1024 \times 1$  pixel images after processing. Given the required proximity of the two devices, the OCT usually blocked the generated ultrasound waves from a few transducers, which likely slightly reduced the generated radiation pressure on the skin.

The OCT and UMH interface were controlled on two different computers (referred as PC1

---

2. <https://www.thorlabs.us/thorproduct.cfm?partnumber=TEL320C1>

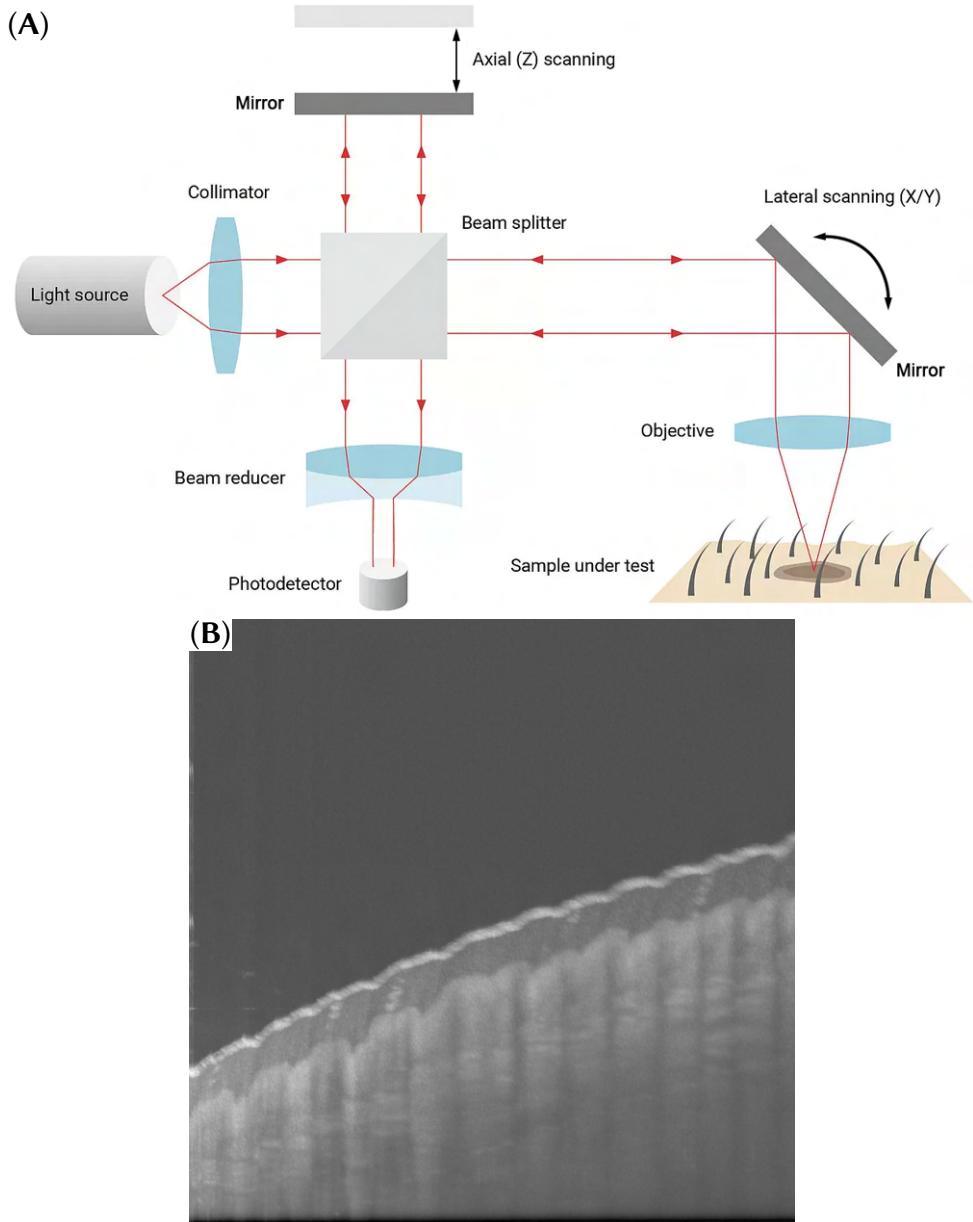


Figure 2.3 – (A) Functional diagram of an SD-OCT: the emitted light is split into a reference beam that follows a controlled path, and a sample beam that reaches the skin. The interference pattern between the reference beam and the back-scattered light from the sample beam is then analyzed to provide in-depth information about the skin. Two mirrors can be adjusted to modify the position of the scanned area. Image adapted from [Physik Instrumente](#). (B) 2D Morphological image obtained from a recording on the index fingertip.

and PC2 respectively). When PC1 sends a trigger to start the OCT recording, the Arduino attached to PC2 detects the trigger and informs PC2, which in turn starts the UMH next UMH

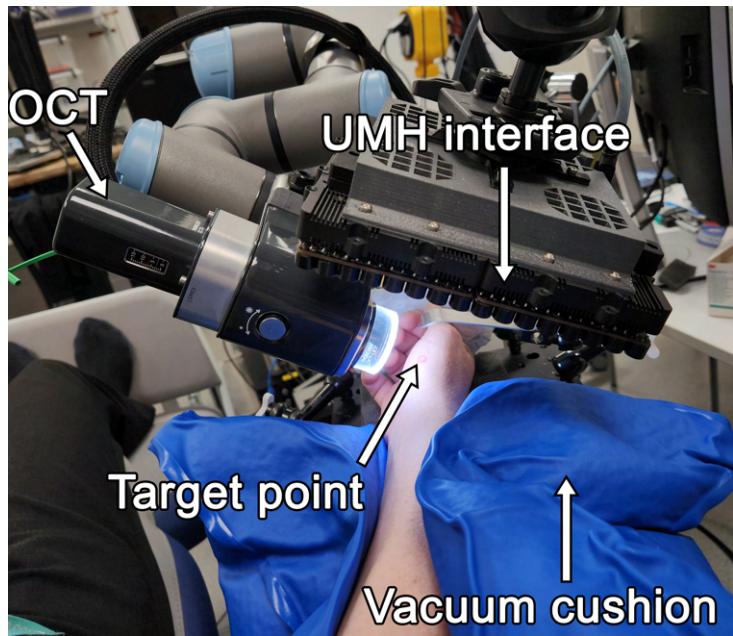


Figure 2.4 – Setup for the OCT recording study. Both the OCT and UMH interface are pointed at the same target point on the participant’s hand, which is maintained still with the vacuum cushion and a hand holder.

stimulation.

Measurements were conducted at two different target locations: one on the index fingertip, and one on the thenar eminence. Acquisitions were made at 10 kHz for 500 ms (including 200 ms stimulation), which lead to resulting images of  $1024 \times 5000$  pixels.

### 2.3.3 Preliminary Results and Discussion

Using the recordings for each condition, we started to generate images of the time-variation for each A-scan (see [Figure 2.5](#)). These images may allow us to visually illustrate the deformation of the skin’s surface. Further analyses will also be conducted to provide quantitative data. For instance, performing a Fourier transform should reveal a deformation at the modulation frequency. The amplitude associated to this peak frequency will also provide insights about the skin deformation amplitude, which can then be compared to afferent thresholds, correlating with the activation or non-activation of different tactile afferents.

As a teaser, we also started looking at the recorded phase signal, rather than the amplitude shown in [Figure 2.5](#). By measuring the instantaneous phase variation, we can highlight areas of rapid change, corresponding to the moment when the focal point reaches the location of

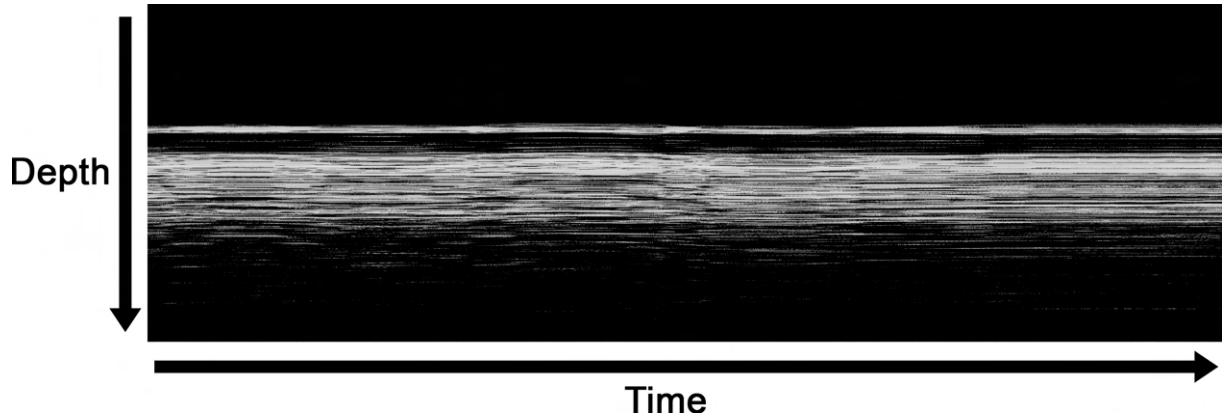


Figure 2.5 – Example OCT result, corresponding to the time variation of an A-scan located on the fingertip. Brighter points indicate that more light is back-scattered.

the A-scan, as shown in [Figure 2.6](#). By counting the number of such rapid changes, we can find the modulation frequency  $f = \frac{n}{\delta t}$ , where  $n$  is the number of rapid changes, and  $\delta t$  [s] = 0.2 s is the stimulation duration. Yet, these analyses are ongoing and will be refined in the future for more quantitative discussions.

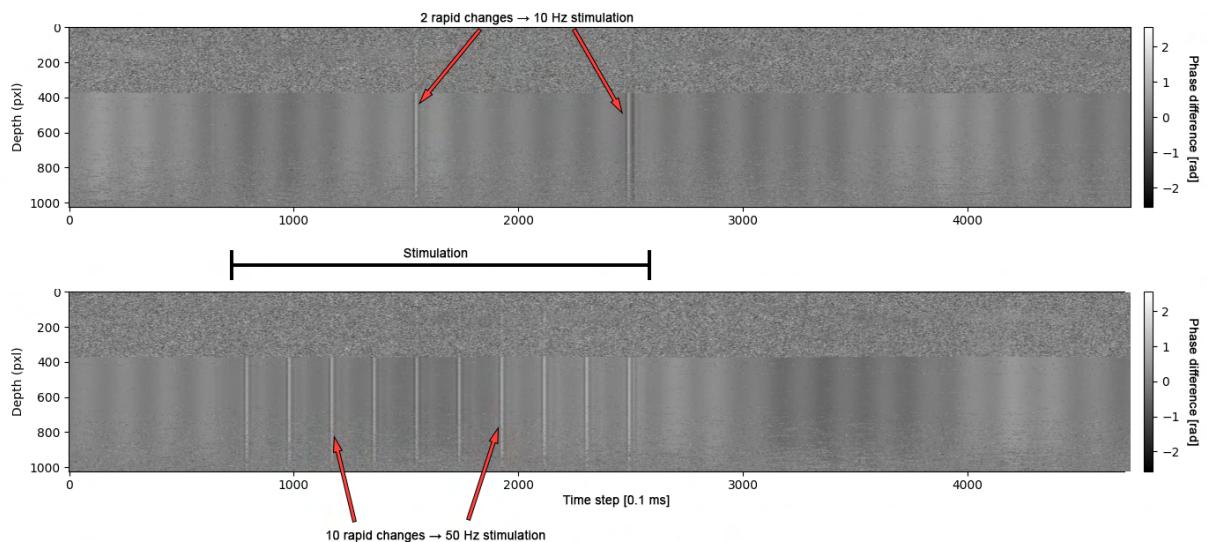


Figure 2.6 – Tentative analysis, looking at the phase changes in the recorded signal. We observe rapid changes each time the focal point passes near the location of the A-scan, which correlates with the stimulus' modulation frequency.

## 2.4 Neurophysiological Recording Study: Tactile Afferent Response in Hairy Skin

As mentioned in [Section 2.2](#), UMH perception seems to be widely different between the glabrous skin of the hand, and hairy skin. Interfaces similar to ours struggle to stimulate hairy skin with standard STM methods, while it has been observed that focal speeds of 1-10 cm/s lead to decently strong stimuli [[Tsumoto et al. 2021](#); [Pittera et al. 2022](#)]. While [Section 2.2](#) highlighted the dominant role of FA2 afferents in the glabrous skin of the hand, they are much rarer in hairy skin [[Corniani et al. 2020](#)], meaning that other receptors are likely at play. This section thus aims at studying afferent response in hairy skin. Specifically, we address the following research question: how do the different afferent classes respond to various UMH stimuli on the hairy skin of the hand and arm, as a function of the modulation frequency?

Note that this study has not yet been conducted. Here, we present the designed protocols and expected results, but more refined insights will be obtained once we collect more data.

### 2.4.1 Materials and Methods

#### Conditions

This study involves two alternative procedures, one of which corresponds to the conditions presented in [Section 2.4.1](#), while the other involves conditions discussed below. For the second procedure, we render two shapes: a 2.5 cm line, and a 2.5 cm diameter circle. Sizes were chosen empirically. Each shape is rendered using STM, with draw frequencies of 0.2, 0.8, 2, 4, and 8 Hz, which corresponds to focal speeds of 0.5, 2, 5, 10, and 20 cm/s for the line. This includes the range of focal speeds that have been shown to provide decent stimulation on hairy skin [[Tsumoto et al. 2021](#); [Pittera et al. 2022](#)].

Empirical testing showed that these sensations felt stronger when relying on the acoustic streaming rather than the focal point's acoustic radiation force (see [Section 1.4.2](#)). This effect was observed to be perceived stronger when the stimulus was rendered around 30 cm from the UMH interface, while the participant's skin is even further away, between 35-40 cm. By having the skin co-localized with the stimulus, it would mostly experience the radiation force resulting from the localized pressure at the focal point, while adding this offset lets the participant experience the resulting turbulent air flow corresponding to the acoustic streaming.

## Apparatus, Setup, and Procedure

We plan on using the same setup as in the first study ([Section 2.2.1](#)), except that we will now record in the radial nerve, which mostly innervates hairy skin.

Given that we cannot predict how long we will be able to perform measurements on a given unit, we came up with two different procedures, with priorities dependent on the unit type. If the unit is a SA1, SA2, FA1, or FA2, then we run the same procedure as in [Section 2.2](#) first, with the same conditions. Otherwise, we first conduct a similar procedure with the conditions presented in [Section 2.4.1](#), with 5000 ms stimulation and 500 ms pause between conditions. The higher stimulation duration was chosen to account for the slower focal motion, allowing at least one full scanning of each shape with the lowest draw frequency. If the unit is not lost after running the first procedure, then we can switch to the other one, but it requires adjusting the distance between the UMH interface and the skin.

### 2.4.2 Expected Results

Since this experiment has not been conducted yet, we do not yet have results to discuss, but we can make informed hypotheses. First of all, it has been observed that humans perceive strong sensations to focal motions around 1-10 cm [[Tsumoto et al. 2021](#); [Pittera et al. 2022](#)]. This speed range is similar to that usually used for social touch experiments [[Hauser et al. 2019](#)], and corresponds to the peak sensitivity of CT afferents in hairy skin [[Löken et al. 2009](#)]. We therefore expect that the acoustic streaming deflects hair, which correlates to CT activation [[Moore et al. 2025](#)]. We also expect to activate FA2 afferents, although they are much rare, which would make them more challenging to find, and less dominant in the overall perception of UMH stimuli on hairy skin.

## 2.5 Conclusion

In this chapter, we started studying the response of single tactile afferents to UMH stimulation, which is the first physiological step of perception. Specifically, our microneurography recordings from 31 afferent units in glabrous skin revealed that only the FA2 afferent responded consistently to UMH stimuli. This finding, despite the low sampling of FA2 units, points to their dominant role in the perception of UMH in this skin region. Yet, it remains possible that stronger UMH interfaces could stimulate other afferent classes. We also performed OCT measurements to correlate our results with in-depth visualization of skin deformation in response to UMH stimulation. Finally, we designed a follow-up microneurography study to extend our

results to the afferent response in hairy skin. While this study has not been conducted yet, we expect CT afferents to play an important role, responding to the acoustic streaming generated beyond the focal point.

It is worth noting that this project is ongoing, meaning that all the results presented are preliminary, although they are very promising. Further studies and analyses will be conducted soon to allow for deeper and more accurate discussions. Such results will help provide a strong base for our understanding of UMH perception, which should help UMH designers optimize their sensations. Given the preliminary result, we expect to confirm the strong domination of FA2 afferents for UMH perception on the glabrous skin of the hand, and to find a domination of CT afferents on the hairy skin of the hand and arm, in response to acoustic streaming sensations, while FA2 may be more responsive to the acoustic radiation pressure sensations. It would also be interesting to further extend our work to measurements on the face. This is particularly interesting as UMH has been shown to be able to stimulate different locations on the face [[Gil et al. 2018](#); [Shen et al. 2022](#); [Lan et al. 2024](#)] while its skin is not innervated by FA2 afferents [[Corniani et al. 2020](#)].

### Chapter Summary

- We provided the first microneurography study for UMH stimulation on glabrous skin, revealing a dominant role of FA2 afferents in UMH perception;
- We performed some OCT recordings to visualize the in-depth deformation behaviour of the skin in response to UMH stimulation;
- We designed another microneurography study for UMH stimulation on hairy skin, which will soon be conducted;
- This project is ongoing, and more analyses will be conducted, while more data will also be gathered.

# Chapter 3

## Perception of Dynamically-Rendered 2D shapes

### Contents

---

3.1	Introduction . . . . .	79
3.2	Spatio-Temporally-Modulated Tactile Pointers . . . . .	80
3.3	User Study: Evaluation of Stimulus Intensity . . . . .	83
3.4	Naive Shape Perception . . . . .	88
3.5	User Study: Clarity of Rendered Shapes . . . . .	92
3.6	Discussion and Perspectives . . . . .	97
3.7	Conclusion . . . . .	98

---

### 3.1 Introduction

As discussed in [Section 1.4.3](#), UMH has been extensively used to render tactile shapes. While STM provides a strong static stimulus [[Kappus et al. 2018](#); [Frier et al. 2018](#)], it also results in rather blurry sensations [[Obrist et al. 2013](#); [Reardon et al. 2023](#)], and poor identification performances [[Howard et al. 2019](#); [Rutten et al. 2019](#); [Hajas et al. 2020](#)]. On the other hand, dynamic tactile pointers (DTP) rely on the slow motion of an AM focal point along the shape's contour, with optional pauses on vertices to highlight them and improve shape identification

performances [Hajas et al. 2020]. DTP thus provide sharper sensations with higher shape identification performances. Yet, the use of AM leads to rather weak sensations compared to STM. Combining both techniques may take advantage of their respective benefits, resulting in a new rendering method that renders both intense and sharp shape stimuli.

In this chapter, we contribute to **Challenge I** by studying the perception of tactile pointers, strengthening our understanding of this promising technique. The new method we propose also leverages perceptual insight to optimize UMH shape perception.

In this chapter, we propose the following contributions:

- Spatio-temporally modulated tactile pointers (STP), a novel rendering method for overcoming the limitations of dynamic tactile pointers and STM ([Section 3.2](#)).
- Two user studies assessing the benefits of STP in terms of perceived intensity ([Section 3.3](#)) and shape identification performance ([Section 3.4](#) and [Section 3.5](#)).

Our work aims to provide new perspectives to UMH stimulus designers by enabling more effective tactile shape rendering with this technology. Furthermore, we discuss our results in the wider context of UMH shape rendering approaches and how they may contribute to further understanding of human mid-air tactile shape perception (see [Section 3.6](#)).

The content of this chapter was adapted from the following publication: [[Mulot et al. 2023b](#)], which is an extension of the following work-in-progress publication: [[Mulot et al. 2023a](#)].

## 3.2 Spatio-Temporally-Modulated Tactile Pointers

We developed a new rendering technique, called STP (Spatio-temporally-modulated Tactile Pointers), with the goal of achieving the high accuracy for shape identification obtained with DTP while conserving the stimulus intensity afforded by STM rendering. Similarly to how DTP [[Hajas et al. 2020](#)] spatially sample down a shape to a representative set of points, and then sequentially render them using an AM focal point, STP splits the shape contour into several small sections which are sequentially rendered as STM line segments or arcs.

### 3.2.1 Rendering Algorithm

Here, we consider a contour to be parameterized by one variable:  $s : [0, 1] \rightarrow \mathbb{R}^3$ . The rendering process of DTP and STP are illustrated in [Figure 3.1](#).

For a DTP rendering [[Hajas et al. 2020](#)], the shape would be represented by a set of  $n$  points

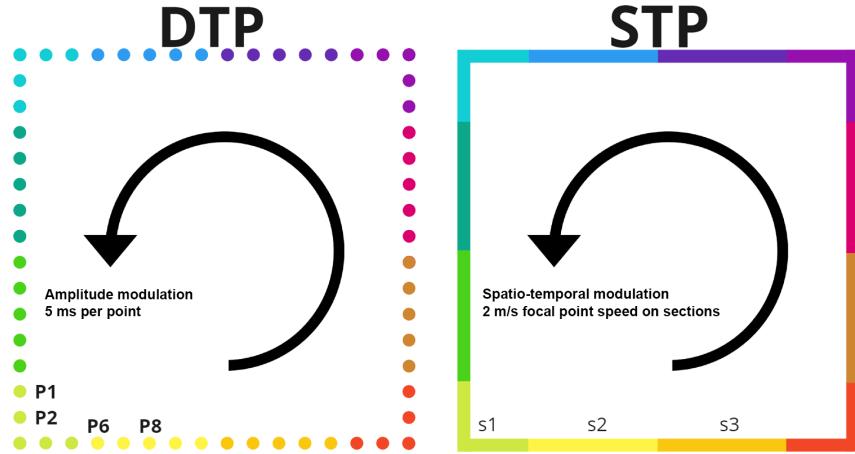


Figure 3.1 – (Left) DTP consists in a succession of amplitude-modulated points along the shape contour (here, a square), while (Right) STP consists in a succession of spatio-temporally-modulated sections along the contour.

$P_1, \dots, P_n$  evenly spaced out, with  $P_i = s(i/n)$ . Then, all of these points would be successively rendered with amplitude modulation (AM). Let  $f$  be the amplitude modulation frequency. The focal point dwells at each position for a duration of  $1/f$  (or a multiple of this), so that a full period of modulation has passed. The points are then rendered sequentially, giving the impression of a pointer drawing the contour.

The design of STP is inspired by that of DTP. We chose to cut the shape into  $n$  sections  $s_1, \dots, s_n$  of identical length. Each section can then be rendered with STM, using a fixed focal speed. This time, there is no amplitude modulation, meaning that the full output power of the haptic interface is always used, achieving higher intensity stimuli. The focal point speed and section size have to be chosen carefully so as to optimize the perceived intensity, as discussed in [Section 3.2.2](#). Finally, the different sections are rendered sequentially, to provide the same impression of a moving pointer as in DTP. Both techniques are illustrated in [Figure 3.1](#).

We propose two variations of our method, depending on whether the shapes to render have vertices or not. The first variation, **STP-SS** (single-stroke), renders each section for an identical duration, giving the feeling of a point moving across the shape contour at a constant speed, in a single-stroke fashion. The second variation, **STP-MS** (multi-stroke), spends more time on sections located around the shape's vertices to highlight them, with the aim of achieving similar improvements to shape identification performances as observed with **DTP-MS** [[Hajas et al. 2020](#)].

Overall, with STP, we aim to benefit from the high perceived intensity of STM, combined with

the high shape perception accuracy from DTP.

### 3.2.2 Stimuli Parameterization for Experimental Validation

The stimuli evaluated in both subsequent user studies ([Section 3.3](#) and [Section 3.5](#)) use the exact same shapes and rendering parameters, described in the following.

**Shapes** We evaluate STP on a set of six different shapes, all inscribed inside a 32 mm-radius circle, so that they can be fully rendered on the majority of hands [[Peebles et al. 1998](#); [Koman-dur et al. 2009](#)]. Their dimensions are summarized in [Table 3.1](#).

**General rendering parameters** To ensure that our results are comparable to those of DTP [[Hajas et al. 2020](#)], we rendered all shapes with a draw frequency of 0.5 Hz (2 seconds for the tactile pointer to perform a full scan of the shape). Our **DTP-MS** and **STP-MS** stimuli were also designed to spend 80% of the time (1.6 s in total) highlighting the vertices. It is worth noting that, by increasing the amount of vertices, we reduce the time spent highlighting individual vertices. Keeping the vertex highlight duration constant would lead to shape-dependent draw frequencies. This constitutes an interesting perspective for future work on this topic. All shapes were rendered for 6 s, corresponding to three full cycles, a value which was empirically chosen to ensure sufficient time to comprehend each stimulus.

**DTP parameters** Both DTP variants use a 200 Hz amplitude-modulated focal point [[Chongyang et al. 2019](#)], and spend 5 ms (one full modulation period) rendering each point, except for the vertices with **DTP-MS**, for which a total of 1.6 s were spent rendering vertices. Per vertex, this equates to 533ms for triangles, 400ms for squares, 267ms for hexagons and 200ms for octagons.

**STP parameters** With **STP-SS** and **STP-MS**, we split shapes into approximately 14 mm-long sections, based on [Takahashi et al. \[2018\]](#)’s optimum results for lateral modulation. There are some variations on this length based on the specific shape being rendered, as we require all sections to be of exact same length, and we want all sides of a regular polygon to feature the same amount of sections. The details for each shape can be found in [Table 3.1](#). Overall, we rendered each section using a 2 m/s focal speed to reach perceivable draw frequencies. It is worth noting that, due to the small size of these sections, the focal speed is lower than usual STM speeds, even subceeding the propagation speed of shear waves in the skin. This means

that no shock waves are created, mitigating the spread of energy that is assumed to cause the blurry sensations of STM [Reardon et al. 2023].

The shape identification experiment (see [Section 3.5](#)) also uses conventional STM rendering, with a focal point speed of 7 m/s [[Frier et al. 2018](#); [Frier et al. 2019](#)]. All STP and STM stimuli were rendered at maximum intensity, which for the device used in our experiments corresponds to a peak sound pressure at the focal point of approximatively 2.5kPa.

Table 3.1 – Summary of the parameters of the stimuli. The circle and ellipse are only rendered with the single-stroke versions, while polygons are drawn with both versions. The first two lines indicate general shape dimensions.  $r$  is the circle’s radius,  $r_a, r_b$  are the ellipse radii along the radial and distal axis respectively, and  $s$  is the side length. The last two lines provide indication about the sampling for STP.

	Circle	Ellipse	Triangle	Square	Hexagon	Octagon
Dimensions (mm)	$r = 32$	$r_a = 32$ $r_b = 15$	$s = 55.4$	$s = 45.3$	$s = 32$	$s = 24.5$
Perimeter (mm)	201	152.5	166.3	181	192	195
Number of sections	14	11	12 (4 / side)	12 (3 / side)	12 (2 / side)	16 (2 / side)
Section length (mm)	14.4	13.9	13.9	15.1	16	12.2

### 3.3 User Study: Evaluation of Stimulus Intensity

Our first experiment aims to validate our objective of designing a shape rendering method which produces stronger stimuli than DTP. To achieve this, we compare perceived intensity for a representative set of shapes (see [Section 3.2.2](#)) rendered either with STP or with DTP in both their variations.

#### 3.3.1 Hypotheses

We hypothesize that:

- **H1-MS:** Shapes rendered with **STP-MS** will be systematically rated as more intense than the same shapes rendered with **DTP-MS**;
- **H1-SS:** Shapes rendered with **STP-SS** will be systematically rated as more intense than the same shapes rendered with **DTP-SS**.

Based on prior results by [Howard et al. \[2019\]](#) when investigating detection thresholds for AM and STM stimuli, we suppose that perceived intensity is independent of exploratory motion.

We therefore perform this experiment only in a passive configuration, where the participant's hand remains immobile throughout each trial.

### 3.3.2 Materials and Methods

#### Apparatus and Setup

Participants were seated in front of a table holding an open-top box with an armrest and an Ultraleap STRATOS Explore UMH device inside. Their non-dominant hand was placed above the hole in the box, as illustrated in [Figure 3.2](#). An STM circle was rendered prior to the experiment to help participants position their hand. Participants were asked to place their hand so that the entire circle fit inside the palm of their hand.

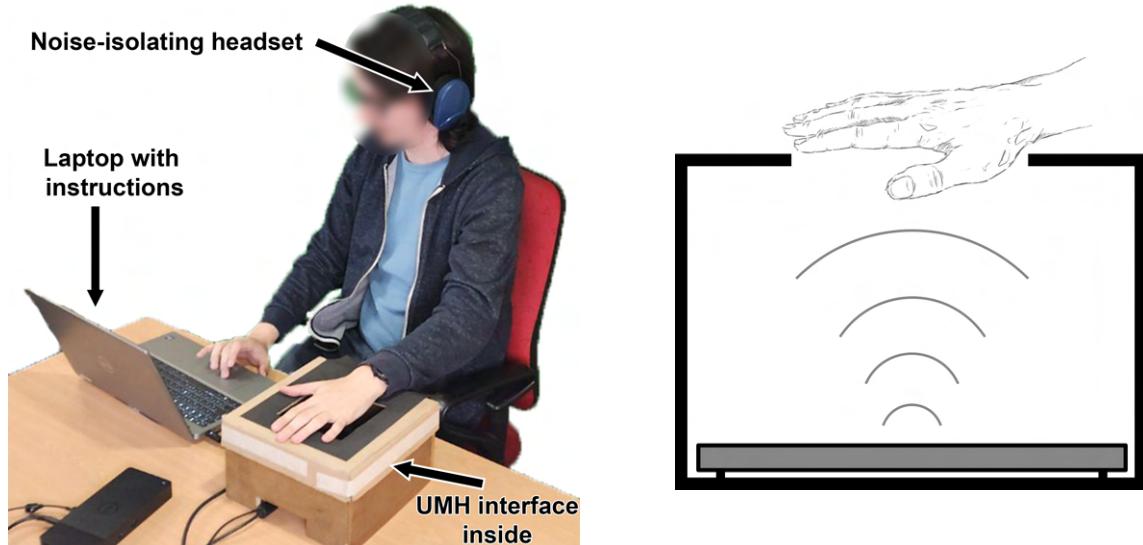


Figure 3.2 – Experimental setup. (Left) Participants are seated in front of the laptop running the experiment and displaying instructions, while their non-dominant hand is resting on a box containing the haptic interface (Right).

We created the experiment logic and user interface using PsychoPy [[Peirce et al. 2019](#)], which triggered the playback of stimuli that were previously designed and hardcoded in C++ using Ultraleap's SDK3.

Participants wore a noise-isolating headset to mask any audio cues from the ultrasound array. They responded by using the laptop's trackpad to click the appropriate buttons on the experiment UI. We also provided each subject with an A4 sheet of paper with numbered 5x5 cm squares and a pen, so they could sketch the shapes they felt after each trial using their dominant hand (See [Section 3.4](#) for further discussion of this point).

## Procedure

Participants were briefed about the experimental procedure prior to the experiment and provided written informed consent to participate. They also filled out a pre-experiment questionnaire assessing basic demographic information, handedness and prior experience with haptics and ultrasound haptics in particular.

The experiment followed a within-subject blocked design, with one block comparing both single-strokes variants **DTP-SS** and **STP-SS**, and the other block comparing the two multi-stroke variants **DTP-MS** and **STP-MS**. The order of blocks was counterbalanced across participants.

Each block tested 4 permutations of rendering technique pairs for each of the 6 and 4 shapes respectively rendered in the (-SS) and (-MS) variations (see [Section 3.2.2](#)): STP vs. DTP, in that order (A); DTP vs. STP, in that order (B); STP vs. STP (C); DTP vs. DTP (D).

Comparisons A and B control for any order bias in the intensity comparisons between rendering methods. Comparisons C and D allow us to assess the potential presence of biases in subject responses. Subjects thus performed a total of 40 trials across both blocks: 24 trials in the single-stroke (-SS) block and 16 trials in the multi-stroke (-MS) block.

Each of these trials followed a 2-IFC protocol [[Jones et al. 2013](#)], where a given shape was rendered successively with two (potentially identical) rendering techniques, and the user was asked to evaluate the intensity of the second stimulus compared to that of the first, by clicking on a button labelled “Stronger” or “Weaker”. The order of presentation of the buttons was randomly chosen for each trial.

In addition, after responding in each trial, we asked subjects to sketch the shape they felt on paper. The aim was to explore to what extent the subject’s mental image of the rendered shape matched the intended shape. A qualitative discussion of these data is provided in [Section 3.4](#).

For each trial, we recorded the participant’s answer (whether the second stimulus was perceived stronger or weaker than the first), along with the shape they sketched on paper.

## Participants

We recruited 18 participants (13 men, 5 women, aged  $m \pm s.d. = 25.8 \pm 9.8$ ). 14 of them were right-handed, 3 of them were left-handed and one was ambidextrous, and completed the

experiment as a left-handed person. While one participant reported extensive experience with haptics, 4 participants had limited experience, and 13 had almost no experience with the topic. One third of participants reported having previously used an UMH interface.

### 3.3.3 Results

In total, we collected data in 432 trials for the single-stroke (-SS) stimulus pairs, and in 288 trials for the multi-stroke (-MS) stimulus pairs.

For both the (-SS) and (-MS) blocks, we calculated the proportions  $P_C$  and  $P_D$  of “second stimulus stronger than the first” answers in the trials where STP and DTP were tested against themselves. These are used to assess the presence or absence of systemic bias in the participant’s answer strategy. On the other hand, we calculated the proportion  $P_{AB}$  of answers where the STP stimulus was rated stronger than the equivalent DTP stimulus in both permutations A and B above combined. These represent objective measures to support **H1-MS** and **H1-SS**.

Shapiro-Wilk tests rejected the normality hypothesis for  $P_D$  ( $p = 0.005$  for the (-MS) block and  $p = 0.002$  for the (-SS) block). Similar tests showed that  $P_C$  did not follow a normal distribution in the (-MS) block ( $p = 0.002$ ) but did in the (-SS) block ( $p > 0.05$ ). Shapiro-Wilk tests failed to reject the normality hypothesis for  $P_{AB}$  for the (-SS) block ( $p > 0.05$ ) but did for the (-MS) blocks ( $p = 0.02$ ).

We compared  $P_{AB}$  and  $P_C$  for the (-SS) block to the 50% chance-level using one-sample t-tests, and all four other distributions to the 50% chance-level using a Wilcoxon signed-rank test. Holm correction was then performed to account for multiple testing. We found no significant difference between  $P_D$  and the 50% chance-level in both blocks, and no significant difference between  $P_C$  and the 50% chance-level in both blocks. This indicates that there likely was no systemic subject bias in responses in the case of uncertainty. Both tests comparing  $P_{AB}$  to the 50% chance-level indicated that  $P_{AB}$  was significantly higher than chance in both the (-SS) ( $t(17) = 9.6, p < 0.001$ ) and (-MS) blocks ( $V = 114.5, p = 0.006$ ).

A summary of these results is shown in [Table 3.2](#), and the distributions are illustrated in [Figure 3.3](#).

Table 3.2 – Mean results of the experiment studying the intensity of STP and DTP stimuli.  $P_{AB}$  is the proportion of answers where a STP stimulus is felt as more intense than a DTP stimulus.  $P_C$  (resp.  $P_D$ ) is the proportion of times the participants felt the second stimulus as more intense than the first, when both of them were rendered using STP (resp. DTP).

	$P_{AB}$	$P_C$	$P_D$
Single-stroke (-SS)	82.41%	56.48%	48.15%
Multi-stroke (-MS)	75.69%	56.94%	51.39%

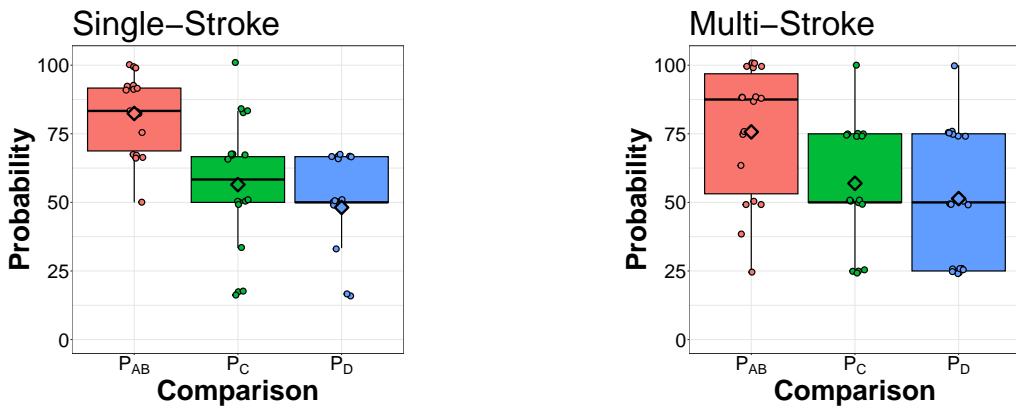


Figure 3.3 – Distributions of  $P_{AB}$ ,  $P_C$  and  $P_D$  for the single-stroke stimuli (-SS) block (Left) and multi-stroke stimuli (-MS) block (Right). For  $P_C$  and  $P_D$ , values close to 50% indicate an absence of systemic bias in participant responses in the case of uncertainty. An average above 50% for  $P_{AB}$  shows that STP stimuli are perceived as more intense than DTP stimuli. Diamonds represent the mean value, and dots represent individual participant results.

### 3.3.4 Discussion

For both blocks, the absence of significant difference between  $P_C$  (resp.  $P_D$ ) and the 50% chance-level indicates an absence of systemic biases in the participants' responses in the case of uncertainty, meaning that the  $P_{AB}$  data likely does not require corrections.

In both blocks,  $P_{AB}$  was shown to be significantly higher than chance level, fully supporting **H1-MS** and **H1-SS**, since STP stimuli were mostly rated as more intense than their DTP counterparts. This validates the main benefit of our method, on par with the literature comparing STM and AM-based stimulus intensity [Kappus et al. 2018; Howard et al. 2019].

### 3.3.5 Limitations and perspectives

This experiment shows that STP produces stronger stimuli than DTP. However, in the future, it would be interesting to precisely quantify the gain in perceived intensity, and to assess

the impact of rendering parameters (shape draw frequency, focal point speed, time spent on vertices) on this gain. Also, a comparison with conventional STM would provide insights into any potential trade-offs between perceived stimulus intensity and tactile shape clarity.

## 3.4 Naive Shape Perception

Here, we present a qualitative analysis and discussion of the naive shape sketches provided by participants for each of the shapes rendered in the intensity comparison experiment ([Section 3.3](#)).

### 3.4.1 Materials and Methods

We collected 18 sets (one per participant) of 40 shape drawings made in standardized 5x5 cm squares on both sides of an A4 sheet of paper. For each subject, we thus obtained a single sketch for each rendering method pair (A, B, C or D in [Section 3.3](#)), shape (see [Section 3.2.2](#)), and block ((-SS) or (-MS)) evaluated in the stimulus intensity comparison experiment ([Section 3.3](#)). An independent experimenter was tasked with classifying all the shape drawings into one of the following categories: “circle”, “ellipse”, “triangle”, “square”, “hexagon”, “octagon”, “other polygonal”, “other curved”, or “other hybrid” (i.e. part polygonal, part curved). Furthermore, the experimenter noted whether the drawn shape was a single closed contour, a single open contour, or multiple disjoint contours. To minimize any potential bias, the experimenter was not aware of the actual rendered shape which led to the subjects’ responses during this classification. [Figure 3.4](#) shows some examples of classifications of actual subject responses.

After classification, we selected only the subset of drawings corresponding to trials where a single rendering method was tested against itself (permutations C and D in [Section 3.3.2](#)), yielding 20 responses per subject. Three subjects provided incomplete responses, with 5 shape drawings missing in the final data set. The shapes drawn by participants in these conditions were then compared to the rendered shape, yielding either a perfect match (e.g. a triangle was rendered and the subject drew a triangle) or a shape category match (e.g. a circle was rendered and the subject drew an ellipsoid, classified as “other curved”). Results are qualitatively discussed below.

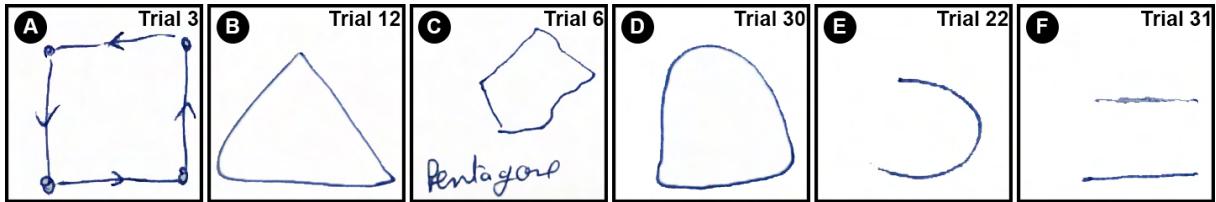


Figure 3.4 – Examples of (A) an unambiguous response (in this case, we are confident that the subject felt a “square”), (B) a response which could match a drawn shape but appears ambiguous (in this case, the subject probably felt a triangle, however the combination of two vertices and one rounded corner is ambiguous, the shape was therefore conservatively classified as “hybrid closed”, i.e. a closed contour combining curved and polygonal elements), (C) a polygon that does not obviously match any drawn shape (here, the subject indicated that they drew a pentagon, albeit poorly, which we thus classified as “polygonal closed”), (D) a curve that does not obviously match any drawn shape (classified as “curved closed”), (E) an open contour (in this case, classified as “curved open”), and (F) multiple disjoint shapes (here, classified as “polygonal disjoint”).

### 3.4.2 Results and Discussion

All considered methods produced a large majority of responses in the form of closed contours (92% for **STP-MS**, 81% for **STP-SS**, 86% for **DTP-MS** and 92% for **DTP-SS**), matching the general type of contours which were rendered.

[Table 3.3](#) summarizes the proportions of apparently correct naive shape representations for each shape and rendering method. [Figure 3.5](#) shows global results for the triangle shape.

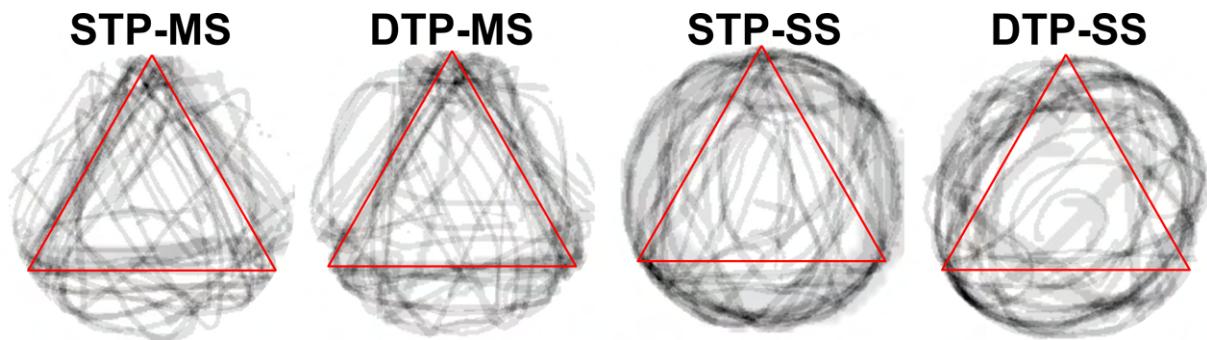


Figure 3.5 – Superposition of all the participant’s sketches for the different rendering techniques, after a triangle was rendered on their hand. Sketches were centered and normalized. The darker the point, the more participants drew it. The red triangle corresponds to the displayed triangle. We observe more confusion with the single-stroke rendering (right) than with the multi-stroke rendering (left).

Both **STP-MS** and **DTP-MS** appear to yield relatively high proportions of correct naive shape

Table 3.3 – The “Correct” columns show percentages of correct naive shape representations, and the “Cat. correct” columns show the percentages of correct shape category (i.e. “curved”, “circle” or “ellipse” for the circle and ellipse, or “polygonal”, “triangle”, “square”, “hexagon” or “octagon” for each of the four polygonal shapes). The “Cat. correct + hybrids” columns show these same percentages when the shapes classified as “hybrid” between polygonal and curved are considered as belonging to the correct shape category, reflecting the upper bound in the uncertainty in shape category classification. The “Closed contours” columns indicate the percentage of shapes drawn as closed contours in each condition, while the “Rounded contours” columns indicate the percentage of rounded shapes drawn in each case.

STP					DTP						
	Correct	Cat. correct	Cat. correct + hybrids	Closed contours	Rounded contours	Correct	Cat. correct	Cat. correct + hybrids	Closed contours	Rounded contours	
(-MS)	tri	56%	72%	72%	94%	22%	61%	72%	83%	89%	17%
	squ	39%	61%	67%	94%	33%	39%	50%	56%	83%	44%
	hex	6%	22%	28%	89%	72%	0%	17%	28%	83%	72%
	oct	0%	17%	17%	89%	83%	0%	6%	17%	89%	83%
(-SS)	circ	39%	61%	61%	89%	67%	33%	83%	89%	94%	83%
	ell	33%	78%	78%	61%	83%	56%	83%	89%	83%	89%
	tri	17%	39%	44%	83%	56%	11%	17%	22%	94%	72%
	squ	17%	28%	33%	89%	67%	6%	6%	11%	94%	89%
	hex	0%	22%	28%	83%	72%	0%	6%	11%	94%	83%
	oct	0%	6%	17%	83%	83%	0%	0%	6%	89%	89%

representations for the triangle (56% and 61% resp.), and somewhat for the square (39% for both), however this performance sharply drops for the more complex polygons (between 0% and 6%). So although both methods seem to be effective at providing relatively clear images of simple shapes, corroborating prior literature results [Hajas et al. 2020], they may inherently be limited to rendering such simple shapes. Both methods appear to provide clearer shape representations than their single-stroke counterparts: **STP-SS** yielded only 17% of correct representations for the triangle and square and 0% correct representations for the more complex polygons, and **DTP-SS** yielded 11% correct representations for the triangle, 6% for the square, and 0% for the more complex polygons.

Both **STP-SS** and **DTP-SS** appear to yield correct naive shape representations for circles and ellipses in a little more than a third of the time (39% and 33% resp. for **STP-SS**, 33% and 56% for **DTP-SS**).

Methods with vertex highlighting (**STP-MS** and **DTP-MS**) yield a low number of rounded contour responses for polygons with low vertex counts (resp. 22% and 33% in **STP-MS**, 17% and 44% in **DTP-MS** for triangles and squares) but a high number of rounded contour responses for polygons with higher vertex counts (resp. 72% and 83% for both **STP-MS** and **DTP-MS** for hexagons and octagons). These misidentification rates appear similar for hexagons and

octagons when using methods without vertex highlighting (resp. 72% and 83% for **STP-SS**, 83% and 89% for **DTP-SS** for hexagons and octagons), although the latter produced notably higher rounded contour responses for polygons with low vertex counts (resp. 56% and 67% for **STP-SS**, 78% and 89% for **DTP-SS** for triangles and squares).

These results appear to indicate that vertex highlighting plays an important role in conveying polygonal shape information, but that this approach is either limited to polygons with low vertex counts, or that rendering parameters we set are suboptimal for such shapes. Additional investigations at different shape draw frequencies and with different proportions of time spent on the vertices could shed some light on this question.

Finally, when the tactile information is unclear to users, it appears that shapes tend to be interpreted as rounded rather than polygonal.

### 3.4.3 Limitations and Perspectives

While they provide insights into naive shape perception with different ultrasound mid-air haptic rendering methods, any conclusions drawn from the above results should be viewed as tentative, as there are several limitations that affect our analyses.

First, we asked subjects to draw shapes, but “shape” may not be the salient attribute for the stimuli they feel so we might be biasing their response here. Thus, participants responses may have to be interpreted as “if you were to assign a shape to the stimulus, what would it be?” rather than a true representation of subjects’ perception.

When asking subjects to describe a shape, we provide them with a pen (which draws thin single lines) so there is a certain chance that we’re biasing them towards drawing shape contours regardless of whether they felt the shape as a contour line, an enclosed surface or something else entirely. In future work, this could be mitigated by providing a larger array of representation tools to the subjects, although this could significantly complicate subsequent interpretation by experimenters.

Although we took steps to limit any bias in the shape classification on the part of the experimenter, there is a certain degree of personal appreciation that necessarily goes into e.g. classifying a shape as having corners (polygonal) or not (curved) as well as judging what constitutes an imperfect drawing of a shape which was presented and what constitutes an “ambiguous” shape. In future studies focusing on this aspect, this may be further mitigated by e.g. having multiple independent experimenters classify the shape drawings and using a majority voting

system to create a final classification. Also, a dedicated experimental protocol where subjects perform trials with repetitions of the same condition, and provide responses in the form of naive drawings, written or verbal descriptions, and then select the best match among a very large set of predefined shapes while describing any perceived mismatches may provide further insight into subjects' actual shape perceptions.

## 3.5 User Study: Clarity of Rendered Shapes

In [Section 3.3](#), we showed that STP shapes are perceived as significantly stronger than DTP shapes. In this section, we investigate whether or not STP conserves shape identification performances obtained in DTP, as the qualitative results in [Section 3.4](#) would lead us to believe.

### 3.5.1 Hypotheses

Based on the qualitative results from [Section 3.4](#) and prior results from [Hajas et al. \[2020\]](#), we hypothesize that:

- **H2:** shapes will be rated as *sharper* (i.e. less blurry) when rendered with **STP-MS (H2-MS)** and **STP-SS (H2-SS)** than when rendered with **STM**.
- **H3:** shapes will be identified correctly more often when rendered with **STP-MS (H3-MS)** and **STP-SS (H3-SS)** than with **STM**.
- **H4:** shapes will be identified as accurately when rendered with **STP** as when they are rendered with **DTP**.

### 3.5.2 Materials and Methods

#### Apparatus and setup

Apart from the experimental user interface, the setup for this experiment is identical to that of the first experiment (see [Section 3.3.2](#) and [Figure 3.2](#)).

#### Procedure

Participants performed the following trial: a shape was rendered with a given rendering technique for 6 s. It was up to the participant to chose whether they wanted to passively experience

the shape on their palm or actively explore it by moving their hand. After the shape was rendered, participants answered three questions. First, they identified the shape they experienced by choosing from the 4 (resp. 6) possible shapes. Then, they rated the perceived sharpness of the shape on a 7-points Likert scale, from “very blurry” to “very sharp”. Finally, they were asked to indicate whether they kept their hand immobile, moved it, or a little of both.

We once again used a within-subject blocked design, with one block evaluating single-stroke (-SS) stimuli and the other evaluating multi-stroke (-MS) stimuli. We used the same set of shapes and rendering techniques as in the previous experiment, with the addition of standard STM rendering [Kappus et al. 2018; Frier et al. 2018] in both blocks. The order of blocks is counterbalanced across participants. Participants were shown graphical representations of all shapes that could be presented with each block prior to completing said block.

Within each block, we ran trials assessing three repetitions of all possible (shape, rendering technique) combinations. The three possible rendering techniques were **STP**, **DTP** and **STM**. We considered 4 shapes in the (-MS) block and 6 shapes in the (-SS) block. The order of trials was fully randomized within each block. In total, participants performed 90 trials.

## Participants

We recruited a total of 20 participants (15 men, 4 women, 1 non-binary, aged  $m \pm s.d. = 24.85 \pm 3.02$ ), who did not participate in the first experiment. 15 of them were right-handed, 4 were left-handed, and one was ambidextrous and performed the experiment as a right-handed person. 11 of them had no prior experience with haptics, 6 had little experience, and 3 had a lot of experience. 16 of them never used an UMH interface prior to the experiment.

## Collected Data

Prior to the experiment, participants were briefed on the experimental procedure, provided written informed consent to participate and filled out a demographics questionnaire. During the experiment, we recorded the shapes identified by the participant for each trial, along with their rated sharpness, in order to assess **H2**. To better discuss the results, participants also indicated the exploration strategy used. Based on each participant’s answers, we computed one recall value for each strategy and for each block (6 recall values per participant in total), which was in turn used to assess **H3** and **H4**.

### 3.5.3 Results

We collected the answers of the 20 participants for all trials, leading to 1080 values per question for the single-stroke block, and 720 for the block with pausing strategies. Confusion matrices for the shape identification are summarized in [Figure 3.6](#), and all average results are in [Table 3.4](#). Recall and sharpness distributions are shown in [Figure 3.7](#).

**STM vs. DTP-SS vs. STP-SS** For the block with the single-stroke stimuli, both the recall and sharpness distributions for all three strategies followed an almost-normal distribution (Shapiro-Wilk,  $p > 0.05$ ), and variances were similar between the three rendering techniques (Levene,  $p > 0.05$ ). A one-way repeated measures ANOVA with recall as the dependant variable, rendering technique as the independent variable, and participant as a random effect variable revealed no significant effect of rendering technique on recall. A similar ANOVA with sharpness as the dependant variable revealed a significant effect ( $F(2, 38) = 16.03, p < 0.001$ ). Post-hoc analyses with Holm-corrected pairwise t-tests showed that both **DTP-SS** and **STP-SS** were rated as producing sharper shapes than **STM** ( $p = 0.001$  in both cases,  $t(19) = 4.01$  for the **DTP-SS** vs **STM** comparison,  $t(19) = -4.27$  for the **STM** vs **STP-SS** comparison).

**STM vs. DTP-MS vs. STP-MS** Recall and sharpness distributions with **STM**, **DTP-MS** and **STP-MS** were almost normal (Shapiro-Wilk,  $p > 0.05$ ) and had similar variances (Levene,  $p > 0.05$ ). We thus applied the same analysis protocol as for the (-SS) block. The ANOVA revealed a significant effect of rendering technique on recall ( $F(2, 38) = 21.08, p < 0.001$ ), and post-hoc pairwise t-tests with Holm corrections revealed that both **DTP-MS** and **STP-MS** yielded significantly better performances than **STM** ( $p < 0.001$  in both cases,  $t(19) = 5.53$  for the **DTP-MS** vs **STM** comparison,  $t(19) = -5.1$  for the **STM** vs **STP-MS** comparison). An ANOVA revealed a significant effect of rendering technique on sharpness ( $F(2, 38) = 46.56, p < 0.001$ ), and the post-hoc analysis similarly showed that **STM** shapes were felt significantly more blurry than the other two strategies ( $t(19) = 7.63$  for the **DTP-MS** vs **STM** comparison,  $t(19) = -7.17$  for the **STM** vs **STP-MS** comparison,  $p < 0.001$  in both cases).

**Exploration** In both blocks, for each participant and each strategy, we measured the percentage of trials that were completed by keeping the hand fully immobile. The distribution for all strategies did not follow a normal distribution (Shapiro-Wilk,  $p < 0.05$ ), so we performed Friedman tests, which revealed a significant effect of rendering technique in both blocks ( $\chi^2(2, N = 20) = 13.4, p = 0.001$  for the single stroke block,  $\chi^2(2, N = 20) = 21.5, p < 0.001$  for the multi-stroke block). Pairwise Wilcoxon signed-rank tests with Holm cor-

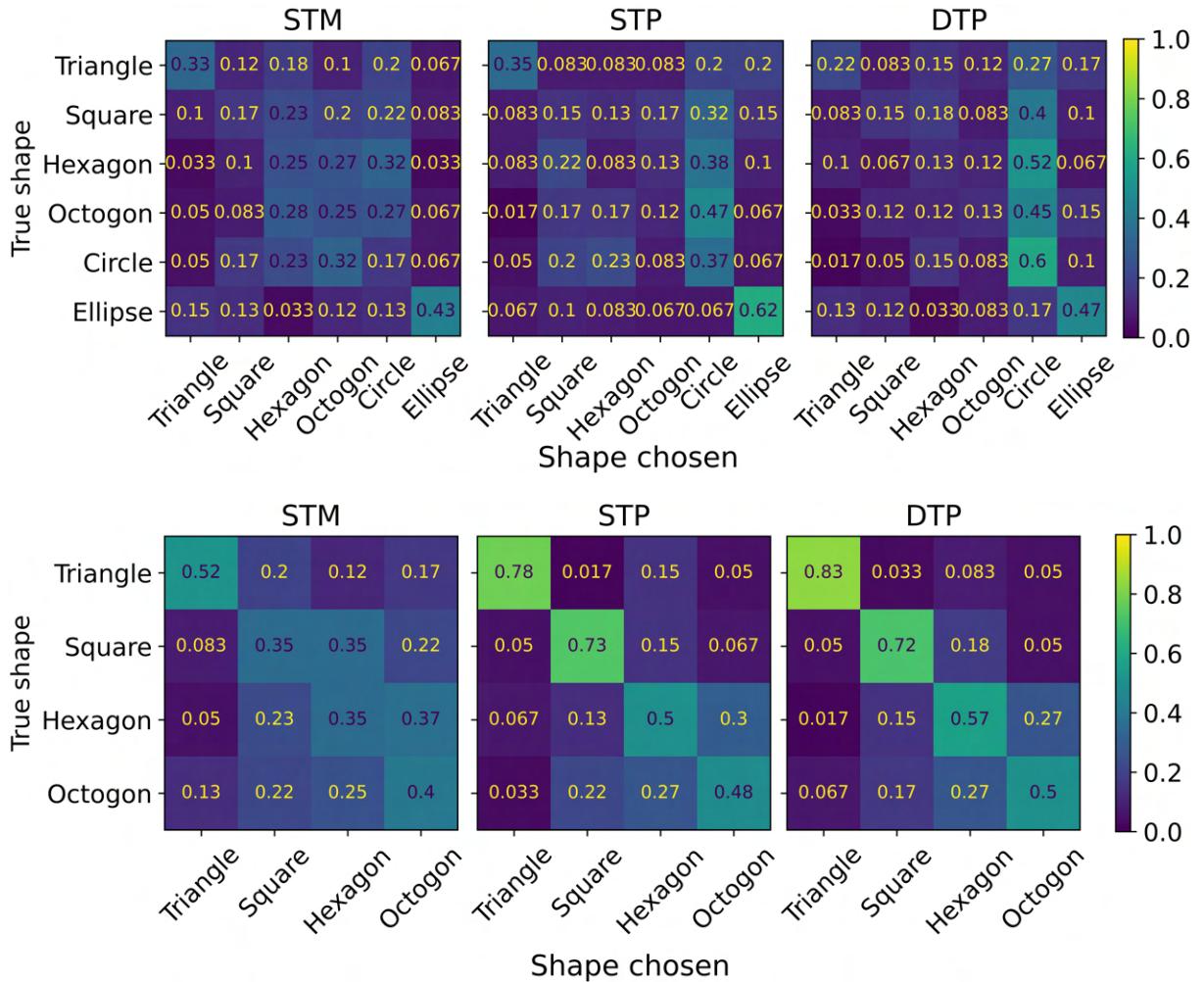


Figure 3.6 – Confusion matrices for the shape identification experiment, for both the (-SS) (Top) and (-MS) (Bottom) blocks.

reactions revealed that participants were significantly more tempted to explore **STM** shapes dynamically ( $V = 196, p = 0.002$  for **DTP-SS** vs. **STM**,  $V = 15, p = 0.002$  for **STM** vs. **STM** vs **STP-SS**,  $V = 204, p < 0.001$  for **DTP-MS** vs. **STM**,  $V = 3, p < 0.001$  for **STM** vs. **STP-MS**).

### 3.5.4 Discussion

The significantly lower sharpness rating for **STM** shapes supports **H2-MS** and **H2-SS**. Shapes rendered using any form of tactile pointer are perceived as sharper.

In terms of shape identification performances, **STP-MS** and **DTP-MS** provided significantly

Table 3.4 – Mean results of the experiment studying shape identification performances with **STM**, **DTP** and **STP** stimuli.

	-SS			-MS		
	STM	DTP	STP	STM	DTP	STP
Recall	0.27	0.28	0.28	0.4	0.65	0.63
Sharpness (/ 7)	2.81	4.11	4.2	2.43	4.53	4.66
Passive exploration rate	0.31	0.78	0.78	0.32	0.90	0.90

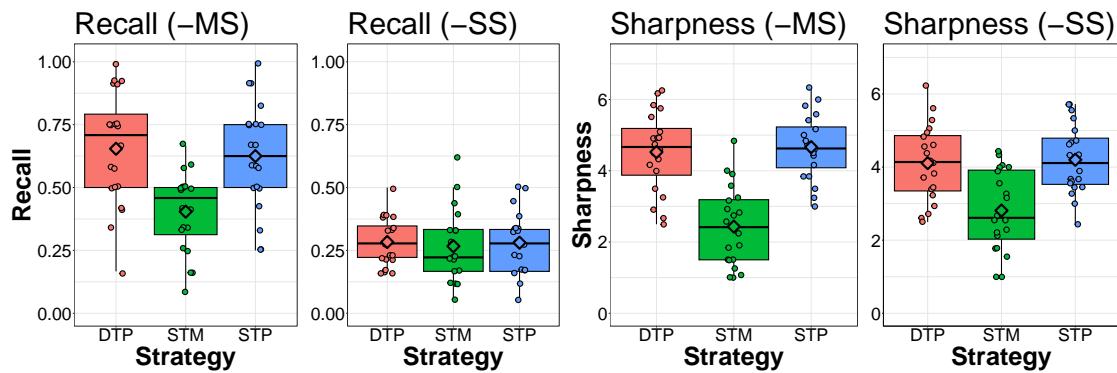


Figure 3.7 – Recall (Left) and Sharpness (Right) distributions for the shape identification experiment. Diamonds represent the mean value, and dots represent individual participant results.

better performances than **STM**, fully supporting **H3-MS**. However, **H3-SS** is not supported, since recall is not significantly different between **STM**, **STP-SS** and **DTP-SS**. Therefore, shapes rendered with any form of tactile pointer are easier to identify than **STM** shapes only when the tactile pointer marks pauses on a shape's vertices.

The absence of significant differences in sharpness ratings and in shape identification performances between **STP-MS** and **DTP-MS** (respectively, **STP-SS** and **DTP-SS**) tends to support **H4**. It therefore appears that **STP** performs on par with **DTP**.

It is also worth noting that participants were used mainly a static exploration strategy for when shapes were rendered with a tactile pointer, while they explore mainly dynamically, or partially dynamically when the shape was rendered statically using **STM**. This shows that the different rendering techniques have uses for different types of applications, although future work is required to study if active exploration could work for tactile pointers if we were to increase the stimulation duration.

## 3.6 Discussion and Perspectives

### 3.6.1 Limitations and Future Work

While 38 people participated in the experiments, the sample population is not sufficiently large or varied to confidently draw general conclusions. Recruiting more participants is thus necessary.

Throughout both experiments, shapes were displayed with a fixed orientation, that was empirically tested before. Yet, it is possible that the shape orientation might play a role in the identification performances, and future work will study this question.

By imposing shape dimensions, we also necessarily introduce biases. Here we chose to have all shapes inscribed inside a fixed-radius circle so that all shapes fit on the palm of the majority of participants, but that leads to shapes with different perimeters and surfaces. This implies differences in area stimulated and focal point speed (see [Table 3.1](#)), which, if perceivable, could be used to identify shapes without using geometry-related features. Yet all of these parameters cannot be fixed altogether. Future work is required to study the effect of these parameters independently. Future work is also required to optimize the different rendering parameters (draw frequency, time spent on vertices, section size), especially for more complex shapes.

### 3.6.2 General Discussions

The measured identification performances for STM and DTP are on par with the results of [Hajas et al. \[2020\]](#). The slightly lower results in our case can be explained by the addition of other answer possibilities, and the slightly suboptimal pause duration for triangles and squares, that were used to enable chunking for the other polygons. Since STP also shows similar performances to DTP, we confirmed that the general concept of tactile pointers improves shape identification.

We also observed improved performances for the naive shape identification ([Table 3.3](#)), in the case of simple shapes. The naive and constrained results ([Figure 3.6](#)) also show that users tend to report circular shapes when the amount of vertices increases. This is on par with our other results [[Mulot et al. 2021b](#)] (see [Chapter 5](#)), where we report that small variations of curvature are mostly unperceivable with DTP.

Apart from the base shapes studied by [Hajas et al. \[2020\]](#), we added three more complex shapes which were more poorly recognized. This can be explained by a suboptimal parametering of

DTP and STP, such as the lower pause duration on vertices. This is coherent with their results for the triangle and square, showing poorer recognition rates with shorter pauses.

The improved results in a constrained answer scenario show the tactile pointers could be very useful for scenarios with visuo-haptic congruence. Future work is necessary to assess the limitations of rendering incoherent visuo-haptic feedback.

More generally, the fact that performances between STP and DTP are very similar shows that we accomplished our goal of reproducing the main features of DTP, while also providing more intense feedback.

### 3.7 Conclusion

In this chapter, we presented **spatio-temporally-modulated tactile pointers** (STP), a novel approach for rendering 2D tactile shapes on the skin using ultrasound mid-air haptics, by splitting them into sequentially rendered spatio-temporally-modulated line segments or arcs. We experimentally demonstrate that this approach produces ultrasound mid-air haptic shapes that are as recognizable as those rendered with the best existing method from the state of the art, dynamic tactile pointers, while significantly increasing the perceived stimulus intensity. We also qualitatively investigate to what extent users perceive shapes that are similar to those being rendered, and derive perspectives for future investigations required to better understand ultrasound mid-air haptic shape perception. Overall, STP show promising results for rendering intense and recognizable mid-air haptic shapes.

#### Chapter Summary

- We introduced Spatio-Temporally-Modulated Tactile Pointer (STP), a new UMH shape rendering method based on Dynamic Tactile Pointers (DTP);
- We conducted two user studies to validate our method. Results showed that STP performs on par with DTP while providing significantly more intense stimulation;
- Tactile pointers seem more suited for the static exploration of simple shapes with a small number of vertices.

---

## **Part II**

# **Optimized Primitives for Improved Rendering Algorithms**

---



# Chapter 4

## Harnessing Shock Waves in the Skin: Spiral-Based Rendering Primitives

### Contents

---

4.1	Introduction . . . . .	101
4.2	Harnessing Shear Shock Formation: Spiral Scanning . . . . .	103
4.3	Spiral-Based Shape Rendering . . . . .	108
4.4	User Study: 3D Object Perception . . . . .	112
4.5	User Study: Haptic Shading . . . . .	116
4.6	Conclusion . . . . .	119

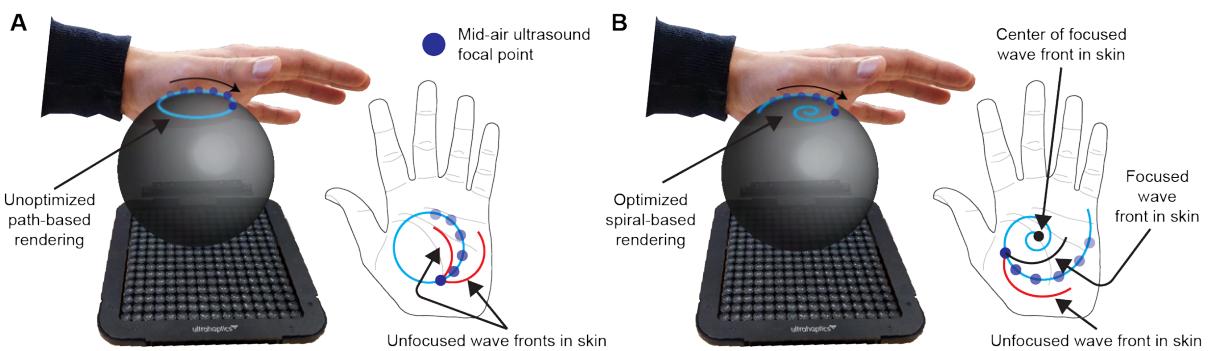
---

### 4.1 Introduction

Part II contributes to **Challenge II**, aiming to provide new and improved rendering primitives to push the boundaries of what UMH can render, while also adding more choice and variation to the UMH designer’s toolbox. More specifically, this chapter follows an approach based on the modelization of wave propagation in the skin to generate optimized focal trajectories.

Reardon et al. [2023] revealed that the high focal speeds often used for STM exceed the prop-

agation speed of the resulting shear waves in the skin. This supersonic motion leads to the formation of a shear shock front that propagates far from the original focal trajectory, spreading acoustic energy along its path. It is therefore assumed that these shear shock waves play a large role in UMH perception, potentially explaining why STM shapes can be perceived as blurry [Obrist et al. 2013]. Instead of spreading this energy and diminishing the sharpness of UMH sensations, our work aims at providing new rendering primitives that leverage these shear shock waves, harnessing them to locations of our choosing to somewhat control the spread of energy in the skin, as illustrated in [Figure 4.1](#).



[Figure 4.1](#) – (A) Current UMH rendering methods rely on the fast scanning of a high-pressure point along a given trajectory, generating shear shock waves in the skin and thus dissipating energy. (B) Our physics-informed method uses spiral trajectories to harness one of the shock fronts, and focus at positions of our choice to recover its energy, and improve 3D object haptic perception.

Adding new rendering primitives expands the design freedom of UMH designers, and enables a new range of rendering methods. Furthermore, to the best of our knowledge, this is the first rendering primitive derived from physics theory, leveraging mechanical phenomena in the skin.

Overall, our contributions are:

- a new physics-based rendering primitive, using logarithmic spiral trajectories to focus shear shock wave to target locations ([Section 4.2.1](#)). Numerical experiments were conducted to assess the validity of our method ([Section 4.2.2](#)).
- a new 3D shape and intensity shading UMH rendering algorithm, based on a combination of spiral primitives ([Section 4.3](#)). This approach was validated through two user studies, highlighting the potential of our technique for holographic haptic displays ([Section 4.4](#), [Section 4.5](#)). Our rendering algorithm allows users to perceive heterogeneous

objects using UMH, paving the way for a novel generation of rendering techniques enriching our 3D interaction with virtual worlds.

These contributions have not yet been published, but correspond to a collaboration between the following researchers (in alphabetical order): Maud Marchal, Lendy Mulot, Claudio Pacchierotti, Gregory Reardon, and Yon Visell.

## 4.2 Harnessing Shear Shock Formation: Spiral Scanning

In this section, we discuss the design of new UMH rendering primitives based on acoustic theory. Then, we validate our approach through numerical experiments, highlighting the robustness and versatility of our method. Providing formalized and validated primitives is an important step towards controllable and predictable UMH rendering methods.

### 4.2.1 Physics-Based Derivation

In holographic haptics, sensations result from dynamic indentations of the skin generated by localized acoustic radiation pressure. The Langevin acoustic radiation force  $\mathbf{F}_L$  [N] is described, to second order, by an expression of the form:

$$\mathbf{F}_L = - \iint_S dS (\langle p_2 \rangle \mathbf{I} + \rho_0 \langle \mathbf{u}_1 \otimes \mathbf{u}_1 \rangle) \cdot \mathbf{n} \quad (4.1)$$

where,  $p$  [Pa],  $\rho$  [kg.m<sup>-3</sup>], and  $\mathbf{u}$  [m.s<sup>-1</sup>] are the fluid pressure, density and velocity fields.  $dS$  [m<sup>2</sup>] is a surface element, while  $\mathbf{I}$  is the identity matrix. The angular brackets denote time-averaged quantities,  $\mathbf{n}$  is the surface normal, and  $\cdot$  represents the outer product. Subscripts 0, 1 and 2 refer to successive terms in an expansion about a quiescent fluid configuration [Doinikov 1994; Danilov et al. 2000; Hasegawa et al. 2000; Settnes et al. 2012]. The acoustic radiation force drives propagating oscillations in the skin in a process described by a viscoelastic wave equation:

$$\mathbf{F}_L(\mathbf{x}, t) = \left( -m \frac{\partial^2}{\partial t^2} + \mu \nabla^2 + ((K + \mu/3) \nabla) \nabla \cdot \right) \xi(\mathbf{x}, t) \quad (4.2)$$

where,  $\xi(\mathbf{x}, t)$  [m] is tissue displacement, and  $m$  [kg.m<sup>-3</sup>],  $\mu$  [N.m<sup>-2</sup>], and  $K$  [N.m<sup>-2</sup>] are the mass density, shear modulus, and bulk modulus, respectively, and  $\mathbf{F}_L(\mathbf{x}, t)$  [N.m<sup>-3</sup>] is the Langevin force density. In the physical regime of interest, the propagation speed  $c$  [m.s<sup>-1</sup>] of waves excited in the skin is well-approximated by  $c = \sqrt{\mu/\rho}$ .

In holographic haptics, a localized acoustic pressure source is dynamically scanned across lo-

cations  $\mathbf{p}(t)$  on the skin surface. We consider the problem of engineering a scanning path  $\mathbf{p}(t)$ , with constant scanning speed, that causes the skin oscillations elicited along the scanning path to converge toward a specified focal location  $\mathbf{p}_0$  at the same time  $t_0$ . Writing the scanning path in polar coordinates parametrizing a plane that locally approximates the volar hand surface,  $\mathbf{p}(t) = (r(t), \phi(t))$ . Without loss of generality, let the focus location be  $\mathbf{p}_0 = (0, 0)$ . Thus,  $r(t)$  is the distance from the path location at  $t$  and the focus location. The scanning velocity vector has components  $\mathbf{v}(t) = \dot{\mathbf{p}}(t) = (\dot{r}(t), \dot{\phi}(t))$ . Since the waves propagate toward the focus location at speed  $c$ , the first wavefront emitted from the scanning source at times  $t$  and  $t + dt$  will reach the focus location simultaneously if  $r(t + dt) - r(t) = dr = -c dt$ .

Thus, for convergence to the origin,  $\dot{r} = dr/dt = -c$ . By induction, this condition ensures that the first wavefront emitted at every point along the path converges to the origin at the same time. The path scanning speed is given by:

$$v = \|\mathbf{v}\| = \sqrt{\dot{r}^2 + r^2\dot{\phi}^2} = c M, \quad (4.3)$$

where  $M = \sqrt{1 + r^2\dot{\phi}^2/c^2}$  is the Mach number. For any non-zero angular velocity  $\dot{\phi}$ , the scanning speed is thus supersonic, with Mach number  $M > 1$ . The conditions for focusing by wave convergence require that the components of scanning velocity  $\mathbf{v} = (\dot{r}, \dot{\phi})$  satisfy:

$$\dot{r} = -c, \quad \dot{\phi} = \frac{\pm c}{r} \sqrt{M^2 - 1} \quad (4.4)$$

These expressions describe a converging vector field that is centered on the origin. The integral curves of this expression are logarithmic spiral trajectories, given by:

$$r(t) = -ct + r_0 \quad (4.5)$$

$$\phi(t) = \alpha \sqrt{M^2 - 1} \log \left( \frac{r(t)}{r_0} \right) + \phi_0 \quad (4.6)$$

where  $\alpha = \pm 1$  dictates the chirality of the spiral (clockwise or counter-clockwise). By scanning a source along such a path, waves excited along the entire path converge at the same time at the focus location. Thus, this technique harnesses wave energy excited along the entire duration of the path for focusing at  $\mathbf{p}_0$ . Given a wave medium with speed  $c$ , and a target focus  $\mathbf{p}_0$ , a trajectory is specified by four free parameters: the Mach number,  $M$  (or equivalently the speed  $v$ ), the initial rotation  $\phi_0$ , the chirality  $\alpha$ , and the initial radius,  $r_0$ . The latter value determines the duration  $\Delta t$  of the spiral primitive, given by  $\Delta t = r_0/c$ . It is also worth noting that this parameterization requires knowledge of the wave speed  $c$ . The effects of adjusting

these parameters are illustrated in [Figure 4.2](#).

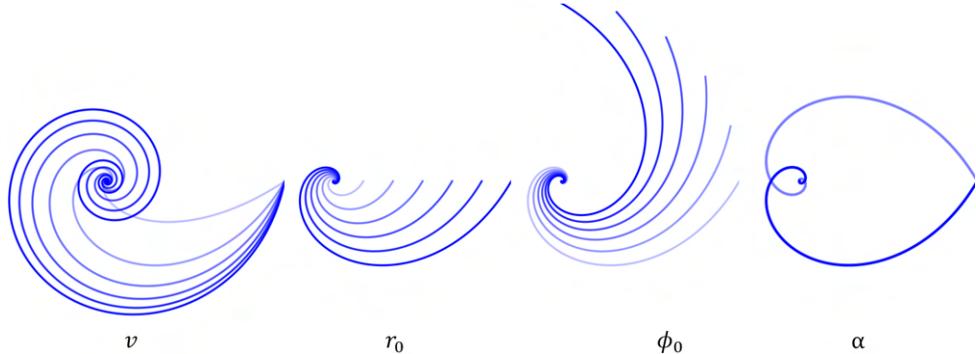


Figure 4.2 – Evolution of the log spiral trajectory when varying different parameters. Increasing  $v$  makes the spiral more swirly, while  $r_0$  makes it longer, and  $\phi_0$  rotates it.  $\alpha = \pm 1$  represents the chirality of the spiral. Darker spirals have higher values of the given parameter.

### 4.2.2 Numerical Experiments

While the [Section 4.2.1](#) presented the theory behind shear shock waves focusing, the end of this section aims to validate the use of logarithmic spirals through numerical experiments.

#### Setup

We used a numerical model that computes the viscoelastic response of a skin-like material, when a focal source is moving at its surface, based on k-Wave, a pseudo-spectral numerical package [[Treeby et al. 2014](#); [Treeby et al. 2014](#)]. This is a similar setup to that of [Reardon et al. \[2023\]](#), using the same parameterization, with  $c = 5$  m/s. This model was experimentally validated by [Reardon et al. \[2023\]](#), through comparisons with vibrometry measurements on gelatin phantom material, and on human skin.

#### Results

We ran simulations of a source moving at different speeds, and with two different assumed wave speeds. [Figure 4.3](#) shows the successful coalescing of the shock front at the focus point, resulting in a high increase in localized energy. [Figure 4.4](#) (A) shows that correct focusing happens at a wide range of focal speed  $v$ , at least up to 25 m/s [ $M = 5$ ]. This provides a large design space when designing haptic sensations, contributing to its versatility.

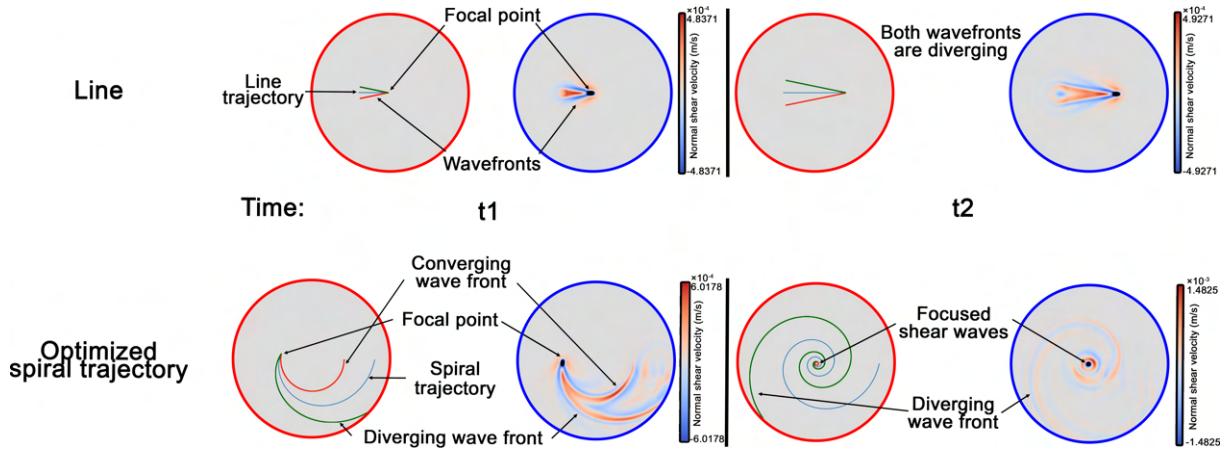


Figure 4.3 – Line (Top) and spiral (Bottom) focal trajectories generating shear shock waves when scanned at high-speed. Naive implementations of trajectory equations (red) and numerical simulations (blue) are shown at two time steps  $t_1 < t_2$ .

## Discussion

While these simulations show great success in coalescing waves to a predefined location, there are several parameters to optimize, and many factors to consider to create optimal spiral rendering. This sub-section aims at discussing the main factors, and evaluating how they may affect the user perception.

**Impact of the Wave Speed** First, the shear wave propagation speed  $c$  depends on the properties of the skin, which implies that there could be some between-participants deviation. An incorrect assumption of the wave speed would lead to a radial speed that does not match the medium's wave speed, meaning that the waves will not reach the target point at the same time, as shown in [Figure 4.4\(B\)](#). This highlights the importance of measuring each participant's skin wave speed to optimize spiral-based rendering to everyone's needs. That said, our numerical experiments showed that the focusing still worked great for errors up to 10-20% in assumed wave speed, highlighting the robustness of our technique.

**Impact of the Hand Geometry** It is also important to note that the obtained equations assume a planar trajectory, which is not strictly true for a human's palm, although an open hand has only little curvature on the palm. If the spiral is rendered in a given plane, this may lead the contact point with the hand to follow a slightly different trajectory, with partially variable speed, thus reducing the quality of the wave focusing. Future work will focus on measuring such effects, and adapting the trajectory equations to work on curved surfaces.

#### 4.2. Harnessing Shear Shock Formation: Spiral Scanning

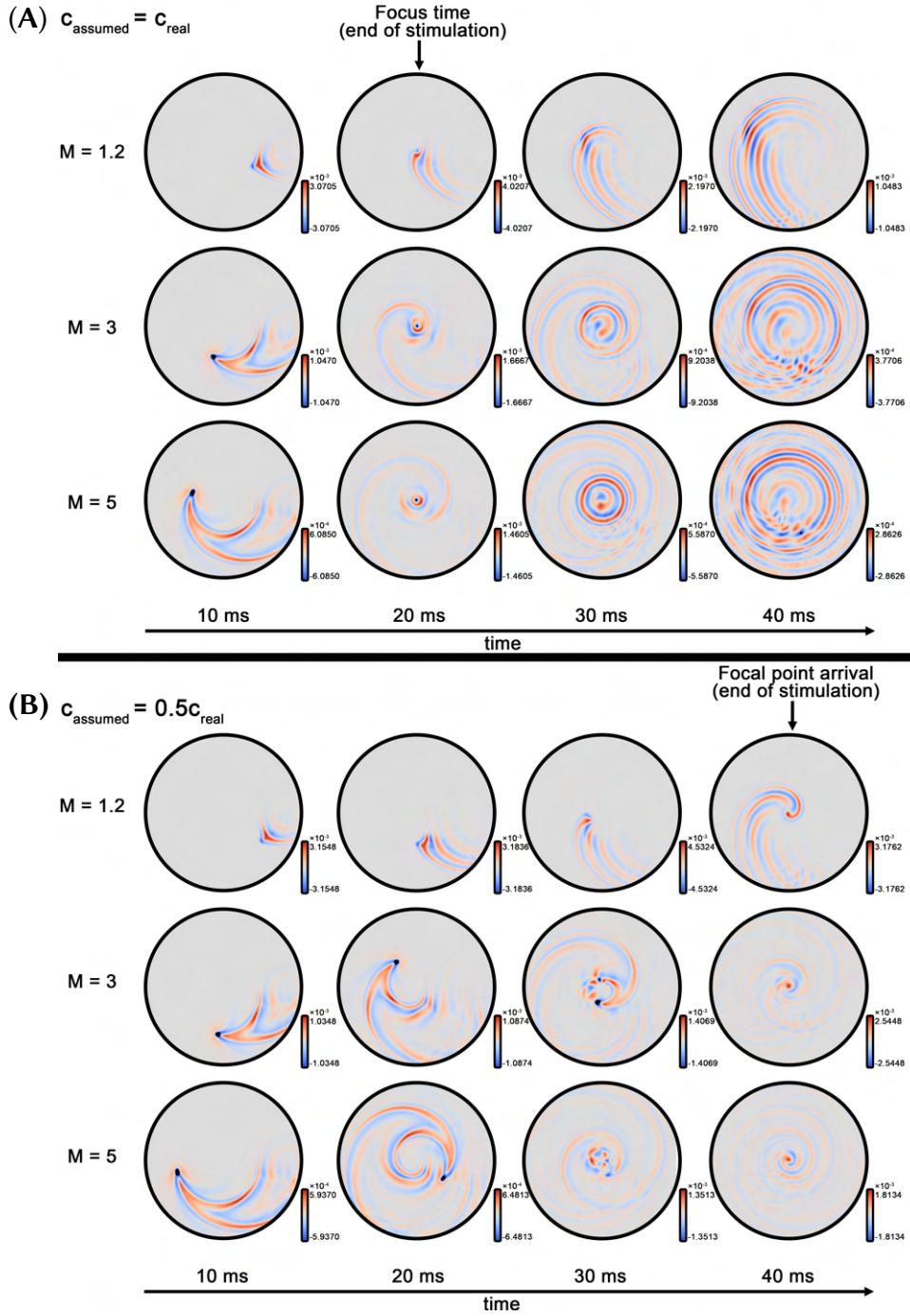


Figure 4.4 – Numerical simulations of a focal point moving along a log spiral trajectory at different speeds. (A) When the assumed wave speed matches the actual medium’s wave speed, we see a clear focusing of the shock front (here at  $t = 20$  ms). (B) But when there is a mismatch (here, we assume the wave speed is half of its actual value), the arrival of the waves at the focus point is spread over a larger period of time (here, from  $t = 20$  to 40 ms). In all cases, we used  $r_0 = 100$  mm,  $\phi_0 = 0$ , and  $M = v/c_{\text{real}}$ .

When combined with precise hand detection tools, this would further improve the stimulation quality.

Moreover, haptic sensations should be designed with exploration location in mind. If the user is supposed to explore using their fingers, one should prefer using small spirals that fit on the finger pad, to avoid wasting energy when rendering sections not in direct contact with the skin.

**Impact of the Damping** The skin is a viscoelastic material, implying that waves dissipate energy while propagating through the medium. This means that waves emitted at a certain location cannot travel further than a certain distance, after which they would fade. Waves emitted at the start of a spiral larger than the fading distance will therefore not contribute to the shock wave focusing, as they would fade before reaching the focus point. This provides an upper limit to  $r_0$ , after which we cannot harness more energy by increasing the size. In such cases, it may be more suitable to render successive smaller spirals instead, to reduce the amount of dissipated energy. Future work will focus on quantifying this wave damping effect, to provide guidelines to haptic designers.

**Perception Limits** Finally, it is important to note that physical and perceptual optima may significantly differ, and that different spiral parameters might have different effects on user perception. For instance,  $r_0$  is inversely proportional to the pattern’s draw frequency, while  $v$  is proportional to the distance between focal points, and also influences the area covered by the spiral, as shown in [Figure 4.2](#). Those parameters have been shown to influence the perception of mid-air haptic stimuli [[Frier et al. 2019](#)]. Similarly,  $\phi_0$  could be randomly modified at each scan to create a sensation that is more diffuse, and less shaped like a spiral. It can also be set to different values depending on the location on the hand, to ensure most of the spiral lies on the skin. Future studies will be conducted to further measure the influence of each rendering parameter, in order to perceptually optimize spiral-based rendering.

## 4.3 Spiral-Based Shape Rendering

When UMH was first introduced, stimuli were rendered using a single amplitude-modulated focal point. Later improvements allowed focal points to be moved in space [[Hoshi et al. 2009](#)], which allowed UMH designers to render a shape by turning it into a point cloud, and moving

the focal point between positions, after one or more modulation periods [Korres et al. 2016; Korres et al. 2017]. Similarly, Long et al. [2014] proposed to render 3D objects by computing the intersection between the object and the user’s hand, leading to a 2D shape intersection that could be rendered with the same process of successive amplitude-modulated points. The amplitude-modulated point was thus a rendering primitive, that could be combined to render higher-level stimuli.

We follow a similar approach, by first proposing a new rendering primitive for UMH (Section 4.2), using logarithmic spiral trajectories to focus shear shock wave to targeted locations. In this section, we propose to combine these primitives together, in order to render 2D and 3D shapes. We also extend this technique to highlight specific locations on the shape, through a process that we call **haptic shading**. This allows us to render more diverse patterns with more complex sensations, introducing the possibilities of rendering 3D heterogeneous objects.

### 4.3.1 2D Shape Rendering

To showcase the novel possibilities offered by our rendering primitive for high-level combination, we chose to mostly focus on rendering the same types of shapes as the initial work of Long et al. [2014]. We therefore rendered a cone, a sphere, a square pyramid, a vertical triangular prism and a cube. That said, our technique can be extended to render more complex shapes (Section 4.3.4). All but the prism were also used in one of our studies [Milot et al. 2025] (see Chapter 9). For interactions with an horizontal hand, this requires the rendering of either a circle, a square, or a triangle cross-section. All rendered patterns are illustrated in Figure 4.5, and discussed below.

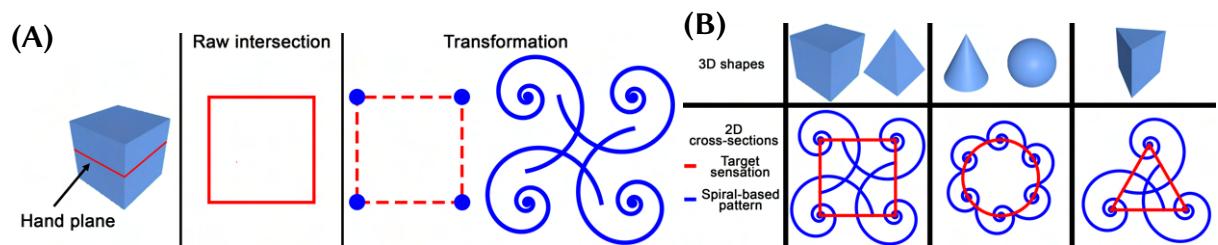


Figure 4.5 – (A) Spiral rendering pipeline. We compute the intersection between the explored object and the hand plane, before extracting vertices and edges to position spirals. The focal point then scans this contour at a fixed speed. (B) (Top) 3D shapes involved in our study. (Bottom) Pattern used for the rendering of a horizontal cross-section of each shape. STM patterns correspond to the target sensation, and are shown in red, while spiral patterns are shown in blue.

Since it has been shown that shape features (e.g. vertices, edges) play a very important role in

discrimination performances [Hajas et al. 2020; Mulot et al. 2023b], we chose to render squares and triangles by using spirals, for which  $\mathbf{p}_0$  is set to the position of each vertex, and  $\phi_0$  is set to orient the spiral along the adjacent edge. The circle requires to render a higher amount of vertices compared to the square and the triangle, so as to provide a smoother contour that will not be confused with them. It has been shown that shapes with many vertices are extremely hard to perceive with static rendering methods, and are very often interpreted as circles [Mulot et al. 2023b] (see [Chapter 3](#)). As such, we chose to render the circle as if it was an inscribed hexagon. Through empirical testings, we set  $r_0 = .7s$ , where  $s$  is the side length. When assessing a value for  $r_0$ , we must ensure that the whole edge feels stimulated, to avoid a feeling of “holes” in the shape. On the contrary, high values for  $r_0$  increase the spiral’s length, reducing the draw frequency, which also likely reduces the perceptual intensity. High values of  $r_0$  would also lead to overlapping spirals, meaning that overlapping areas would be rendered twice as much as other places, which may be considered a waste of energy, depending on the objective. Thus, 70% of the side’s length was empirically found to be a good trade-off. In terms of rendering, each spiral was scanned with a focal speed of  $v = 25$  m/s, or equivalently,  $M = 5$ , as we assumed  $c = 5$  m/s, following Reardon et al.’s results [[Reardon et al. 2023](#)]. To render the whole pattern, spirals are rendered sequentially, and the process is repeated when each of them has been scanned once. These value also led to very similar draw frequencies between our spiral-based method, and a classical STM contour-rendering method at 8 m/s focal speed.

### 4.3.2 Haptic Shape Shading

After using the spiral primitives to render shapes, we propose a new process, called **haptic shading**, where a portion of the shape is emphasized. This allows the presentation of heterogeneous surfaces, which is a step towards detailed textured objects.

Specifically, once the shape has been turned into a set of spirals, we proposed in [Section 4.3.1](#) to render them one after the other. Instead, here we propose to select a single spiral, and render it several times, before moving on to the next one. With our spiral-based primitive, the more we want to emphasize on a specific spiral, the more times we render it before moving on to the next spiral. This introduces a novel algorithm for rendering objects with various properties.

This novel method leads to a change in the provided energy distribution, where the shaded region receives  $n$  times more energy than the other spirals during a full scanning of the shape, where  $n$  is the amount of times the shaded area was scanned. It is worth noting that, by increas-

ing  $n$ , we make the shaded area more salient, but we also reduce the amount of energy delivered to all other spirals, which may negatively impact user shape recognition performances. After empirical testings, we used  $n = 4$  for our experiments. For a circle with 6 spirals, that means 44% of the time is spent on the shaded area, compared to 11% on every other spiral.

### 4.3.3 3D Shape Rendering

To add the third dimension to our interactions, we first represent the virtual hand as a plane, aligned with the user's palm. Users were instructed to maintain their hand parallel to the UMH interface in order to avoid sudden changes in focal speed, and to provide optimal stimulation. We therefore rendered these 3D objects using an intersection strategy (see [Chapter 6](#)) consisting in a *raw intersection* between the object and hand plane. Our spiral-based rendering method used the optional *transformation* step to turn the contour into a set of spirals, according to the aforementioned process ([Section 4.3.1](#)). This means that when moving their hand up and down, the user was experiencing a horizontal slice of the object, at the same height as their hand.

### 4.3.4 Generalization and Discussion

While we discussed the rendering of specific shapes that fit our study objectives, our process could easily be extended to render any polygon, by using one spiral per vertex and its associated edge. Yet, in a general scenario, it may be worth representing some edges as a set of smaller spirals, instead of a single very large one.

It is also worth noting that the values for each rendering parameters were chosen through empirical testings. While they make sense when compared to usual pattern sizes and draw frequency, they are probably not optimal. Our goal here is to propose a new rendering primitive, and show its potential, but future work will be focused on assessing the perceptual impact of each rendering parameter, to better optimize rendering techniques.

Specifically, it is worth investigating the impact of the amount of spirals. For instance, a large edge would result in a large spiral, which may lead to a larger perceived area. It would be possible to instead use  $n$  spirals, positioned evenly along the path, that are  $n$  times smaller than the initial big spiral. This process would not change the draw frequency, as the total path length remains the same, but it would reduce the "bounding box" of the rendered trajectories, which may lead to a sharper or more accurate shape perception. On the other hand, adding more spirals generates more points of discontinuity in the focal point motion, which may

create some unwanted audible harmonics, and reduced perceived intensity. Finally, the energy conveyed through the shock front is fading with distance, due to damping effects. By reducing the size of these spirals, we may increase the amount of energy actually reaching the spiral focus points.

## 4.4 User Study: 3D Object Perception

We first want to highlight the potential and versatility of spiral-based rendering, by demonstrating its ability to render 3D objects.

Here, we hypothesize that shape discrimination performances for 3D objects are similar when using spiral-based feedback, or state-of-the-art STM (**H1**).

### 4.4.1 Materials and Methods

#### Apparatus and Setup

The setup for this experiment is illustrated in [Figure 4.6](#). During the experiment, the participant is seated in front of a computer screen displaying the scene. They interact using their dominant hand above an Ultraleap STRATOS Explore <sup>1</sup> UMH interface, located to the side. An Ultraleap Stereo IR 170 <sup>2</sup> hand tracker is attached to the interface. The user also wears a noise-isolating headset to partially hide the operating noise of the interface. The experiment was conducted using Unity <sup>3</sup> 2022.3.18f1.

During the main interaction phase, the virtual scene ([Figure 4.6\(A\)](#)) was composed of a semi-transparent cube, indicating where the user needed to explore. Text and timers were also present to show the progress and recap the current task. Answer boxes were placed behind the interaction cube, showcasing the possible answers. Finally, a sphere was placed in front of the virtual user to validate their answer. When interacting, the users see a virtual replica of their hand, moving according to their movements.

#### Conditions

This first study was inspired from experiments conducted by [Long et al. \[2014\]](#) and one of our own studies (see [Chapter 9](#)), extending the ever-expanding study of 3D object rendering with UMH to our new spiral-based method.

- 
1. [https://www.ultraleap.com/datasheets/STRATOS\\_Explore\\_Development\\_Kit\\_datasheet.pdf](https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf)
  2. <https://www.ultraleap.com/product/stereo-ir-170/>
  3. <https://unity.com/>

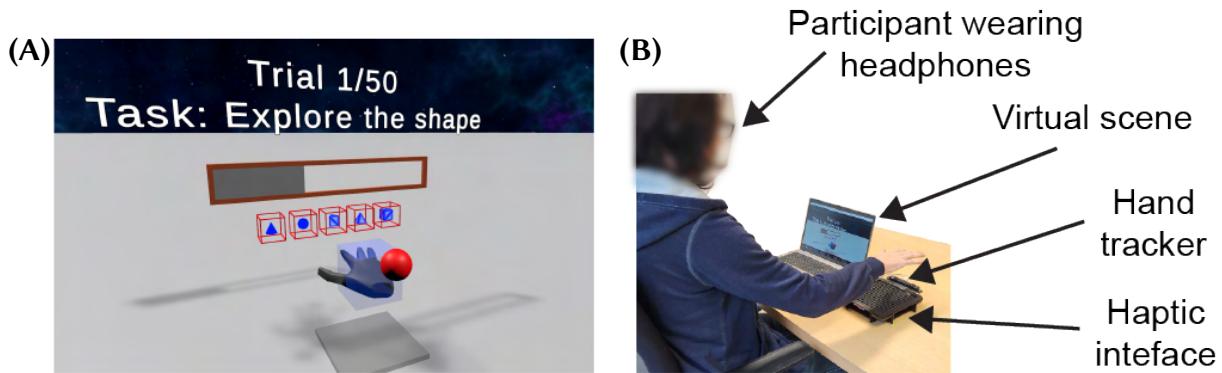


Figure 4.6 – Virtual scene (A) and physical setup (B) for both experiments ([Section 4.4.1](#)).

We used ten different conditions, corresponding to the combination of five shapes (cone, sphere, pyramid, vertical prism, and cube), and two rendering methods (spiral-based, and STM, as presented in [Section 4.3](#), and illustrated in [Figure 4.5](#)). When using state-of-the-art STM, we render the contour of these shapes with a focal speed of 8 m/s [[Frier et al. 2019](#)]. All shapes were positioned at the center of the interaction cube, and scaled to fit inside a 6 cm-sided cube. This way, most of any planar section of the object could fit on the hand of the vast majority of users [[Peebles et al. 1998](#)].

### Procedure

Each of the 10 conditions was repeated 5 times, for a total of 50 fully-randomized trials. When a trial started, the user was prompted to move their hand to a starting position, above the interaction area. Then, they could explore the invisible 3D object inside the interaction area by moving their hand up and down. They were prompted to avoid any rotation of the hand to keep a consistent focal point speed on the skin. The shape was rendered for a maximum of 20 s, after which the user had to select the shape they thought they experienced, out of the five possibilities. They did so by touching the corresponding answer box with their index finger, before validating by touching the validation sphere (see [Figure 4.6\(A\)](#)). If the users found the shape before the end of the 20 s, they were allowed to register their answer earlier. Overall, this experiment lasted between 20 and 30 minutes.

Prior to this experiment, the users underwent a familiarization phase to get accustomed to the setup. To limit potential biases, this phase used different conditions, rendering only a 2D square, triangle, and circle, using state-of-the-art STM. Each of those were repeated twice, for a total of 6 trials. Here, there were only 3 answer boxes for these 3 shapes.

## Collected Data

Prior to the experiment, we collected informed consent from the participants, and gathered demographics data through a questionnaire. During each trial, we collected the user’s answer. After the experiment, they were also asked to rate the perceived difficulty of the task, on a 7-points Likert scale.

At the end of the experiment session (including both experiments from [Section 4.4](#) and [Section 4.5](#)), the participants were asked to reflect on their experience, by answering the following questions when discussing with the experimenter:

- Did you feel differences between stimuli? If so, what kind of differences?
- What strategy did you used?

## Participants

We recruited 19 participants (13 men, 5 women, and 1 non-binary, aged  $m \pm s.d. = 32.4 \pm 9.7$ ), all right-handed. Nine of them reported having limited prior experience with haptics, and 4 never experienced UMH stimuli before.

### 4.4.2 Results

We gathered data about 950 trials (19 participants, 50 trials per participants). First, we computed each participant’s accuracy (proportion of correct response) for each of the two rendering technique, resulting in averages of  $(62.7 \pm 16.9)\%$  for STM, and  $(66.5 \pm 19.0)\%$  for spiral rendering. Shapiro tests on these two groups of accuracy values failed to reject the normality hypothesis ( $p > 0.05$  in both cases). One-sample t-tests showed that accuracy results are significantly higher than the 20% chance-level ( $t(18) = 11, p < 0.001$  for STM,  $t(18) = 10.7, p < 0.001$  for spiral-based rendering). A paired-samples t-test indicated no significant accuracy differences between the two rendering techniques ( $t(18) = 1.1, p > 0.05$ ).

Confusion matrices for the two rendering techniques are shown in [Figure 4.7 \(A\)](#). For each (shape, rendering technique) pair, we computed each participant’s accuracy. Shapiro tests revealed a significantly non-normal distribution in most groups. We therefore performed Friedman tests for both rendering techniques, with accuracy as dependant variable, shape as independent variable, and participant as random effect variable. This revealed a significant effect of the shape when using STM ( $\chi^2(4, N = 19) = 24.7, p < 0.001$ ), but no significant effect when using spiral-based rendering ( $\chi^2(4, N = 19) = 2.6, p > 0.05$ ). Pairwise Holm-corrected

Wilcoxon signed-rank tests for the STM data revealed that the pyramid lead to significantly lower results than the prism ( $V = 143, p = 0.014$ ), the cube ( $V = 105, p = 0.01$ ), and the cone ( $V = 133, p = 0.008$ ).

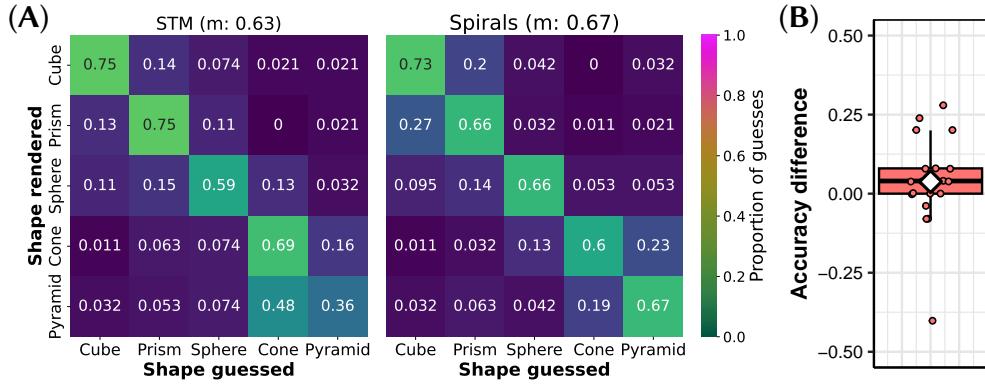


Figure 4.7 – (A) Confusion matrices for the 3D shape discrimination experiment, when using STM (left) and spiral-based (right) shape rendering techniques (Section 4.4). Higher values along the diagonal indicate greater success. (B) Accuracy difference between participant’s spiral and STM results for the first experiment. Dots represent individual data, and the diamond represents the mean. Positive values indicate better results when using spiral rendering.

#### 4.4.3 Discussion

This experiment showed no significant difference in terms of shape recognition accuracy between state-of-the-art STM, and spiral-based rendering, which fully supports **H1**. When looking at the confusion matrices (Figure 4.7 (A)), we observe very similar results between the two techniques, except for the pyramid, which is more easily differentiated from a cone when using spiral-based rendering. This may be due to the increased pattern area, which can accentuate differences between shapes, although edges become less sharp. Future work will need to be conducted to verify this hypothesis. Overall, this is extremely promising as we are able to provide similar performances to state-of-art methods, while using a focal path that is very different from the actual shape. So the benefits of spiral-based rendering (even for this non-perceptually-optimized version) outweigh the geometrical inaccuracies.

On average, participants rated the difficulty of the task  $5.2 \pm 1.1$ , out of 7, indicating a quite difficult task. Despite this, they reached quite good accuracies, with 66.5% success on average for spiral-based rendering. This highlights a low confidence in participant’s answer, despite decent results.

In terms of reported strategy, most participants first explored the shape vertically to under-

stand how the pattern size evolves. For instance, if they felt no change in pattern when moving down, the users could infer that it was either a cube or triangular prism. Then they experienced a horizontal portion of the shape, trying to find its 2D intersection. For instance, if they feel a square intersection, shrinking as the hand goes down, then it must be a pyramid.

Interestingly, many participants reported perceiving differences between stimuli of the same shape, highlighting differences between classical STM and spiral-based stimuli. Many discussed differences in size, likely due to the larger area covered by spirals. Others reported differences in terms of “frequency”. Finally, many described STM patterns as sharper but hollow, while spiral-based patterns gave the impression of more continuous or filled shapes.

Also, as illustrated in [Figure 4.7](#) (B), there was a single participant who had an accuracy difference between rendering conditions that was way larger than the other participants, with STM scoring 40 points higher. This may be an outlier behaviour. When removing this participant from the data, the t-test between rendering accuracies becomes significant ( $t(17) = 2.5, p = 0.02$ ). This indicates a trend of the spiral-based feedback slightly outperforming STM, although deeper investigations would be required to confirm this trend.

Overall, we showed that spiral-based rendering is at least as good as state-of-the-art STM, showing high potential for the technique. Future work will look into optimizing the rendering parameters for spirals, which may then best STM.

## 4.5 User Study: Haptic Shading

While the first experiment showed that spiral-based rendering can render 3D shapes with at least as much precision as state-of-the-art methods, this section showcases the versatility of spiral-based rendering. To do so, we propose to use it to highlight specific parts of the shape, through a process we call **haptic shading**. Here, we hypothesize that users can effectively perceive shaded locations (**H2**) when using UMH. Perceiving such differences in the UMH stimuli could pave the way for the rendering of 3D heterogeneous shapes, enriching the possibilities of interaction with holographic displays.

This second experiment used the same setup as the first experiment ([Section 4.4.1](#)). Both experiments were run consecutively in a single session. Therefore, this experiment shares the same participants as the first one ([Section 4.4.1](#)).

### 4.5.1 Materials and Methods

#### Conditions

For this experiment, participants always experienced a 2D circle, rendered using shaded spirals, as presented in [Section 4.3.2](#). The circle was rendered as shown in [Figure 4.5](#). We used four different conditions, corresponding to four shaded areas, being one of the spirals on the left or right. For convenience, we refer to these conditions as **NE** (North-East), **NW** (North-West), **SE** (South-East), and **SW** (South-West), as shown in [Figure 4.8](#). For instance, right-handed person experiencing the circle centered on their palm would perceive the **SW** location around the thenar eminence.

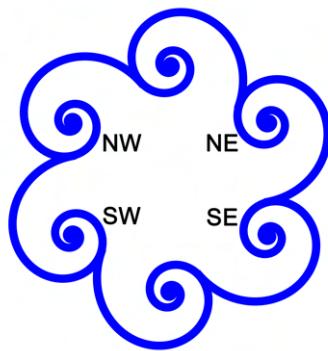


Figure 4.8 – Conditions for the haptic shading experiment. One of the four designated spiral is rendered more often than the others.

#### Procedure

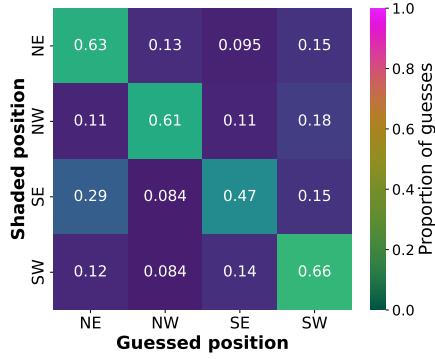
The procedure was almost identical to the one of the previous experiment ([Section 4.4.1](#)), except that participants had to locate the shaded area, and pick the correct answer out of the four possibilities. They could explore the shaded circle for up to 30 s before stimulation was stopped. Once again, each condition was repeated 5 times, leading to 20 fully-randomized trials, lasting for around 10 minutes.

#### Collected Data

We gathered the participant's answer for each trial. After the experiment, participants rated the perceived difficulty of the task on a 7-points Likert scale. They were also asked to discuss their strategy with the experimenter.

### 4.5.2 Results

We collected data about 180 trials (19 participants, 20 trials per participant). The confusion matrix gathering the location answers is shown in [Figure 4.9](#).



[Figure 4.9](#) – Confusion matrix for the haptic shading experiment ([Section 4.5](#)). Higher values along the diagonal indicate greater success.

Similarly to the previous experiment, we computed each participant’s global accuracy ( $m: 59.5\%$ ,  $s.d.: 15.6\%$ ). A Shapiro test supported a normal distribution of the data ( $p > 0.05$ ), and a one-sample t-test indicated that the data was significantly higher than the 25% chance-level ( $t(18) = 9.6, p < 0.001$ ).

We also computed participant’s accuracies for each shaded position, resulting  $(63.2 \pm 26.0)\%$  for the **NE** condition,  $(61.1 \pm 33.6)\%$  for **NW**,  $(47.4 \pm 21.3)\%$  for **SE**, and  $(66.3 \pm 29.9)\%$  for **SW**. Apart for the **NE** condition, all groups followed a non-normal distribution (Shapiro,  $p < 0.05$ ). We thus performed a Friedman test, with the accuracy as dependent variable, the shaded position as independent variable, and the participant as random effect variable, which revealed no significant difference between locations.

### 4.5.3 Discussion

We showed that participant accuracy was significantly higher than chance-level, indicating that participants can effectively locate the shaded region, which fully supports **H2**.

Participants rated this task’s difficulty  $4.7 \pm 1.3$ , indicating a decently difficult task but easier than shape discrimination, which shows a higher confidence.

Although no significant differences were found between shaded locations, [Figure 4.7](#) (C) shows a tendency of **SE** (47%) having worst results than the other locations (61-66%). This is coherent

with the hand's tactile sensitivity, having higher tactile sensitivity around the thenar prominence (**SW**) and near the metacarpophalangeal joints (**NE/NW**), but lower sensitivity around the **SE** location.

Overall, we showed that haptic shading is possible using spiral-based rendering. This further highlights the potential and versatility of this rendering technique. Haptic shading is an important step towards detailed and textured UMH sensations.

## 4.6 Conclusion

While UMH technologies present a high potential for holographic haptics, the quality of rendered stimuli is lessened due to the fast scanning of a high-pressure focal point generating shear shock waves in the skin. This dissipates energy, and makes the sensations more blurry. To tackle this issue, we proposed new focal trajectories based on mechanical theory of the wave propagation in the skin. These trajectories aimed at focusing the shock front to a location of our choosing, harnessing some of the energy that would otherwise have been lost. We conducted numerical experiments using state-of-the-art models, showing that we successfully manage to focus shear wave energy to single locations, proving the effectiveness of our new rendering primitive. We then expanded this work, combining spiral primitives to render 2D and 3D shapes, resulting in performances at least as good as state-of-the-art. We also introduced haptic shading, a process that lets us emphasize specific portions of the shape. Through the introduction of new rendering primitive, our work opens a new range of possibilities for holographic rendering.

## Chapter Summary

- Based on mechanical wave propagation theory, we derived a family of logarithmic spiral trajectories, a new **rendering primitive** to focus the shear waves generated by the motion of the focal point along this path onto a single location;
- We provided numerical evidence of the effectiveness of our method, along with insights about the perceptual implications of different rendering parameters;
- We used these spiral primitives to design new intersection strategies for the rendering of 3D virtual objects;
- We introduced UMH **shading**, using these primitives to highlight specific parts of a shape;
- Two user studies were conducted to assess the effectiveness of our method. Despite the complex geometry of the resulting patterns, we observed shape identification performances that are at least as good as state-of-the-art STM rendering.

---

## **Part III**

# **UMH Rendering Pipeline: Formalization**

---



# Chapter 5

## 2D Rendering: Sampling Strategies

### Contents

---

5.1	Introduction . . . . .	123
5.2	Stimulus Design Formalization . . . . .	124
5.3	The DOLPHIN Framework . . . . .	125
5.4	Proof of Concept . . . . .	132
5.5	Interoperability . . . . .	138
5.6	Conclusion . . . . .	139

---

### 5.1 Introduction

Part III focuses on formalizing the overall UMH rendering pipeline, it attempts to address Challenge III, providing common grounds to design, discuss, or replicate UMH studies. As discussed in Section 1.4, UMH rendering uses a decently large amount of rendering parameters, each having their own physical and perceptual effects. Providing a common ground for discussions and presentations would benefit designers and researchers altogether.

Specifically, this chapter can be contextualized by the lack of good generic open-source tool to design and evaluate UMH sensations, as discussed in Section 1.4.4. Here, we choose to address this issue from the bottom-up, by prioritizing full control over low level rendering parameters,

with perceptual evaluations in mind. Overall, we propose the following contributions:

- A formalization of **sampling strategies**, representing the rendering of 2D shapes ([Section 5.2](#)).
- DOLPHIN, an open-source framework to help designers create, reuse, and evaluate UMH sensations, with full control over the shape and sampling parameters ([Section 5.3](#)).
- A psychophysics study of curvature discrimination performances. This study was designed using DOLPHIN, and also serves as a proof of concept for the usefulness of the framework ([Section 5.4](#)).

This work is then extended in [Chapter 6](#) for 3D rendering.

The content of this chapter was adapted from the following publication: [[Mulot et al. 2021b](#)], which is an extension of the following work-in-progress publication: [[Mulot et al. 2021a](#)].

## 5.2 Stimulus Design Formalization

An UMH stimulus has two main aspects. The first aspect is spatial, and concerns the stimulus' geometry as well as its position in space, as shown in [Figure 5.1\(A\)](#). The shape can be represented as a parametric function  $f : [0, 1]^k \rightarrow \mathbb{R}^3$ . The number of parameters  $k$  may differ depending on the shape. A time parameter can be added for time dependant (dynamic) shapes, such as a moving line.

The second aspect is encompassed under the term of **sampling strategy**, i.e. the shape's spatial sampling and temporal evolution of the stimulus, illustrated in [Figure 5.1](#). We thereby define a set of  $N$  sampled points  $P^s = \{q_i, 0 \leq i < N\}$  s.t.  $\forall i, \exists t_{1,i}, \dots, t_{k,i}, q_i = f(t_{1,i}, \dots, t_{k,i})$  is a point on the shape. The temporal evolution in turn encompasses the order and timing with which the sampled points on the shape are displayed, as well as the evolution of the focal point's amplitude with respect to time. We can formally define the sampling order as a finite series  $P^{s,o} = ((P_j, T_j))_j$  where  $P_j \in P^s$  for all  $j$  and  $T_j$  is the dwell time on point  $P_j$ . By adding the amplitude overlay and update rate information, we obtain a configuration for each time step  $C_k = (P_{j_k}, I_k)$  where  $0 \leq j_k < N$  and  $I_k = \frac{p(P_j)}{p_{peak}(P_j)}$  represents the pressure to peak pressure ratio for the device at the position  $P_j$ , with  $p : \mathbb{R}^3 \rightarrow \mathbb{R}^+$ .

A sampling strategy can thus be represented as the transformation  $f \mapsto (C_k)_k$ , given an update rate and peak pressure. This means that a same sampling strategy can be applied to different shapes. A sampling strategy transforms a possibly continuous shape ([Figure 5.1\(A\)](#)) into a discrete and finite representation ([Figure 5.1\(B\)](#)). Since the same set of spatially sam-

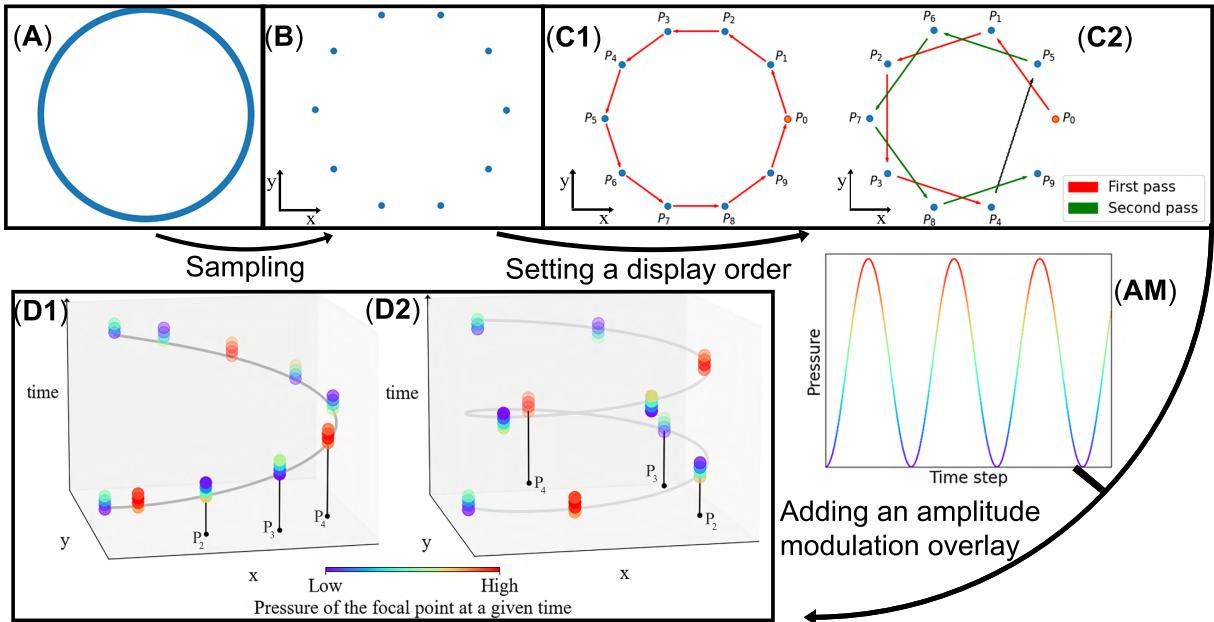


Figure 5.1 – Illustration of the three steps for applying a sampling strategy to a shape. The process begins with (A) an abstract shape, here a circle. The shape is then (B) spatially sampled. (C1) and (C2) show examples of possible display orders of the sampled points. Finally, an intensity modulation overlay (AM) is applied. (D1) and (D2) show the application of (AM) to the display orders respectively shown in (C1) and (C2). The vertical axis corresponds to time. Points are color-coded with respect to their intensity at a given display time step, with weak intensities shown as blue and high intensities as red. Dwell time is a third of the AM period.

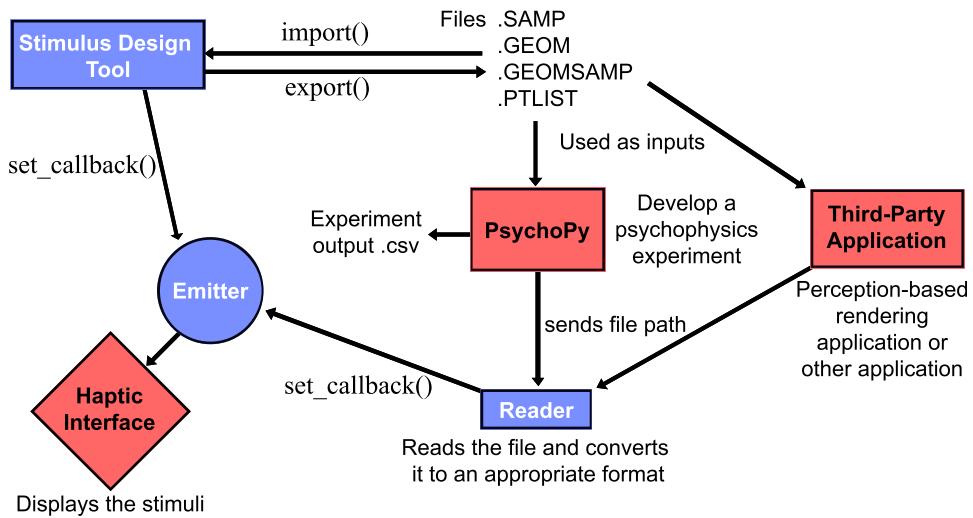
pled points  $P^s$  can be displayed in different orders and with different timings, the sampling strategy also provides rules to define both these aspects (Figure 5.1(C1, C2)). This step leads to the creation of  $P^{s,o}$ . In addition to this, the sampling strategy must provide information about the peak pressure of the focal point at each time step (Figure 5.1(AM)). For instance, a simple STM approach would generate the maximum possible pressure at each time step, while an AM approach generates a pressure varying as a (usually sinusoidal) function of time. This data, combined with the previous one, finally leads to the computation of the different configurations  $C_k$  at each time step (Figure 5.1(D1, D2)).

## 5.3 The DOLPHIN Framework

We developed DOLPHIN, an open-source framework<sup>1</sup> to help users create haptic stimuli by customizing both shape and sampling strategy, and to use them for perception studies. A stim-

1. URL: <https://gitlab.com/h-reality/dolphin>

ulus design tool provides a visual representation of the stimulus being created along with physical information about generated pressure distributions. Users can use DOLPHIN to directly test the stimulus and export it for later reuse in the framework's design tool or in a third-party application using DOLPHIN's reader for the ultrasound interface. [Figure 5.2](#) shows DOLPHIN's components and possible use cases, and [Figure 5.3](#) presents a UML diagram of the framework.



[Figure 5.2](#) – DOLPHIN's architecture. Rectangles represent programs, the diamond represents a device and the circle represents a class. Blue elements are provided as part of the framework, while red ones are external. Export from the design tool is possible in four formats corresponding to the shape (.GEOM), the sampling strategy (.SAMP), or both (.GEOMSAMP, .PTLIST) with the generic .PTLIST format intended for use in external tools. Stimulus files can be used in experiments or third-party applications thanks to a reader program which sends data to the emitter for rendering.

### 5.3.1 Design Tool

The design tool architecture can be split into four parts. The "geometry" component handles the representation of mathematical shapes. The "sampling strategy" component implements different user-customized ways of displaying a shape. An "import and export" component handles saving and loading of shapes and strategies for reuse within the framework or by external applications executing the framework's Reader. Finally, the "user interface" links everything together. The tool also uses the Emitter library presented in [Section 5.3.2](#) to display shapes with according sampling strategies on a hardware interface. The stimulus design tool is devel-

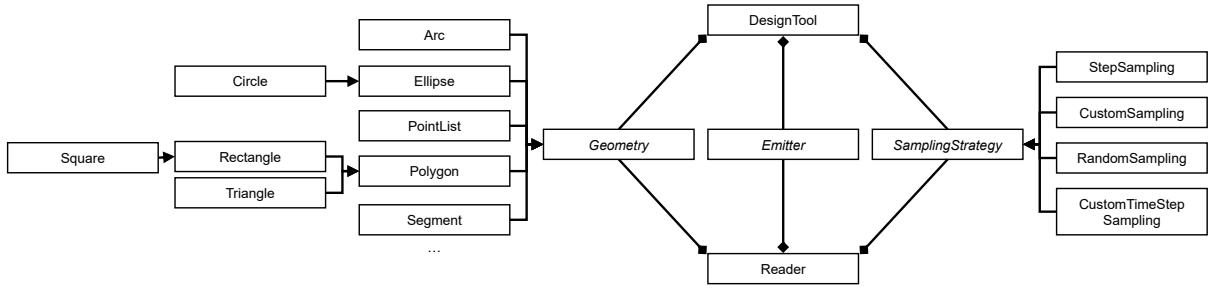


Figure 5.3 – Simplified UML representation of the DOLPHIN framework. The DESIGNTOOL and READER represent the two high-level components that users will mostly be interacting with.

oped in C++ and uses the Qt framework<sup>2</sup>.

## Geometry

This component handles the mathematical representation of shapes to be displayed to the user. Every shape class inherits from the Geometry virtual class. The tool currently handles circles, ellipses, squares, rectangles, triangles, segments, point clouds and arcs. Users can create new shapes by creating sub-classes of Geometry, adding support for their configuration in the user interface and expanding the import and export operations to deal with the new shape.

Each shape can be customized using parameters specific to the shape, such as e.g. the radius and center position for a circle.

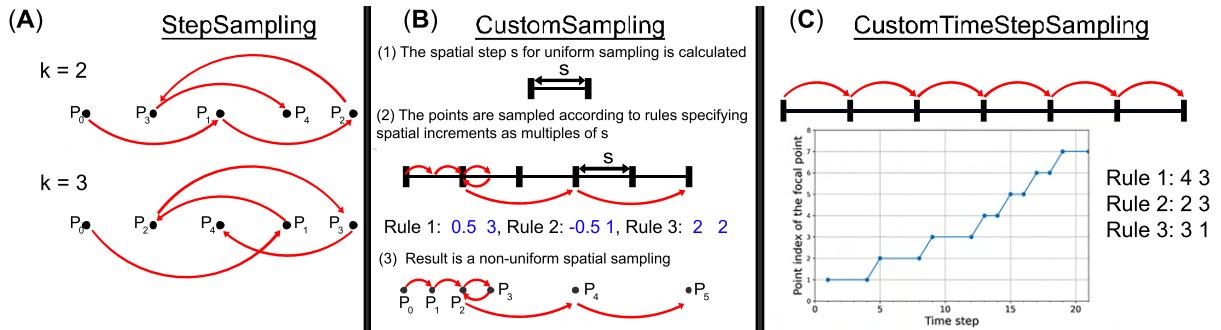
These shapes allow the on-screen display of the stimulus to the user and the computation of its spatial sampling as per the selected sampling strategy. To do so, each shape implements an evaluate\_at method defining the parametric representation  $f$  of the shape. While the 2D plane in which the shapes are defined could theoretically be tilted, the software currently only supports a horizontal plane parallel to the device, as tilt would highly reduce the perceived intensity [Awatani 1955; Hoshi 2022a]. The tool currently only displays shapes defined as continuous curves or point clouds, and will be extended to surface and volume rendering in the future.

## Sampling Strategies

The SamplingStrategy virtual class presents the main characteristics needed for a strategy. Combined with a Geometry it is used to create a list of 2D or 3D points corresponding to the

2. <https://www.qt.io/>

sampled points, along with a list representing the intensity at each point. The following four strategies, illustrated in [Figure 5.4](#) are currently available.



[Figure 5.4](#) – Three sampling strategies available in DOLPHIN’s design tool. (A) `StepSampling` is a uniform sampling with controllable rendering step. (B) `CustomSampling` allows for a non-uniform sampling using a sequence of rules, all corresponding to a `StepSampling` for a given amount of points. (C) `CustomTimeStepSampling` is the equivalent of `CustomSampling` for temporal control instead of spatial control: rules control how long the focal point should remain at each location. DOLPHIN also implements a `RandomSampling` strategy, where sampled positions are randomly picked. But this strategy is only really useful as a baseline for comparisons.

`StepSampling` corresponds to a uniform division of the shape into  $n$  points with a step parameter  $k$ . It leads to the sequential display of point 1, then point  $k + 1$ ,  $2k + 1$  and so on, modulo  $n$ . For instance, with  $k = 2$  and an odd number of points  $n$ , we first display the points with even indices and then display the ones with odd indices.

`CustomSampling` enables the user to create a non-uniform sampling. It is configured using a series of  $(p, k)$  instructions, meaning that the next  $k$  points will be at a distance of  $p$  times the uniform sampling distance. The value of  $p$  can be negative, meaning that the focal point will travel in the opposite direction. Potentially unspecified points will be dispatched so that the remaining space is uniformly sampled. A `StepSampling` can be applied on top of this by setting up the `skip_step` parameter.

`RandomSampling` corresponds to a random sampling of  $n$  points. It is mainly intended to be used as a reference to compare with the performances of other strategies.

`CustomTimeStepSampling` is similar to `CustomSampling` except that the control focuses on temporal sampling. It is also parameterized by a series of instructions of the form “ $p\ k$ ”, meaning that the focal point will spend  $p$  time steps on each of the next  $k$  points, leading to a dwell time of  $T = \frac{p}{s}$  with  $s$  [Hz] being the update rate. While dwelling on one point, the amplitude

is modulated following a sine curve with configurable frequency. An AM frequency can also be set, allowing the reproduction of tactile points, where the user feels the tactile sensation slowly moving along the shape [Hajas et al. 2020]

Each of the spatial sampling strategies can be complemented with additional AM information. Currently, two methods are available. The first one sets the intensity of each point to 100%, corresponding to a classical STM approach, while the other one follows a sinusoidal variation with configurable frequency, which can be useful for displaying dynamic stimuli. The tool can also be used to make pure AM stimuli by setting a long dwell time on each point along with an AM overlay.

## Import and Export

The previous two components presented how stimuli are internally represented. Users may want to use such stimuli for other purposes unrelated to the design tool, such as experiments. This requires a way to export stimuli to a convenient file format. It is also important to provide the possibility of importing stimuli to enable design over several sessions or the reuse of pre-made shapes or strategies.

The design tool can independently create geometry or sampling strategy files. These are custom comma-separated values (CSV) files, starting with the name of the geometry or strategy, followed by the different parameters required to recreate the object. A second line serves as a comment letting users know what each element represents, ensuring external compatibility of the file format.

Users can also store a complete stimulus, corresponding to a combination of shape and strategy. In this case, two files are created. The first one is a concatenation of the shape and strategy files described above, and is intended for reuse in the design tool. The second one is a CSV-like file starting with the number of time steps along with the hardware update rate and the maximum intensity of the device used to create the file, followed by a list of (x y z i) corresponding to the position and intensity of the point at each time step. The position is given in the system coordinate associated with the direct frame of reference centered on the device with the z axis perpendicular to the array plane, pointing outward. This format is voluntarily simple, allowing users to easily create external readers to translate the file and send information to an emitter.

In terms of performance, the export time for a .PTLIST file is proportional to the number of time steps, saving about 60,000 time steps per second, running on a single thread on an Intel

Core i7-10750H CPU. Importing in the design tool is almost instant, as the creation of geometry and sampling strategy objects do not require heavy computations. Since these operations are done in the design tool, the time does not impact any rendering activity.

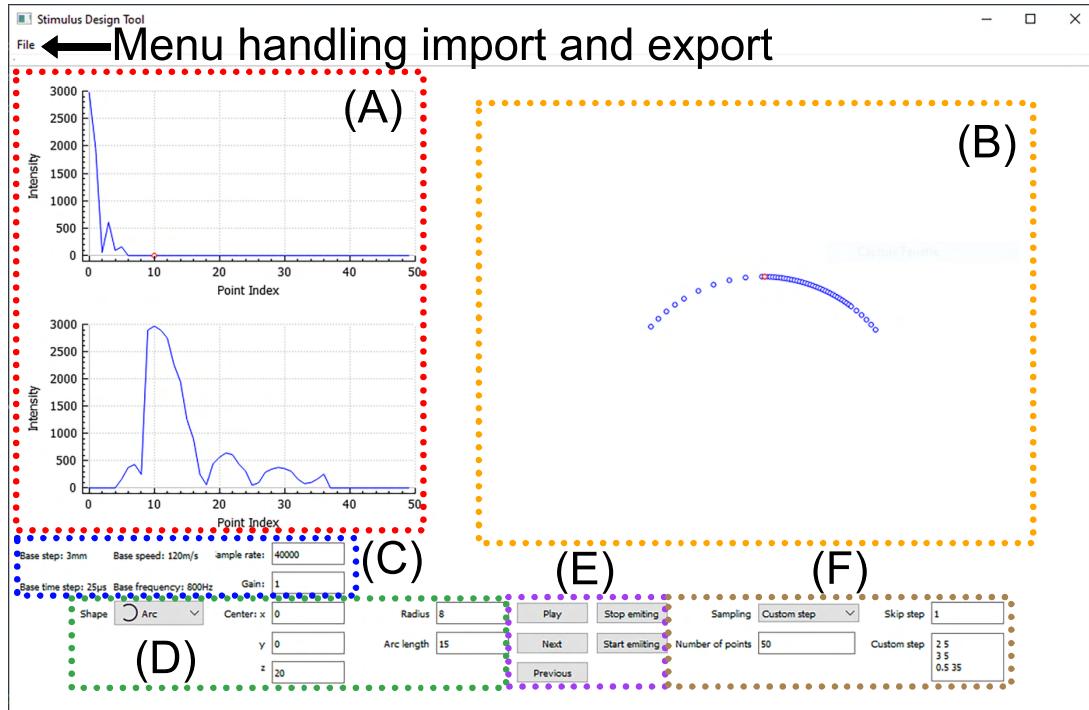


Figure 5.5 – Stimulus Designer User Interface. (A) Graphs of the peak pressure as a function of position indices along the shape at a given time step. The top graph shows the pressure distribution with  $x=0$  corresponding to the current point of emission, the bottom graph shows the same pressure distribution but with  $x=0$  fixed at a reference point on the shape (red dot in the top graph). One can think of it as (top) the pressure at each point at this exact timestep, or (bottom) the time-variation of the pressure at this exact position. (B) Visual representation of the spatial sampling. (C) Basic information concerning the interface used (update rate, peak pressure output). (D) Shape dimensions and position in space (values in cm). (E) Animation and emission. Animation triggers a slowed down visual display of the sampling strategy, moving the red dot in (B) along the path corresponding to the strategy and updating the pressure graphs in (A). (F) Sampling strategy parameters.

## User Interface

The User Interface (see [Figure 5.5](#)) links everything together. On the right is a preview of the spatially sampled shape. On the left side are two graphs corresponding to the pressure distributions as a function of spatial sampling point indices at each timestep. Data for these graphs

was obtained using Ultraleap’s HandyBeam simulation tool<sup>3</sup>. The bottom part allows the user to set stimulus parameters, with the shape parameters on the left and sampling parameters on the right. Preview animation and emission options are located in the middle. The user interface also provides basic information about modulation frequencies, spatial and temporal steps and focal point speed. The top menu provides access to import and export operations.

Given that the software is open-source, users can freely use individual components of the stimulus design tool in third-party applications or scripts.

### 5.3.2 Emitter

As we aim to make the software usable with different UMH interfaces, an abstraction of the actual device is required. Currently, only Ultraleap’s STRATOS Explore interface is supported, as it is the device used during initial development of DOLPHIN. Having a platform-agnostic representation of an emitter is the next logical step toward more universality in the stimulus use and representation. The representation is quite simple as each emitter just needs to implement a start and stop option along with an operation to set a callback and getters for some device characteristics such as the hardware update rate and the maximum achievable pressure. Thus, nothing in the framework is specific to a particular device. We chose to go with a callback approach in order to better separate the different operations. The design tool runs parallel to the callback, which sends the appropriate data to the device displaying the focal points. Any user may add support for a new interface by creating a sub-class of `Emitter` and implementing its methods.

If the emitter’s update rate is different from the one used to create the stimuli, an upsampling or downsampling strategy needs to be applied to adapt the stimuli to the new device while conserving the initial temporal display and modulation characteristics. Resampling presents some limitations that the user must be aware of, thus each saved file contains enough information to recreate an accurate version of the initial stimulus. Users can also include the `ImportExport` class in their program along with the file representation initially made for the design tool to avoid resampling the intensity. This recomputes the intensities for the correct time steps.

---

3. <https://github.com/ultraleap/HandyBeam>

### 5.3.3 Interfacing for Perception Evaluation

DOLPHIN provides an interface for perception evaluation tools such as PsychoPy [Peirce et al. 2019] thanks to the stimulus file format. PsychoPy is a tool for conducting studies on perception which lets the experiment designer trigger audio, visual and haptic stimulus display hardware, deal with precise timing, ask the user custom questions and record responses and multiple objective metrics. While DOLPHIN is not a multimodal tool, it can be used in the study of multimodal perception as Psychopy can trigger and synchronize different types of stimuli. The tool automatically generates data files logging responses, information regarding stimuli and experiment progress for subsequent analyses.

A user designing an experiment begins by designing the individual stimuli required using DOLPHIN’s stimulus design tool. PsychoPy is then provided with the paths to the exported stimuli, allowing it to trigger the appropriate stimulus during an experiment. To do so, PsychoPy executes the Reader component of DOLPHIN with the appropriate file path. The reader acts as a sort of intermediate server, that in turn uses the emitter library and optionally libraries from the design tool to recreate the stimulus and send it to the haptic interface.

In terms of performances, loading a .PTLIST file with the reader program took about 11 ms per 1000 time steps on an Intel Core i7-10750H CPU. While this does not affect stimulus display quality or interactivity of the program, it can be an issue for stimuli with large numbers of time steps required to be played in sequence without pauses between them. A workaround for this issue is to load all stimuli to be displayed in sequence prior to the experiment. Whenever possible, the user should use the *ImportExport* class to import a .GEOMSAMP file instead of parsing the .PTLIST file, as it only takes about 0.3 ms to recreate the sampling objects and generate the same 1000 time steps on the same hardware.

## 5.4 Proof of Concept

In order to validate DOLPHIN’s usefulness, we used it to design and conduct a user study investigating the effect of sampling strategy parameters on curvature discrimination for dynamic tactile pointers [Hajas et al. 2020].

Stimuli were generated with DOLPHIN’s design tool, then fed into the experiment designed with PsychoPy [Peirce et al. 2019], which in turn used the framework’s reader component. All stimulus files and experiment scripts are available in DOLPHIN’s repository<sup>4</sup> to allow

---

4. <https://gitlab.com/h-reality/dolphin>

replication of this experiment.

### 5.4.1 Materials and Methods

This study aims at measuring the effect of sampling strategy parameter (especially the number of sampled positions  $n$ ) on human curvature discrimination performances.

#### Stimuli

The stimuli were dynamic tactile pointers [Hajas et al. 2020] displaying arcs of four possible curvature radii on the participant's hand for a fixed duration of two seconds. The arc length was set to the participant's hand width for all conditions. This ensures that the stimulus reaches similar parts of the hand for all participants, regardless of their hand size.

The stimuli were designed using a `CustomTimeStepSampling` strategy. The four different ratios of curvature radii to hand width were equal to  $\frac{1}{\pi}$  (semi-circle),  $\frac{3}{2\pi}$ ,  $\frac{5}{2\pi}$ ,  $\infty$  (straight line). For the analyses, the radius to hand width ratio for the straight line was set to 1.99 which is the smallest ratio ensuring the height difference at the edge of the hand remains smaller than the radius of a focal point. We evaluated five sampling strategies with  $n = 10, 50, 200, 400, \text{ and } 800$  points respectively, resulting in dwell times on each point of  $t = \frac{2}{n}$  seconds. This corresponds to a single rule  $\frac{80000}{n}$  as the Ultraleap STRATOS Explore's hardware update rate is 40 kHz. The amplitude modulation frequency was set to 200 Hz as in [Hajas et al. 2020]. Each stimulus was displayed 10 cm above the array, and the tactile pointer always moved from the thumb side towards the pinkie.

DOLPHIN allowed us to generate the required stimuli with a simple script, and a similar result could be achieved without any programming by manually exporting from the design tool. All the generated .PTLIST files were then sent to the reader program of DOLPHIN by PsychoPy as needed.

#### Apparatus and Setup

**Figure 5.6** shows the experimental setup. Users positioned their hand horizontally over the box containing the array, parallel to the device plane, with the palm centered on the stimulus presentation location. The users wore noise-cancelling headphones playing pink noise to hide the device's operating noise.

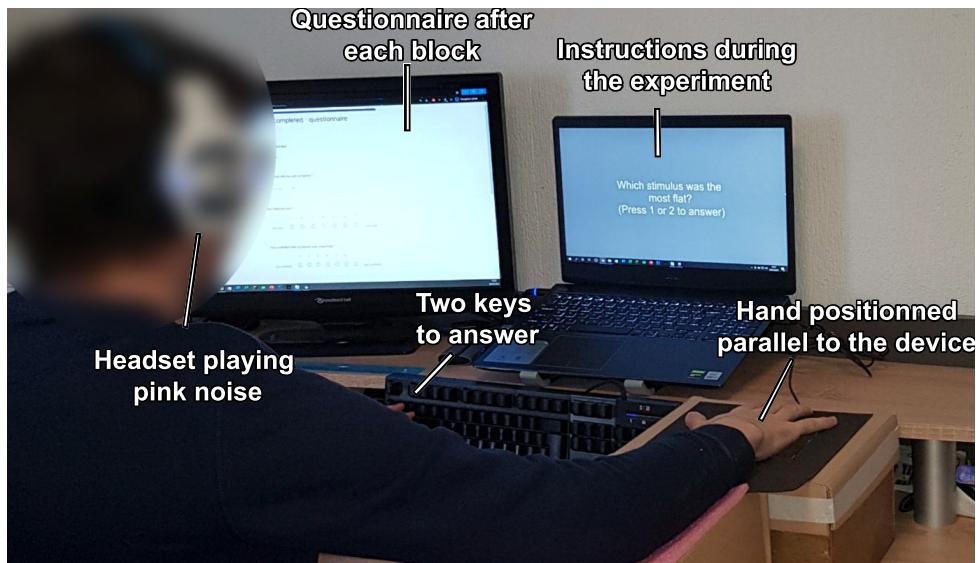


Figure 5.6 – Experiment setup: Subjects sit facing the computer with their dominant hand centered 10 cm above the array, using their other hand to answer the experiment questions. A headset plays pink noise to cover device noise.

### Procedure

Two test stimuli were first presented, to allow the user to get used to the feel of the tactile stimuli and enable them to correctly center their hand position.

The experiment was then divided into five blocks corresponding to the five studied sampling strategies. Each of the five blocks followed a two-intervals forced choice (2-IFC) protocol [Jones et al. 2013], where pairs of stimuli with different curvatures were presented with a 1.5 s break between them, after which users had to indicate which felt flattest (i.e. closest to a straight line). Users responded with their non-dominant hand, using two keys mapped to the two stimuli. Stimulus pairs were repeated three times per block (accounting for order, each pair was thus presented 6 times), yielding thirty-six trials per block. The order of the pairs within blocks as well as the order of blocks were randomized using a latin square.

### Participants

19 volunteers (2 women, 16 men, 1 non-binary, aged  $m \pm s.d. = 23 \pm 3.4$ ), all right-handed participated in the experiment. For each  $(n, r_1, r_2)$  triplet, we then computed each participant's correct response rate.

## 5.4.2 Results and Discussion

### Data Processing

For each of the 19 participants, we gathered one response per trial, leading to 3420 trials ( $19 \text{ participants} \times 5 \text{ sampling strategies} \times (4 \times 3) \text{ radii pairs} \times 3 \text{ repetitions}$ ). This allowed us to plot the proportion of 'stimulus is flatter than the reference' answers against the relative difference in curvature for each of the four references, as shown in [Figure 5.7\(A\)](#). For a reference stimulus with radius  $r_1$  and test stimulus with radius  $r_2$ , the x value corresponding to the pair was  $\frac{c_2 - c_1}{c_1}$ , where  $c_i = \frac{1}{r_i}$  corresponds to the curvature of an arc with radius  $r_i$ .

We added the hypothetical point at  $(0, 0.5)$  corresponding to a fully random answer when both stimuli are identical. For each participant, we then fitted the set of observations with a cumulative gaussian centered on 0 and used the curves to calculate the 75% JND estimate in curvature for this (user, strategy, reference curvature) triplet. The obtained JND is thus expressed as the Weber fraction in curvature. Out of the 360 JNDs (19 participants, 5  $n$ , 4 reference  $r$ ), 56 could not be computed due to convergence issues.

31 outlier JND values (values greater than  $Q_{75\%} + 1.5IQR$  where  $Q_{75\%}$  is the 75% quartile and  $IQR$  is the interquartile) were also removed.

For each (participant,  $n$ ), we then computed the average JND across all reference curvatures. Four participants had at least one strategy for which no JND could be calculated (or were all outliers), and were therefore removed from the dataset. This lead to 15 JND values for each strategy, as shown in [Figure 5.7\(B\)](#).

We also applied the same process, but averaging for all reference curvatures across sample rates instead. This time, only a single participant had to be removed, leading to 17 JNDs per radius.

### Analysis

The JND distribution of the  $n = 200$  group did not follow a normal distribution (Shapiro-Wilk test,  $p = 0.002$ ). We therefore performed a Friedman test, using the JND as dependent variable, the sample rate as independent variable, and the participant as random effect variable, which showed no significant effect ( $\chi^2(2, N = 15) = 2.88, p > 0.05$ ). We conclude that when designing dynamic tactile pointers, the number of points used for the sampling strategy does not have a significant impact on users' ability to discriminate arc curvatures. The mean JNDs obtained after data processing were 1.38, 1.36, 1.34, 1.20 and 1.35 for the strategies with

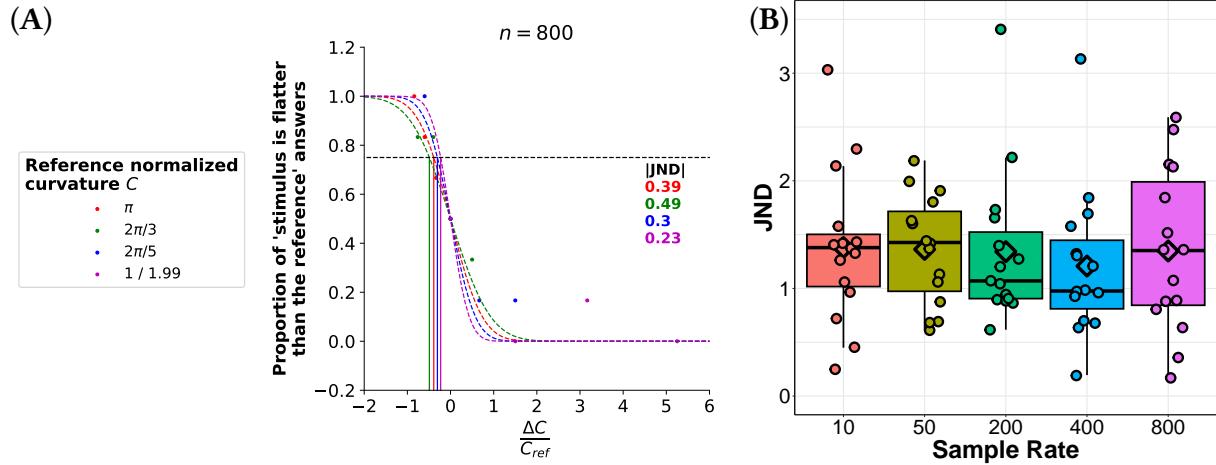


Figure 5.7 – Experiment results. (A) Cumulative Gaussian curves fit to the results for **one user** for the  $n = 800$  points strategy, with one curve fit per reference stimulus. The proportions of "stimulus flatter than reference" answers from the user are plotted against the relative difference in curvature. JND estimates for this user and strategy are reported on the right. (B) Box plot of the mean JNDs obtained for each strategy. Each point corresponds to the average JND across all reference radii for a single participant. Diamonds correspond to the average value for the corresponding group.

$n = 10, 50, 200, 400$ , and  $800$  points respectively. These results mean that designers have more freedom when designing such stimuli, as the strategy will likely not adversely impact stimulus discrimination.

The JND distribution of the  $r = \frac{1}{\pi}$  group did not follow a normal distribution (Shapiro-Wilk test,  $p = 0.005$ ). As before, we followed a Friedman test, but using the reference radius as independent variable, resulting in rejection of  $H_0$  ( $\chi^2(3, N = 18) = 27.5, p < 0.001$ ). Post-hoc Wilcoxon signed-rank tests with Holm corrections were then performed, highlighting significant differences between the  $r = (\frac{1}{\pi}, \frac{5}{2\pi}), (\frac{1}{\pi}, 1.99)$ , and  $(\frac{3}{2\pi}, 1.99)$  groups ( $p < 0.001$  for all three comparisons). The mean JNDs obtained after data processing were  $0.74, 1.09, 1.66, 2.24$  for the reference curvature radii  $r = \frac{1}{\pi}, \frac{3}{2\pi}, \frac{5}{2\pi}$ , and  $1.99$  respectively. Interestingly, this shows that the Webber fractions are not constant when the reference curvature is varied. This can be explained by the "non-linearity" of the curvature metric. Indeed, when increasing the curvature radius, the obtained arc converges very quickly towards the flat line. This leads to the Webber fraction increasing with the curvature radius, as illustrated in [Figure 5.8](#)

While this method is effective to use statistics to measure the effect of different conditions, having to calculate a JND for each condition, for each participant, is prone to generate a lot

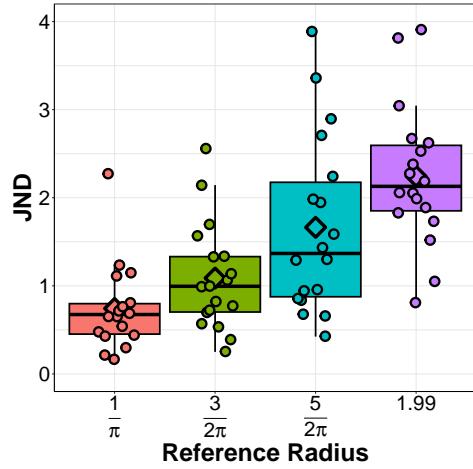


Figure 5.8 – Obtained JND distribution for each normalized reference radius. Each point corresponds to the average JND for a single participant, while the diamond represents the average across all participants.

of outliers, thus biasing the obtained JNDs. If we are more interested in the JND values, a more robust approach consists in measuring the correct response rate across the results of all participants together, before applying the same curve-fitting method. The larger amount of data behind each obtained point reduces deviation, thus allowing a good fit for all conditions. [Figure 5.9](#) presents the resulting global fit, and the resulting Webber fractions are summarized in [Table 5.1](#), where we can easily observe the absence of significant effect from the sample rate, and positive effect of the curvature radius.

Table 5.1 – Global curvature JNDs (expressed as Webber fractions) for each reference stimulus, obtained from the more robust method on globally averaged data.

$r \setminus n$	10	50	200	400	800	mean
$1/\pi$	0.67	0.69	0.72	0.79	0.73	0.72
$3/2\pi$	0.85	0.94	0.74	1.19	1.21	0.99
$5/2\pi$	1.80	1.47	1.73	1.79	1.97	1.75
1.99	3.46	3.18	2.73	4.16	4.84	3.67
mean	1.70	1.57	1.48	1.98	2.194	1.78

Overall, we found no effect of the number of sample points on the perception of curvature, which provides more freedom for UMH designers. Measured Webbed fractions are relatively high, revealing difficulties for humans to discriminate between curvatures of UMH stimuli.

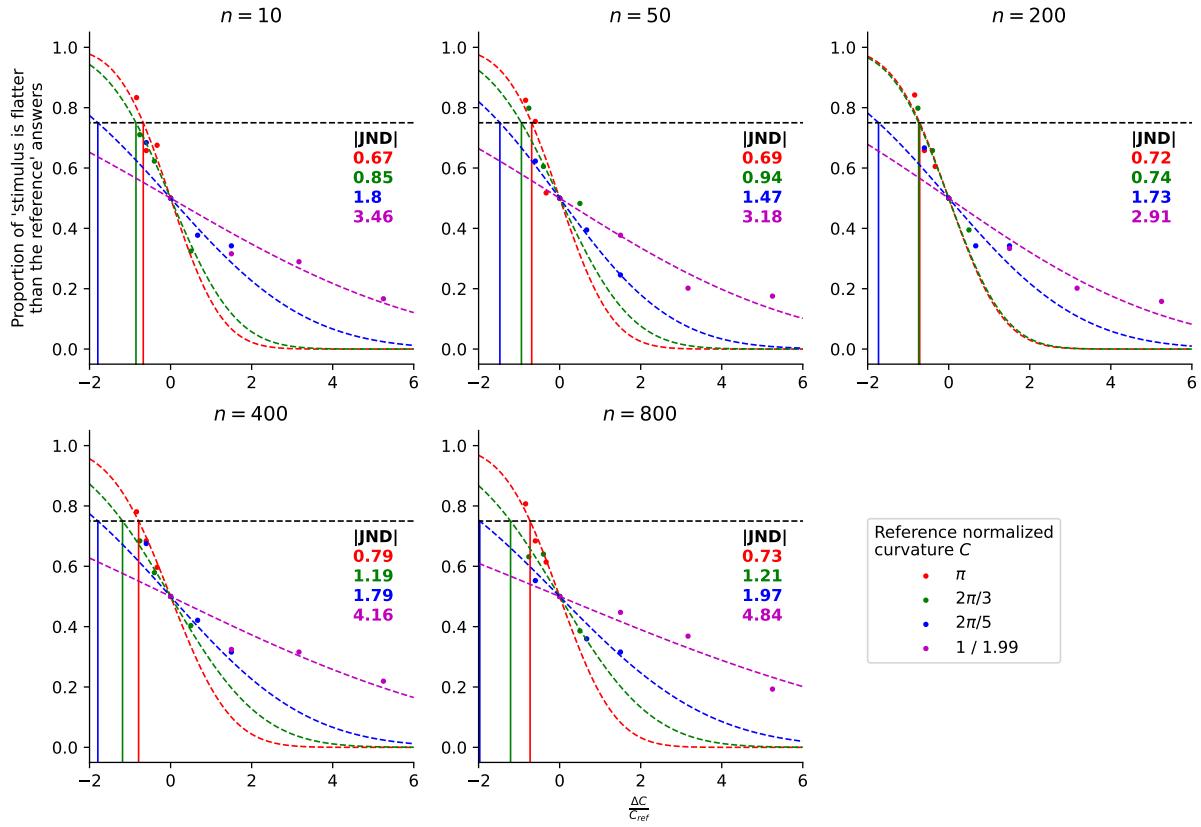


Figure 5.9 – Similar process to Figure 5.7, but for data averaged across all participants. Obtained JNDs (expressed as Webber fractions) are reported on the right.

## 5.5 Interoperability

While our work is built around the Ultraleap STRATOS Explore interface because of its commercial availability (see Section 5.3.2), DOLPHIN is designed to be platform-agnostic. We aim to make it compatible with other arrays (e.g. Ultraino [Marzo et al. 2018], or the AUTD [Iwamoto et al. 2008b]).

While the geometries, user interface and sampling strategies are not inherently dependent on the device, the emitter object is. The export format also has to contain enough data so that any stimulus can be recreated accurately enough using any device. There is no strict assurance that the stimulus will feel exactly the same on different platforms since differences in computation of the acoustic field may introduce perceptual discrepancies, but DOLPHIN aims to provide the highest reproduction fidelity possible given this limitation using the information from the exported files and possible resampling. Any device owner can create a sub-instance of the

Emitter class to make their device compatible with DOLPHIN. The user then implements the callback function which runs parallel to the design tool, sending appropriate data to the device to generate the focal points. All the required information can be recovered from the shape and sampling strategy, regardless of whether the phases and amplitude for each transducer are computed by the device upon receiving the data or beforehand, within the callback function, and independently of the computation model used for this. This makes the emitter model essentially platform-agnostic.

The exported file formats are made platform-agnostic by adding the necessary information to either recreate the shape or strategy object, or to resample the stimulus accordingly while warning the user about potential differences in the stimuli if the hardware characteristics do not match. The coordinate system used by each device may also differ from the one used by the framework. The emitter class thus implements a coordinate change. An inherent limitation to interoperability comes from the fact that different devices can have different sized workspaces, meaning that devices with a smaller workspace may not be capable of displaying stimuli designed on devices with a larger workspace.

## 5.6 Conclusion

In this chapter, we presented a new open-source framework called DOLPHIN, enabling easy mid-air ultrasound tactile stimulus design for rendering and experimental evaluation. In order to highlight our framework’s strengths, we use DOLPHIN’s to design and generate tactile pointer stimuli for a psychophysics study, investigating the effect of sampling parameters on the tactile perception of curvature. We observed no significant effect of rendering parameters on discrimination performances, thus providing a lot of freedom to UMH designers. We hope this framework will ease research on mid-air haptic stimuli and improve study replicability.

While we chose to approach the rendering issue by prioritizing full control over the rendering pipeline, another tool was proposed two years after our initial publication, providing a very good design tool, but more targeted towards high-level designers [Seifi et al. 2023]. This means that there are now two great tools available, depending on the designer’s specific needs and approach.

While this chapter formalized 2D sampling strategy, the following chapter extends this work to 3D object rendering.

### Chapter Summary

- We provided a formalization of **sampling strategies**, representing the process of turning a 2D shape into a UMH stimulus;
- We presented DOLPHIN, an open-source framework to design and evaluate UMH sensations;
- We conducted a psychophysics study to show that the number of sampled positions on an arc circle does not significantly impact curvature discrimination performances.

# Chapter 6

## 3D Rendering: Intersection Strategies

### Contents

---

6.1	Introduction . . . . .	141
6.2	Intersection Strategies: Definition . . . . .	143
6.3	Using Intersection Strategies to Generate Haptic Rendering Schemes	147
6.4	Pilot Study: Impact of Haptic Rendering Schemes on Perceived Shape Properties . . . . .	151
6.5	User Study: Modifying the Perception of Infill Density . . . . .	155
6.6	Conclusion . . . . .	159

---

### 6.1 Introduction

In [Chapter 5](#) we formalized **sampling strategies**, which represent the UMH rendering process for 2D shapes, and more generally curves. Yet expanding to the rendering of 3D objects is not straightforward, due to the larger areas to render (see [Section 1.4.3](#)). Therefore, standard techniques usually attempt to reduce the rendering problem to a 2D curve, by only rendering a representation of the hand-object intersection, as illustrated in [Figure 6.1](#). This chapter thus focuses on this intersection aspect, to provide a better understanding of how it works, and how it can affect the perception of virtual 3D objects.

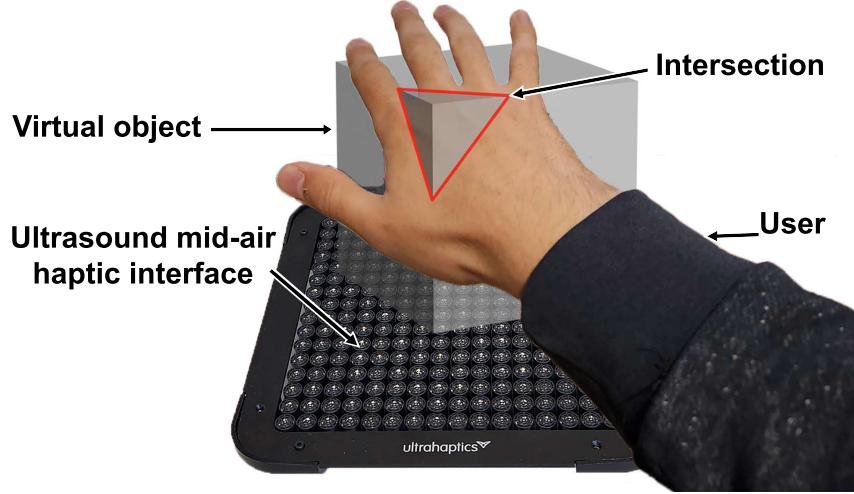


Figure 6.1 – Exploration of a virtual 3D object rendered with an ultrasound mid-air haptic interface. When representing 3D objects with mid-air technologies, we generally render their intersection with the user’s hand.

In this chapter, we propose the following contributions:

- We formalize **intersection strategies**, describing the process of computing intersection representations based on hand and object information, regardless of the hand-tracking methods used (Section 6.2). Combined with **sampling strategies** (Chapter 5), this creates an end-to-end pipeline for designing 3D object rendering algorithms. We discuss implications for rendering performance (e.g. computation time and operating noise) and perception of virtual 3D objects (Section 6.3).
- We illustrate the importance of studying intersection strategies, researching whether differences between strategies influence the perception of virtual object properties. To that end, we discuss a state-of-the-art 3D rendering technique through the prism of intersection strategies. We also designed an example intersection strategy, following the intersection strategy pipeline (Section 6.3). Based on user’s feedback (Section 6.4), we evaluated through a user study the impact of the state-of-the art strategy and our new example strategy on users’ perception of the infill density of virtual 3D objects (Section 6.5). The results demonstrate that well-designed intersection strategies can render challenging properties such as infill density.

Our work aims both to provide tools for standardizing the algorithmic aspects of rendering 3D objects with UMH, and to experimentally demonstrate that intersection strategy design can have a significant impact on the perception of virtual object properties. By normalising reporting and studying methods, we hope to facilitate further exploration of perceptual effects

related to 3D object rendering using UMH within the haptic community, but also the virtual reality and HCI communities.

This content of this chapter was adapted from the following publication: [[Mulot et al. 2024](#)].

## 6.2 Intersection Strategies: Definition

A dynamic UMH rendering scheme for 3D objects requires two inputs at all times: the object to render, and the configuration of the hand exploring it. Based on these inputs, we propose a generalized two-part process: an **intersection strategy** first interprets the hand-object intersection as a contour, then a **sampling strategy** [[Frier et al. 2019](#); [Mulot et al. 2021b](#)] (see [Chapter 5](#)) uses the result to generate the spatiotemporal evolution of a focal point, defining an UMH stimulus. In this section, we formally define the representation of inputs and outputs ([Section 6.2.1](#)) and the generation and parameterization of an intersection strategy ([Section 6.2.2](#)).

### 6.2.1 Component Representations

The intersection strategy algorithm requires three components to be represented: the virtual object (input), the hand (input), and the resulting intersection (output).

The virtual object and hand are both virtual 3D objects, for which many geometrical and numerical representations, either implicit or explicit, exist [[Lin et al. 1998](#)]. Implicit representations often offer easy geometric collision computation, while explicit representations may be more convenient to generate and more intuitive to work with. But geometric collisions and intersections can be computed between any two representations, with state-of-the-art computer graphics algorithms [[Lin et al. 1998](#)], meaning that intersection strategies are agnostic regarding this aspect. If needed, classical object representations can also be augmented to contain additional information that could be useful to the rendering (e.g. local texture or physics-based representation).

It should be noted that regardless of the chosen data structure, it can often be computationally efficient to separate the visual representation (potentially high-resolution) from a collocated simplified representation used for intersection processing. For the object, these simplifications can be made on the basis of rendering resolution, e.g. there is no need to represent object details smaller than the size of a focal point [[Rakkolainen et al. 2021](#)]. This can serve as a guide to define an approximate implicit surface representation of an object, to choose voxel size or to

set the level of mesh detail, for instance. For the hand, simplifications can be made based on the required areas of interaction. While some applications have used mesh representations of the hand [Long et al. 2014; Maunsbach et al. 2024], it is often more efficient and just as satisfactory to represent the hand through simple implicit surfaces, e.g. a plane [Mulot et al. 2023b] or set of planes [Martinez et al. 2019; Mulot et al. 2025]. This is because UMH interaction mostly focuses on the glabrous skin of the palm and fingers, with interactions happening mostly with more open hands.

The representation choices will impact computational load, haptic rendering detail, as well as the resulting representation of the intersection. Regarding the latter, any resulting intersection representations need to be converted to a 3D curve to be usable as an input for the downstream sampling strategy (see [Chapter 5](#)). This is commonly done using sets of polylines, i.e., continuous sets of line segments connected at their extremities [Martinez et al. 2019]. These are particularly convenient when working with meshes, as they provide an exact representation of the intersection.

### 6.2.2 Definition of Intersection Strategies

We define the **intersection strategy** as the process of computing a contour which semantically represents the intersection between a hand and a 3D virtual object. This process follows three steps (see [Figure 6.2](#)):

- The **raw intersection** computes the collision between the hand and virtual object.
- Then, the **pruning** removes unwanted parts of the pattern.
- Finally, the **transformation** step modifies the obtained contours.

These steps have been carefully chosen to facilitate the discussion of state-of-the-art 3D rendering techniques under the viewpoint of **intersection strategies**, while still allowing for the representation of any type of intersection technique.

Once the resulting contour has been computed using the intersection strategy, a **sampling strategy** ([Chapter 5](#)) can be applied to generate the sequence of focal points that will be rendered by the UMH interface. As sampling strategy's parameters can impact the user's haptic perception of a 2D stimulus (see [Part I](#)), we hypothesize that the parameters of the intersection strategy may also impact the user's perception of 3D objects, as well as other meta-parameters (e.g. computation time, noise produced).

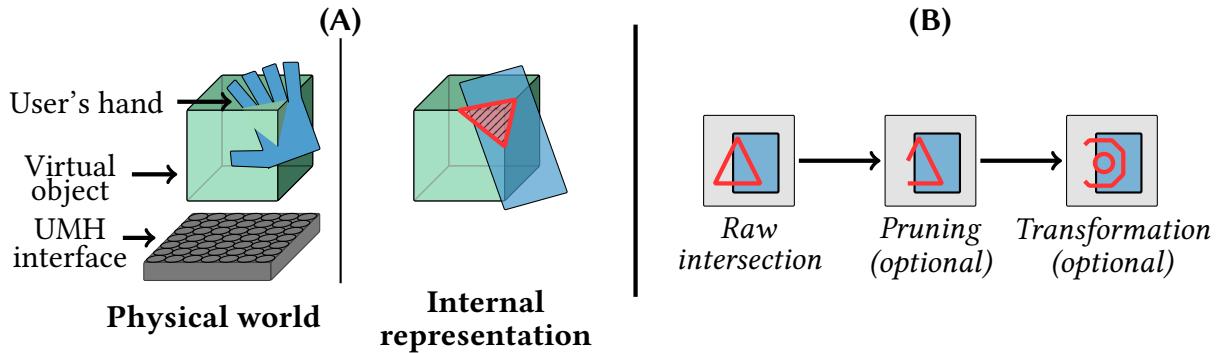


Figure 6.2 – Overview of UMH rendering of a 3D Object. (A) Raw intersection (red) between a virtual object (green cube) and the hand’s internal representation. (B) The three steps for applying an intersection strategy, given a virtual object and hand. First, we compute the outline of the intersection between the virtual object and the hand. Then, we optionally prune parts of the obtained contour, if they are not required. Finally, we optionally apply a transformation to the contour.

### Raw Intersection

In the case where hand and object are represented either both as implicit functions or both as voxels, evaluating the Boolean intersection is relatively straightforward. In other cases, either objects can be converted into convenient representations (e.g. voxelized), or a specific intersection algorithm must be selected based on the nature of the component representations. For example, considering palm and fingers as bounded planes in 3D space and virtual objects as meshes, a mesh-plane intersection algorithm (e.g. [Minetto et al. 2017]) can be used to produce a set of curves lying within the planes of the glabrous skin of the hand. If instead the hand and object representations generate a surfacic intersection, it is possible to turn it into one or several polylines, e.g. by following the algorithm from Maunsbach et al. [2024]

This raw intersection step (Figure 6.2(B-Left)) is mandatory, as it extracts the geometric features of the hand-object intersection. If the data structure for the object allows it, this step could also extract additional information (e.g. texture, feature) that could be used for the next intersection strategy steps, or downstream for the sampling strategy.

### Pruning

The computed raw intersection may contain information that is not relevant when aiming for efficient rendering. An optional contour pruning step can therefore be useful to reduce the handled intersection contours down to the bare minimum required for efficient rendering and perception. For example, parts of the raw intersection may end up lying at a distance from the

skin (e.g. the circle in [Figure 6.3](#)), where even if they were rendered, they would not be felt by the user. These parts could thus be removed. Also, as proposed by [Martinez et al. \[2019\]](#), and implemented by [Mulot et al. \[2025\]](#) (see [Chapter 9](#)), we could remove segments that are too close to each other (e.g. when intersecting an object with small width), if users would not be able to distinguish them.

It should be noted that applying this step may transform continuous contours into discontinuous open patterns (see [Figure 6.2\(B-Center\)](#)) which in turn could lead to discontinuities in the focal point motion. This is a well-known source of operating noise for UMH interfaces [[Georgiou et al. 2022a](#)].

## Transformation

Finally, the raw or pruned contour may not be the most effective way of conveying the properties of an explored object. Haptic designers may want to transform the contour obtained after the previous two phases while conserving some of the obtained geometric information. The aim of this may be to use specific patterns (e.g. circles) whose perception is well-studied [[Frier et al. 2019](#)], to take advantage of interference effects [[Reardon et al. 2023](#)], or to highlight certain object features [[Martinez et al. 2019](#)]. Thus, a final optional transformation step completes the general definition of the intersection strategy ([Figure 6.2\(B-Right\)](#)).

## Generalizability

We formalized the intersection strategy rendering pipeline in an effort to provide a baseline for future discussion about 3D object rendering with UMH. The underlying steps were also chosen to fit previous work in the field. This shows that the steps are well chosen, and ensures easy comparisons of rendering techniques. For instance, the 3D aspect in the work of [Long et al. \[2014\]](#), and [Matsabayashi et al. \[2019\]](#) was mostly centered on the raw intersection, while the proposed techniques by [Martinez et al. \[2019\]](#) could be interpreted as a basic raw intersection, combined with different transformation steps.

Furthermore, given that collisions can be computed between any two object representations, and since the last step can represent any arbitrary transformation, the generalizability of this rendering pipeline is ensured.

Overall, in this section we formalised **intersection strategies**, the first part prior to the **sampling strategy** when rendering a 3D object with UMH. **Intersection strategies** are the succession of a **raw intersection**, **pruning**, and **transformation** step, ensuring its generaliz-

ability. In doing so, we enable easier discussions and exploration of 3D object rendering with UMH, and deeper exploration of their perceptual applications.

## 6.3 Using Intersection Strategies to Generate Haptic Rendering Schemes

To illustrate the intersection strategy process, we designed examples of haptic rendering schemes using the previously described formalization, to explore whether differences between strategies affect the perception of virtual object properties (Section 6.4, Section 6.5). We followed a **user-centered approach**, designing these example strategies and asking participants how they would interpret the haptic feedback difference, in order to find the direction of the following user study.

### 6.3.1 Intersection Strategies

We first consider a basic intersection strategy, which uses only the default raw intersection step and a pruning step that removes contours that are outside the boundaries of the hand. With this strategy, partial intersections of an object by the hand causes the user to feel sets of open contours (see Figure 6.3) which users are free to interpret in any way they wish. This is equivalent to some state-of-the-art rendering techniques for 3D object rendering (e.g. strategy A by Martinez et al. [2019]), and serves here as a comparison baseline.

We then design a derived intersection strategy where we perform an additional transformation step to systematically close intersection contours so that they encapsulate the “inside” of the intersected virtual object (see Figure 6.3). In the remainder of this paper, we refer to the first intersection strategy as **OC** (Open Contour) and the second intersection strategy as **CC** (Closed Contour).

While **OC** is a state-of-the-art baseline, presented here under the intersection strategy pipeline, **CC** is a novel haptic rendering scheme, the purpose of which is to illustrate how different intersection contours may affect a user’s interpretation of virtual object properties (e.g. geometry or fill density).

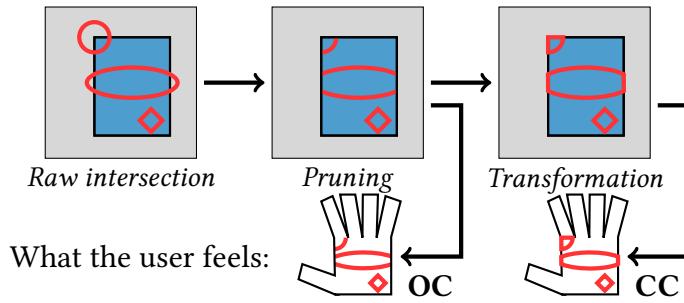


Figure 6.3 – We designed an intersection strategy, following the process of [Figure 6.2](#). After the plane-mesh intersection has been computed (red), parts that are outside of the hand rectangle (blue), are removed during the pruning step. Then, segments are added along the rectangle contour to close the open contours. We call **OC** and **CC** the feedback obtained after the pruning and transformation steps, respectively ([Section 6.2.2](#)).

### 6.3.2 Chosen Component Representations

For these haptic rendering schemes, we represented virtual objects as meshes. We chose to instruct users to explore objects with their palm only. This approach simplifies the hand to a single rectangle covering the palm, which could easily be implicitly represented. This decision was made to reduce the computational cost and ensure a consistent feedback between participant, regardless of the shape of their hand. However, it is important to note that the concepts presented in this paper still apply to other representations, although more work is required to understand how they impact the user’s perception.

Once the final pattern is obtained, we apply a sampling strategy that uses STM with a focal point speed of 7 m/s [[Frier et al. 2018](#); [Frier et al. 2019](#)].

### 6.3.3 Effects on Non-Haptic Factors

By closing the contour, we render some additional segments that are inside the virtual object, which could have several implications beyond a user’s perception of the object. For instance, given the constant-speed sampling strategy we used, rendering these longer patterns means that the **CC** feedback will be rendered at a lower draw frequency than the equivalent **OC** feedback.

#### Computation Time

In general, the use of different algorithms or parameters for the different steps of the intersection strategy may impact the computation time. For **CC** and **OC**, with a naive C# implementa-

tion running on a 12th Gen Intel Core i7-12700H CPU, we measured that the raw intersection step takes about 1 ms per 100 mesh faces, while the pruning step requires about 0.2 ms per 100 mesh faces, and the transformation step for **CC** requires less than 0.1 ms in total. That said, performances could be greatly improved by using a GPU, but we do not expect that a significant difference would appear between the **CC** and **OC** computation time. For reference, a naive plane-mesh intersection can be computed in  $\mathcal{O}(f)$  time, for a mesh with  $f$  faces. This results in a set of  $n$  small segments, each corresponding to one intersected face. The segments can then be reordered into the actual contour in  $\mathcal{O}(n^2)$  time with a naive implementation. Our pruning and transformations steps can also be computed in  $\mathcal{O}(n)$  time. It is worth noting that these complexities represent naive implementations of our specific strategies. Other input structures or rendering objectives may heavily alter the resulting complexities.

## Operating Noise

When an open contour is drawn using STM, the spatial discontinuities of the focal point path generate high-pitch audible harmonics, which may not be pleasant. By closing the contour, the number of spatial discontinuities can be reduced, thus reducing the generated noise. To assess this, we positioned a Shure SM58 microphone 7 cm above and 15 cm to the side of the haptic interface, connected to a Focusrite Scarlett 2i2 sound card, with 96 kHz sample rate. Measurements were conducted on a cube, tube, and dog shape, scaled to fit inside a 13 cm-wide virtual cube, centered 20 cm above the interface. The setup is illustrated in [Figure 6.4](#). We recorded the noise produced by a 30 s exploration of the object with both the **CC** and **OC** feedback. These recordings were repeated ten times each. After that, we measured the differences in peak and RMS sound pressure levels for the **OC** and **CC** recordings. This resulted in **OC** being on average 0.56 dB (RMS) louder than **CC**, while having peaks 0.7 dB lower. That said, these differences were not significantly different from 0 (Shapiro-Wilk,  $p > 0.05$ , one-sample t-test,  $p > 0.05$ ).

We also computed the Fourier spectrum for each recording, as illustrated in [Figure 6.5](#). From these, we observe that the discontinuities in the **OC** feedback lead to the generation of higher frequency harmonics in the 300 Hz - 10 kHz band, explaining the higher perceived pitch for the noise generated by **OC**. This is also coherent with the slightly higher RMS level for **OC**. Since humans perceive frequencies between 300 Hz and 10 kHz as louder than lower frequencies [[Suzuki et al. 2004](#)], the **OC** feedback appears louder and less pleasant [[Patchett 1979](#)].

Since we used a sampling strategy that used a constant focal speed of 7 m/s, the added segments in **CC** increase the total length of the pattern, thus reducing the draw frequency. We

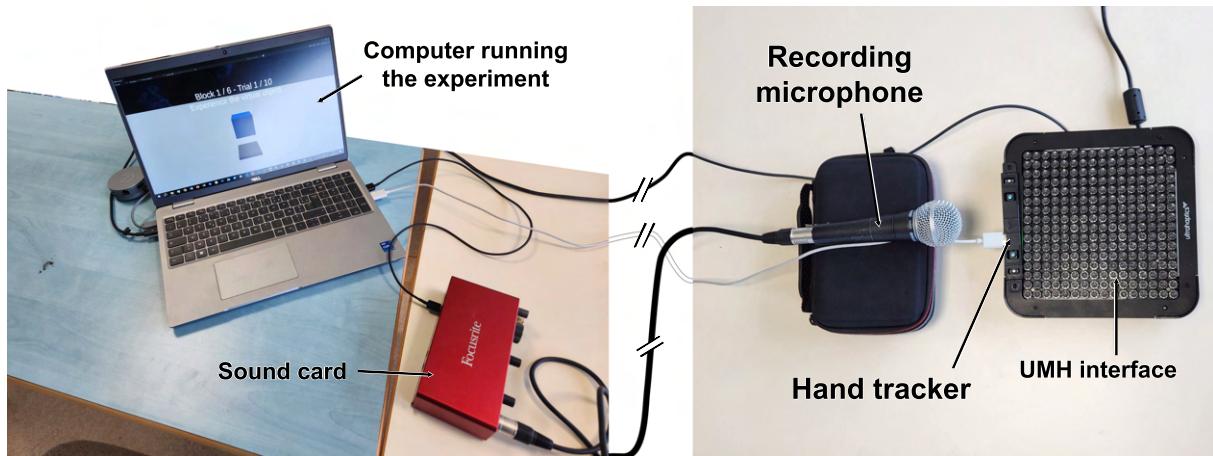


Figure 6.4 – Illustration of the experimental setup for recording the operating noise of UMH interfaces when exploring a virtual object. The explored virtual object is positioned 20 cm above the UMH interface.

thus performed the same measurements, but using a sampling strategy with a constant draw frequency of 100 Hz. Thus, longer CC pattern thus had a higher focal speed to compensate. The resulting spectra are shown in [Figure 6.6](#). We can see that the range of higher sound level for OC is narrowed down around the 3 kHz range.

In both cases, the difference between OC and CC fades out as the experience shape increases in complexity and irregularity. Indeed, the obtained pattern after OC's pruning is likely to be a single polyline when intersecting with the cube. CC thus closes the contour, leaving no discontinuities. But more complex contours are more likely to generate non-connected patterns, meaning that the focal point still has to jump between the different connected components, whether they are closed or not.

In the end, we showed that even regardless of the tactile implications, carefully designed UMH intersection strategies can significantly alter the quality of the user experience. These noise concerns have to be particularly taken into account, so as to avoid any safety or exposure risk [[Battista 2022](#); [Hoshi 2022b](#); [Morzyński et al. 2024](#)].

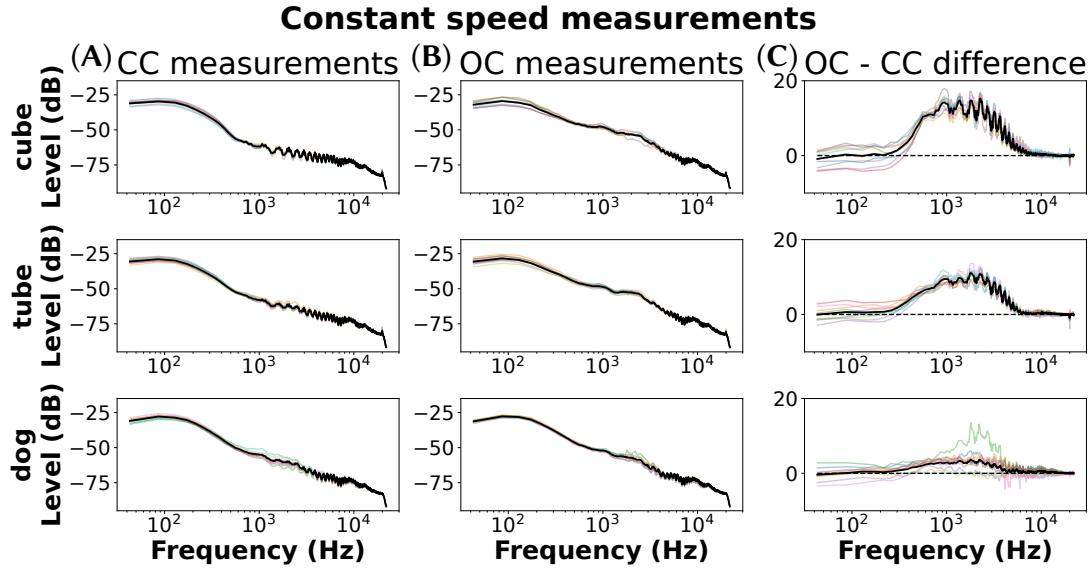


Figure 6.5 – (A) (resp. (B)) Fourier transforms of the audio recordings during an interaction with **CC** (resp. **OC**) feedback, with a fixed focal speed sampling strategy. (C) Difference between the (A) and (B) spectra: **OC** generates more energy than **CC** in the 300 - 10 kHz band and less energy below 300 Hz, corresponding to a mid and high pitch range. The three rows represent measurements when experiencing the cube, tube, and dog shapes respectively. Semi-transparent colored curves represent individual measurements, while the dark opaque line represents the average.

## 6.4 Pilot Study: Impact of Haptic Rendering Schemes on Perceived Shape Properties

To highlight the importance of studying intersection strategies, we used the previously described **OC** and **CC** rendering schemes in a pilot study to initiate the exploration of the perception and interpretation of these stimuli. It is worth noting that **OC** and **CC** were designed to **illustrate** the intersection strategy pipeline rather than to render specific properties. So this pilot aims to understand whether participants feel a difference between the two rendering schemes, and how they would interpret this difference in terms of object properties. This will then allow us to study the impact of the two feedbacks on these relevant properties ([Section 6.5](#)).

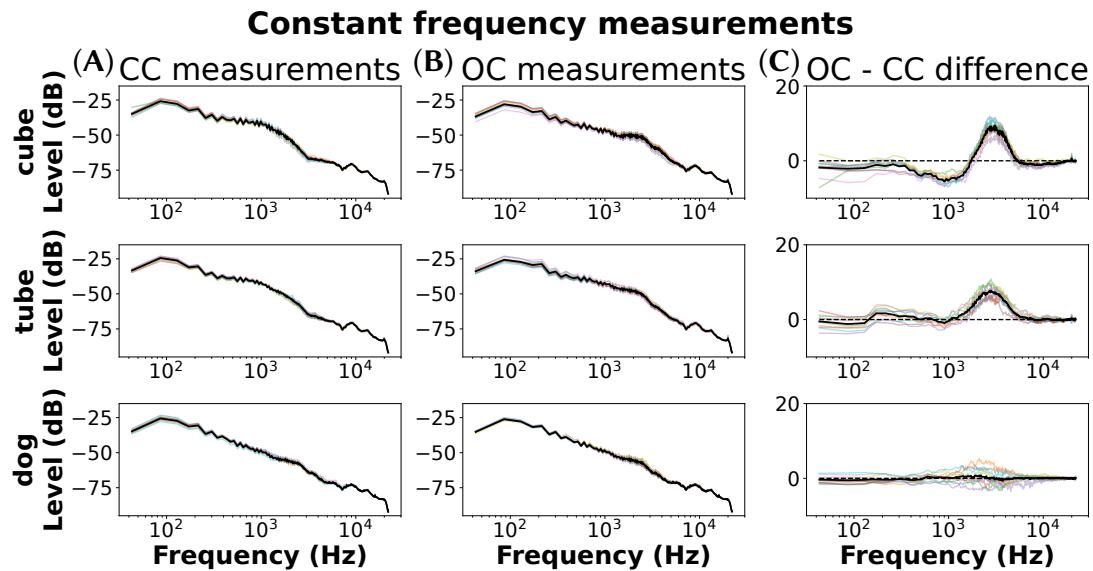


Figure 6.6 – Same measurements as [Figure 6.5](#), but with a constant draw frequency sampling strategy.

#### 6.4.1 Materials and Methods

##### Apparatus and Setup

Participants were seated facing the laptop running the experiment software, and an Ultraleap STRATOS Explore <sup>1</sup> haptic interface was placed next to the laptop on the side of the participants' dominant hand, ([Figure 6.7\(A\)](#)). The participant's dominant hand, with which they explored the virtual objects in the experiment, was tracked using an Ultraleap StereoIR 170 camera <sup>2</sup>. The experiment software was designed and run using Unity <sup>3</sup> 2022.3.9f1.

##### Conditions

This experiment involved two independent variables:

- The **shapes** being experienced: a tube, a sphere, a dog, a cube, or a rock (see [Figure 6.7\(B\)](#)). They were all designed to fit inside a 13 cm wide cube whose centroid was positioned 20 cm above the UMH interface. This ensured that the entirety of the shape was inside the workspace of both the hand tracker and the haptic interface.
- The **intersection strategy** used to render the object: **OC** or **CC**.

1. [https://www.ultraleap.com/datasheets/STRATOS\\_Explore\\_Development\\_Kit\\_datasheet.pdf](https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf)  
 2. <https://www.ultraleap.com/product/stereo-ir-170/>  
 3. <https://unity.com/fr>

#### 6.4. Pilot Study: Impact of Haptic Rendering Schemes on Perceived Shape Properties

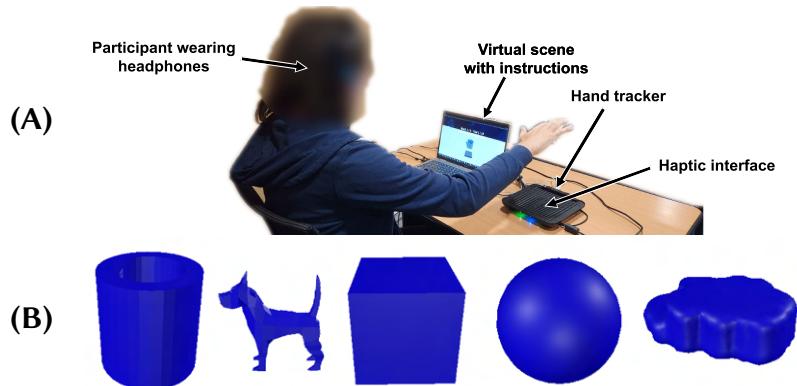


Figure 6.7 – (A) Experimental setup for the pilot study (Section 6.4) and user study (Section 6.5): participants seat facing the screen, interacting with their dominant hand above the UMH interface. They wear headphones to mask exterior noises. (B) Participants explored different 3D objects, namely a tube, a dog, a cube, a sphere, and a rock. The first three were used in both the pilot (Section 6.4) and user study (Section 6.5), while the last two were only used in the pilot.

#### Procedure and Collected Data

This preliminary study comprised 5 trials, for five virtual shapes, corresponding to the 5 shapes, presented in the same order as in Figure 6.7(B). For each trial, the participant performed two successive 20 s free explorations of the virtual object displayed on the laptop's screen, rendered once with OC and once with CC.

After these two explorations, they then gave free answers to the following three open questions:

- 1) Did the haptic stimuli feel different? If so, how?
- 2) Did the virtual objects feel physically different? If so, how?
- 3) Was there a difference in terms of perceived coherence between the tactile sensation and the object's visual aspect?

#### Participants

Six participants (6 men, aged  $m \pm s.d. = 24.2 \pm 2.2$ ) took part in this pilot experiment. Four were right-handed and two were left-handed. One was very experienced with haptics, four reported limited experience with haptics, and one had almost no experience. Five of them had already felt an ultrasound haptic stimulus in the past.

### 6.4.2 Results and Discussion

We observed that two participants were always able to perceive a difference between the two rendering schemes, while two participants failed to perceive a difference once, and two participants failed to perceive a difference twice (out of 5 trials). This seems to reveal a relatively clear apparent perceptual difference between the two rendering schemes, although deeper investigation in a larger and more diverse population would be necessary to evaluate how the different parameters influence the perception of these stimuli. For instance, some participants reported “differences in terms of geometry”, as the **CC** scheme renders segments that are not part of the original shape’s contour, while other participants mentioned differences in terms of perceived “frequency”, as the difference in length between the patterns computed by the two schemes led to different haptic draw frequencies. Some participants also interpreted this change in draw frequency as a textural difference between the two patterns, with **OC** feeling “smoother” than **CC**. This is coherent with the results of [Wojna et al. \[2023\]](#) and [Ablart et al. \[2019\]](#), who showed that texture roughness and draw frequency are negatively correlated.

We also observed differences between shapes. Most participants perceived the biggest feedback differences while intersecting with the cube or the sphere. Since both these shapes have convex geometries with larger inner volumes, the differences in length between both intersection patterns may be more pronounced. This leads to a larger difference in stimulated skin area and in draw frequency. The combination of these factors could explain this observed effect.

When interpreting the haptic differences in terms of physical properties of the virtual objects, the participants mentioned a variety of parameters. Two participants mentioned “geometric differences,” while one participant reported that **OC** gave a sensation of a contour rendering, interpreting the shape as hollow, while **CC** rendered surfaces, leading to the sensation of something filled. They also mentioned an impact on the “solidity,” “density,” and “smoothness” of the object. Overall, the most cited parameter was the density and filling of the object, which was mentioned at least once by every participant. This is coherent with the larger stimulated area from **CC**, which can be interpreted as the presence of more physical matter. In the rest of this presentation, we refer to this property as **infill density**.

The majority of answers from five of the six participants indicated that the **CC** rendering scheme was perceived as more coherent with the virtual aspect of the virtual objects.

Given the motivations for this study and low amount of participants, these results are to be

interpreted as a way to guide our next study, presented in [Section 6.5](#), with hints towards potential perceptual effects that would need further investigation.

## 6.5 User Study: Modifying the Perception of Infill Density

In light of the results of the pilot experiment ([Section 6.4](#)), we designed a user study to further study the impact of the rendering scheme on the perception of an object’s infill density, as it was the most referenced property in the pilot study. We hypothesize that shapes rendered with **CC** will be felt as more densely filled than shapes rendered with **OC**.

### 6.5.1 Materials and Methods

#### Apparatus and Setup

We used a similar setup to that of the pilot experiment ([Section 6.4, Figure 6.7\(A\)](#)), with participants seated facing the experiment laptop and haptic interface. They wore a headset playing pink noise to mask the UMH interface operating noise.

#### Conditions

This experiment involved three independent variables:

- The **shapes** being experienced: a cube, a tube, and a dog. The cube and tube used the same mesh for the visual and haptic representations, composed of 12 and 256 triangles respectively, while the dog used a relatively detailed (1160 triangles) visual representation, and a simplified (136 triangles) collocated haptic representation. They were positioned and scaled as in the pilot study [Section 6.4](#).
- The intersection strategy used to render the object: **OC** or **CC**.
- The visual **opacity** of the virtual object: opaque or semi-transparent ( $\alpha = 70\%$ ). We used this transparency factor as it may be a way to visually suggest the infill density of the object, and its effect may interact with that of the haptic feedback.

#### Procedure

The experiment was divided into six blocks, corresponding to all (shape, opacity) pairs.

Blocks followed a 2-IFC protocol [Jones et al. 2013]: participants explored the object for 7 s with either the **CC** or **OC** rendering scheme, before exploring it a second time with the other one. After that, they answered the following question: “Which one of the two shapes felt the most densely filled?” using the keyboard’s left and right arrow keys. The answer corresponding to each key (“First”/“Second”) randomly changed between trials and was also written on the screen for the participants to see. Five repetitions of each rendering order were presented in a random sequence, yielding 10 trials per block. The block order was randomized using a Latin square.

### Collected Data

For each trial, we collected the participant’s answer. After each block, participants also rated their confidence in their answers and the perceived difficulty of the task on a 7-point Likert scale.

During each of the explorations with **CC**, we also measured the ratio  $r$  between the average length of the patterns computed using the **CC** intersections strategy, and those that would be computed using **OC** during the same exploration. Given that **CC** only adds contours compared to **OC**,  $r$  is always greater than 1. This measure relates to how physically different the two schemes are, with  $r = 1$  indicating identical stimulation, and higher values indicating more significant changes.

### Participants

We recruited 18 participants (15 men, 2 women, 1 non-binary, aged  $m \pm s.d. = 26.9 \pm 5.8$ ), who had not participated in the pilot study. One was left-handed, and all others were right-handed. Four participants were experienced with haptics, while eight participants reported having limited experience, and six had little to no prior experience. Ten participants had never used an UMH device before.

#### 6.5.2 Results

For each participant and each block, we computed the proportion  $p_{CC}$  of answers indicating that the **CC** feedback gave the impression of a more densely filled object. The results for each shape and visual levels are illustrated in [Figure 6.8](#).

As the  $p_{CC}$  data did not follow a normal distribution (Shapiro-Wilk test,  $p < 0.001$ ), we performed a one-sample Wilcoxon signed-rank test, showing that  $p_{CC}$  is significantly higher than

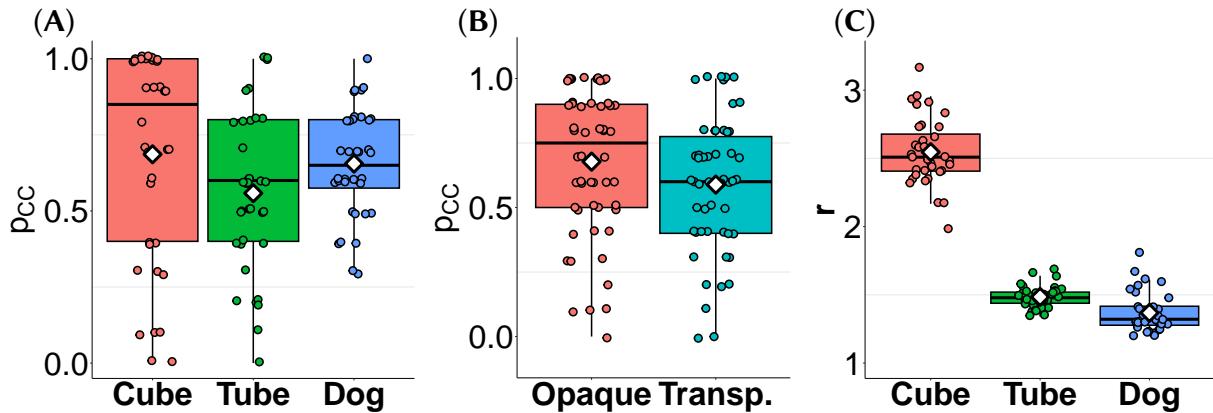


Figure 6.8 – (A) Distribution of the proportion of answers  $p_{CC}$ , reporting that the object rendered with **CC** felt more densely filled than the one rendered with **OC**. (B) Distribution of  $p_{CC}$  per visual representation. (C) Distribution of the ratios  $r$  between the average **CC** pattern length and their **OC** equivalent. Dots represent measurements per participant per block, and diamonds represent the global average.

the 50% chance-level ( $V = 3696.5, p < 0.001$ ). Pairwise Wilcoxon tests with Holm corrections revealed that the cube had significantly higher  $p_{CC}$  scores than the tube ( $V = 378, p = 0.03$ ). A Wilcoxon test also revealed that  $p_{CC}$  was significantly higher for opaque shapes than for transparent ones ( $V = 585, p = 0.007$ ).

Shapiro-Wilk tests revealed that the distribution of  $r$  values for the dog did not follow a normal distribution. We therefore performed Holm-corrected pairwise Wilcoxon tests, revealing significant  $r$  differences between all object pairs ( $p < 0.001$  for all comparisons). Specifically,  $r$  was higher for the cube ( $r = 2.5 \pm 0.2$ ) than for the tube ( $r = 1.5 \pm 0.08$ ), which is itself higher than for the dog ( $r = 1.4 \pm 0.1$ ).

Looking at the individual participant data, 12 participants had at least 5 out of 6  $p_{CC}$  results at 0.5 or higher, showing a strong tendency to report the **CC** feedback as giving the impression of a more densely filled object. 3 participants had opposite results, and 3 participants had ambiguous results revolving around the 0.5 chance level.

Holm-corrected pairwise Wilcoxon tests showed that participants were more confident in their answers for the cube than for the dog ( $V = 340, p = 0.005$ ) or the tube ( $V = 281, p = 0.012$ ). Similarly, they also found the task easier with the cube than with the dog ( $V = 109, p = 0.004$ ), or the tube ( $V = 87, p = 0.004$ ). On the contrary, the opacity had no significant impact on the confidence and perceived difficulty (Wilcoxon,  $p > 0.05$  for both comparisons).

### 6.5.3 Discussion

In this study, we hypothesized that shapes rendered with **CC** will be perceived as more densely filled than shapes rendered using **OC**, as hinted by the pilot study (Section 6.4). Since  $p_{CC}$  is significantly higher than the chance level, this hypothesis is fully supported.

The  $p_{CC}$  distribution for the cube shows a high variance. However, it is important to note that the  $p_{CC}$  data appear to form a bimodal distribution, hinting at the presence of two subgroups within the population who interpret the relationship between intersection contour and shape infill density in opposite manners. The significantly high  $r$  for this shape indicates a particularly strong difference between **OC** and **CC** feedback, which was thus easy to perceive. Therefore, depending on the participant's interpretation of the feedback, their  $p_{CC}$  value tends to be closer to 0 or 1, but further away from the 0.5 chance level, thus increasing the variance. This is also coherent with the higher  $p_{CC}$ , and confidence level observed for the cube. To better understand this effect, it would be interesting to conduct a similar experiment with other sampling strategies (e.g. constant draw frequency) to study whether the perceived difference is mostly due to the difference in terms of actual drawn contour, or to the difference in resulting draw frequency.

On a similar topic, it is important to note that by altering the properties of the rendered pattern (e.g. size, geometry), we also alter the range of different sampling parameters (e.g. we can likely use slower focal point patterns on shorter path). Future work will be dedicated to study how intersection strategies and sampling strategies influence each other.

The per-participant results also seem to hint towards the presence of two distinct subgroups, depending on their interpretation of the haptic difference between **OC** and **CC**. Most participants appear to easily discriminate **CC** and **OC**, but their interpretations vary. This was partly confirmed by free comments given by participants after completing the experiment. Also, similarly to what was observed in the pilot, some participants reported feeling differences in other properties than infill density (e.g., roughness or stimulus area).

A larger and more diverse sample would be required to confirm the existence of subgroups and to assess their proportions within the population. While this experiment showed that haptically rendering virtual objects with different intersection strategies can create differences in perceived object properties, future work could investigate the factors determining participants' interpretations of intersection strategies.

Very interestingly, the opaque visual led to higher  $p_{CC}$  values, but did not significantly impact

the user's confidence or perceived difficulty, which seems to indicate that the effect may be unconscious. This could also emerge from a visuo-haptic effect, where the opaque object may be expected to be more densely filled, confirming the participant's haptic perception. That said, further investigations with more participants would be required to test those hypotheses.

Given the low intensity provided by UMH devices, rendering solid objects has always been a struggle. This experiment showed that by altering the intersection strategy, we can improve the perception of such properties. Furthermore, the pilot study also seemed to indicate that this feedback is also felt as more coherent with the visual of the object, despite the rendering of segments that are not part of the initial shape's outline. That said, deeper experiments should be conducted to confirm this last result.

## 6.6 Conclusion

The 3D workspace of UMH devices makes them particularly suited to rendering 3D virtual objects. In an effort to formalize the process of 3D object rendering with UMH, we proposed the concept of **intersection strategies**: a high-level algorithmic approach to compute sets of curves representing a hand-object intersection. Paired with the existing formalism of sampling strategies for rendering curves ([Chapter 5](#)), this chapter provides an end-to-end process for generating UMH stimuli based on the exploration of 3D virtual objects. We hope this formalism can assist future UMH interaction designers, as well as provide a basis for studying and discussing the perceptual impacts of intersection strategy design choices.

To emphasize the importance of studying the different steps of intersection strategies, we designed a new haptic rendering scheme, and followed a user-centered methodology to compare it to a state-of-the-art scheme. We showed that acting on the transformation step can affect users' perception of virtual object's infill density. Furthermore, we assessed the impacts of these rendering schemes on UMH interface operating noise.

We observed that a majority of users interpreted the abstract feedback in the same way, however more detailed user studies would be required to study whether subgroups exist in the population as well as to understand the parameters affecting their interpretations of virtual object properties.

### Chapter Summary

- We formalized **intersection strategies**, representing the process of turning a 3D shape and a hand representation into a set of curves representing their intersection;
- A **sampling strategy** can be applied to this output to generate the final UMH stimulus;
- We designed **CC**, a new intersection strategy to illustrate intersection strategies;
- A pilot study led us to a user study showing that **CC** increases the perceived infill density of a virtual object.

---

## **Part IV**

# **Enhancing Virtual Reality Interactions Using UMH**

---



# Chapter 7

## Hand Guidance

### Contents

---

7.1	Introduction	163
7.2	UMH Strategies for Hand Guidance	164
7.3	User Study: Evaluation in a Static VR Environment	167
7.4	User Study: Evaluation in a Dynamic VR Environment	173
7.5	Conclusion	176

---

### 7.1 Introduction

In Part IV, we study how we can leverage UMH technologies to support, complement, and improve VR interactions. Specifically, this chapter focuses on the use of UMH stimulation to guide the user's hand towards a given pose using position and orientation cues. In such cases, UMH acts as a complementary channel to the visual rendering of the VR environment. Such uses of tactile feedback can allow the user to receive additional information without overloading the visual sensory channel.

Indeed, providing guidance feedback in VR can be beneficial in a wide range of applications, including industrial, sport, and medical training. While guidance cues are often conveyed visu-

ally [Harada et al. 2022], there is a growing interest for the use of haptic interfaces to convey guidance information, whether it is from human navigation guidance [Kappers et al. 2022; Cabaret et al. 2024], or single limb guidance [Sarac et al. 2019]. Guidance feedback can be even more useful in immersive VR scenarios, as the user would not be able to see their own physical body. Furthermore, VR interaction can benefit from unobtrusive complementary haptic feedback techniques which leave users unencumbered and free to engage in concurrent interactions. In this respect, UMH is of particular interest as it provides haptic sensations without the need for the user to wear or grasp any interface [Howard et al. 2022].

While some techniques have been proposed for hand guidance coupling schemes with UMH [Freeman et al. 2019; Suzuki et al. 2019] (discussed in [Section 1.5.2](#)), they only proposed position guidance, and evaluated their techniques in real environments. Here, we build upon their work, refining position techniques and proposing orientation feedbacks, specifically evaluated in VR.

Specifically, we propose the following contributions:

- we present four 2 DoF position, and two 2 DoF orientation coupling schemes for hand guidance purposes ([Section 7.2](#)).
- we evaluate these techniques in both a static ([Section 7.3](#)) and dynamic ([Section 7.4](#)), proving the usefulness of UMH guidance feedback, and highlighting the determining role of the task difficulty on the guidance feedback effectiveness.

The content of this chapter was adapted from the following publication: [[Mulot et al. 2023c](#)].

## 7.2 UMH Strategies for Hand Guidance

We designed 6 UHM guidance strategies in order to study user performances: four providing 2 DoF position guidance cues and two providing 2 DoF orientation cues, summarized in [Figure 7.1](#) and [Figure 7.2](#) respectively. The resulting patterns were rendered using STM, with a focal speed of 8 m/s so as to maximize its perceived intensity [[Frier et al. 2018](#)]. The pattern shapes were chosen to expand on the work of [Freeman et al. \[2019\]](#) and [Suzuki et al. \[2019\]](#).

**Position guidance** Considering the user’s hand position  $\mathbf{H} \in \mathbb{R}^2$  and the target position to reach  $\mathbf{T} \in \mathbb{R}^2$ , the error vector between the user’s hand and the target is defined as  $\mathbf{T} - \mathbf{H}$ . We devised four position guidance techniques:

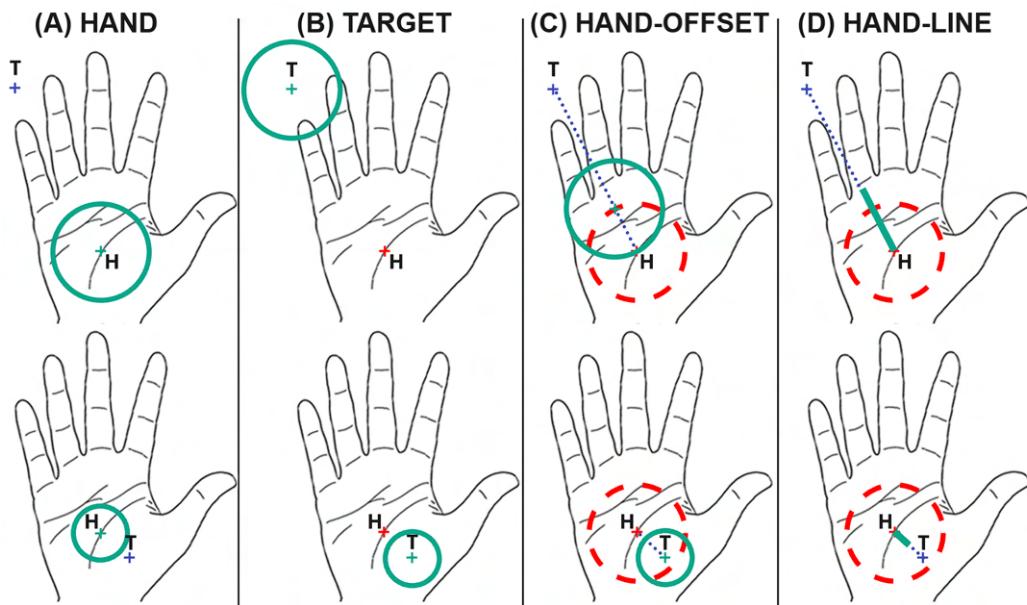


Figure 7.1 – Representation of the four position guidance strategies for two different target positions (top and bottom). Green solid circles and lines represent the haptic pattern rendered by the UMH interface. In **HAND-OFFSET** and **HAND-LINE**, the red dashed circle shows the circumference on which the haptic circle is rendered when the target is far away. In all but **HAND-LINE**, the green cross shows where the haptic circle is centered. **T** and **H** crosses correspond to the target and hand positions, respectively.

- **HAND** (see Figure 7.1(A)) provides information about the distance to the target,  $d = \|\mathbf{T} - \mathbf{H}\|$ , with no information regarding direction. To do so, we render a haptic circle with a varying radius at the centre of the palm, **H**. The radius of the circle varies proportionally to the distance  $d$  to the target, with a maximum of  $r = 4$  cm, to fit inside most hands [Peebles et al. 1998], when  $d = 10$  cm. The circle is always rendered centered on the user’s palm **H**
- **TARGET** (Figure 7.1(B)) is similar to the **HAND** strategy, except that the haptic circle with a varying radius is always centered at the target **T**. As before, the radius varies proportionally to the distance  $d$ , with a maximum of  $r = 4$  cm when  $d = 10$  cm. Since the circle is centered on the target, users thus do not receive any stimulation if they are further than 4 cm from it.
- **HAND-OFFSET** (Figure 7.1(C)) provides information about both distance and direction to the target. Similarly to **HAND**, we render a haptic circle with a varying radius on the hand. However, instead of always positioning it at the center of the palm, this time we offset it, to indicate the direction towards the target. Specifically, if  $d > 3$  cm, the circle

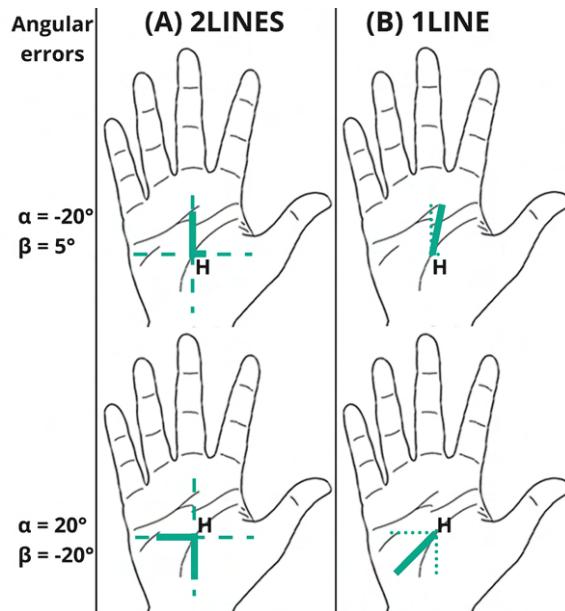


Figure 7.2 – (Right) Representation of the two orientation guidance strategies for three different target positions (left, center, right). In the top images, black and gray vectors represent the normal vector to the participant’s hand, and target hand, respectively. Orientation errors with respect  $\alpha$  and  $\beta$  requires movements along the proximal-distal and radial-ulnar axis of the hand, respectively.

is centered at the intersection between a 3 cm-radius virtual circle centered on  $\mathbf{H}$ , and the  $\mathbf{T} - \mathbf{H}$  segment. On the other hand, if  $d \leq 3$  cm, the circle is centered directly on the target  $\mathbf{T}$ . As before, the radius of the circle varies proportionally to the distance  $d$ , with a maximum of  $r = 4$  cm when  $d = 10$  cm.

- **HAND-LINE** (Figure 7.1(D)) also provides information about both the distance and direction to the target, but it uses a line instead of a circle. To do so, we render a segment of varying length along the direction  $\mathbf{T} - \mathbf{H}$ , starting from  $\mathbf{H}$  and growing proportionally to the distance  $d$ , with a maximum length of  $l = 6$  cm when  $d = 10$  cm. We used a wider pattern to make up for the fact that a circle covers a larger area of the hand.

**Orientation guidance** The orientation of the hand can be considered as a composition of two rotations, one defined by a flexion/extension of the wrist and one by its pronation/supination. Angles  $\alpha$  and  $\beta$  indicate the orientation error with respect to flexion/extension and pronation/supination movements, respectively. To match the 2 DoF rendering of positioning strategies, and keep a simple feedback, we do not consider rotations defined by radial/ulnar deviations of the wrist. We devised two orientation guidance techniques:

- **2LINES** (Figure 7.2(A)) provides information about  $\alpha$  and  $\beta$  errors by rendering two orthogonal segments of varying length, starting at  $\mathbf{H}$ . One segment runs along the proximal-distal hand axis and one along the radial-ulnar hand axis. The length of the segments is proportional to  $|\alpha|$  and  $|\beta|$ , reaching a maximum length of  $l = 4$  cm when the angular error is more than  $30^\circ$  around the considered axis. If the first segment grows towards the distal direction of the hand (towards the fingers), the user should extend the wrist; if it grows towards the proximal direction of the hand (towards the wrist), the user should flex the wrist. Similarly, if the second segment grows towards the radial direction of the hand (towards the thumb), the user should supinate the wrist; if it grows towards the ulnar direction of the hand, the user should pronate the wrist.
- **1LINE** (Figure 7.2(B)) provides the same information in a composite way, by rendering the vector addition of the two vector segments described in the **2LINES** strategy.

The threshold values of 10 cm for the position and  $30^\circ$  for the orientation strategies were chosen empirically. Higher values would allow a description of a bigger workspace. Yet, this would lead to slower variations of the stimulus, requiring larger hand motion to feel any change. If participants go beyond this value, they still receive feedback, but it will not grow any larger. Reaching the optimal configuration reduces the pattern to a point, which will no longer be perceptible, indicating that the target has been reached.

## 7.3 User Study: Evaluation in a Static VR Environment

Our objective is to evaluate the proposed ultrasound haptic strategies in providing 2 DoF guidance information to users in VR. We seek to answer two main questions:

- Can we guide the user to a certain hand position and orientation using UMH feedback?
- What is the best guidance strategy among the ones we designed?

### 7.3.1 Materials and Methods

#### Apparatus and Setup

The experimental setup is shown in Figure 7.3(A). Participants were seated facing the table, and wore an HTC VIVE Tracker strapped to the back of their dominant hand. They observed the virtual environment through a HTC VIVE Pro Eye VR headset<sup>1</sup>, with the headphones

---

1. <https://developer.vive.com/resources/hardware-guides/vive-pro-eye-specs-user-guide/>

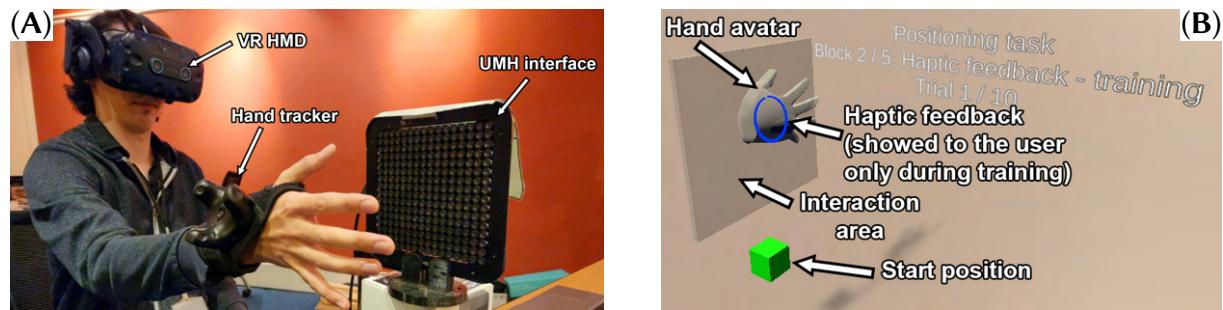


Figure 7.3 – (A) Experimental setup. Participants sit near the ultrasound haptic interface, with their dominant hand facing the device. They wear a VR HMD and a tracker. They hold a controller in their other hand to validate trials and start blocks. (B) VR scene, where the interaction zone is indicated by a translucent squared area. During training only, the scene also showed the rendered haptic sensations (blue circle in this case).

playing pink noise to mask any audio cues from the haptic interface. Tracking errors were found to be way below the size of a focal point [Bauer et al. 2021; Rakkolainen et al. 2021], and were thus considered negligible. Participants were asked to place their dominant hand facing the haptic interface. They could rest the arm on an armrest between trials.

The UHM feedback was provided by an Ultraleap STRATOS Explore<sup>2</sup> interface. The haptic device was housed on a custom stand, keeping the array in a vertical position, perpendicular to the table supporting it.

The VR environment (developed in Unity<sup>3</sup>) is illustrated on Figure 7.3(B). It was composed of a translucent 28 cm-edge square representing the area where the participants should interact. This square area was placed 20 cm away from the haptic device to ensure good haptic perception [Howard et al. 2020]. A small green cube, located 25 cm under the center of the interaction area, acted as a button allowing users to proceed to the next trial. A virtual hand avatar mimicked the movements of the user's hand in the virtual environment. During the training phase, the scene also showed a visual rendition of the haptic cues in the form of blue circles or lines, similar to that in Figure 7.1 and Figure 7.2.

### Task and Protocol

Participants were immersed in the VR scene described above. At the beginning of each task, the target hand pose is displayed for 1 s thanks to an additional virtual hand appearing in the

2. [https://www.ultraleap.com/datasheets/STRATOS\\_Explore\\_Development\\_Kit\\_datasheet.pdf](https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf)

3. <https://unity.com/>

scene. This gives participants time to create a mental image of the target. The visual indication of the target then disappears and participants are asked to move their own hand to match this target pose, both in terms of position and orientation, being as accurate as possible, and relying on the mental image they constructed. They were given up to 15 s to refine their hand pose, or could validate early by pressing using a controller held with the non-dominant hand.

Participants performed the task using each of the six guidance strategies presented in [Section 7.2](#) and under a baseline condition without haptic feedback. For this baseline, participants relied only on the visual information provided by the target virtual hand briefly showed at the beginning of each trial. Participants were told whether the current strategy provided position or orientation guidance (see [Section 7.2](#)), but they had to match the target hand pose both in terms of position and orientation for all trials.

The experimental protocol was divided in two main parts, evaluating the position and orientation guidance strategies, respectively. Each part was divided into blocks, evaluating one strategy each. For both parts, we carried out an additional set of tasks with no haptic guidance. The order was balanced between participants. This design led to a total of 4 (position strategies) + 1 (no haptic feedback) + 2 (orientation strategies) + 1 (no haptic feedback) = 8 blocks. The block order for each participant was computed using a balanced Latin square. Each block started with a training phase, where the feedback strategy under consideration was explained orally by the experimenter and textually with an illustration in the VR scene. After training, participants carried out 15 guidance trials, composed of 3 repetitions of 5 predefined target hand positions and orientations. These poses covered as much of the available workspace as possible, both in terms of position and orientation.

## Collected Data

Before the beginning of the experiment, participants were asked to fill a demographics questionnaire. During the experiment, we recorded the trial completion times, as well as the final errors in position and orientation. After each of the two parts of the experiment (position and orientation guidance), participants were asked to rank the haptic strategies they just experienced based on perceived effectiveness.

## Participants

20 participants took part in the study (17 men, 2 women, 1 non-binary, aged  $m \pm s.d. = 25.8 \pm 5.56$ ). 7 of them were very experienced with VR, while 13 had little to no experience

with it at all. Similarly, 8 had a lot of experience with haptic interfaces, while 13 had little to no experience. 7 participants had already used an UMH device at least once. 17 participants were right-handed, 2 were left-handed, and 1 was ambidextrous (and used their right hand for the experiment).

We observed little variability in the participant’s hand sizes. However, if needed, all the strategies can be adapted to different hand sizes by modifying the circle radii or line length. The personalization of haptic rendering is an open and interesting topic in the field [Young et al. 2020].

### 7.3.2 Results

**Table 7.1** summarises the main results of this experiment.

Table 7.1 – Results of the static user study. **H-O**, **H-L**, and **N-H** stand for the **HAND-OFFSET**, **HAND-LINE**, and no haptic feedback (**no-haptics**) strategies, respectively. We report the mean value for the ranking and the median for the other metrics. Lower values indicate better performances.

Evaluation in a static VR environment								
Strategy	HAND	TARGET	H-O	H-L	N-H	2LINES	1LINE	N-H
Error	11 mm	10 mm	8.7 mm	16.7 mm	20 mm	9.8°	10.6°	10.6°
Duration	5.9 s	5.5 s	5.4 s	3.5 s	2.1 s	5.2 s	3.8 s	2.2 s
Ranking	2.95 / 4	2.95 / 4	1.63 / 4	2.47 / 4		1.37 / 2	1.63 / 2	

### Comparisons Based on Objective Data

We gathered a total of 1500 trials with during blocks with position guidance feedback, and 900 trials with orientation guidance feedback. Error and duration data were significantly non-normal, as assessed by Shapiro-Wilk tests. We therefore applied a logarithmic transformation to the data. **Figure 7.4(A, B)** shows the errors in position and orientation, while **Figure 7.4(C, D)** shows the completion times.

**Position/orientation errors** We modelled the position and orientation errors with mixed linear models with respect to the haptic strategies (4 and 2 degrees of freedom, df) and considering the participants as a random effect. With these models, we then performed an analysis of deviance.

The analysis of deviance showed significant differences in position errors between the position strategies ( $F(4, 1476) = 30.6, p < 0.001$ ). Post-hoc analysis using a Tukey test showed

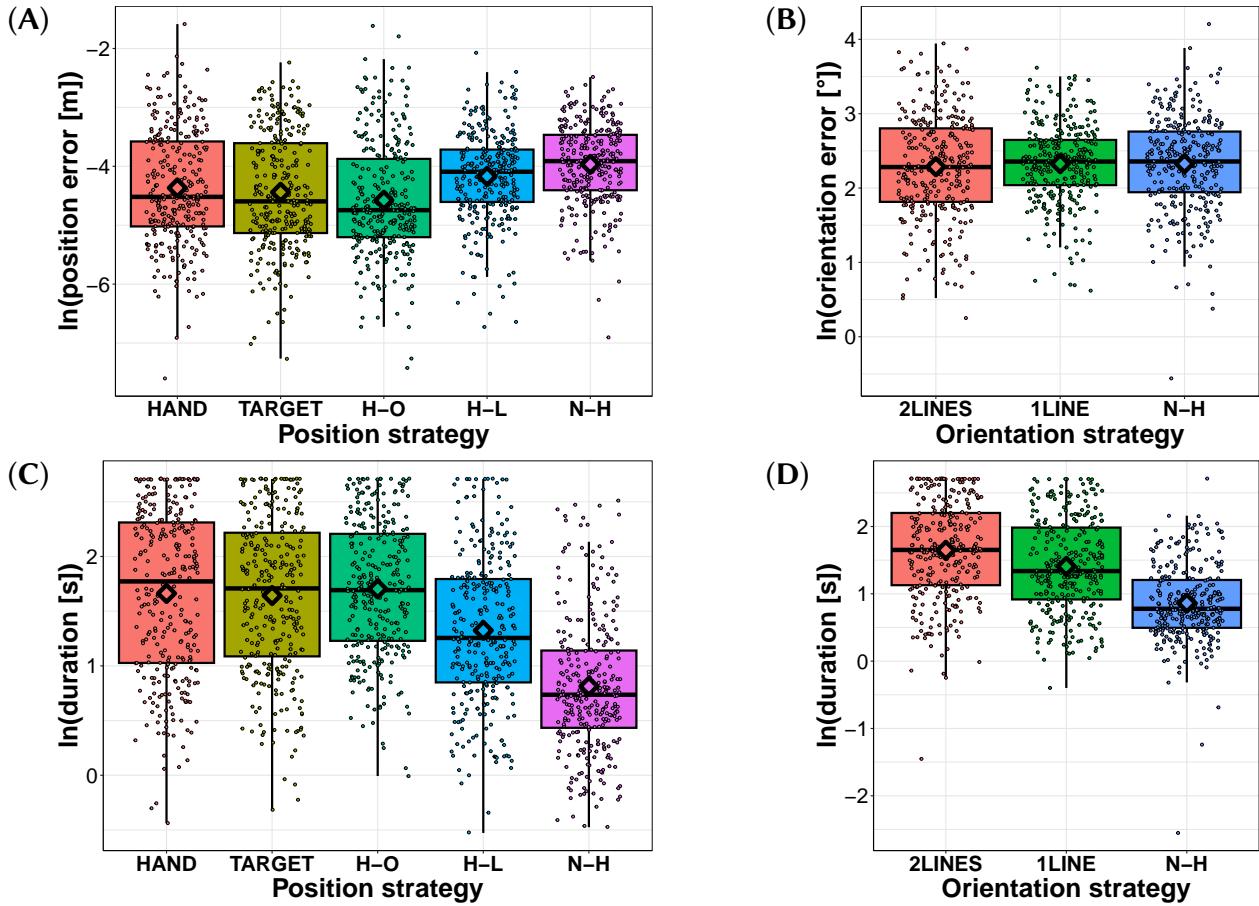


Figure 7.4 – Distribution of the errors in reaching the target pose. (A) Position errors are computed as the distance [m] of the 2D vector between the hand and the target. (B) Orientation errors are computed as the absolute angle [°] between the user’s and the target’s rotation. (C, D) Box plot of the completion time [s] for each position (C) and orientation (D) strategy. From left to right, strategies are **HAND**, **TARGET**, **HAND-OFFSET**, **HAND-LINE**, **NO-HAPTICS**, **2LINES**, **1LINE**, **NO-HAPTICS**. Each point represents a trial, while the diamond represents the average for the corresponding strategy.

that all pairwise comparisons were significant ( $p < 0.05$ ), except for **HAND-OFFSET** vs. **TARGET** and **HAND** vs. **TARGET**. This result shows that all the proposed haptic strategies allowed improved position performances compared to the no-haptics scenario. Considering the pairwise comparisons and median results, the **HAND-OFFSET** and **TARGET** strategies were considered to be the best.

We carried out the same comparisons for the orienting strategies. An ANOVA ( $F(2, 878) = 0.4949, p > 0.05$ ) showed no significant effect of the haptic feedback on the user’s orientation performance for this experiment.

**Completion time** We followed the same methodology as in [Section 7.3.2](#) to investigate the effects of employing different haptic strategies on the duration of the trials.

The analysis showed once again a significant difference across the position guidance strategies ( $F(4, 1476) = 153.84, p < 0.001$ ). A post-hoc analysis using a Tukey test revealed that all pairwise comparisons were significant ( $p < 0.05$ ) except for **HAND** vs. **HAND-OFFSET**, **HAND** vs. **TARGET**, and **HAND-OFFSET** vs. **TARGET**. This result shows that the three strategies rendering a circle (**HAND**, **TARGET**, **HAND-OFFSET**) take a similar amount of time, while the one rendering a line is significantly faster, although still significantly slower than **NO-HAPTICS**.

Significant differences across the orientation guidance strategies ( $F(2, 878) = 213.37, p < 0.001$ ) were also detected by the analysis of deviance. A post-hoc analysis using a Tukey test revealed that all comparisons were significant ( $p < 0.001$ ), meaning that **NO-HAPTICS** is faster than **1LINE** strategy, which is in turn faster than **2LINES**.

## Subjective Rankings

We analysed the subjective rankings that users submitted at the end of each experimental part (position and orientation), shown in [Table 7.1](#). The **NO-HAPTICS** was not included in the ranking. We had to remove the ranking data from one participant who did not fill the form correctly.

A Friedman test showed a significant difference across the position guidance strategies ( $\chi^2(3, N = 19) = 13.2, p = 0.004$ ). A post-hoc analysis using Holm-corrected pairwise Wilcoxon signed-rank tests revealed that the **HAND-OFFSET** strategy was ranked significantly better than the **TARGET** ( $V = 22, p = 0.015$ ) and **HAND** strategies ( $V = 170, p = 0.014$ ).

A Friedman test did not find a significant difference across the orientation guidance strategies ( $\chi^2(1, N = 19) = 1.32, p > 0.05$ ).

## Discussion

Results showed that the **TARGET** and **HAND-OFFSET** strategies offered better performances, while the **HAND-OFFSET** and **HAND-LINE** were perceived as being the most intuitive and effective by our participants. Thus, we can consider the **HAND-OFFSET** strategy as being the best strategy out of the four position strategies. This result is in accord with the work of

Freeman et al. [Freeman et al. 2019] and Suzuki et al. [Suzuki et al. 2019], who showed that hand guidance was possible with circular and linear patterns in a non-virtual environment. On the other hand, the only significant comparison with respect to the orientation strategies (completion time) showed that **1LINE** allows comparable performance and level of intuitiveness than **2LINES**, but it requires less time. Therefore, we can consider the **1LINE** strategy as being better than the **2LINES** one.

## 7.4 User Study: Evaluation in a Dynamic VR Environment

Following the first experiment’s promising results, we extended our investigation to a dynamic virtual environment.

### 7.4.1 Materials and Methods

#### Apparatus and Setup

The experimental setup is the same as described in [Section 7.3.1](#) and shown in [Figure 7.3\(A\)](#), except for the VR environment, which is composed of a ping pong table. The participants are immersed in the environment at one end of the table, and the virtual representation of their hand in VR is a paddle. Similarly to the first experiment, we placed a  $28 \times 28$  cm translucent rectangle where the user-controlled paddle should move to ensure good haptic perception [[Howard et al. 2020](#)]. A “ghost” paddle represents the projection of the user-controlled paddle onto this area to encourage users to stay within it.

#### Task and Protocol

For this experiment, participants were provided with either the **HAND-OFFSET** or **1LINE** strategies, as we considered them to be the best for the position and orientation tasks, respectively (see [Section 7.3.2](#)).

Similarly as before, the experiment is divided into two parts, one for evaluating position guidance and one for evaluating orientation guidance. For the position guidance tasks, a ping pong ball is launched towards the participant, who has to intercept it by moving the paddle around. In this case, the orientation of the paddle is not important, the only objective is to catch the ball. For the orientation guidance tasks, a ping pong ball is again launched towards the participant. However, this time, the ball is always shot towards the same fixed location, and participants need to adjust the orientation of their hand/paddle to bounce the ball towards a specific target

above the net. In both cases, users can construct a mental image of how the ball will behave, in a similar way as in the first experiment. As the environment is dynamic, this image will be refined as the ball moves.

Each part (position and orientation) is divided into two blocks, one providing the considered haptic feedback strategy and one baseline providing no haptic feedback. Similarly as before, each block started with a training phase, where the feedback strategy under consideration is explained. The task and block order are randomized to avoid order effects. After the training, participants carried out 15 guidance trials, considering three different ball speeds and random target positions and orientations. These poses are again designed to cover as much as possible the available workspace, both in terms of position and orientation. Once participants were confident with their position or orientation, they could press the controller trigger to speed up the ball and proceed to the next trial.

## Collected Data

We gathered the same pre-experimental data as in the first experiment (see [Section 7.3.1](#)). During the experiment, at the end of each trial, we registered the trial duration as well as the position and angular errors. After each of the two parts of the experiment (position and orientation guidance), participants were asked to rate the perceived difficulty of the task, how well they understood the haptic stimuli, and what cues they relied on (from visual only to haptics only) on a 7-points Likert scale.

## Participants

8 participants participated in the study (8 men, 1 woman, aged  $m \pm s.d. = 25.8 \pm 4.6$ ). 4 of them had plenty of experience with VR and 4 of them had little experience with VR. 4 of them had plenty of experience with haptics and 4 of them had little experience with haptics. 7 participants were right-handed and 1 was ambidextrous, so everyone used their right hand. All participants had already used an UMH device at least once prior to the experiment. 5 of them also participated in the first experiment.

### 7.4.2 Results

[Table 7.2](#) shows the main results of this experiment and the data distribution is illustrated in [Figure 7.5](#).

Table 7.2 – Results of the dynamic user study. **H-O** and **N-H** stand for the **HAND-OFFSET** and **NO-HAPTICS** feedback strategies, respectively. We report the mean value for the difficulty and the median for the other metrics. Lower errors and durations indicate better performances.

Evaluation in a dynamic VR environment				
Strategy	H-O	N-H	1LINE	N-H
Error	17.8 mm	19 mm	7.2°	9.3°
Duration	4.88 s	5 s	4.7 s	3.6 s
Difficulty	2.375 / 7		4.75 / 7	

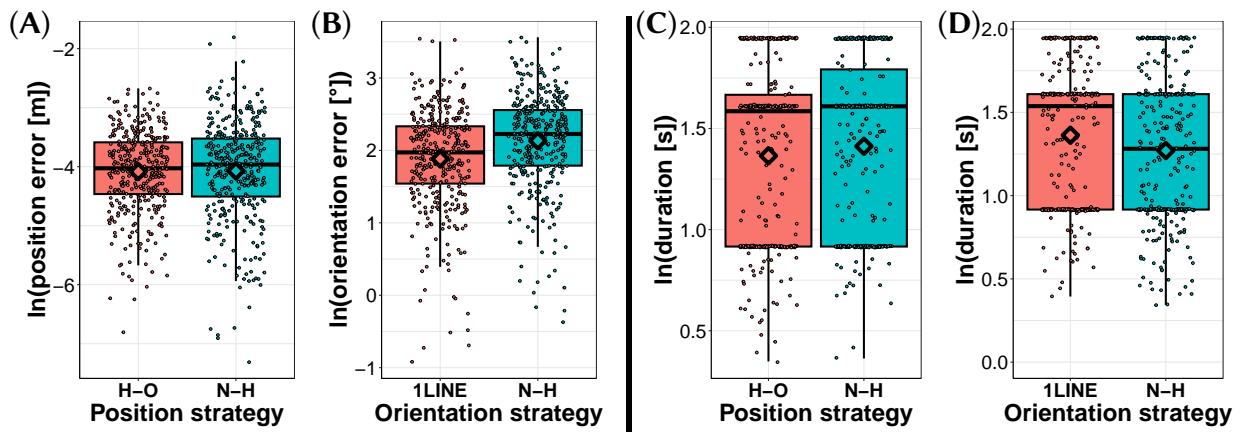


Figure 7.5 – Results for the user study in a dynamic environment. (A) Positioning [cm] and (B) orientation [°] errors. (C, D) Trial completion times [s]. Each point represent a trial, while the diamond represent averages.

## Performance Analysis

We gathered a total of 720 position trials and 720 orientation trials.

As the data did not follow a normal distribution (Shapiro-Wilk test,  $p < 0.001$ ), we used the same analysis protocol as in [Section 7.3.2](#). The ANOVA revealed a significant impact of the guidance strategy on the orientation error ( $F(1, 711) = 26.1, p < 0.001$ ). This proves that the orientation guidance feedback does indeed help participants orient their hand. The tests also revealed a significant effect on the trial duration during the orientation tasks ( $F(1, 711) = 10.9, p < 0.001$ ). We also observed a significant effect of ball speed on the orientation performances ( $F(2, 710) = 11.36, p < 0.001$ , with lower speeds leading to better results, from 9.7° median error in a 2.5 s trial, to 7.2° in a 7 s trial), showing that participants need some time to understand and follow this cue.

On the other hand, we observed no significant effect of the feedback on the position error, nor

on the duration of position trials.

## Discussion

Contrary to the previous study of [Section 7.3.2](#), the position feedback did not improve position performances, while the orientation strategies did. This result shows that orientation guidance is possible with UMH technologies, in line with previous studies in UMH [[Howard et al. 2022](#)] and with other technologies [[Drewing et al. 2005](#)]. This difference between our two studies can be explained by the difference in task difficulty. On a 1 to 7 Likert scale, participants ranked the difficulty of the position task with an average of 2.375, and the orientation task with an average of 4.75. A Wilcoxon signed-rank test showed that this difference was statistically significant ( $V = 0, p = 0.012$ ). Similarly, the position task revealed to be easy enough to be carried out without haptic feedback.

## 7.5 Conclusion

In this chapter, we presented different UMH guidance strategies designed to help users position and orient their hand in VR environments. We evaluated them in static and dynamic VR environments, revealing significant effects of the strategies on the performances, depending on task difficulty.

Position and orientation feedback were both shown to be potentially beneficial in tasks where they are pertinent. Positioning feedback reduced errors in the static task, yet orientation feedback was found to be ineffective in this task. On the other hand, orientation feedback was found to be effective in the dynamic task, where the complexity of the orientation task was greater than that of the position task. These results lead us to hypothesize that task complexity is determining in the effectiveness of a provided guidance coupling scheme. We also showed that haptic feedback often significantly slowed task execution, as participants are able to refine their position. This result means that these techniques can be beneficial in scenarios requiring high precision while not being extremely time-sensitive. Yet, the current study design does not allow us to draw any inference regarding what task difficulty thresholds may influence the effectiveness of UMH guidance feedback. Overall, our work opens new opportunities for UMH in VR, highlighting that using UMH as a complementary feedback is a viable option.

## Chapter Summary

- We proposed 4 UMH coupling schemes to render position guidance cues, and 2 for orientation feedback in VR;
- We evaluated these techniques in a static VR environment, showing that **HAND-OFFSET** and **1LINE** are the best techniques for position and orientation guidance respectively;
- We evaluated these best techniques in a dynamic VR environment, highlighting the crucial role of the task's difficulty on the feedback's effectiveness.



# Chapter 8

## Visuo-Tactile Cues for Contact Rendering

### Contents

---

8.1	Introduction . . . . .	179
8.2	Contact Rendering Techniques . . . . .	180
8.3	Pilot Experiment: Haptic Trigger Calibration . . . . .	186
8.4	User Study: Generating Coupling Schemes . . . . .	191
8.5	User Study: Comparing Coupling Schemes . . . . .	196
8.6	General Discussion . . . . .	205
8.7	Conclusion . . . . .	206

---

### 8.1 Introduction

Despite the large versatility of UMH, they cannot generate forces high enough to push a user's hand. This means that they cannot directly prevent users from moving into locations where they should not. This is especially important in VR, where the usual lack of tactile sensations when touching an object makes it hard to naturally touch or grasp.

This is why this chapter explores the design and use of visuo-haptic coupling schemes for rendering interpenetration cues in VR. Combining visual with tactile cues reinforces the message, helping users interact with virtual objects while remaining at its surface. Yet there is a very

large design space for visual and UMH feedback techniques, and therefore an even wider design space for visuo-haptic coupling schemes. This chapter thus contributes to **Challenge IV**, expanding our study of UMH coupling schemes and UMH-based VR interactions. In contrast to [Chapter 7](#), which uses haptics as a **complementary** channel to provide new information, this chapter uses it as a **redundant** channel. Here, it aims to reinforce the visual cues of contact to improve interaction fidelity.

More specifically, we propose the following contributions:

- A design and implementation of four visual and three UMH feedback techniques for rendering interpenetration with a virtual object ([Section 8.2](#));
- Two user studies investigating the user experience and contact accuracy for different visuo-haptic coupling schemes ([Section 8.4](#), [Section 8.5](#)).

These contributions have not yet been published, but correspond to a collaboration between the following researchers (in alphabetical order): Maud Marchal, Lendy Mulot, and Claudio Pacchierotti.

## 8.2 Contact Rendering Techniques

To pursue this goal of studying visuo-haptic coupling for interpenetration, we implemented different visual ([Section 8.2.2](#)) and haptic ([Section 8.2.3](#)) feedback techniques to provide multi-modal interpenetration cues when interacting with a virtual object. These techniques were carefully designed, taking inspiration from related works [[Sreng et al. 2007](#); [Zilles et al. 1995](#); [Punpongsanon et al. 2015](#); [Frier et al. 2019](#)], and ensuring a wide coverage of the design space ([Section 8.2.4](#)).

### 8.2.1 Interpenetration Computation

As this chapter initiates the study of UMH-based visuo-haptic coupling schemes for contact rendering, we restrict our focus to interactions between the index fingertip and a virtual cube. This is because the index finger is the most used finger for pointing and probing the environment. Yet this is without loss of generality, as one could compute independent interpenetration distances for each finger using the following methodology. A small collider was virtually placed at the end of the fingertip, and used both to detect the collision with the cube, and as the reference point when computing the interpenetration with the object.

Then, it is worth noting that interpenetration distance is an ambiguous metric, that could be

computed in many different ways, such as the distance from the fingertip to its closest point on the object's surface, or to a given hand proxy that follows a given set of physical rules. Here, we consider directional interpenetration, computed as the signed distance from the fingertip to the plane of each cube face, measure along that face's normal vector. Formally, if  $\mathbf{p} \in \mathbb{R}^3$  represents the position of the fingertip, and  $\bar{\mathbf{p}}_{face}$  its projection onto a given face of the cube, then the interpenetration distance relative to that face would be computed as

$$d_{face} = s(\mathbf{p}) \cdot \|\mathbf{p} - \bar{\mathbf{p}}_{face}\|$$

with  $s(\mathbf{p}) = 1$  if  $\mathbf{p}$  is inside the cube, and  $-1$  if it is outside. We consider the interpenetration inside the cube to be positive when the finger is touching the cube and negative otherwise.

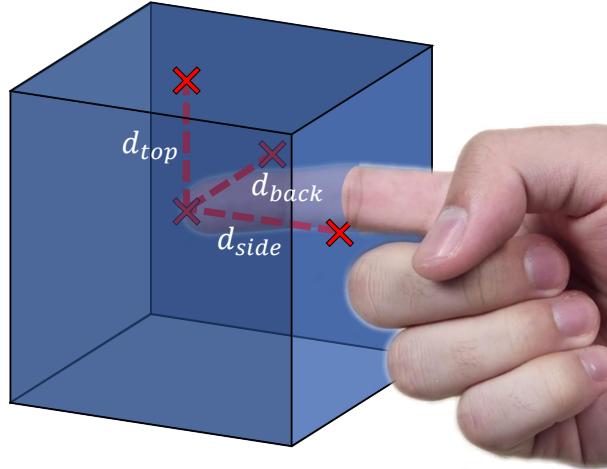


Figure 8.1 – Finger colliding with a virtual cube, illustrating the ambiguity of computing Interpenetration distances. It can be computed relative to any face, by measuring the distance between the fingertip and its projection onto the face.  $d_X$  represents the interpenetration relative to the face X.

### 8.2.2 Visual Rendering

We implemented four visual feedback techniques, to provide interpenetration cues, illustrated in [Figure 8.2](#), and presented in more details hereafter. For any interaction, the visual and haptic feedback are computed as a function of the reference cube face, and corresponding interpenetration distance  $d_{face}$ .

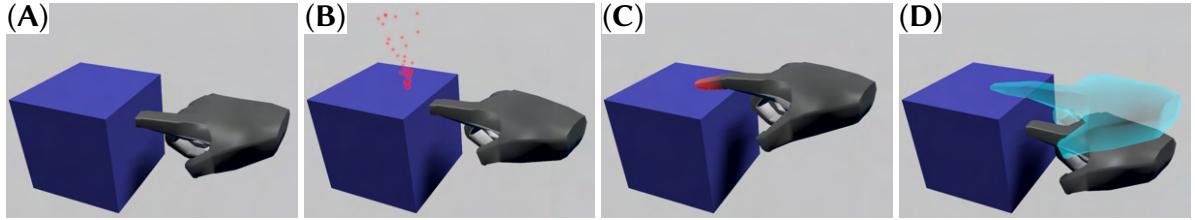


Figure 8.2 – We implemented four visual interpenetration feedbacks. (A) **Default Hand**: the virtual hand follows the participant’s real hand and interpenetrates the virtual object with no resistance or additional feedback (Section 8.2.2). (B) **Visual Particles**: particles are emitted at the surface of the cube, above the fingertip (Section 8.2.2). (C) **God-Object and Finger Coloration**: interpenetration with the cube is prevented, and the finger gets progressively more red as the real hand descends (Section 8.2.2). (D) **Ghost Hand**: the classical virtual hand follows the user’s real hand with no restrictions, and a “ghost” hand appears, constrained at the surface of the object (Section 8.2.2).

### Default Hand

During any of our VR interactions, users have a virtual hand avatar, moving in response to the user’s real movements. Our first visual technique serves as a baseline: it uses one of the default hand avatars from Ultraleap’s Unity plugin<sup>1</sup>. This hand follows exactly the motion of the user’s real hand, which means it can interpenetrate visual objects without any sort of resistance (see Figure 8.2(A)).

### Visual Particles

Our second feedback is inspired from the work of Sreng et al. [2007]. Specifically, the virtual hand is controlled identically as for **Default Hand**, and particles are emitted when the finger is in contact with the virtual object (see Figure 8.2(B)). We used the same parameterization as the one proposed by Sreng et al. [2007] for impact rendering, except that we use the interpenetration distance  $d_{face}$  as input instead of the fingertip speed. This is because the impact is an event best described by discussing the energy during impact, which directly relates to the finger speed, while interpenetration is a continuous state less dependent on speed (e.g. the hand can remain static at any level of interpenetration). Specifically, when the finger is in contact with the virtual cube, a source of particles is placed in the virtual scene at  $\bar{\mathbf{p}}_{face}$ . Particles are emitted with random directions inside a cone that is aligned with the normal to the cube’s face, and of angle 15°. Particles are emitted with a rate of 200 particles per second, and they disappear 0.2 s after being emitted. By default, each individual particle speed is chosen

1. <https://docs.ultraleap.com/xr-and-tabletop/xr/unity/index.html>

randomly between 0.1 m/s and 1 m/s, and then multiplied by a factor linearly increasing from 0.1 at the surface of the cube, to 1 when reaching the maximum interpenetration threshold. This results in particles going up to 2 cm outside the cube when the user initiates contact, to up to 20 cm when interpenetration increases. These parameters were chosen by [Sreng et al. \[2007\]](#), and their suitability for interpenetration rendering was empirically verified.

### God-Object and Finger Coloration (G-OFC)

Our G-OFC feedback follows the concept of god-object [[Zilles et al. 1995](#)] to constrain the virtual hand at the surface of the virtual object, and thus prevent it from penetrating inside the object. This is achieved by translating the virtual hand provided by **Default Hand** by a vector of  $\bar{\mathbf{p}}_{face} - \mathbf{p}$  when it would be colliding with the virtual object.

Additionally, when in contact with the virtual object, the index finger emits a red glow, in a similar way as in the work of [Punpongsanon et al. \[2015\]](#). When initiating the contact, the color is not very bright and only covers part of the fingertip. But when the user continues to press on the cube, the color becomes brighter and covers a larger area, until it covers the whole finger (see [Figure 8.2\(C\)](#)).

### Ghost Hand

Finally, when in contact with the virtual object, the **Ghost Hand** feedback shows both the **Default Hand**, and a god-object [[Zilles et al. 1995](#)] hand (at the same position as G-OFC). This second hand is rendered using a shader which gives it a “ghost” appearance (see [Figure 8.2\(D\)](#)).

#### 8.2.3 Haptic Rendering

We also designed and implemented three UMH feedback techniques to provide interpenetration cues, as illustrated in [Figure 8.3](#). Each technique is the continuous looping of a dynamic stimulus until the contact is broken. The duration to complete one loop decreases linearly with interpenetration, from 2 s at the surface of the cube, to 0.5 s when reaching the maximum interpenetration threshold. These values were chosen empirically to be easily discriminable, being long enough to be able to perceive dynamic aspects of the feedback, and short enough so that the user can rapidly understand that something is moving. This allows us to provide continuous dynamic feedback, matching the continuous interpenetration state, and dynamic aspects of the visual feedbacks. This temporal mapping, where the loop duration shortens with deeper

interpenetration, is designed to create a feeling of rising urgency or pressure, analogous to the increasing resistance one would feel when pushing into a physical object. Secondly, naive feedbacks could e.g. increase the focal point intensity as interpenetration increases to simulate stiffness or resistance, but such feedback would provide a low intensity at the surface, which would make it harder to detect, and thus provide suboptimal interpenetration cues.

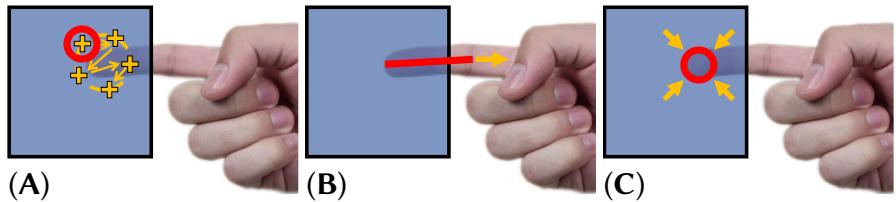


Figure 8.3 – We implemented three haptic interpenetration feedbacks, in response to a contact with a virtual cube (blue). The haptic feedback is represented in red, and orange arrows indicate the evolution of the feedback with respect to time. (A) **Haptic Particles**: a small haptic circle is rendered on successive random positions around the fingertip. (B) **Line**: a line is growing from the fingertip to the metacarpophalangeal joint. (C) **Circle**: a circle is shrinking on the fingertip.

### Haptic Particles

Our first haptic feedback, called **Haptic Particles** consists on the random motion of a small radius circle on the finger pad. Specifically, we sample 5 random positions on a 2 cm radius virtual circle, centered on the user’s fingertip. Then, we render a 1 cm radius circle, whose center cycles between these random positions at regular intervals. The circle is rendered using STM, with a focal point speed of 8 m/s [Frier et al. 2019].

### Line

The **Line** technique generates a growing polyline under the user’s index finger. When initiating contact, the line is reduced to a single point at the fingertip, and it then grows at a constant speed until it reaches the metacarpophalangeal joint. If the finger is bent, the polyline will be composed of three segments, for each of the phalanges. This line is rendered using STM, with a draw frequency of 250 Hz. This value was chosen empirically as it appeared to provide the strongest sensation.

### Circle

Finally, with the **Circle** haptic feedback, we render a small shrinking circle on the user’s fingertip. The circle’s radius starts at 1 cm, which is enough to cover the whole fingertip for the

vast majority of people [Peebles et al. 1998], and then linearly shrinks until it reaches a single point. Similarly to **Haptic particles**, the circle is rendered using STM, with a focal point speed of 8 m/s [Frier et al. 2019].

#### 8.2.4 Design Discussion

It is worth noting that these feedback techniques were chosen or designed to cover a large part of the design space.

Namely, **Default Hand** and **Visual Particles** show a virtual hand that follows exactly the user's hand, meaning they have no impact on proprioception. On the contrary, **G-OFC** fully prevents the virtual hand from penetrating the cube, which may have a strong proprioceptive effect, while the **Ghost Hand** is a hybrid technique, showing both the normal and redirected hands.

Similarly, **Visual Particles** and **G-OFC** provide cues that are very soft when the contact starts, and get stronger when the interpenetration increases, while **Ghost Hand** makes the whole ghost hand appear when the contact starts, generating a large cue from the start.

The different techniques also provide cues at different locations. Visually speaking, **Visual Particles** provides cues directly on the cube, while the coloration from **G-OFC** affects the whole finger and the redirection affects the whole hand. **Ghost Hand** also provides cues that cover the whole hand. Haptically, **Haptic Particles** and **Circle** both stimulate the fingertip, with **Haptic Particles** mostly focused on its edges, and **Circle** on its center, while **Line** stimulates the whole finger.

Similarly, **Haptic Particles** and **Circles** use small haptic circles, which allow for a high localized concentration of energy, while **Line** spreads it over a larger area.

**Haptic Particles** provides a non-continuous motion, with the circle being rendered at the same position for a certain amount of time, before being moved to a completely different location, without any transition between the two. On the opposite, **Circle** is always centered on the same position, but its size is varying. **Line** also mostly changes its size, but on a much larger scale.

## 8.3 Pilot Experiment: Haptic Trigger Calibration

To calibrate our experiments, we needed to ensure that users perceived visual and tactile contact cues at exactly the same time. While one could trigger the tactile sensation when the finger collides with the object, factors such as latency, stimulus integration, user expectations and training could impact their perception of synchronicity of multi-modal stimuli [Di Luca et al. 2019].

Specifically, we designed a pilot study to find out when we should start the haptic stimulation when the virtual hand is approaching each face of the cube.

### 8.3.1 Materials and Methods

#### Apparatus and Setup

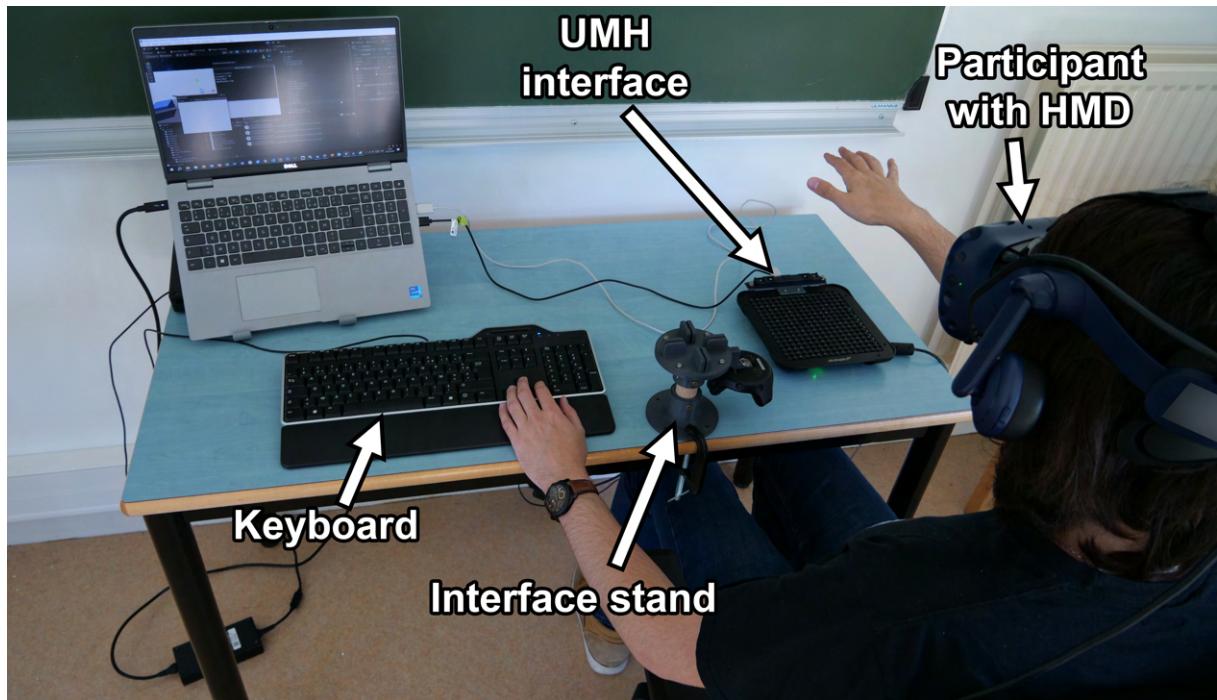


Figure 8.4 – Experimental setup for all experiments. Participants are seated in front of the UMH interface and wear a VR HMD. For the pilot study (Section 8.3) and first experiment (Section 8.4), users use their left hand on the keyboard. When interacting with the **Side** or **Back** face of the cube, the UMH interface is positioned on a 3D-printed stand. Here, the interface is in the **Top** configuration.

The experimental setup is illustrated in [Figure 8.4](#), and inspired from the work of [Tinguy et al.](#)

[2021] and Vizcay et al. [2021].

During the experiment, the user was sitting on a chair, wearing an HTC Vive Pro Eye<sup>2</sup> head-mounted display. Pink noise was playing from the headset to hide any exterior audio cue.

An Ultraleap STRATOS Explore<sup>3</sup> UMH interface with and attached and Ultraleap Stereo IR 170<sup>4</sup> hand tracker could be moved between the table and a custom 3D-printed stand to adapt to the different faces of the cube, as illustrated in [Figure 8.5](#).

Users could use their right hand facing the UMH interface to interact with the virtual cube. They had their left hand positioned on the keyboard's arrow keys.

The experiment was designed and run using Unity<sup>5</sup> 2022.3.18f1. The virtual scene, shown in [Figure 8.5](#) was composed of a 10 cm wide virtual cube that the user was able to touch. Invisible planes were also placed parallel to each of the cube's face, and the haptic feedback was triggered when the user's virtual fingertip collided with the plane. This allows for the late or early rendering of the stimulus. Using the keyboard's arrow keys, the user could move the plane outwards or inwards by increments of 1 mm, thus altering the synchronicity of the visuo-haptic collision. The user was also able to see a virtual representation of the haptic interface, to avoid hitting it by accident. They also saw a virtual representation of their hands.

## Conditions and Feedback

Users experienced three faces of the cube, as illustrated in [Figure 8.5](#):

- **Top:** the cube is positioned right in front of the participant, slightly lower than their head, and then they can bring their index finger down to hit the cube.
- **Back:** with the same view as **Top**, the participant's arm goes around the cube, and they move their finger towards them to hit the cube. Here, they cannot see the contact point due to occlusion from the cube.
- **Side:** the cube is positioned at the same height as the user's head, and the right face is positioned at the center of their field of view so that they can barely see the face, but can clearly see the contact point when they hit the cube by moving their finger to the left.

---

2. <https://developer.vive.com/resources/hardware-guides/vive-pro-eye-specs-user-guide/>

3. [https://www.ultraleap.com/datasheets/STRATOS\\_Explore\\_Development\\_Kit\\_datasheet.pdf](https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf)

4. <https://www.ultraleap.com/product/stereo-ir-170/>

5. <https://unity.com/>

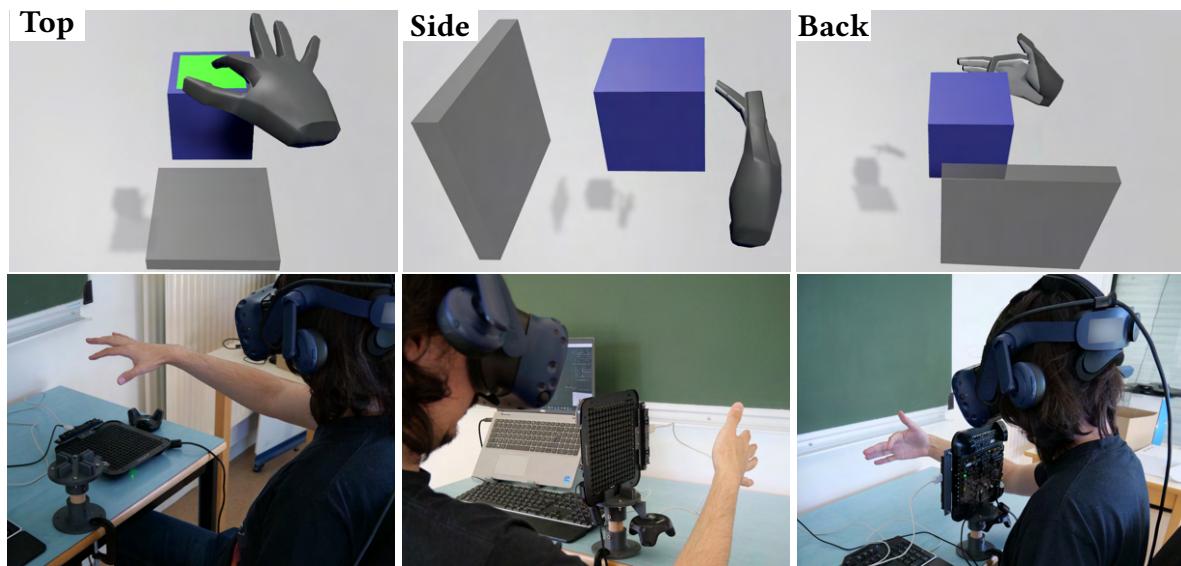


Figure 8.5 – Experimental configuration (Bottom) and associated virtual point of view (Top) when interacting with each face of the cube.

We also used two different starting positions for the invisible plane that triggers the haptic feedback:

- **Out**: the plane is positioned 2 cm away from the cube’s face, outside of the cube.
- **In**: the plane is positioned 2 cm away from the face, inside of the cube.

When a collision was detected between the fingertip and the invisible plane, the UMH interface rendered a circle of radius 1 cm on the user’s fingertip for 1 s. The circle was rendered using STM, with a focal point speed of 8 m/s [Frier et al. 2019].

### Procedure

The experiment was composed of three blocks, for the three faces, and each participant experienced a different block order. Each block was composed of six trials, corresponding to three repetitions of the task, for each starting position of the invisible plane. Trials alternated between **In** and **Out** conditions, with half the participants starting each block with **In**, and the other half with **Out**. This led to 18 trials in total, and the experiment usually lasted between 15 and 20 minutes.

For each trial, the users were asked to tap on the cube’s face with their right index finger. Then, depending on whether they perceived the haptic feedback before or after the visual collision between the finger and the cube, they could move the invisible plane inwards or outwards

respectively. Once the haptic feedback feels synchronized, they can press a validation button with their virtual finger to end the trial. Otherwise, the interaction was free: there was no time limit, nor indication to follow a specific motion or speed.

### Collected Data

At the end of each trial, we collected the distance between the virtual plane and the cube's face.

Participants also filled a demographics questionnaire prior to the experiment. After the experiment, participants also rated the difficulty of the task for each face of the cube, on a 7-points Likert scale.

### Participants

For this preliminary study, we recruited 6 participants (6 men, aged ( $m \pm s.d. = 25.8 \pm 5.3$ ), all right-handed. Three of them had a lot of experience with VR and four had a lot of experience with haptics. Five of them had already experienced UMH stimuli at some point in the past. It is important to note that this study is a pilot study, and only serves as a calibration step. This is why the relatively low amount of participant, and the gender and experience skew are sufficient for our needs. For further studies on these topics, a larger and more diverse population would be required.

### 8.3.2 Results and Discussion

Measured results are shown in [Figure 8.6](#), and summarised in [Table 8.1](#). Questionnaire results are also summarised in [Table 8.1](#).

Table 8.1 – Result summary for the pilot study ([Section 8.3](#)). (Left) Distance from the haptic triggering plane to the cube's face. (Right) Self-reported difficulty rating (lower is easier). Values are reported as mean  $\pm$  standard deviation.

Face	Start Position	Distance (mm)	Face	Difficulty Rating (/7)
Top	Out	$-7.6 \pm 7.2$	Top	$3.2 \pm 1.2$
Top	In	$7.8 \pm 6.8$	Side	$2.5 \pm 1.0$
Side	Out	$-7.2 \pm 6.4$	Back	$4.8 \pm 1.2$
Side	In	$8.9 \pm 4.9$		
Back	Out	$-11.6 \pm 5.2$		
Back	In	$13.3 \pm 3.9$		

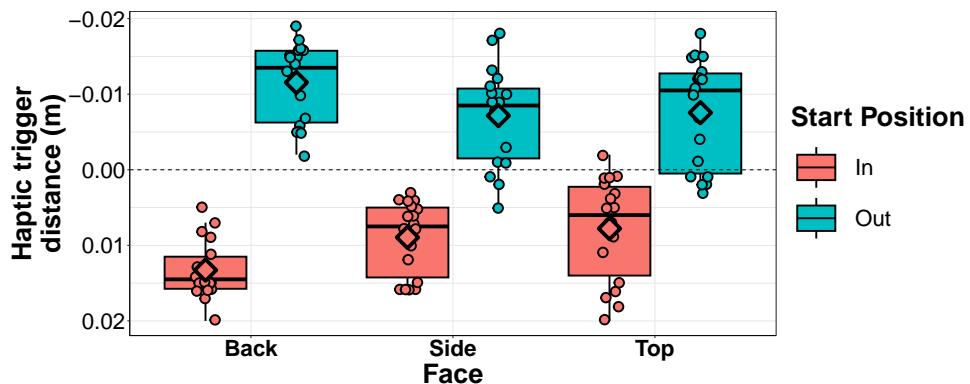


Figure 8.6 – Position of the invisible plane at the end of each trial of the pilot study ([Section 8.3](#)). Each dot represents one trial, and the diamond represents the mean. The horizontal line at  $y = 0$  represents the actual position of the cube’s face. Note that the y axis is inverted to have a more coherent representation for the top face, for which negative interpenetration would be above the face.

After averaging each participant’s data per start position, a Shapiro-Wilk test failed to reject the normality hypothesis for both the **In** and **Out** data, we performed a t-test, revealing a significant effect of the start position ( $t(5) = 5.88, p = 0.002$ ). This shows that synchronicity does not seem to have a very important role here. As long as the haptic feedback is triggered within about 1 cm of the visual collision, participants seem to perceive it as synchronous. Yet, it is worth noting that we would expect this range to become narrower if participants were to move very slowly. To fit most scenarios for the following experiments, we decided to set the position of the haptic triggering plane at the same position as the cube’s face, meaning that the haptic feedback is emitted at the exact moment the virtual finger collides with the cube.

After averaging the data per face, Shapiro tests also failed to reject the normality hypothesis for each of them. An ANOVA did not reveal any significant effect of the face on the haptic trigger distance. A Friedman test revealed a significant effect of the face on the difficulty rating ( $\chi^2(2, N = 6), p = 0.006$ ), but pairwise Wilcoxon signed-rank tests revealed no significant differences due to the Holm-correction and low amount of participants. Yet, by looking at the data ([Table 8.1](#)), the **Back** face seem to be slightly more difficult, likely due to the occlusion.

It is worth noting that these results should be used as a way to calibrate future experiments, and hint towards some trends, but more participants should be recruited to strengthen these results.

## 8.4 User Study: Generating Coupling Schemes

While our end goal is to find the best visuo-haptic feedback for interpenetration, trying every pair against each other would result in a very long experiment. We therefore divided this process in two parts. First, we find the best haptic feedback for each visual feedback, resulting in a subset of candidate pairs. Then (see [Section 8.5](#)) we compare the performances of these selected visuo-haptic pairs.

### 8.4.1 Hypotheses

For this first study, we make the following hypotheses:

- **H1.a:** people prefer coherent visuo-haptic cues.
- **H1.b:** people prefer having visual cues (i.e. **Normal Hand** is less appreciated than the other visual cues).

### 8.4.2 Materials and Methods

#### Conditions

For this experiment, participants experienced all pairs of visuo-haptic feedback presented in [Section 8.2](#). Namely, they experienced four visual feedbacks: **Normal Hand**, **Visual Particles**, **G-OFC**, and **Ghost Hand**. They also experienced three haptic conditions: **Haptic Particles**, **Line**, and **Circle**. For this experiment, only the **Top** face of the cube was considered.

#### Apparatus and Setup

The physical setup for this experiment was similar to that of the pilot study (see [Section 8.3.1](#)), except that we did not need the stand for the UMH interface, as interactions were only performed with the top face.

The virtual scene was slightly different: three virtual buttons, labelled “A” to “C”, were positioned above the cube, corresponding to the three haptic feedback techniques. Their order was randomly chosen at the start of the trial, and users could change the selected haptic feedback by pressing the top arrow key with their left hand. They could also change the ranking of the selected feedback using the left and right arrow keys. There was also a validation button, that could be pressed with their virtual hand to lock their results and proceed to the next trial.

## Procedure

This study was composed of 5 repetitions of 4 trials, whose order followed a Latin square. Each trial corresponded to one of the visual feedback techniques.

Participants were asked to tap on the cube with each of the three haptic feedbacks, and rank them by order of preference (left to right). They were allowed to freely alternate between haptic feedbacks, and experience each of them for as long as they wanted. Similarly to the pilot study ([Section 8.3.1](#)), the interaction with the cube was mostly free, as users were only asked to tap with their index finger.

As the haptic feedback order was randomized between each trial, button “A” did not always refer to the same feedback. This was done to reduce potential learning effects. Overall, the experiment was composed of 20 trials, and lasted from 10 to 20 minutes.

## Collected Data

After each trial, we gathered their final haptic ranking, and the amount of time they were stimulated by each feedback.

Participants also filled a demographics questionnaire before the experiment. After the experiment, they were also asked to evaluate each visual feedback on a 7-points Likert scale regarding the precision of contact, the aestheticism, and the difficulty of the task.

## Participants

We recruited 16 participants (14 men, 1 women, aged  $m \pm s.d. = 34.4 \pm 9.4$ ), all right-handed. Six of them had a lot of experience with VR, while ten had little to no prior experience. Similarly, seven had a lot of experience with haptics, and 13 had already experienced UMH at least once before.

### 8.4.3 Results

Overall results are summarized in [Table 8.2](#).

After averaging the data over all repetitions, we performed a Shapiro-Wilk test on the ranking (shown in [Figure 8.7](#) (Top)) values for each (Visual, Haptic) subset of the data, showing that most of them did not follow a normal distribution ( $p < 0.05$ ). Therefore, for each visual feedback, we performed a Friedman test, with the haptic feedback as independent variable,

Table 8.2 – Result summary for Experiment 1 (Section 8.4). (Top) Data obtained during the VR interaction: ranking of each haptic feedback for a given visual (1-3, lower is better), and time spent stimulated with each haptic feedback. (Bottom) Results of the post-experiment questionnaires, about the difficulty of the task with each visual feedback, their aestheticism, and the perceived precision of contact. “S” values represent the given score on a 7-points Likert scale (lower is better for the difficulty), and “R” values represent the ranking (1-4, lower is better). Values are reported as mean  $\pm$  standard deviation.

Visual Feedback	Haptic Feedback	Rank (/3)	Exploration Duration (s)
<b>Normal Hand</b>	<b>Circle</b>	$1.3 \pm 0.6$	$6.1 \pm 6.1$
	<b>Haptic Particles</b>	$2.1 \pm 0.5$	$7.2 \pm 8.0$
	<b>Line</b>	$2.6 \pm 0.7$	$6.5 \pm 6.8$
<b>G-OFC</b>	<b>Circle</b>	$1.4 \pm 0.7$	$4.7 \pm 4.0$
	<b>Haptic Particles</b>	$2.1 \pm 0.6$	$6.1 \pm 4.6$
	<b>Line</b>	$2.5 \pm 0.7$	$5.6 \pm 4.4$
<b>Ghost Hand</b>	<b>Circle</b>	$1.4 \pm 0.7$	$8.0 \pm 5.9$
	<b>Haptic Particles</b>	$2.1 \pm 0.7$	$8.3 \pm 7.0$
	<b>Line</b>	$2.5 \pm 0.7$	$7.4 \pm 5.7$
<b>Visual Particles</b>	<b>Circle</b>	$1.4 \pm 0.7$	$5.9 \pm 5.3$
	<b>Haptic Particles</b>	$2.1 \pm 0.6$	$6.7 \pm 5.6$
	<b>Line</b>	$2.5 \pm 0.7$	$5.7 \pm 4.9$
Visual Feedback	Difficulty	Aestheticism	Contact Precision
<b>Normal Hand</b>	$S : 2.9 \pm 1.3$	$S : 3.8 \pm 1.8$	$S : 3.9 \pm 1.7$
	$R : 2.8 \pm 1.2$	$R : 3.3 \pm 0.9$	$R : 3.5 \pm 0.9$
<b>G-OFC</b>	$S : 2.8 \pm 2.0$	$S : 6.3 \pm 0.8$	$S : 6.1 \pm 1.3$
	$R : 1.7 \pm 1.0$	$R : 1.4 \pm 0.6$	$R : 1.4 \pm 0.8$
<b>Ghost Hand</b>	$S : 3.4 \pm 1.5$	$S : 4.7 \pm 1.6$	$S : 5.1 \pm 1.2$
	$R : 2.9 \pm 0.9$	$R : 2.6 \pm 0.8$	$R : 2.9 \pm 0.7$
<b>Visual Particles</b>	$S : 3.0 \pm 1.6$	$S : 4.5 \pm 1.9$	$S : 4.9 \pm 1.8$
	$R : 2.6 \pm 1.0$	$R : 2.7 \pm 1.2$	$R : 2.1 \pm 1.0$

the participant as random variable, and the feedback’s ranking as dependent variable. Each of them showed a significant effect of the haptic feedback ( $\chi^2(2, N = 16) = 10.1, p = 0.006$  for **Normal Hand**,  $\chi^2(2, N = 16) = 15.1, p < 0.001$  for **G-OFC** and **Ghost Hand**,  $\chi^2(2, N = 16) = 9.1, p = 0.01$  for **Visual Particles**). Post-hoc analyses were performed using pairwise Wilcoxon signed-rank tests on paired data, with Holm corrections. These resulted in **Circle** being always statistically significantly ranked better than **Line** and **Haptic Particles**. **Haptic Particles** was also ranked significantly better than **Line** for the **Normal Hand** conditions.

Similarly, a non-parametric ANOVA based on the aligned rank transform (ART) similarly revealed a significant effect of the haptic feedback, but no significant effect of the visual feedback, nor any interaction between both types of feedbacks. Post-hoc analyses using the ART-C procedure (a method suitable for our non-parametric factorial design) and Holm corrections showed significant differences in ranking between all three haptic feedback, with **Circle** being preferred to **Visual Particles** ( $t(165) = -7.0, p < 0.001$ ), which is also preferred to **Line** ( $t(165) = -5.0, p < 0.001$ ).

Shapiro-Wilk tests also indicated that exploration duration data did not follow a normal distribution for most groups ( $p < 0.05$ ). Similarly to before, we performed an ART-based ANOVA, revealing a significant effect of the visual ( $F(3, 165) = 12.4, p < 0.001$ ), and haptic ( $F(2, 165) = 4.0, p = 0.02$ ) feedback, but no significant interactions. Similar post-hoc analyses revealed that **G-OFC** was used for a significantly longer time than the other three visual feedbacks ( $t(165) = 5.8, p < 0.001$  for **Normal Hand**,  $t(165) = 4.6, p < 0.001$  for **Ghost Hand**, and  $t(165) = 3.8, p < 0.001$  for **Visual Particles**). All other visual pairwise comparisons were non-significant. In terms of haptics, **Circle** was used for significantly less time than **Haptic Particles** ( $t(165) = 2.7, p = 0.02$ ).

Regarding the final questionnaires, we used the same analysis protocol for each metric: Friedman tests were conducted, using the participant as random variable, and if significant results were observed, we performed additional Holm-corrected pairwise Wilcoxon signed-rank tests on paired data. Specifically, we observed a significant effect of the visual feedback on the perceived aestheticism ( $\chi^2(3, N = 16) = 19.0, p < 0.001$  for the rank,  $\chi^2(3, N = 16) = 16.5, p < 0.001$  for the score), with **G-OFC** being preferred to the other three feedbacks, both in rank and score. In terms of difficulty of the task, no significant difference was observed in terms of score, but there was one in terms of rank ( $\chi^2(3, N = 16) = 9.2, p = 0.03$ ). Specifically, **G-OFC** was ranked as easier than **Ghost Hand** ( $V = 16, p = 0.04$ ). Finally, a significant effect was observed on both the score ( $\chi^2(3, N = 16) = 17.5, p < 0.001$ ) and ranking ( $\chi^2(3, N = 16) = 23.6, p < 0.001$ ) related to the perceived precision of contact ( $p < 0.001$  for both). Specifically, **Visual Particles** ( $V = 14, p = 0.02$ ) and **G-OFC** ( $V = 6, p = 0.006$ ) were ranked better than **Normal Hand**, and **G-OFC** was ranked better than **Ghost Hand** ( $V = 4, p = 0.005$ ). In terms of score, only **G-OFC** was scored better than **Normal Hand** ( $V = 97, p = 0.03$ ).

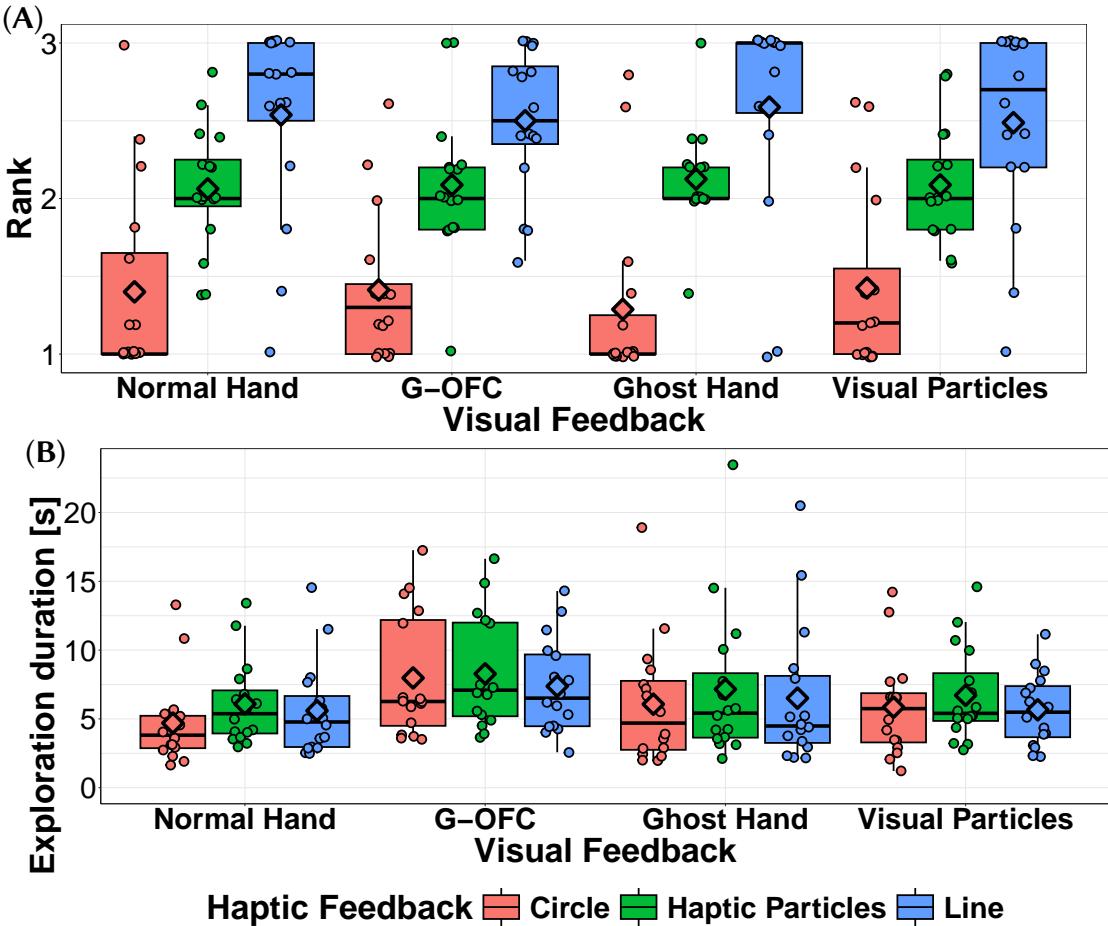


Figure 8.7 – Data obtained from Experiment 1. (A) Haptic feedback ranking for all visual conditions. Lower results indicate a preferred feedback. (B) Time spent in collision with the cube, receiving feedback. For both graphs, each point represents a participant’s average value over all repetitions of the same conditions, and diamonds represent the overall mean value for the whole condition.

#### 8.4.4 Discussion

Interestingly, **Circle** was reported to be the preferred haptic feedback with every visual feedback, which has several implications.

First, this does not support **H1.a**. Indeed, we believe that **Visual Particles** and **Haptic Particles** have a more coherent dynamic, with small elements moving in a random way. Similarly, we expect **G-OFC** to be more coherent with **Line** as they stimulate the same part of the finger. Yet, the consistent preference for the **Circle** feedback, regardless of the visual condition, suggests that its higher perceived intensity at contact onset was the dominant perceptual factor,

overriding any potential effects of visuo-haptic coherence. Indeed, the **Line** starts small, leading to a weaker perceived intensity at first. Similarly, the haptic circles from **Haptic Particles** feel weaker than **Circle** because they do not dwell at the same position for as much time, and because part of each small circle lays outside the finger, and does not contribute to the received energy by the user. This reveals that the impact phase is likely the most important part of contact for the users, when it comes to interactions with UMH. Although **H1.a** is not supported with our data, we should not reject it either. Further studies should be conducted with UMH feedback controlled to have the same perceived intensity. This way, the effect of intensity should be nullified, and the effect of visuo-haptic coherence may be stronger.

Secondly, the strong preference for the **Circle** feedback means that our candidates for the best visuo-haptic coupling scheme will all use **Circle** as their haptic component. Then, Experiment 2 ([Section 8.5](#)) explores which of these coupling candidates is the best for interpenetration rendering.

When it comes to appreciation of the visual feedback, **G-OFC** stands out as being found more aesthetically pleasing than all other visuals. **Normal Hand** was also ranked as providing a worth precision of contact than **G-OFC** and **Visual Particles**. This partially supports **H1.b**, and looking at the data ([Table 8.2](#)), we expect the effect to be more significant with more participants.

It is worth noting that the exploration duration also correlates with the previously discussed results. First, **G-OFC** was reported as being used for a longer time than all other visual feedbacks, which correlates with the higher aesthetics scores it received. As the participants found the feedback pleasant, they spent more time experiencing it. Similarly, the lower stimulation time for **Circle** may indicate that it was easy to recognize, or that users found it easier to rank it, compared to the other haptic feedbacks.

## 8.5 User Study: Comparing Coupling Schemes

Since the Experiment 1 ([Section 8.4](#)) allowed us to find a few candidates, we now need to compare them to find the best visuo-haptic coupling for interpenetration rendering with UMH. Specifically, we want to evaluate the user's precision of contact when experiencing each coupling, which would serve as an objective performance metric to compare them.

### 8.5.1 Hypotheses

For this final experiment, we made the following hypotheses:

- **H2.a:** Delivering visuo-haptic feedback improves performances (i.e. reduces interpenetration).
- **H2.b:** The back face of the cube will lead to the worst performances.
- **H2.c:** The side face of the cube will lead to the best performances.
- **H2.d:** Performances will be better when users try to remain at the surface of the cube, than when they tap on the face.

### 8.5.2 Materials and Methods

#### Conditions

Here, participants interacted with the three faces of the cube: **Top**, **Side**, and **Back**.

They interacted while receiving feedback from one of five techniques, corresponding to the four best visuo-haptic coupling schemes resulting from the first experiment ([Section 8.4](#)), and a fifth technique consisting of the **Normal Hand** visual with no additional haptic feedback. This last condition serves as a control condition, and represents what a basic VR interaction would feel like. We refer to the no-haptic condition as **Normal Hand + None**, and all others as **X + H**, where **X** is the name of the visual feedback.

Finally, participants experience these feedbacks under two different tasks: either they are asked to quickly **Tap** five times on the surface of the cube, or to remain for 5 s at the surface of the cube (**Hold**).

For the training phases, participants experienced the **Normal Hand** with **Circle** visuo-haptic feedback. This pair was chosen as **Normal Hand** is the simplest visual feedback, while **Circle** is the only haptic feedback that was selected as a result of the first experiment.

#### Apparatus and Setup

The physical setup for this experiment was also similar to that of the pilot study (see [Section 8.3](#)), except that users no longer needed to use their left hand or the keyboard.

The virtual scene was also very similar, with just the addition of some progress indicators, to let the user know how much of the task has been performed.

## Procedure

The experiment was divided into two parts, corresponding to the two tasks. Half of the participants experienced **Tap** before **Hold**, and the other half experienced the reversed order. At the very beginning of each part, users underwent a training phase, composed of 5 trials, where they completed the task while receiving the **Normal Hand** and **Circle** visuo-haptic feedback.

Each part is composed of three blocks, corresponding to the three faces of the cube. All six block orders were experienced by the same amount of participants. The block order for the first task was identical to that of the second task. The experimenter took care of changing the orientation of the haptic interface at the beginning of each block. Each block was composed of three repetitions of five trials, corresponding to the five visuo-haptic conditions. The order of trial was fully randomized. During each trial, users were asked to interact in a natural way, while trying to keep interpenetration as low as possible.

After completing each trial, users had to move their hand away from the cube, to make sure that they have to find the position of the face at each trial. Overall, the experiment was composed of 90 trials, for a duration of about 30 to 40 minutes, including trainings and questionnaires.

## Collected Data

During this experiment, we measured interpenetration with cube through the whole interaction. From this, we computed the average, and maximum interpenetration per trial. We also measured the movement speed of the fingertip, and computed the average and maximum per trial.

Participants filled a demographics questionnaire before the experiment. After each block, they were asked to rank all visuo-haptic feedback, and to score them on a 7-point Likert scale regarding the difficulty of the task when experiencing each of them. After the experiment, they were asked to rank and score:

- the faces regarding the difficulty of the task.
- the four visual feedbacks by aestheticism.
- the five visuo-haptic pairs by coherence between the visual and haptic cues.

## Participants

We recruited 18 participants (16 men, 2 women), aged 18 to 33 (mean =  $24.2 \pm 3.1$ ), with 16 of them being right-handed. 11 had a lot of experience with VR, while 9 had a lot of experience with haptics, and 8 had already experienced an UMH stimulus in the past.

### 8.5.3 Results

#### Tap

Obtained data during the **Tap** part of the experiment are illustrated in [Figure 8.8](#), and summarized in [Table 8.3](#).

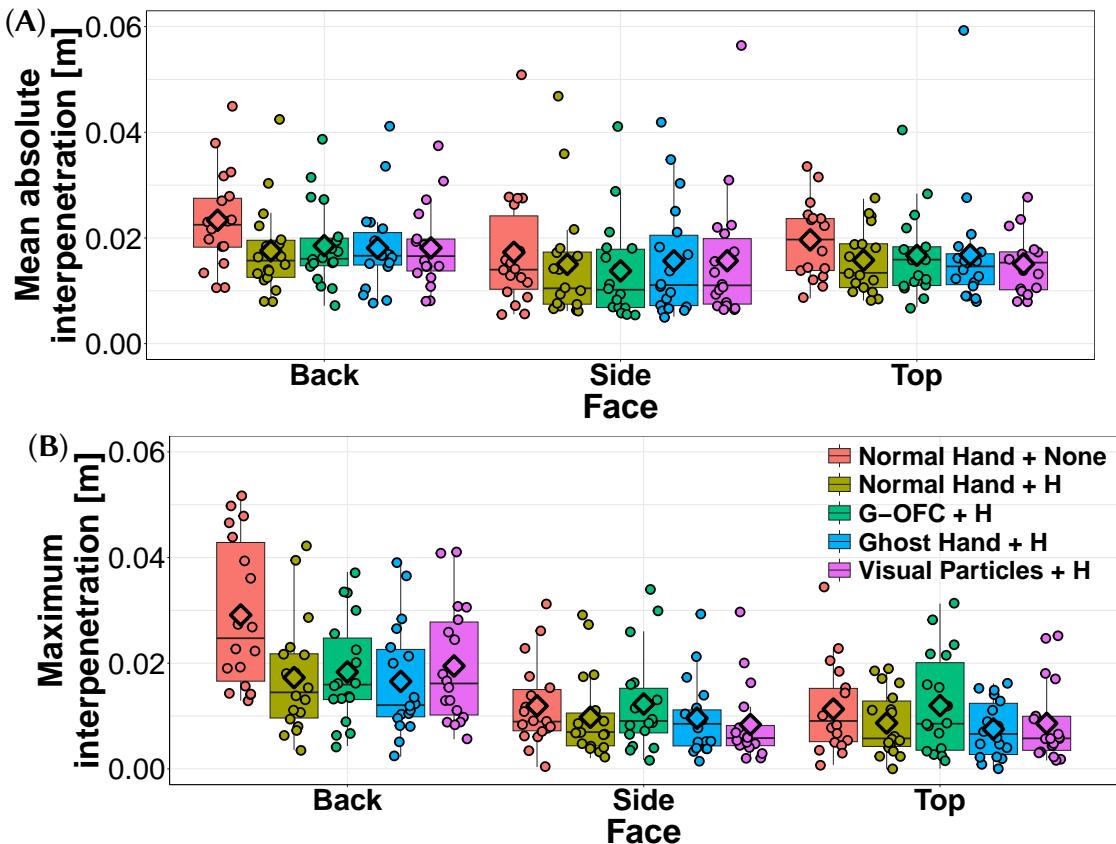


Figure 8.8 – Distribution of the mean absolute interpenetration (A) and maximum interpenetration (B) obtained during the **Tap** part of Experiment 2 ([Section 8.5](#)). For (A), six points between 6 mm and 8.2 mm are not shown for readability purposes. Each point represents a participant's individual results averaged over all repetitions. Diamonds represent the overall average for each condition.

Table 8.3 – Result summary for the **Tap** part of Experiment 2 (Section 8.5). (Top) Average absolute interpenetration, and maximum interpenetration distances for each cube’s face. Note that the maximum can be lower than the mean absolute, if the user goes far outside the cube, which corresponds to a negative interpenetration. (Bottom) Interpenetration as for (Top), and rated difficulty for each coupling scheme. “S” values represent the score (/ 7), and “R” values represent the ranking (/ 5), with lower values indicating an easier task. All values are represented as mean  $\pm$  standard deviation.

Face	Mean Absolute Interpenetration (mm)	Maximum Interpenetration (mm)	
<b>Top</b>	$19.2 \pm 14.0$	$9.6 \pm 7.6$	
<b>Side</b>	$16.6 \pm 12.9$	$10.4 \pm 7.9$	
<b>Back</b>	$19.1 \pm 8.5$	$20.2 \pm 12.1$	

Coupling	Mean Absolute Interpenetration (mm)	Maximum Interpenetration (mm)	Difficulty
<b>Normal Hand + None</b>	$21.2 \pm 12.7$	$17.5 \pm 13.3$	$S : 5.5 \pm 1.3$ $R : 5.0 \pm 0.1$
<b>Normal Hand + H</b>	$18.0 \pm 13.0$	$11.9 \pm 9.2$	$S : 3.5 \pm 1.2$ $R : 3.7 \pm 0.6$
<b>G-OFC + H</b>	$18.1 \pm 12.3$	$14.2 \pm 9.9$	$S : 2.6 \pm 1.3$ $R : 2.3 \pm 1.1$
<b>Ghost Hand + H</b>	$16.9 \pm 10.3$	$11.2 \pm 8.7$	$S : 2.2 \pm 1.1$ $R : 1.7 \pm 0.8$
<b>Visual Particles + H</b>	$17.4 \pm 11.8$	$12.2 \pm 10.1$	$S : 2.8 \pm 1.1$ $R : 2.3 \pm 0.8$

First, we averaged the measured interpenetration data over all three repetitions. As the interpenetration data did not follow a normal distribution (Shapiro-Wilk,  $p < 0.05$ ) for many conditions, we performed an ART-based ANOVA, revealing a significant effect of both the face and feedback on the mean absolute interpenetration ( $F(2, 238) = 24.4, p < 0.001$  for the Face,  $F(4, 238) = 8.0, p < 0.001$  for the feedback), but no significant interaction effect. Post-hoc analyses using ART-C with Holm corrections revealed that there was significantly higher interpenetration with the **Back** face than with the **Top** one ( $t(238) = 2.3, p = 0.02$ ), which also had higher interpenetration than with the **Side** face ( $t(238) = 4.5, p < 0.001$ ). A similar method was used to show that **Normal Hand + None** led to significantly more interpenetration than all other conditions ( $t(238) = 4.3$  against **Normal Hand + H**,  $t(238) = 4.1$  for **G-OFC + H**,  $t(238) = 4.8$  for **Ghost Hand + H**, and  $t(238) = 4.5$  against **Visual Particles + H**,  $p < 0.001$  for all four comparisons).

By applying the same methodology to the maximum interpenetration distance, we found significant effects of the face ( $F(2, 238) = 88.6, p < 0.001$ ) and feedback ( $F(4, 238) = 9.7, p < 0.001$ ), as well as a significant interaction effect ( $F(8, 238) = 3.0, p = 0.003$ ). The post-hoc analysis revealed that the **Back** face led to more interpenetration than **Side** ( $t(238) = 11.0, p < 0.001$ ) and **Top** ( $t(238) = 12.0, p < 0.001$ ). Similar to the mean absolute interpenetration, **Normal Hand + Hand** led to a significantly higher maximum interpenetration than all other conditions ( $t(238) = 4.8, p < 0.001$  against **Normal Hand + H**,  $t(238) = 2.9, p = 0.03$  for **G-OFC + H**,  $t(238) = 5.4, p < 0.001$  against **Ghost Hand + H**, and  $t(238) = 4.7, p < 0.001$  for **Visual Particles + H**).

We also used the same methods on the difficulty scores and rankings obtained from the questionnaires after each block. These tests revealed significant effects of the the face ( $F(2, 238) = 15.5, p < 0.001$ ) and feedback ( $F(4, 238) = 97.0, p < 0.001$ ) on the scores, as well as significant interaction effects ( $F(8, 238) = 2.9, p = 0.004$ ). Pos-hoc tests showed that the **Back** face led to more difficult ratings than **Side** ( $t(238) = 5.2, p < 0.001$ ) and **Top** ( $t(238) = 4.4, p < 0.001$ ). All feedback pairwise comparisons were also significant, apart from the **G-OFC + H** vs. **Visual Particles + H** one. Overall, this shows that the task is easier with **Ghost Hand + H**, slightly harder for both **G-OFC + H** and **Visual Particles + H**, even harder for **Normal Hand + H**, and even harder for **Normal Hand + None**. On the other hand, only the feedback had a significant effect on the difficulty ranking ( $F(4, 238) = 150.0, p < 0.001$ ), and the pairwise analysis gave the same results as for the difficulty scores.

## Hold

Obtained data during the **Hold** part of the experiment are illustrated in [Figure 8.9](#), and summarized in [Table 8.4](#).

We applied the same methodology as for the **Tap** analysis. We observed significant effects of the face ( $F(2, 238) = 121.9, p < 0.001$ ), feedback ( $F(4, 238) = 34.1, p < 0.001$ ), and interaction effects ( $F(8, 238) = 8.1, p < 0.001$ ) on the mean absolute interpenetrations. Specifically, all pairwise face comparisons were significant, with **Back** leading to more interpenetration than **Side** ( $t(238) = 12.3, p < 0.001$ ), which itself led to more interpenetration than **Top** ( $t(238) = 2.1, p = 0.03$ ). All pairwise feedback comparisons were significant, apart from the **Normal Hand + None** vs. **G-OFC + H**, and **Normal Hand + H** vs. **Visual Particles + H**. Overall, **Ghost Hand + H** leads to the lowest interpenetration, followed by **Virtual Particles + H** and **Normal Hand + H**, followed by **G-OFC + H** and **normal Hand + None**.

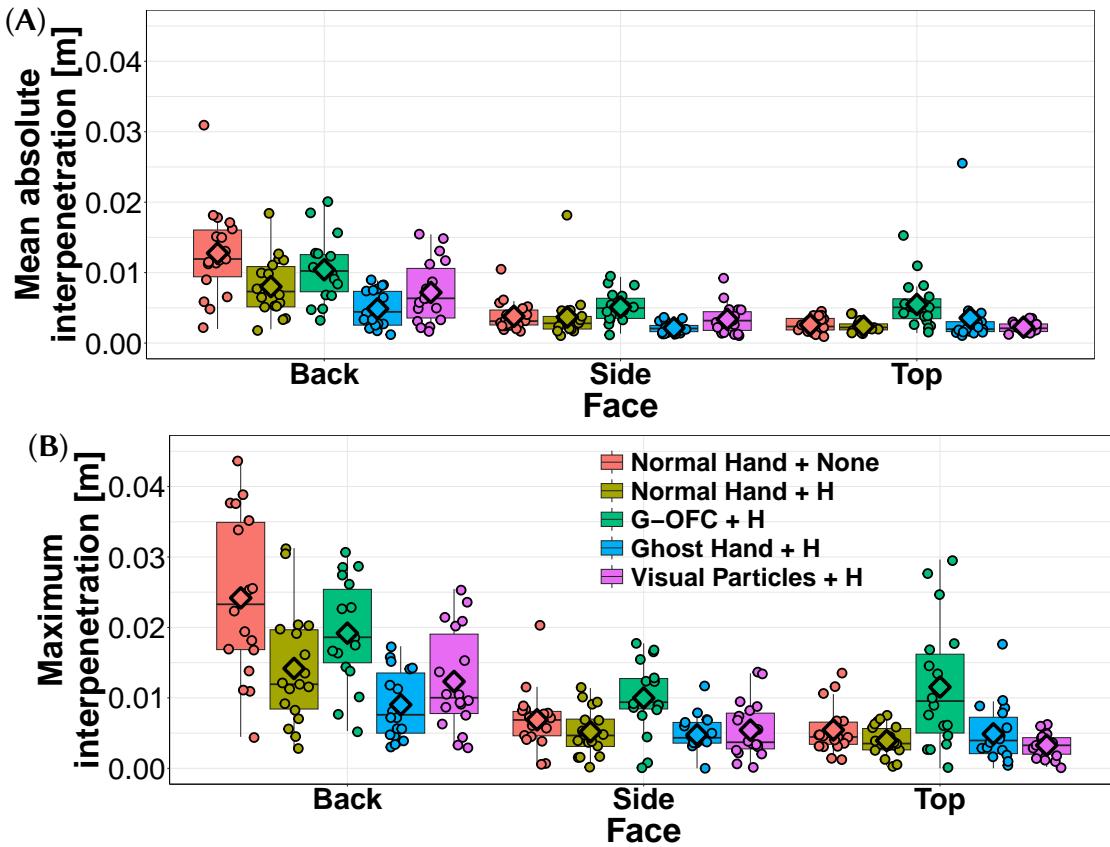


Figure 8.9 – Distribution of the mean absolute interpenetration (A) and maximum interpenetration (B) obtained during the **Hold** part of Experiment 2 (Section 8.5). Each point represents a participant’s individual results averaged over all repetitions. Diamonds represent the overall average for each condition.

For the maximum interpenetration, we observed exactly the same significant effects, except that the **Ghost Hand + H** vs. **Visual Particles + H** is not significant.

Significant effects of the face ( $F(2, 238) = 33.1, p < 0.001$ ), feedback ( $F(4, 238) = 138.9, p < 0.001$ ), and interaction effects ( $F(8, 238) = 2.1, p = 0.04$ ) were observed on the reported difficulty scores. Specifically, the **Back** face was again reported as significantly harder than the other two faces ( $t(238) = 7.1$  against **Side**,  $t(238) = 7.0$  against **Top**,  $p < 0.001$  in both cases), and all pairwise feedback comparisons were significant, which orders the feedback as follows: **Ghost Hand + H**, **Visual Particles + H**, **G-OFC + H**, **Normal Hand + H**, **Normal Hand + None** by increasing difficulty. For the difficulty ranking, only the feedback had a significant effect ( $F(4, 238) = 244.5, p < 0.001$ ), and all pairwise comparisons were also significant, revealing the same order as the scores.

Table 8.4 – Result summary for the **Hold** part of Experiment 2 (Section 8.5). All metrics are the same as Table 8.3.

Face	Mean Absolute Interpenetration (mm)	Maximum Interpenetration (mm)	
<b>Top</b>	$3.3 \pm 3.1$	$5.8 \pm 5.6$	
<b>Side</b>	$3.6 \pm 2.5$	$6.4 \pm 4.3$	
<b>Back</b>	$8.6 \pm 5.3$	$15.8 \pm 9.5$	
Coupling	Mean Absolute Interpenetration (mm)	Maximum Interpenetration (mm)	Difficulty
<b>Normal Hand + None</b>	$6.4 \pm 6.0$	$12.2 \pm 11.2$	$S : 5.5 \pm 1.2$ $R : 5.0 \pm 0.2$
<b>Normal Hand + H</b>	$4.7 \pm 4.0$	$7.8 \pm 6.9$	$S : 3.6 \pm 1.2$ $R : 3.7 \pm 0.5$
<b>G-OFC + H</b>	$7.0 \pm 4.2$	$13.6 \pm 8.2$	$S : 3.1 \pm 1.3$ $R : 2.9 \pm 1.0$
<b>Ghost Hand + H</b>	$3.5 \pm 3.7$	$6.2 \pm 4.4$	$S : 1.9 \pm 0.9$ $R : 1.4 \pm 0.7$
<b>Visual Particles + H</b>	$4.3 \pm 3.5$	$7.0 \pm 6.1$	$S : 2.6 \pm 1.2$ $R : 2.1 \pm 0.7$

## Final Questionnaires

Results from the final questionnaires are summarized in Table 8.5.

Friedman tests revealed a significant effect of the face on the reported overall difficulty scores ( $\chi^2(2, N = 18) = 28.8, p < 0.001$ ) and rankings ( $\chi^2(2, 18) = 26.7, p < 0.001$ ), with pairwise Holm-corrected Wilcoxon tests on paired data showing that **Back** was perceived as significantly more difficult than **Side** ( $V = 171, p < 0.001$  for both) and **Top** ( $V = 171$  for the score,  $V = 153$  for the rank,  $p < 0.001$  for both).

With the same methodology, we observed a significant effect of the visual feedback on the aestheticism scores ( $\chi^2(3, N = 18) = 18.0, p < 0.001$ ) and rankings ( $\chi^2(3, 18) = 21.4, p < 0.001$ ), with **Normal Hand** being worse than all other techniques.

We also observed a significant effect of the coupling scheme on the reported visuo-haptic coherence, with **Normal Hand + None** being worse than all the others.

Table 8.5 – Result summary for final questionnaires of Experiment 2 (Section 8.5). (Left) Global difficulty scores (/ 7) and rankings (/ 3) for each face of the cube. Lower values represent easier tasks. (Right) Aestheticism and visuo-haptic coherence scores (/ 7) and rankings (/ 4 for aestheticism, / 5 for coherence) for all coupling schemes. “S” values represent the score, and “R” values represent the ranking. Higher scores and lower rankings indicate a better feedback. All values are represented as mean  $\pm$  standard deviation.

Face	Difficulty	Coupling	Aestheticism	Visuo-Haptic Coherence
<b>Top</b>	$S : 1.2 \pm 2.2$	<b>Normal Hand + None</b>	$S : 3.2 \pm 5.4$	$S : 2.2 \pm 1.6$
	$R : 1.4 \pm 0.6$		$R : 3.7 \pm 0.7$	$R : 4.7 \pm 0.8$
<b>Side</b>	$S : 2.6 \pm 1.0$	<b>Normal Hand + H</b>		$S : 4.1 \pm 1.4$
	$R : 1.7 \pm 0.5$			$R : 3.4 \pm 1.1$
<b>Back</b>	$S : 5.4 \pm 0.9$	<b>G-OFC</b>	$S : 5.4 \pm 1.6$	$S : 5.0 \pm 1.8$
	$R : 3.0 \pm 0.0$		$R : 1.8 \pm 1.0$	$R : 2.4 \pm 1.3$
		<b>Ghost Hand</b>	$S : 5.3 \pm 1.3$	$S : 5.4 \pm 1.2$
			$R : 2.3 \pm 0.8$	$R : 2.2 \pm 1.2$
		<b>Visual Particles</b>	$S : 5.1 \pm 1.6$	$S : 5.4 \pm 1.1$
			$R : 2.3 \pm 1.0$	$R : 2.2 \pm 0.8$

## Other Results

Two Wilcoxon signed-rank tests showed that both the mean absolute ( $V = 0, p < 0.001$ ), and maximum interpenetration ( $V = 23, p < 0.001$ ) were higher during the **Tap** task (18.3 mm mean absolute, and 13.4 mm maximum interpenetration on average) than during the **Hold** task (5.2 mm mean absolute, and 9.3 mm maximum interpenetration on average). Yet, another Wilcoxon signed-rank test revealed no significant effect of the task on the difficulty score (on average 3.3 / 7 for both). A fourth Wilcoxon signed-rank test ( $V = 0, p < 0.001$ ) revealed that users interacted with a significantly higher speed during the **Tap** condition (12.8 cm/s fingertip speed on average) than during the **Hold condition** (2.6 cm/s speed on average).

### 8.5.4 Discussion

First, we showed that **Normal Hand + None** always performs more poorly than everything else, and is perceived by participants as a more complicated condition. Given that **Normal Hand + H** always shows better performances, we can conclude that adding even just haptics can improve the contact experience with virtual objects. Yet the other coupling schemes further outperform it, which fully supports **H2.a**.

Results show at numerous times that the **Back** face leads to increased interpenetration, and

higher perceived difficulty, which fully supports **H2.b**. This is due to the self-occlusion of the virtual cube. Similarly, the visual cues from **G-OFC** and **Visual Particles** may be fully occluded for a low interpenetration, forcing users to go further, or to only rely on haptic cues.

In most comparisons, no effect was observed between the **Top** and **Side** faces. We only observed that the mean absolute interpenetration was slightly higher on the **Top** during the **Tap** task, and slightly lower during the **Hold** task. But no significant differences were observed in terms of self-reported difficulty. Therefore, **H2.c** is not supported.

Finally, results showed more interpenetration for the **Tap** task than for the **Hold** task, which fully supports **H2.d**. Given that we observed no significant effect of the task on the reported difficulty, and that speed was on average almost five times higher for the **Tap** task, we believe that speed increases the user's range of acceptable interaction. This is why we observe poorer performances, but users do not seem very sensitive to it. This is also likely why the pilot study ([Section 8.3](#)) reported this wide acceptable range for the synchronicity of the haptic feedback.

## 8.6 General Discussion

### 8.6.1 Limitations

First, it is important to note that these results are to be interpreted carefully. Indeed, despite our best efforts, the recruited population is largely male-skewed. Deeper studies with larger and more diverse populations would be required to strengthen our results.

On the other hand, we chose to have relatively natural interactions, by only asking them to tap the cube or to stay at its surface. While this is a scenario that is closer to what would happen in a real VR interaction, this may also introduce different behavioral strategies between participants, which may impact the results. Future studies may want to focus on more constrained scenarios.

Finally, these experiments require the user to keep their arm in the air for a certain amount of time, or to twist their arm and wrist in unusual ways, which may lead to discomfort and fatigue. While we attempted to reduce these issue by using chairs with an armrest, and by ensuring that the experiments are not too long, eye and muscle fatigue may also happen. Specifically, users self-reported their level of fatigue on a 7-point Likert scale before and after

each experiment. For the second experiment, the average fatigue went from 3.2 before the experiment, to 3.7 after the experiment, showing a slight increase.

### 8.6.2 Discussion

Our main objective was to determine which visuo-haptic coupling scheme should be used for UMH interpenetration rendering in VR. When regarding only the performance metrics and perceived difficulty, the **Ghost Hand + H** coupling seem to stand out as providing the best performances. This correlates with free comments given by the users, who reported that the immediate appearance of the whole ghost hand at impact generates a large visual change that is easy to perceive. The fact that it covers the whole hand also makes it easier to see in most configurations, despite self-occlusion of the cube. That said, **G-OFC + H** was reported as having a high aestheticism score in both experiments. This also correlates with the higher exploration duration in Experiment 1, indicating that users liked using this feedback. This appreciation also correlates with some free comments of the participants who found it very natural to use, despite its observed poorer performances (which may be due to the god-object providing a misleading sense of hard limit, encouraging users to apply more force than necessary). So overall, our results reveal a clear trade-off between objective performance and subjective experience. The **Ghost Hand + H** scheme, which explicitly visualizes the user's error, yielded the lowest interpenetration and is thus ideal for precision-critical tasks (e.g. teleoperation). Conversely, **G-OFC + H**, rated as the most aesthetically pleasing despite its weaker objective performance, is better suited for applications prioritizing immersion over task accuracy (e.g. entertainment).

Interestingly, the pilot study and second experiment both indicate that higher speed interactions lead to users being less sensitive to interpenetration. In such high-speed scenarios, providing accurate cues may thus be less important, and accuracy could be traded for a more immersive experience, using **G-OFC + H**.

## 8.7 Conclusion

Despite the absence of any resistance to user motion, UMH devices have gained popularity due to their large design space and airborne nature. Specifically, UMH has been shown to be able to improve user experience for various types of feedback and interactions in VR.

In this paper, we went a step further and combined visual and tactile cues to create interpen-

etration feedback in VR. We explored different visual and haptic feedback, and evaluated the underlying coupling schemes, both in terms of user appreciation, and performances.

Results showed that intense impact-like haptic feedback was always preferred, which suggests that the initial moment of contact is the most perceptually critical phase of the interaction for users. Coupled with intense visual cues, this feedback led to the most accurate contact interactions, while more subtle and more natural cues were more appreciated.

### Chapter Summary

- We proposed visual and UMH feedback techniques, inspired by related works, to provide visuo-haptic interpenetration cues in VR;
- We conducted a pilot study to calibrate the triggering time for the haptic feedback, revealing a low human sensitivity to stimulation latency;
- We conducted a user study to find the best haptic feedback for each of our visual feedbacks, revealing that **Circle** was always preferred, probably due to its higher perceived intensity;
- We conducted a second user study to assess the performances of each of the resulting visuo-haptic coupling scheme. We showed that **G-OFC + H** was generally preferred, while **Ghost Hand + H** led to the least interpenetration.



# Chapter 9

## Bimanual Interactions

### Contents

---

9.1	Introduction	209
9.2	Bimanual UHM Manipulation System	210
9.3	User Study: Object Grasping and Holding with Bimanual Haptics	214
9.4	User Study: 3D Object Shape Exploration with Bimanual Haptics	223
9.5	Global Discussion	230
9.6	Use Cases	231
9.7	Conclusion	233

---

### 9.1 Introduction

The vast majority of UMH interactions are performed using a single-hand, partly due to the limited workspace of UMH interfaces, as discussed in [Section 1.5.2](#). Yet, bimanual interactions are fundamental, while also often providing cognitive and performance benefits [[Talvas et al. 2014](#)]. For instance, many interactions (e.g. writing on paper) often use one hand as a reference point leading to an improved integration of the other hand's position. This leads to more precise motions, and overall improved performances.

This chapter aims at tackling this issue, studying whether (1) unimanual UMH feedback can benefit bimanual manipulations, and if (2) bimanual UMH feedback provides an advantage

over unimanual feedback during bimanual tasks. [Figure 9.1](#) illustrates our contributions.

Our contributions are:

- A novel system and its associated haptic coupling schemes, which allow multi-device bimanual UMH feedback during bimanual object manipulation in VR ([Section 9.2](#)).
- A user study evaluating the benefits of bimanual UMH feedback in a bimanual virtual object grasping and holding task ([Section 9.3](#)).
- A user study investigating the advantages of bimanual UMH feedback for the perception of virtual object shapes, where the feedback was designed based on a state-of-the-art 3D shape rendering algorithm [[Martinez et al. 2019](#)], that we customized for bimanual shape rendering ([Section 9.4](#)).
- Three use-cases illustrating bimanual UMH for different manipulation tasks in VR environments. ([Section 9.6](#)).

The content of this chapter was adapted from the following publication: [[Mulot et al. 2025](#)].

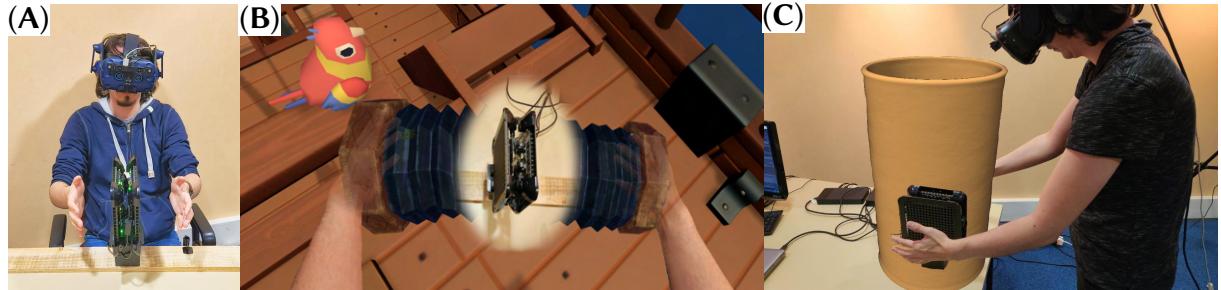


Figure 9.1 – (A) We propose a novel approach to bimanual UMH in VR using two ultrasound devices mounted back-to-back. This enables bimanual haptic feedback during bimanual virtual object manipulation tasks, such as (B) the virtual concertina playing interaction or (C) the virtual pottery scenarios pictured here.

## 9.2 Bimanual UHM Manipulation System

We propose a novel system using a pair of UMH devices mounted back to back, around which the virtual environment is positioned so as to enable bimanual interactions with bimanual haptic feedback when the hands are non-coplanar (see [Figure 9.2](#)).

The proposed system is composed of five parts: the virtual environment (VE), hardware and software for simulating and viewing the VE, hardware and software components for tracking the user, hardware for delivering UMH feedback, and a software coupling scheme for linking tracking data to visual and haptic feedback generation (see [Figure 9.3](#)).

### 9.2.1 System Hardware

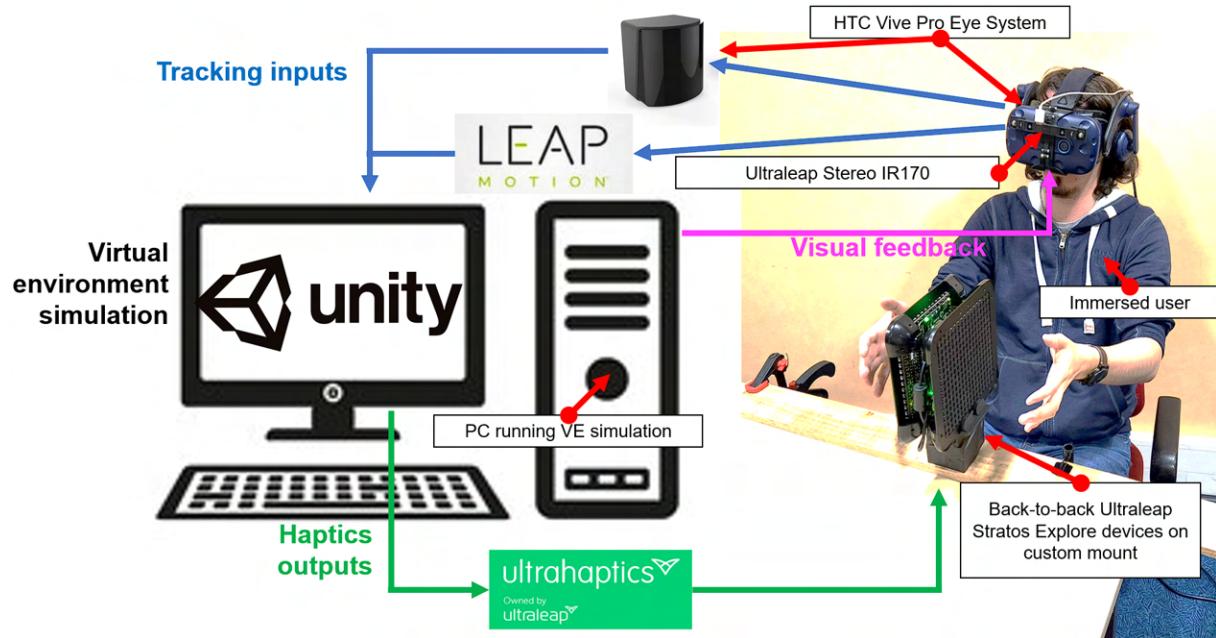


Figure 9.2 – Our proposed system for bimanual UMH manipulation in VR. Users view the virtual environment through an HTC Vive Pro Eye HMD. Their hands are tracked using an Ultraleap Stereo IR170 camera mounted on the headset, allowing the user to interact with virtual objects haptically rendered using a pair of Ultraleap STRATOS Explore devices mounted vertically back-to-back with the array planes parallel to the user’s sagittal plane.

For these studies, users view the virtual environment through an HTC Vive Pro Eye<sup>1</sup> HMD (pink arrows in Figure 9.2).

Tracking components are represented with blue arrows in Figure 9.2. The user’s viewpoint is tracked using the HTC Vive’s lighthouse system. An Ultraleap Stereo IR 170<sup>2</sup> camera mounted to the front of the HMD tracks the user’s hands within the scene. The fixed camera pose with respect to the HMD enables a conversion of the hand tracking data to the virtual environment’s world coordinate system defined around the HTC Vive’s tracking region.

UMH feedback is provided by a pair of Ultraleap STRATOS Explore<sup>3</sup> arrays mounted vertically back-to-back on a custom 3D-printed mount placed on a table, such that the array planes face outward and lie approximately parallel to the user’s sagittal plane during the interactions. The mount is designed to leave minimum space between the array faces (64 mm) so as

1. <https://developer.vive.com/resources/hardware-guides/vive-pro-eye-specs-user-guide/>  
 2. <https://www.ultraleap.com/products/>  
 3. [https://www.ultraleap.com/datasheets/STRATOS\\_Explore\\_Development\\_Kit\\_datasheet.pdf](https://www.ultraleap.com/datasheets/STRATOS_Explore_Development_Kit_datasheet.pdf)

to limit physical obstructions of the interaction region as much as possible. Software methods for dealing with this issue are described in [Section 9.2.2](#). The mount raises the arrays 10 cm above the table, to ensure that the full lateral workspace of the arrays remains usable and unobstructed. The haptic interaction workspace's dimensions are approximately 120x50x50 cm [[Howard et al. 2020](#)].

Mounted in this way, the UMH devices are effectively independent, each one delivering haptic stimuli to one half of the interaction region, with no overlap between the acoustic fields generated by each device. Extending our proposed method to other arrangements of non-coplanar arrays mounted at angles with possibly overlapping workspaces would not fundamentally change the approach proposed here, but would require additional algorithms for generating haptic stimuli in the overlap between both devices' workspaces. Such an approach could be implemented using a priority system, or with algorithms for joint generation of focal points, as e.g. in the work by [Matsubayashi et al. \[2019\]](#).

### 9.2.2 Coupling Schemes

At the heart of the system lies a software coupling scheme linking the hand tracking data to the state of the virtual environment and to the generation of UMH feedback stimuli. Here, we consider two forms of manipulation interactions: those where the user can physically act upon the virtual environment, receiving visual and haptic feedback according to the effective changes (orange arrow path in [Figure 9.3](#)), and those where the user can only explore the virtual environment, receiving haptic feedback of the encountered virtual objects (green arrow path in [Figure 9.3](#)). Each of these interaction types involves a slightly different coupling scheme, the specifics of which are respectively described below.

**Redirection** The first common element in both coupling schemes is the optional hand redirection component.

Set back to back on the custom-designed mount, the UMH devices physically occupy a region which is 64 mm wide. In addition, UMH stimuli of a reasonable quality can only be achieved beyond a minimum distance of around 50 mm from the front face of the device. Therefore, a 164 mm wide “dead space” exists at the center of the interaction region. This could be dealt with by strategically placing a wide solid virtual object in the center of the workspace so as to keep users from bringing their hands together and going below the minimum haptic rendering distance or colliding with the arrays (as in our “pneumatic box” use case, see [Section 9.6](#)). Another approach which poses fewer constraints on the design of the virtual environment is

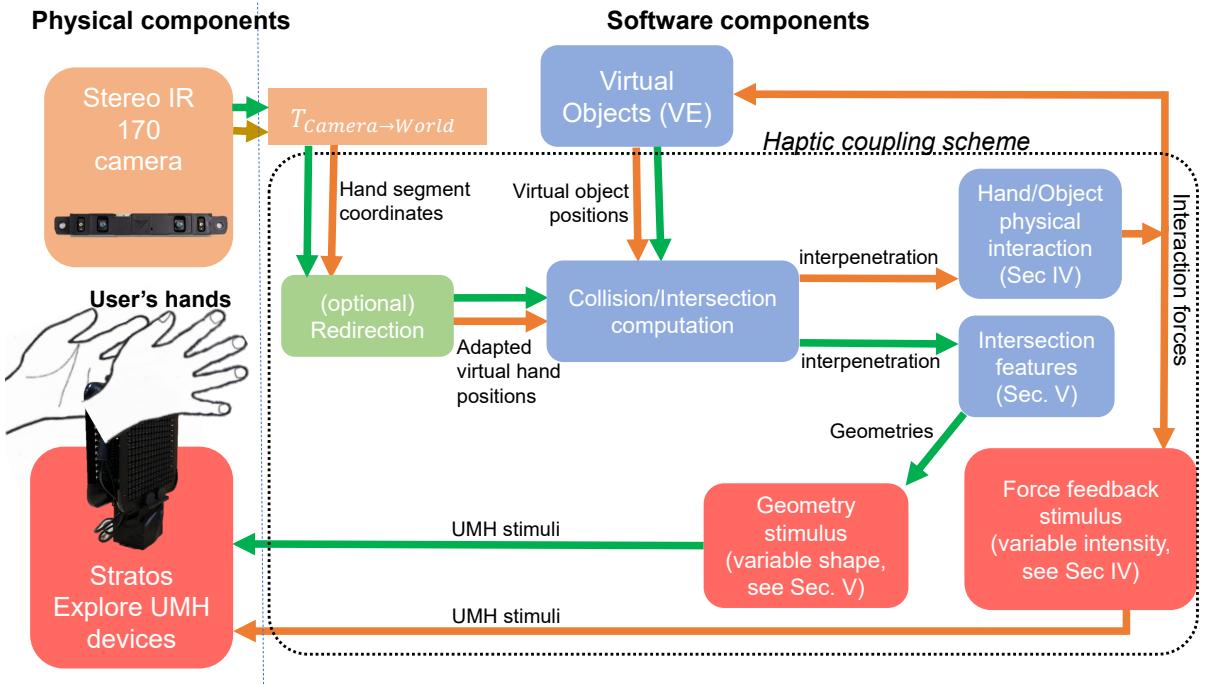


Figure 9.3 – General overview of the coupling schemes relating bimanual hand tracking to bimanual UMH rendering of interactions with the VE. In this chapter, we explore two cases: the case where the hands physically act on the VE, and thus reciprocally act on each other ([Section 9.3](#)), and the case where the hands do not physically act on the VE and purely serve the purpose of exploring it ([Section 9.4](#)). Tracking components are orange, haptics components are red, optional redirection components are green and VE objects and physics components are blue.

to use redirection. For our experiments (see [Section 9.3](#) and [Section 9.4](#)), we offset the virtual hands by a fixed vector of 70 mm towards the sagittal plane. Virtual hands were thus 140 mm closer to each other than the participant's real hands were. In this way, even when the participants brought both their virtual hands in contact, the physical hands remained about 40 mm away from the ultrasound arrays, and could thus receive UMH feedback. Having a fixed offset ensured that hand motions perceived visually matched those perceived kinaesthetically, which we expect will not disturb proprioceptive mechanisms involved in bimanual haptic perception.

**Definition of Hand Interaction Zones** Once the position of the virtual hand is computed, we can define a set of areas that will be able to interact with the VE and receive haptic feedback. For our experiments, we defined 6 areas approximated as bounded planes, corresponding to the palm and five fingertips (see [Figure 9.5](#)), i.e. the hand regions where ultrasound feedback is felt

with the highest intensity [Chongyang et al. 2019; Rakkolainen et al. 2021]. Haptic feedback computed with these virtual hand interaction regions can then be shifted back to the position of the non-redireceted virtual hand, and then transformed into the reference frame of the haptic interfaces to be rendered.

**Hand-Object Physical Interaction** We consider physical interactions with a deformable object, assuming that the object can break if squeezed too tightly between hands. The object undergoes elastic deformation as soon as one or both hands interpenetrate it (see Figure 9.5). The computed spring force causes a visual deformation at the level of the contacting hand and serves as input for controlling the intensity of the UMH stimulus. We consider two modes for this object. If both hands are in contact, the object is attracted to the center of the hands using a critically damped spring-damper system, allowing the user to move a grasped object. But if at least one hand is not in contact, the object remains static and cannot break, although it can still be deformed, and can still generate feedback for the hand in contact. A formalization of this behaviour, corresponding to the orange path in Figure 9.3, is presented in Figure 9.4, and its application is illustrated in Figure 9.5 for a bubble manipulation (see Section 9.3).

**Exploration: Computing Intersection Geometries** To render the explored 3D shapes with UMH, we use an adaptation of an algorithm proposed by Martinez et al. [2019] who showed that rendering both hand-object intersection contours as well as salient object features such as edges and vertices yielded good shape perception results. This algorithm, corresponding to the green path in Figure 9.3, is formalized in Figure 9.6. Applications of this algorithm are illustrated in Figure 9.7. Overall, it uses an intersection strategy (see Chapter 6) composed of a raw intersection step between the hand planes and the object mesh, followed by a pruning of the contours lying outside the bounded planes, and contours that are too close from one another. Finally, the transformation step adds representations of the shape’s features (edges and vertices).

### 9.3 User Study: Object Grasping and Holding with Bimanual Haptics

As discussed in Section 9.2.2 a first category of direct haptic manipulation in VR concerns interactions where the user physically acts on virtual objects, and thus the feedback reflects the reciprocal action of those virtual objects on the user. We performed an experiment to investigate whether our proposed system for delivering bimanual UMH feedback was functional in a

**Algorithm:** Physical feedback

---

**Input:** The object  $obj$  interacted with  
 The two hands  $leftHand$ ,  $rightHand$   
 The threshold force  $f_{max}$  that makes the object break  
 The minimum haptic intensity  $i_{min} \in [0, 1]$  rendered when initiating contact  
 The spring-damper system  $sd$

---

```

If  $leftHand.CollidesWith(obj)$  and  $rightHand.CollidesWith(obj)$  do
    /* Move the object */
    targetPosition  $\leftarrow (leftHand.position + rightHand.position) / 2$ 
    obj.ApplyForce(sd.ComputeForce(obj.position, targetPosition))
    /* Compute the haptic feedback */
    leftForce  $\leftarrow$  AppliedForce(leftHand, obj)
    rightForce  $\leftarrow$  AppliedForce(rightHand, obj)
    meanForce  $\leftarrow (leftForce + rightForce) / 2$ 
    If  $meanForce > f_{max}$  do
        Return /* The object broke, we do not render anything */
    Else do
        intensity  $\leftarrow$  Lerp( $i_{min}$ , 1,  $meanForce / f_{max}$ )
        RenderFeedbackOnHand(Circle(leftHand.size, intensity), leftHand)
        RenderFeedbackOnHand(Circle(rightHand.size, intensity), rightHand)
    Else do /* At most one hand is in contact */
        closestHand  $\leftarrow$  ClosestHand(obj, leftHand, rightHand)
        closestForce  $\leftarrow$  AppliedForce(closestHand, obj)
        If  $closestForce = 0$  do /* No hand is in contact */
            Return
        Else do
            intensity  $\leftarrow$  Lerp( $i_{min}$ , 1,  $closestForce / f_{max}$ )
            intensity  $\leftarrow$  Clamp(intensity, 0, 1)
            RenderFeedbackOnHand(Circle(closestHand.size, intensity), closestHand)

```

---

Figure 9.4 – Algorithm for rendering the haptic feedback for a physical interaction with an object. Here, `ApplyForce`, `CollidesWith`, and `AppliedForce` are parts of the simulation of the VE, which respectively apply a physical force to an object, determine if a hand is colliding with an object, and compute the force applied by the hand on said object. `ClosestHand` returns the hand that is the closest to the object. `RenderFeedbackOnHand` takes the computed geometry, expressed in the reference frame of the virtual hand, and converts it into haptic coordinates so that it can be rendered on the corresponding physical hand.

bimanual object grasping and holding task, and whether bimanual haptic feedback provided a benefit over unimanual haptic feedback in this context. The experiment made use of the system described in [Section 9.2](#), following the coupling scheme presented by the orange arrows in [Figure 9.3](#).

### 9.3.1 Hypotheses

We hypothesized that:

**H1 a)** Manipulation accuracy will be better in the presence of haptic feedback, and **b)** will be

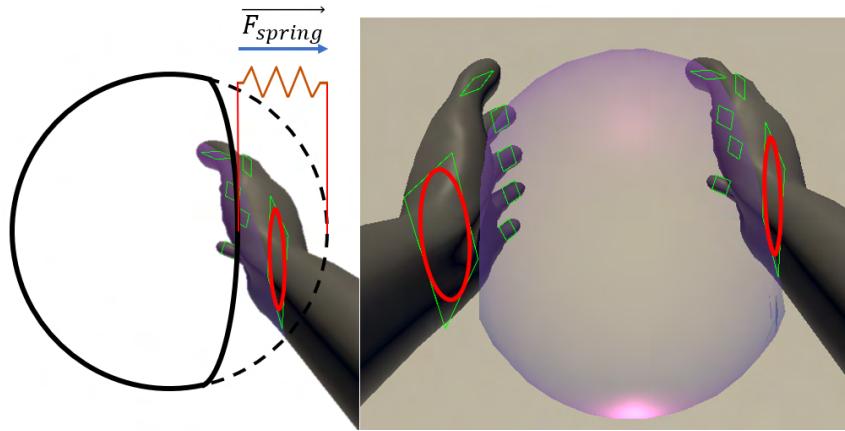


Figure 9.5 – Tactile force feedback: when a hand interpenetrates a deformable object, a spring force  $\overrightarrow{F_{spring}}$  is computed based on the interpenetration between the hand and non-deformed object, the magnitude of which is used to proportionally adjust the intensity of a spatio-temporally modulated UMH circle (in red) projected on the hand interaction region (green rectangle) at the location of interpenetration.

better when bimanual haptic feedback is provided than when unimanual haptic feedback is provided.

**H2** **a)** Bimanual manipulation will be more efficient, i.e. faster at equal accuracy, in the presence of haptic feedback, and **b)** will be more efficient with bimanual haptic feedback than with unimanual haptic feedback.

**H3** **a)** Manipulation task workload will be lower in the presence of haptic feedback, and **b)** be lower in the presence of bimanual haptic feedback than in the presence of unimanual haptic feedback.

**H4** **a)** Participants will prefer performing manipulations in the presence of haptic feedback, and **b)** will prefer when the haptic feedback is bimanual.

**H5** **a)** Participants will rate the manipulation experience as more coherent when haptic feedback is provided, and **b)** will find bimanual haptic feedback more coherent than unimanual haptic feedback.

### 9.3.2 Materials and Methods

**Task** To test these hypotheses, participants performed a grasping and holding task in VR (see [Figure 9.8](#)). Participants began with both their hands in a fixed starting position, close to their body and far from the haptic interaction region. A virtual bubble appeared at a short distance in front of them, and they had to carefully bring their hands in contact with it in order to grasp

---

**Algorithm:** Object exploration

---

**Input:** The object *obj* interacted with  
 The two hands *leftHand*, *rightHand*  
 The trigger distance *trigDist* for the feature rendering  
 The minimum required distance *betweenDist* between parts of the feedback

---

```

Foreach hand ∈ {leftHand, rightHand} do
    rawFeedback ← ∅
    Foreach part ∈ hands.parts do
        /* Dealing with the features */
        Foreach feature ∈ obj.features do
            rawFeedback ← rawFeedback ∪ CutFurther(feature, part, trigDist)

        /* Dealing with the intersection */
        rawFeedback ← rawFeedback ∪ OutlineIntersection(obj, part)

    /* Removing the parts that are too close to each other */
    cleanFeedback ← PruneFeedback(rawFeedback, betweenDist)
    RenderFeedbackOnHand(cleanFeedback, hand)

```

---

Figure 9.6 – Algorithm for rendering the haptic feedback for an object exploration interaction. Here, features represent the polylines associated to all of the object’s features. For our experiments, linear and circular edges were represented as lines and circles of the same sizes, and vertices were represented as small circles with a 1 cm radius. parts contains all of the hand interaction parts. CutFurther(a, b, c) removes the parts of a that are at a distance greater than c from b. OutlineIntersection(obj, part) computes the outline of the intersection between the object obj and the hand interaction part part. PruneFeedback(f, d) optimizes the rendering by removing parts of the feedback f that are at a distance lower than d from each other. RenderFeedbackOnHand takes the computed geometry, expressed in the reference frame of the virtual hand, and converts it into haptic coordinates so that it can be rendered on the corresponding physical hand.

it. Contact between a virtual hand and the bubble applied a force to the bubble, which caused it to visually deform. If both hands were in contact with the bubble and applied excessive force, causing a deformation of more than 7% of the bubble’s radius, the bubble burst. In this case, participants had to return their hands to the starting position, so as to respawn the bubble and continue the task. After pilot testing, we set the haptic intensity when initiating contact with the bubble to 10% of the device capabilities, yielding a sound pressure of around 790 Pa. This is a good compromise between lower values that may not be perceivable, and higher values that would lead to a harder perception of the intensity change. This value is superior to the mean STM pattern detection thresholds for at least 75% of the population [Howard et al. 2019]. This deformation limit was chosen empirically to ensure that both haptic and visual deformations of the bubble were sufficiently perceivable, without making the task too easy. The trial ended if participants successfully held the bubble between both hands, without bursting it, for a

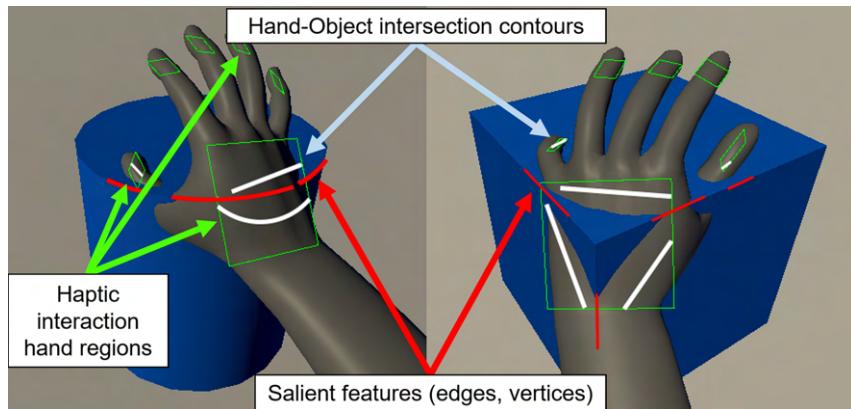


Figure 9.7 – Visualization of the haptic shape rendering approach. Proximity of the hand interaction regions (green rectangles) to salient features (red) are coupled with intersection contours of the interaction regions with the shape (white) to form sets of line segments rendered with spatio-temporal modulation.

consecutive duration of 4 s. This duration was also empirically chosen to be sufficiently long to ensure that participants had to make an effort to succeed, while keeping the overall experiment duration reasonably short. After explaining how a trial works, the experimenter asked the participants to grasp the bubble as fast as possible, while trying to avoid bursting it.

**Procedure** participants performed the task in three blocks, one for each of the following feedback conditions, whose order was counterbalanced across participants:

- N** No haptic feedback, in which only visual deformation and motion of the bubble informed about the force applied.
- U** Unimanual haptic feedback, in which the participant's dominant hand received additional tactile feedback in the form of a STM circle [Frier et al. 2018] rendered on their palm, whose intensity was proportional to the force applied to the bubble (see [Section 9.2.2](#)).
- B** Bimanual haptic feedback, in which both participants' hands received simultaneous tactile feedback identical to that delivered in the **U** condition.

Within each block, participants performed 18 trials which randomly alternated between a large (22 cm radius) and a small (12 cm radius) sphere, each appearing 9 times. These object sizes were chosen to cover a sufficiently wide span of bimanual grasping motions within the haptic interaction region, without compromising hand tracking performance.

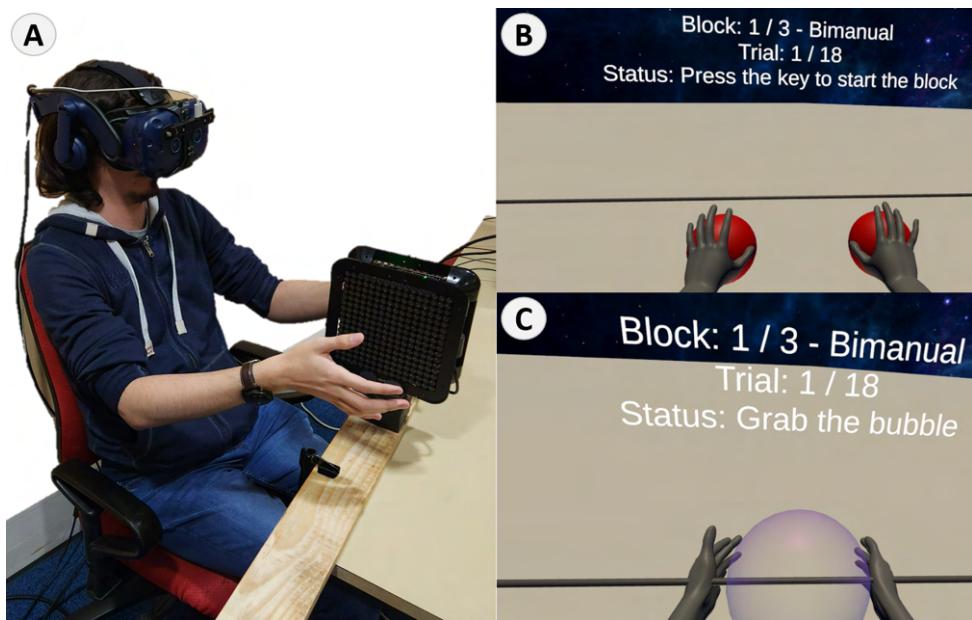


Figure 9.8 – (A) Participants were seated in front of the system described in [Section 9.2](#), viewing the VE in which they performed the task. (B) At the start of each trial, participants placed their hands at fixed starting positions indicated by a pair of red spheres. (C) A virtual bubble which participants were to grasp and hold then spawned in front of them. Instructions and feedback on progress were provided via a text panel in the VE. A horizontal bar also indicated the preferred axis along which to grasp the bubble to ensure both hands were used in a repeatable manner across trials and participants.

**Collected Data** During each trial, we measured manipulation accuracy by recording the number of bubble bursts as well as the total time during which participants accidentally break contact with the bubble (release time) while trying to hold it. As measures of manipulation efficiency, we recorded the total time of the task and the time users were holding the bubble. The lower these two times, the more we considered the manipulation to be efficient. It should be noted that, by design, both times were necessarily greater than 4 s, which was the minimum consecutive holding time required to succeed in the task.

At the end of each block, participants filled out a raw NASA-TLX questionnaire [[Hart 2006](#)], from which a global workload measure was calculated. Participants also responded to a subjective questionnaire assessing perceived visuo-haptic coherence during the manipulation task for two criteria: perceived coherence between the sensory feedback and expectations given the scene appearance (referred to as **c-FB/EXP**), and perceived coherence between the applied force and bubble deformation (referred to as **c-F/D**), both using 7-point Likert scales. They were asked the following two questions: “Rate the coherence between the haptic feed-

back and what you would expect from the virtual scene”, and “Rate the coherence between the deformation and the “force” you felt”.

At the end of the experiment, participants were asked to rank the feedback conditions in order of overall preference, as well as based on how easily they were able to establish a strategy for performing the task. We also asked participants to rate the perceived coherence between the motion of their real and virtual hands after each block to ensure that the hand redirection method described in [Section 9.2.2](#) did not have a detrimental impact on the interaction.

**Participants** The study involved 12 participants (10 men, 1 woman, 1 non-binary, aged  $m \pm s.d. = 24.9 \pm 3.1$ ). Ten participants were right-handed, 1 left-handed and 1 ambidextrous. Half of the participants were experienced in VR, while the other half were not. Three participants had extensive experience in haptics, while 7 only had limited haptics experience.

### 9.3.3 Results

Mean results for each evaluated criterion are summarized in [Table 9.1](#). Key results for total task time, number of bubble bursts, and subjective workload estimates are also shown in [Figure 9.9](#).

[Table 9.1](#) – Mean results and standard deviation for the object grasping and holding experiment. For the first four metrics, standard deviations are computed on the dataset containing all trials, including repetitions. “Bubble bursts” is the number of times the participants applied too much force during grasping, resulting in the bubble bursting (lower is better). “Release time” is the cumulative time during which participants accidentally break contact with the object during grasping (lower is better). “Holding time” is the cumulative time during which participants grasped the object before successfully completing the task (lower is better). “Total time” is the total task execution time, i.e., release time + holding time + time to get hands into position (lower is better). “Workload” is the TLX score (lower is better). “c-FB/EXP” is the coherence rating between sensory feedback and expectations given the scene’s appearance (7-point Likert scale, higher is better). “c-F/D” is the coherence rating between the force applied and the bubble deformation (7-point Likert, higher is better). Finally, we consider two rankings of the conditions based on “overall preference” for performing the task and “ease of establishing a (grasping and holding) strategy” (lower is better).

Feedback	Bubble bursts	Release time (s)	Holding time (s)	Total time (s)	Workload (/ 100)	c-FB/EXP (/ 7)	c-F/D (/ 7)	Ranking: overall preference (/ 3)	Ranking: ease of establishing a strategy (/ 3)
<b>B</b>	$0.33 \pm 0.75$	$0.51 \pm 1.86$	$5.80 \pm 5.52$	$11.27 \pm 11.32$	$31.25 \pm 12.12$	$6.17 \pm 0.83$	$4.92 \pm 1.44$	$1.25 \pm 0.62$	$1.17 \pm 0.58$
<b>U</b>	$0.44 \pm 1.56$	$0.70 \pm 1.76$	$5.95 \pm 4.40$	$12.48 \pm 12.56$	$39.93 \pm 14.35$	$5.42 \pm 1.51$	$5.42 \pm 1.00$	$1.92 \pm 0.51$	$2.08 \pm 0.29$
<b>N</b>	$1.12 \pm 2.95$	$2.51 \pm 9.43$	$7.67 \pm 6.68$	$21.07 \pm 25.42$	$53.82 \pm 16.45$	$1.25 \pm 0.87$	$1.42 \pm 1.44$	$2.83 \pm 0.39$	$2.75 \pm 0.62$

**Objective Metrics** For each participant, we averaged all the objective metrics per type of feedback, leading to 36 (3 blocks, 12 participants) values per metric. Total interaction time and bubble burst distributions are illustrated in [Figure 9.9](#) (B, C).

For each of the duration metrics (release time, holding time, and total interaction time), either the normal distribution hypothesis was not rejected (Shapiro-Wilk test,  $p > 0.05$  for each feedback), or the Q-Q plot was good enough to perform an ANOVA, given its robustness to moderately non-normal data. Homogeneity of variances was also verified for all three metrics (Levene test,  $p > 0.05$ ). We thus performed a one-way repeated measures ANOVA, with the duration as dependant variable, the feedback as independent variable, and the participant as random effect variable. Significant effects on holding time ( $F(2, 22) = 5.6979, p = 0.01$ ) and total interaction time ( $F(2, 22) = 13.816, p < 0.001$ ) were found, but none was found on release time ( $p > 0.05$ ). Post-hoc Holm-corrected pairwise t-tests showed that **U** led to significantly lower holding time than **N** ( $t(11) = 3.33, p = 0.02$ ), and that **B** and **U** yielded significantly lower interaction time than **N** ( $t(11) = 4.02, t(11) = 3.87$  respectively,  $p = 0.006$  for both comparisons).

On the other hand, Shapiro tests and Q-Q plots rejected the normal distribution hypothesis ( $p < 0.05$ ) for the number of bubble bursts. We therefore applied a Friedman test which revealed a significant effect of the feedback condition ( $\chi^2(2, N = 12) = 11.5, p = 0.003$ ). Post-hoc pairwise Wilcoxon signed-rank tests with Holm corrections revealed that both **B** and **U** led to significantly less bubble bursts than **N** ( $V = 74.5, V = 65$  respectively,  $p = 0.015$  for both comparisons).

**Subjective Metrics** For all subjective metrics, we gathered one value per participant, per feedback condition, leading to 36 values per metric. The workload score distribution is shown in [Figure 9.9](#) (Right).

After verifying its assumptions (Shapiro-Wilk test, Levene test,  $p > 0.05$ ), an ANOVA showed a significant effect of the feedback on the score ( $F(2, 22) = 23.51, p < 0.001$ ), and a post-hoc analysis with pairwise Holm-corrected t-tests revealed that the workload under condition **B** was significantly lower than that under condition **U** ( $t(11) = 3.31, p = 0.007$ ), which in turn was significantly lower than that under condition **N** ( $t(11) = 4.76, p = 0.001$ ).

For the **c-FB/EXP** and **c-F/D** perceived coherence scores, we performed a Friedman test, which revealed a significant effect of the feedback ( $\chi^2(2, N = 12) = 21.3, \chi^2(2, N = 12) = 19.6$  respectively,  $p < 0.001$  in both cases). In both cases, pairwise Holm-corrected Wilcoxon signed-

rank tests revealed that **B** and **U** were rated as more coherent between visuals and haptics than **N** ( $V = 0, p = 0.007$  for both **c-FB/EXP** comparisons,  $V = 0, p = 0.009$  for both **c-F/D** comparisons).

Friedman tests ( $\chi^2(2, N = 12) = 15.2, p < 0.001$ ) followed by pairwise Holm-corrected Wilcoxon signed-rank post-hoc tests on the feedback condition ranking data showed that **B** and **U** were significantly preferred to **N** ( $V = 76, p = 0.008$  and  $V = 72, p = 0.01$  respectively). Participants also found it significantly easier (Friedman test,  $\chi^2(2, N = 12) = 15.2, p < 0.001$ ) to establish a grasping strategy in the **B** condition than in the **U** condition ( $V = 72, p = 0.02$ ). Establishing a grasping strategy was also significantly easier in the **U** condition than in the **N** condition ( $V = 65, p = 0.02$ ).

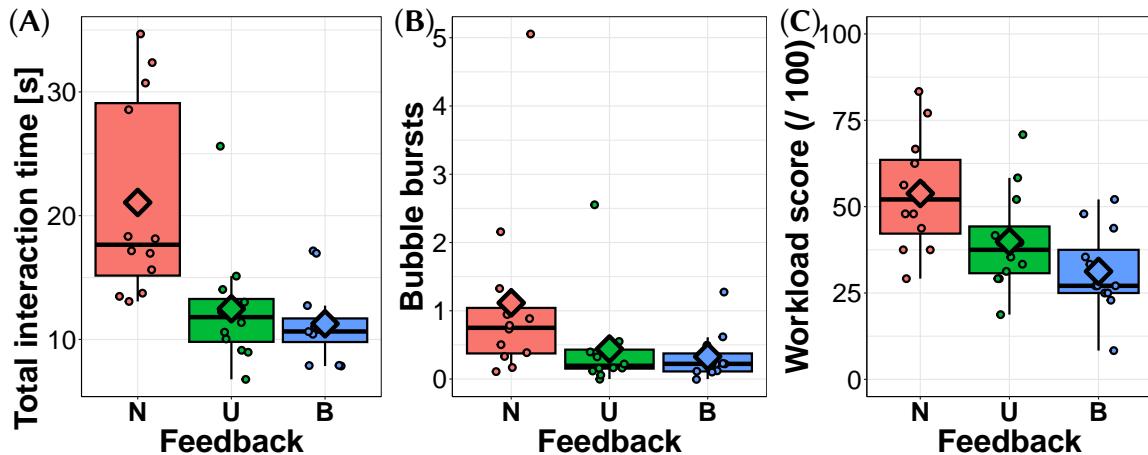


Figure 9.9 – Distributions of the mean interaction time per trial (A), mean bubble burst count per trial (B), and workload (C) across participants, for each condition. Dots represent mean single participant results, and diamonds represent the overall mean value for each condition.

Finally, according to the questionnaires, the hand redirection method employed in the experiment does not appear to have disturbed users (hand coherence rated 6.17 / 7 on average).

### 9.3.4 Discussion

Regarding hypotheses **H1** (manipulation accuracy), we observed no difference in release time, but significantly more bubble bursts in **N** than in **B** and **U**. These results tend to support **H1.a**. Post-hoc testing revealed no significant differences between **B** and **U**, thus **H1.b** is not supported.

When it comes to manipulation efficiency, both **B** and **U** required significantly less total time

than **N**. This was mostly because of the difference in approach time which can be traced back to the significantly larger number of bubble bursts in **N**, although the significant difference in grasp times between **U** and **N** also contributed. These results tend to support **H2.a**. However, **H2.b** was not supported, as we found no significant differences between **B** and **U**.

Looking at the NASA-TLX scores, **B** scored significantly lower than **U**, which in turn scored significantly lower than **N**. This supports **H3** entirely, indicating that bimanual haptic feedback yielded the lowest subjective workload for the task at hand.

Participants significantly preferred **B** and **U** over **N**. This tends to support **H4.a**, but does not support **H4.b**. They also found it significantly easier to establish a grasping strategy in **B** than in **U**, while also finding it significantly easier to do so in **U** than in **N**, fully supporting **H4**. In summary, **H4.a** appears fully supported, but **H4.b** is only partially supported by our results.

Conditions **B** and **U** both yield significantly higher perceived coherence than **N**, supporting **H5.a**. However, our results do not support **H5.b**.

Overall, for all tested metrics it appears that the bimanual haptic interaction performed as good as, or better than unimanual haptic feedback, and consistently better than no haptic feedback. We can therefore confidently conclude that our proposed novel approach to bimanual UMH is functional in this type of 3D object manipulation task with tactile force feedback. The benefit of bimanual UMH over unimanual UMH is mainly apparent when it comes to subjective ease of use and workload measures. We can therefore conclude that our approach can be particularly suited to improving user experience in VR manipulation, but does not necessarily change much in terms of objective performance metrics when compared to simpler approaches to UMH feedback. However, bimanual UMH appears to conserve all performance benefits that can also be obtained through unimanual UMH, making it at least as good of a candidate for a haptic feedback method as the state of the art.

## **9.4 User Study: 3D Object Shape Exploration with Bimanual Haptics**

A second category of direct haptic manipulations in VR concerns interactions where the user merely explores the virtual environment, receiving feedback but not physically acting on virtual objects. We performed an experiment to investigate whether our proposed bimanual UMH

feedback system was functional in bimanual 3D object shape exploration tasks, and whether such haptic exploration with bimanual UMH feedback provided any benefit over what could be achieved with a single UMH device. The experiment once again made use of the system described in [Section 9.2](#), following the coupling scheme presented by the green arrows in [Figure 9.3](#).

#### 9.4.1 Hypotheses

We hypothesized that:

- H6** Shape identification accuracy will be higher when simultaneous bimanual haptic feedback is provided than when unimanual haptic feedback is provided.
- H7** Shape identification speed will be higher when simultaneous bimanual haptic feedback is provided than when unimanual haptic feedback is provided.
- H8** Shape identification task workload will be lower than when simultaneous bimanual haptic feedback is provided than when unimanual haptic feedback is provided.
- H9** Participants will prefer performing the task when bimanual haptic feedback is provided rather than when unimanual haptic feedback is provided.
- H10** Participants will perceive the virtual shapes as equally coherent when exploring them unimanually or bimanually.
- H11** In alternating unimanual exploration with both hands (pseudo-bimanual, see **P** below), shape identification accuracy will be higher due to redundancy in perception between hands [[Talvas et al. 2014](#)]. Shape identification speed will be lower than in unimanual exploration because of the added sequential exploratory motions. For all other metrics, pseudo-bimanual feedback will perform similarly to unimanual feedback, since the haptic exploration task remains essentially unimanual in both cases.

#### 9.4.2 Materials and Methods

**Task** To test these hypotheses, participants performed a virtual 3D shape exploration task. A virtual glovebox was placed in front of the user, as shown in [Figure 9.10\(B, C, D\)](#). Two holes on the front side of the glovebox enabled the users to insert their hands into the box to haptically explore its content. An invisible 3D object was generated within the glovebox. The ultrasound haptic interfaces were located at the center of the glovebox, mounted vertically back-to-back (see [Section 9.2](#) and [Figure 9.10\(A\)](#)), and providing haptic feedback of the invisible object to the exploring hand(s). Participants received feedback on the exploring hand(s) according to the

rendering scheme discussed in [Section 9.2.2](#). The task was to identify the shape from a set of four possible shapes as accurately as possible, and as fast as possible. Participants responded by selecting a 3D shape from a menu of the four possible shapes using the index finger of their dominant hand as a pointer. The order of shapes in the menu was randomized at each trial. To cap the duration of the experiment to a reasonable time, no more haptic feedback was provided after a duration of 30 s, forcing participants to make a choice if they had not yet responded. They then similarly rated their level of confidence on a 7-points Likert scale. After being informed of the trial sequence, participants were instructed to answer as quickly as possible once they had a reasonable idea of what the shape was.

**Procedure** Participants performed the task in three blocks, one for each of the following exploration and feedback conditions, which were counterbalanced across participants:

- U** Unimanual, in which the participants' dominant hand was used alone, and received tactile feedback. This condition represents the current state-of-the-art for UMH virtual shape exploration [[Long et al. 2014](#); [Martinez et al. 2019](#); [Matsubayashi et al. 2019](#)].
- P** Pseudo-bimanual, in which participants' hands **alternately** explore the shape, and can only receive tactile feedback one at a time. This condition was chosen to investigate the effectiveness of involving both hands in a way where UMH feedback could be provided using conventional single-board approaches (e.g. [[Howard et al. 2020](#); [Brice et al. 2019](#)]).
- B** Bimanual haptic feedback, in which both participants' hands simultaneously explore the shape, both receiving tactile feedback at the same time.

It is worth noting that the **P** condition was not present in the first experiment ([Section 9.3](#)). Indeed, this experiment required the interactions of both hands simultaneously. While alternating feedback could have been implemented, any alternating rule would have been rather arbitrary and unnatural. For this reasons, it was excluded from the first experiment. Similarly, the **N** condition from the first experiment was not considered for this second experiment, since in the absence of visual cues, the task could not be performed without haptic cues. Future work can focus on comparing bimanual feedback to pseudo-bimanual feedback for other manipulation tasks.

Four possible 3D objects (Sphere, Cylinder, Cone, Cube) were presented, such that they were symmetrical around the seated participant's sagittal plane. In this way, the same quantity of information about the objects was accessible in all conditions. We presented each shape in two sizes: large (fits inside a 20 cm wide cube) or small (fits inside a 13 cm wide cube). To ensure

all relevant features of each shape could be perceived in all conditions, the shape orientations were held constant and corresponded to those of the visual shapes in the response menu. Participants were informed that the objects could have one of two different sizes, but that they only needed to indicate which shape they felt, not its size. Within each block, each shape was presented four times in total, twice in the small size and twice in the large size. The order of presentation of shapes was fully randomized.

**Collected Data** For each trial, we recorded task completion time as the sum of time spent in contact with the shape. This value was capped at the maximum exploration time of 30 s, and lower values indicate better performances.

For each condition, we measured precision and recall for shape identification, with higher values indicating better performances. At the end of each block, participants filled out a raw NASA-TLX questionnaire [Hart 2006] to assess task workload. We also assessed perceived coherence (**c-H/V**) between the explored haptic shapes and the visual representation that was provided in the response menu. Finally, we inquired about the strategies that participants developed to perform the task, and whether they changed their strategies during a block or between blocks.

At the end of the experiment, participants were asked to rank the feedback conditions in order of overall preference, as well as based on how easily they were able to establish a strategy for performing the task. We also asked participants to rate the perceived coherence between the motion of their real and virtual hands after each block to ensure that the hand redirection method described in [Section 9.2.2](#) did not have a detrimental impact on the interaction.

**Population** The study involved 18 participants (16 men, 2 women, aged  $m \pm s.d. = 25.8 \pm 7.3$ ). 16 participants were right-handed, 1 left-handed and 1 ambidextrous. Four participants had extensive VR experience, while 11 only had limited VR experience. Two participants had extensive experience in haptics, while 10 only had limited haptics experience.

### 9.4.3 Results

Mean results for each evaluated criterion are summarized in [Table 9.2](#).

**Objective Metrics** One precision and recall value was computed for each condition and each participant, leading to 54 values in total. One participant was removed from the precision data, as their data contained a block where one shape was never chosen, meaning that the precision

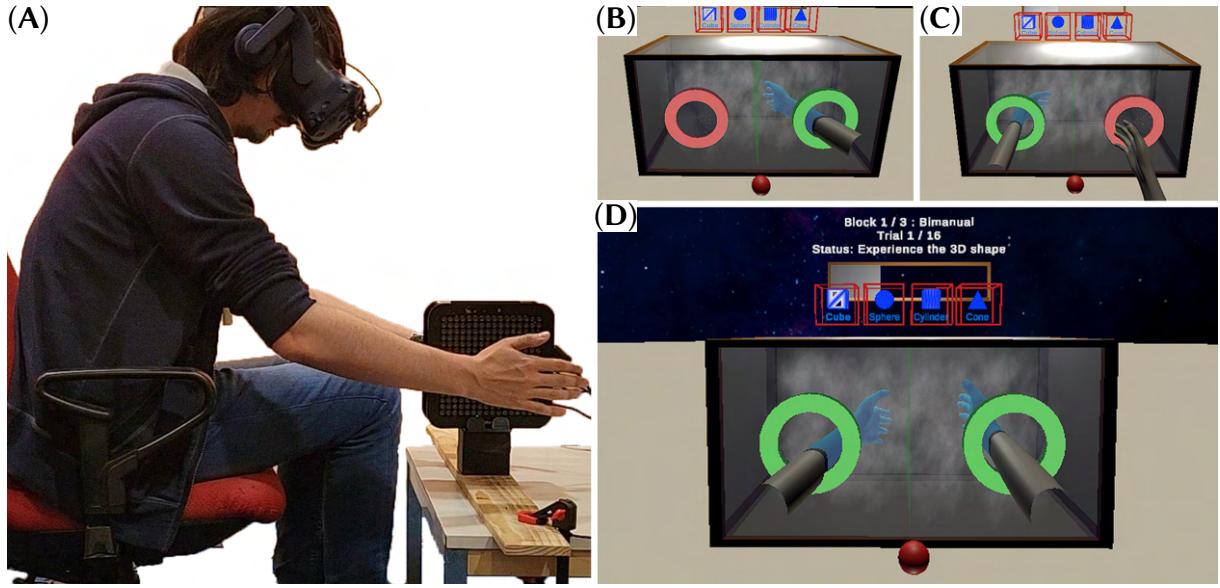


Figure 9.10 – (A) Participants were seated in front of the setup described in [Section 9.2](#), and viewed a virtual glovebox in which an invisible 3D object could haptically be explored either (B) unimanually, (C) with both hands alternating one at a time (pseudo-bimanual), or (D) with both hands simultaneously (bimanual). Above the glovebox, participants could view instructions for the trial as well as use a menu of shapes to identify.

computation was not possible. [Figure 9.11](#) shows the confusion matrices for each condition. Task execution times were averaged across all trials within each condition, for each participant, resulting in a total of 54 values (see [Figure 9.12\(A\)](#)).

The ANOVA assumptions were verified (Shapiro-Wilk test, Levene test,  $p > 0.05$ ), but the ANOVA failed to reveal any effect of feedback on precision and recall ( $p > 0.05$ ). However, Holm-corrected one-sample t-tests revealed that both these values were significantly above the 25% chance level ( $t(16) = 7.46$  for **U**,  $t(16) = 4.7$  for **P**,  $t(16) = 6.15$  for **B**,  $p < 0.001$  for all three measurements).

Q-Q plots and a Levene test ( $p > 0.05$ ) on the task completion times (see [Figure 9.12\(A\)](#)) allowed us to perform an ANOVA which revealed significant effects of feedback ( $F(2, 34) = 3.59, p = 0.04$ ). Holm-corrected pairwise t-tests showed that participants performed significantly faster in the **U** condition than in the **P** condition ( $t(17) = -3.37, p = 0.01$ ).

**Subjective Metrics** We recorded one confidence value per trial, and averaged them per block. For all other subjective metrics, we gathered one value per participant, per feedback condition. In all case, this lead to 54 values per metric (3 blocks, 18 participants).

Table 9.2 – Mean results and standard deviation for the 3D object shape exploration experiment. For the interaction time, the standard deviation is computed on the dataset containing all trials, including repetitions. “Precision” and “Recall” measure the shape identification accuracy (higher is better). “Interaction time” is the time spent close to the experienced shape (lower is better). “Confidence” is the score given by the participant about the confidence they have in each of their answers (7-point Likert scale, higher is better). “Workload” is the TLX score (lower is better). “Coherence” is the coherence rating between the perceived stimulation and the visually rendered shape (higher is better). Finally, we consider two rankings of the conditions based on “overall preference” for performing the task and “ease of establishing an exploration strategy” (lower is better).

Feedback	Precision	Recall	Interaction time (s)	Confidence (/ 7)	Workload (/ 100)	Coherence (/ 7)	Ranking: overall preference (/3)	Ranking: ease of establishing a strategy (/3)
<b>B</b>	$0.52 \pm 0.18$	$0.50 \pm 0.16$	$24.32 \pm 6.57$	$4.39 \pm 0.81$	$52.08 \pm 15.41$	$4.28 \pm 1.23$	$1.39 \pm 0.70$	$1.44 \pm 0.70$
<b>U</b>	$0.53 \pm 0.16$	$0.52 \pm 0.15$	$23.32 \pm 6.54$	$4.25 \pm 0.94$	$54.28 \pm 14.05$	$3.78 \pm 1.22$	$2.39 \pm 0.70$	$2.33 \pm 0.84$
<b>P</b>	$0.50 \pm 0.22$	$0.48 \pm 0.19$	$25.11 \pm 6.33$	$4.16 \pm 0.81$	$54.05 \pm 16.87$	$4.06 \pm 1.11$	$2.22 \pm 0.73$	$2.22 \pm 0.65$

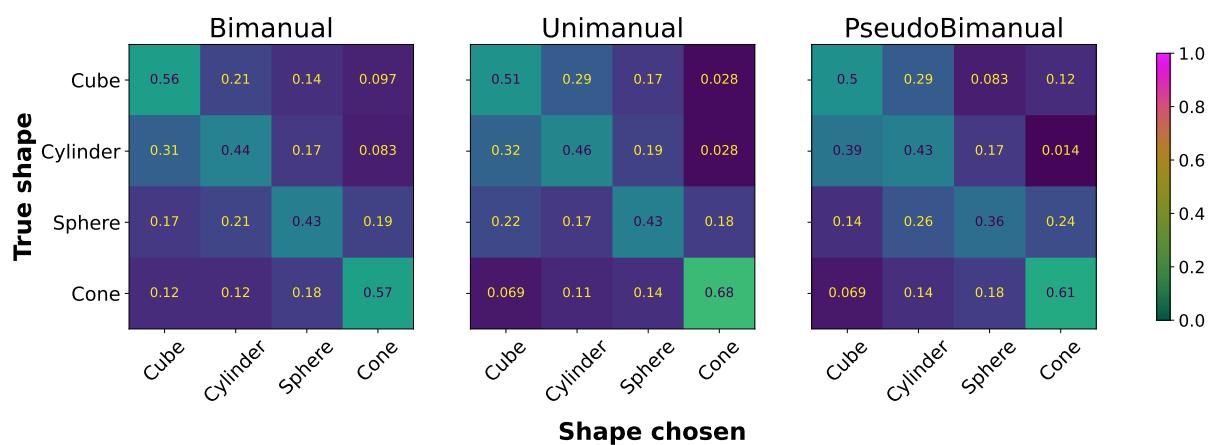


Figure 9.11 – Confusion matrices for the 3D shape selection. Values correspond to recall for the three conditions.

Confidence and workload values had a good-enough Q-Q plot, and similar variances between feedbacks (Levene test,  $p > 0.05$ ), allowing us to perform an ANOVA, which showed no significant effect of the feedback over any of the two factors ( $p > 0.05$ ).

A Friedman test also failed to identify any significant difference in perceived visuo-haptic shape coherence ( $p > 0.05$ ).

That being said, we found a significant effect of the feedback on the participant’s preferences (Friedman test,  $\chi^2(2, N = 18) = 10.3, p = 0.006$ ). Post-hoc Holm-corrected Wilcoxon signed-rank tests showed that **B** was preferred to both **U** ( $V = 146, p = 0.04$ ) and **P** ( $V = 138, p =$

0.02). Participants also found it easier to establish a strategy in the **B** condition, than in the **U** and **P** conditions ( $V = 140$ ,  $p = 0.04$  for both comparisons).

Finally, according to the questionnaires, the hand redirection method employed in the experiment was not perceived as too disturbing by users (rated 5.22 / 7 on average).

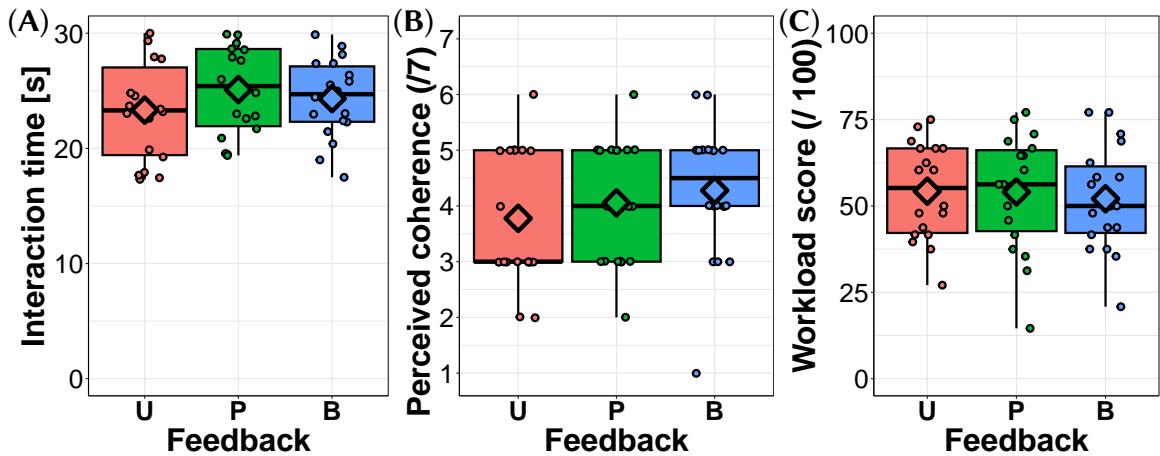


Figure 9.12 – Distributions of the mean interaction time per trial (A), perceived coherence between the shapes and haptic feedback (B), and workload (C) across participants, for each condition. Dots represent individual participant results, and diamonds represent the overall mean value of each condition.

#### 9.4.4 Discussion

All conditions yielded performances significantly above chance levels, meaning all methods allowed effective shape identification. However, no significant differences were observed between conditions when it comes to precision or recall. Therefore, we cannot accept **H6**, as simultaneous bimanual haptic feedback appears to yield identical performance to unimanual and pseudo-bimanual haptics.

The only significant difference in shape identification speed was observed between **U** and **P**. Thus, **H7** is not supported.

The NASA-TLX scores showed no significant differences between conditions, failing to support **H8**.

In terms of participant preferences, **B** was deemed significantly better than both **U** and **P**, both in terms of overall preference for accomplishing the task and in terms of ease of establishing an exploration strategy. **H9** is therefore fully supported.

No significant differences were found between conditions in terms of perceived coherence between the haptic and visual representation of shapes, supporting **H10**.

**H11** appears partly supported, as **U** significantly outperformed **P** in terms of shape identification speed and otherwise no significant differences were observed between both. However, **P** did not perform significantly better than **U** with respect to shape identification accuracy.

Similarly to the experiment on object grasping and holding ([Section 9.3](#)), fully bimanual UMH, as enabled by our system, appears to mainly be beneficial in terms of subjective preference, and not in terms of objective performance in haptic shape exploration tasks. However, it should be noted that the lack of observed differences may however mask a potential advantage of bimanual haptics in a more realistic scenario, where objects are not necessarily symmetrically arranged around the user’s sagittal plane.

Contrary to our expectations, involving both hands with a feedback scheme which could be implemented using existing mobile single-board UMH systems [[Howard et al. 2020](#); [Brice et al. 2019](#)] does not improve shape identification performances when compared to conventional unimanual UMH using a fixed array. Overall, it appears that if both hands are to be involved in UMH shape exploration interactions in VR, they should be involved in such a way that they can both receive simultaneous tactile feedback so as to maximise benefits in terms of user experience.

## 9.5 Global Discussion

Our system enabled simultaneous generation of UMH feedback on both hands. Our results showed that, both for a grasping and exploration task, bimanual feedback led to performances at least as good as with unimanual feedback. Therefore, these techniques are at least as good as the state-of-the-art, and they show that simultaneous UMH feedback on both hands does not seem to degrade one’s integration of the stimuli.

The absence of significant differences between the bimanual and unimanual feedback may be explained by the relatively low difficulty of the tasks, as evidenced by the low amount of bubble pops in the grasping task ([Section 9.3](#)). We expect that more difficult tasks would accentuate such differences.

That said, bimanual UMH feedback offers two significant advantages. First, our results show that UMH feedback significantly improves the user experience, by reducing the workload ([Sec-](#)

[tion 9.3](#)), being preferred by participants ([Section 9.4](#)), and making the establishment of an interaction strategy easier ([Section 9.3](#), [Section 9.4](#)). Secondly, bimanual feedback enables the completion of less constrained tasks that would be impossible to perform with only unimanual feedback. For instance, let us consider our exploration task ([Section 9.4](#)), but with a different set of shapes. We could for example consider two-parts shapes where the right side is composed of a cylinder, and the left part is composed of either of a cone, a cylinder, or a semi-sphere, with both parts being joined at their base, lying on the sagittal plane. Given that all these shapes have the same right-part, any unimanual exploration with the right hand would fail to determine with certainty which of these three shapes is being explored. But bimanual feedback would enable the rendering and perception of the left part of the shape, which could then make discrimination possible.

It is worth noting that our experiments used populations composed of 10 out of 12 and 16 out of 18 men respectively. This is not representative of the global population, meaning that our results are to be interpreted and used carefully.

Then, the presence of our system in the interaction space is in itself a potential limitation. When interacting with small objects in front of the user, they may need to have both hands close to each other, which result in contacts with the system. To avoid this issue, we proposed to use hand redirection ([Section 9.2.2](#)). Our results show a strong hand coherence perception, which leads us to believe that redirection was not an issue in our case. That said, the redirection offset would need to be increased when interacting with even smaller objects, which would enlarge the proprioceptive discrepancy, and may then be perceived by users. Future work will be conducted to assess the range of acceptable redirection offset, as well as to design interaction techniques to allow a wider workspace with lower risk of colliding with the system.

Overall, this first bimanual UMH system appears very promising to improve user experience in haptic VR manipulation tasks, while also enabling a whole new range of interactions.

## 9.6 Use Cases

We tested the proposed bimanual UMH approach in three immersive VR scenarios (see [Figure 9.13](#)) described below.

**Pneumatic Puzzle Box** In this first use-case, users face a pneumatic puzzle, box where the aim is to direct a ball through a circuit and into a goal using combinations of air jets. Users

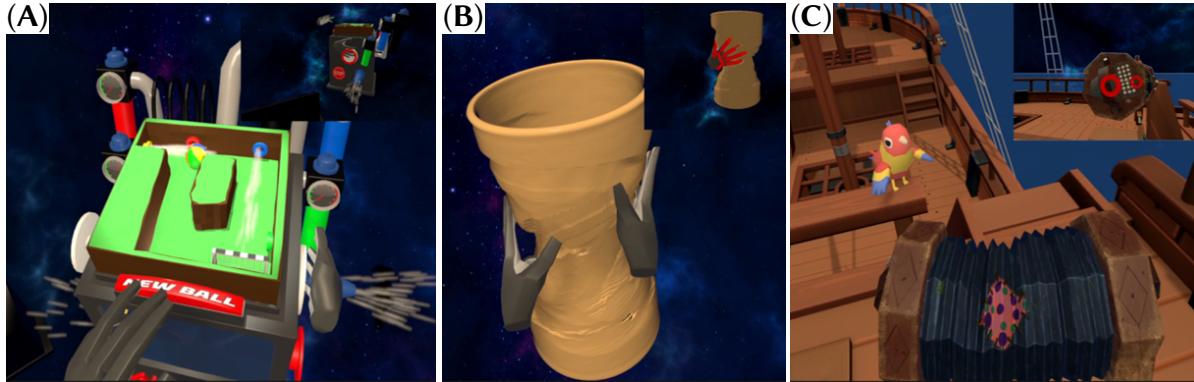


Figure 9.13 – Three VR scenarios exemplifying variations on applications and implementations of our approach. The figures show the VR view in first-person perspective, while the miniatures in the top right corners show a third person view with the tactile feedback patterns highlighted in red. (A) Pneumatic puzzle box use-case: users receive diverse tactile feedback while exploring various air-jets, pressing on pistons and manipulating nozzles. (B) Virtual pottery: tactile force feedback is provided across the hand as users mould a block of virtual clay into shape. (C) Virtual concertina: tactile rhythm cues are provided simultaneously to both hands to guide a user’s motions when playing a tune on the virtual instrument.

can press different pistons and connect sets of nozzles and stoppers to manipulate the air flow in order to achieve this objective.

This use-case exemplifies a VR interaction where the optional hand redirection components discussed in [Section 9.2.2](#) are not necessary, as a wide solid virtual object obstructs the region occupied by the pair of UMH devices. This naturally guides users away from the region where collisions between their hand and the devices could occur.

In this interaction, the hands can perform independent haptic manipulations, showing how our system can also be used to render different haptic sensations due to different interactions between each hand and the virtual environment. For example, a spatio-temporally modulated circle of varying intensity can be rendered on the right hand as it presses down on a piston (similar to the tactile force feedback scheme discussed in [Section 9.3](#), while the left hand explores an outgoing air jet, receiving a spatio-temporally modulated circle whose radius and intensity randomly oscillate around fixed values, simulating turbulent air flow.

**Virtual Pottery** In this second use-case, users can rotate a virtual block of clay and shape it as they desire by pressing their hands into it and deforming it. Tactile feedback is provided based on contact and interpenetration between the hand interaction regions (see [Section 9.2.2](#)) and the virtual clay. The edges of the hand interaction regions which come into contact with

the clay are added to a pool of path points for a global spatio-temporally modulated polyline drawn across all hand parts in contact with the clay. In addition, the intensity of each line segment is increased from a baseline value, proportionally to the local interpenetration between the hand and clay.

This tactile force feedback scheme is similar to that described in [Section 9.2.2](#) and used in the experiment presented in [Section 9.3](#), although the manipulated virtual object in this case is fixed in the virtual environment and can only incur plastic deformation instead of being both mobile in the environment and incurring elastic deformation.

While the realism of the haptic feedback may not be perfect, partly due to the relative softness of ultrasound-based vibrotactile feedback compared to the physical properties of actual clay, it has been shown that using UMH feedback for clay modelling can actually provide very rich interactions, while helping the participants better control their hand movements [[Barreiro et al. 2021](#)].

**Virtual Concertina** In this last use-case, we provide an example of how our approach to UMH feedback can be applied to guiding a user’s motions. Users manipulate a virtual concertina and must move their hands in rhythmical back-and-forth motions to trigger bars of music. Spatio-temporally modulated circles which grow and shrink in accordance with the ideal opening and closing motions to be made by the hands are projected onto the users’ palms, providing a tactile guide for playing the melody with the correct rhythm.

## 9.7 Conclusion

In this chapter, we described a novel approach for providing bimanual ultrasound mid-air haptic (UMH) feedback during bimanual manipulation in VR. By using a pair of UMH devices back-to-back, we enabled simultaneous bimanual manipulation of virtual objects in opposing hand configurations. We investigated the benefits of bimanual UMH feedback in two representative human participant experiments studying a bimanual object grasping and holding task, and a bimanual 3D shape exploration and identification task.

Results show that our approach yields performances at least on par with state-of-the-art unimanual UMH feedback, and significantly improves user experience (lower workload, higher preference, greater ease to establish an interaction strategy) during such bimanual manipulation when compared to unimanual UMH feedback. These results confirm the viability of

our approach, complementing the current state-of-the-art for bimanual ultrasound haptics in VR, which thus far has been limited to simple demonstrations of interactions with coplanar hands. In addition, while many tasks would be impossible with unimanual feedback, our system would enable their completion.

### Chapter Summary

- We introduce a system and its associated coupling scheme to provide UMH bimanual stimulation during VR interactions;
- A grasping and holding experiment, using UMH to provide force feedback about the ongoing manipulation. Another shape exploration and identification experiment was conducted, using UMH to haptically render the shape of an invisible object. Both experiment showed promising bimanual performances that are at least on par with state-of-the-art unimanual methods, while significantly improving the user experience;
- We provided three use cases of bimanual interactions enabled by our contributions.

# Chapter 10

## Conclusion

### Contents

---

10.1 Summary and Short-Term Future Work . . . . .	235
10.2 Perspectives . . . . .	239

---

### 10.1 Summary and Short-Term Future Work

This thesis explores the design behind UMH rendering, including contributions from different points of view, such as algorithmics, neurosciences, biomechanics, and engineering. Our work revolved around four research axes: first, we studied **human perception** of UMH stimuli (**Challenge I, Part I**), before providing new **rendering primitives** (**Challenge II, Part II**). We then formalized the UMH **rendering pipeline** to provide common baselines for discussing UMH (**Challenge III, Part III**). Finally, we investigated the use of UMH to complement or improve **VR interactions** (**Challenge IV, Part IV**).

Specifically, in **Chapter 2** we present measurements of the fundamental response of LTMRs to UMH stimuli, using microneurography recordings on the median nerve. Doing so, we aim to characterize the afferent response, highlighting the dominant role of FA2 afferents in the perception of UMH stimuli on the palm and fingers, contributing to **Challenge I**. We also performed some OCT recordings to have in-depth measurements of skin deformation in re-

sponse to UMH stimulation. While further analyses will soon be conducted, we provide the first measurements at the afferent level for UMH, which is an important step towards a better understanding of UMH perception.

This chapter presents promising early results of a project that remains largely ongoing. We are also in the process of recording afferent response in the radial nerve, innervating hairy skin, where we assume CT afferents to have an important role in responding to acoustic streaming. We will also conduct formal analyses to properly characterize the afferent's response rate for the different stimulus conditions. We will also work on relating these results to the observed deformations visualized by the OCT.

Then, [Chapter 3](#) explores the perception of 2D shapes rendered using dynamic stimuli. We propose spatio-temporally-modulated tactile pointers (STP), a new algorithm for rendering these shapes, improving upon dynamic tactile pointers (DTP). These techniques leverage the cognitive benefits of chunking by breaking the shape perception into different chunks, corresponding to each side of the shape. Yet, we improved upon DTP by replacing the weak AM used by STM, providing significantly stronger sensations. But the design of STP imposes a relatively slow focal speed, thus avoiding the generation of shear shock waves in the skin that were usually responsible for the blurry sensations of STM. In doing so, we provide a rendering technique that is both intense and sharp, achieving shape identification performances on par with the best state-of-the-art rendering techniques. In addition to providing the new best rendering technique for dynamic shape stimuli, our studies provide valuable insights into the perception of dynamic UMH stimuli, thus contributing to **Challenge I**.

Future work will be focusing on optimizing the different rendering parameters, and explore variants of this method, such as using overlapping segments in STP. We will also focus on investigating whether the rendering of more complex shapes is possible, and determining what is required from an ultrasound haptic rendering algorithm to do so effectively. As STP and DTP both render dynamic stimuli, it would also be beneficial to apply this vertex highlighting concept to static stimuli. To obtain a static stimulus, pauses would not be possible, implying that the rendering will need to leverage other techniques (e.g. higher intensity around the vertices), and should be evaluated during similar shape identification studies.

In [Chapter 4](#) we dig deeper into the biomechanics of skin deformation in response to UMH stimulation. Specifically, we leverage the generated shear shock waves generated by the fast focal motion during STM stimulation, in order to focus them at a location of our choosing, thus enabling a reuse of the acoustic energy. We did so using logarithmic spiral focal trajectories, which is a new rendering primitive, based on physics-based models. By providing new possi-

bilities to UMH designers, [Chapter 4](#) contributes to **Challenge II**. Specifically, we illustrated the use of these new primitives for the rendering of 2D shapes, and haptic shading of specific areas. We obtained very positive results, showing that despite the non-geometrically accurate focal path and the non-optimized rendering primitives, we already reach identification performances on par with state-of-the-art methods. This is extremely promising for future uses of this method, while also inciting researchers in anchoring future methods in practical physical or perceptual models.

Future studies will be conducted to optimize spiral parameters and measure their perceptual impact (e.g. perceived intensity, stimulated area). We will also continue our exploration of the design space to compare different spiral-based rendering techniques, to further highlight the versatility of our method.

[Chapter 5](#) provides the first contributions to **Challenge III** by formalizing sampling strategies, representing the transformation from a 2D shape or set of curves into a UMH stimulus. We developed DOLPHIN, a framework for the low-level design of UMH stimuli, specifically suited for the study of sampling strategies. We used DOLPHIN in a proof of concept user study investigating the effect of the sample rate on curvature discrimination performances. These contributions highlight the usefulness of such frameworks, and the importance of common baselines for the reproducibility and discussion of our works. Additionally, the open-source nature of our framework allows for a wide and free use, and can generate feedback from the community.

Short term future work should focus on minor features for the design tool such as support for new shapes, new sampling strategies, and other hardware platforms, along with a detailed evaluation of cross-platforms compatibility. Support for rendering multiple simultaneous focal points should also be included in the future. We intend to interface the tool with simulations of the acoustic fields and spatio-temporal behaviour of skin indentation. This should provide designers with more information on how to achieve specific effects and help explain perception results with a physical model.

[Chapter 6](#) directly expands on the previous chapter, formalizing intersection strategies, the process of computing a set of curves based on the intersection between a hand and a 3D object being experienced. The resulting curves can then be fed into a sampling strategy to generate the final UMH stimulus. Overall, the combination of intersection strategies and sampling strategies form an end-to-end formalized framework for the rendering of 3D objects, contributing to **Challenge III**. To highlight the importance of studying intersection strategies, we designed example strategies and conducted a user study that revealed significant

differences in perceived object properties. Specifically, the example intersection strategy we designed allowed for the rendering of more filled or solid objects.

Future studies could study the potential impacts of other more complex intersection strategy designs on the perception of other virtual object properties, and look into the rendering of other volumetric properties. We will also investigate the dependencies between intersection strategies and sampling strategies. It would also be interesting to expand DOLPHIN’s capabilities to incorporate intersection strategies in the stimulus design framework.

Chapter 7 begins Part IV, contributing to **Challenge IV**. In this chapter, we proposed feedback techniques that are complementary to the VR interaction. These techniques aim at guiding the user’s hand towards a target pose, either in terms of position or orientation (2 DoF). We evaluated these techniques both in static and dynamic virtual scenarios to assess their performances. This allowed us to find the best position and orientation technique, while also observing a strong correlation between the feedback’s usefulness and the task’s difficulty. This difficulty influence will be the topic of future investigations. We will also explore the optimization of the guidance strategy designs by investigating the perception of dynamic shapes. Notably, we will also look into extending these techniques to 3 DoF and continuous motion guidance.

While Chapter 7 uses haptics as a complementary channel to provide new information, Chapter 8 provides redundant feedback, combining visual and UMH cues to render the interpenetration with a virtual object during VR interactions, as a contribution to **Challenge IV**. We reused and developed both visual and tactile feedback, before conducted user studies to find the best visuo-haptic feedback pair. Interestingly, our study showed that participants prefer an immediately strong tactile sensation, highlighting the predominant importance of the impact phase. We also showed that the best technique may vary depending on the context. Participants perform better with a very strong visuo-haptic feedback, while they seem to appreciate more the use of a subtler and more natural feedback. The feedback thus has to be chosen carefully, depending on whether the interaction values more performances or user experience. Future work should study these coupling schemes during more complex scenarios that include multiple fingers, the palm, and ultimately the bimanual interactions investigated in the following chapter. We should also evaluate our techniques with more diverse objects, for which interpenetration may be harder to perceive. We will additionally explore the rendering of the other components of contact, by first focusing on the contact, which has been shown to have an important role in our first experiment.

Finally, Chapter 9 provides our last contributions to **Challenge IV** by generalizing UMH cou-

pling schemes for bimanual manipulations in VR. Specifically, we provided algorithms for rendering the physical manipulation or exploration of virtual objects. We evaluated the impact of bimanual interactions and bimanual stimulation during both types of VR scenarios. Doing so revealed that bimanual interactions and bimanual feedback that are always at least as good as unimanual, while often significantly improving the user experience by reducing the workload and increasing the appreciation.

Future work will be focused on further improving the system, while also studying its performances for other types of interactions and other hand configurations. We also plan to work on key limitations of this system, in particular its workspace, and intend to provide the ability to enable both coplanar and opposing hand configuration, e.g. by actuating the pair of UMH devices.

## 10.2 Perspectives

While this thesis contributed to UMH advancements in each of the four challenges mentioned in the [Introduction Chapter](#), UMH remains a relatively novel technology, with most of its scientific future still ahead.

As discussed in [Part I](#), a good understanding of human perception is crucial to design efficient rendering techniques. Studies must therefore dig deeper into the mechanisms behind UMH perception. While we showed that microneurography is a very interesting physiological approach to directly measure the response of single afferent units to UMH stimulation, our proposed contributions should be largely extended to many skin locations. This would allow the derivation of a perceptual map of skin response across the human skin. Stronger interfaces may also enable the stimulation of other types of receptors, adding a new layer of complexity and possibilities.

It will also be beneficial to go a step further in the perceptual process to understand how responses from different units are combined into a complex percept. This is especially important for UMH due to the richness of sensations it can provide and the size of its design space compared to other tactile technologies (e.g. vibration motors being constrained to single location stimulation).

Going even further, psychophysics studies have been used more extensively for the evaluation of higher-level perception of UMH stimuli. These techniques should be more systematically used to better understand how simple patterns are perceived. For instance, we started a project investigating the purely geometric perception of UMH shapes, by trying to eliminate biases caused by non-geometric cues (in alphabetical order, contributions of Maud Marchal, Eileen

Mathey, Lendy Mulot, and Claudio Pacchierotti). While this project is still at its early stages, which is why it has not been included in this thesis, we believe that this type of studies would be greatly beneficial for UMH designers.

Part II introduced a whole new research directions for UMH rendering. AM was initially introduced as the simplest possible modulation, similar to what is used for other vibrotactile devices. Later, STM/LM were introduced to counter the intensity limitations of AM, with the added spatial motion coming as a very natural and elegant solution to the non-perceivable static focal point problem. Yet, we showed that we can also root our rendering primitives in physical processes, opening the way for a new category of rendering techniques. Specifically, while our work derived focal motion to focus wavefronts to single locations, the future of UMH rendering may be expanded in a more general method that could focus shear waves along more complex patterns (e.g. a shape). Achieving such goal also requires refinements of current skin deformation models, which are the basis for such studies.

One could also take a parallel approach to design perception-based rendering methods. Such method would adapt to the topology and type of skin stimulated, attempting to stimulate specific receptors in order to elicit target sensations. Yet, achieving such goals would require strong advancements to **Challenge I**.

Now that Part III provided a formalized baseline for UMH rendering, it becomes important to systematically map the design and perceptual space of sampling strategies and intersection strategies. This can be done by considering all the rendering parameters, their relations to each other, and how they can affect perception. UMH is particularly complex in this regard, as many of its rendering parameters are interdependent (e.g. changing the focal speed can alter the sample rate and inter-sample distance, or the dwell time, while simultaneously altering the draw frequency). This is why having standardized frameworks like DOLPHIN can help in this systematic pursuit of perception knowledge. It should therefore be extended to support intersection strategies, and an overall wider set of input rendering parameters.

Visualization is also a crucial component when designing UMH stimuli. Since their airborne nature makes them invisible to the high, and since visual stimuli or expectations may take priority over the tactile stimuli, it is therefore important to have a good visualisation of the actual rendered stimulus. While a few methods have been discussed in Section 1.4.4, most of them require complex setups to be used. This is why we (in alphabetical order, contributions of Maud Marchal, Lendy Mulot, and Visaki Nithianantham) chose to reproduce and adapt the device of Hasegawa [2023]. The resulting tool is a small hand-held device that can display the acoustic field at its location, providing significant help for debugging, designing UMH patterns,

and for scientific dissemination purposes.

Finally, UMH should continue to be used to improve user interactions. While we provided contributions to a few specific interactions, future studies should focus on a larger range of 3D interactions. Systematic assessments of user experience, appreciation, and performances should be conducted to evaluate the corresponding UMH coupling schemes. For instance, deformable objects add a new layer of complexity when interacting with them, while also providing coupled visual feedback when deforming. Interactions with such objects should therefore be studied in depth.

Yet, the range of available interactions is very limited by the size of the UMH workspace. This is why we started working on a new device, improving upon the work of [Howard et al. \[2020\]](#) to provide bimanual or unimanual feedback in a large workspace, supporting a large range of hand configurations. Yet, this project is currently at its early stages.

While this thesis strongly focused on UMH rendering to provide tactile stimulation, it could be very beneficial to use multimodal stimulation for richer tactile sensations during interactions. This would require new hardware for combining two or more modalities, as well as software coupling schemes to synchronize these modalities and ensure they remain congruent.



## Bibliography

- [Abdouni et al. 2019] Abdenaceur Abdouni, Rory Clark, and Orestis Georgiou. « Seeing is believing but feeling is the truth: Visualising mid-air haptics in oil baths and lightboxes ». In: *ACM International Conference on Multimodal Interaction*. 2019, pp. 504–505.
- [Ablart et al. 2019] Damien Ablart, William Frier, Hannah Limerick, Orestis Georgiou, and Marianna Obrist. « Using ultrasonic mid-air haptic patterns in multi-modal user experiences ». In: *IEEE International Symposium on Haptic, Audio and Visual Environments and Games*. IEEE. 2019, pp. 1–6.
- [Adel et al. 2018] Alaa Adel, Mina M Micheal, Mohamed Abou Self, Slim Abdennadher, and Islam SM Khalil. « Rendering of virtual volumetric shapes using an electromagnetic-based haptic interface ». In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2018, pp. 1–9.
- [Adilkhanov et al. 2022] Adilzhan Adilkhanov, Matteo Rubagotti, and Zhanat Kapassov. « Haptic devices: Wearability-based taxonomy and literature review ». In: *IEEE Access* **10** (2022), pp. 91923–91947.
- [Ahmaniemi et al. 2010] Teemu Ahmaniemi, Juha Marila, and Vuokko Lantz. « Design of dynamic vibrotactile textures ». In: *IEEE Transactions on Haptics* **3**.4 (2010), pp. 245–256.

- 
- [Allevato et al. 2022] Gianni Allevato, Matthias Rutsch, Jan Hinrichs, Christoph Haugwitz, Raphael Müller, Marius Pesavento, and Mario Kupnik. « Air-coupled ultrasonic spiral phased array for high-precision beamforming and imaging ». In: *IEEE Open Journal of Ultrasonics, Ferroelectrics, and Frequency Control* **2** (2022), pp. 40–54.
- [Allevato et al. 2023] Gianni Allevato, Christoph Haugwitz, Matthias Rutsch, Raphael Müller, Marius Pesavento, and Mario Kupnik. « Two-scale sparse spiral array design for 3D ultrasound imaging in air ». In: *IEEE Open Journal of Ultrasonics, Ferroelectrics, and Frequency Control* **3** (2023), pp. 113–127.
- [Andersen et al. 2017] Hjalte H Andersen, Jesper Elberling, Silvia Lo Vecchio, and Lars Arendt-Nielsen. « Topography of itch: evidence of distinct coding for pruriception in the trigeminal nerve ». In: *Itch* **2.1** (2017), e2.
- [Arafsha et al. 2015] Faisal Arafsha, Longyu Zhang, Haiwei Dong, and Abdulmotaaleb El Saddik. « Contactless haptic feedback: State of the art ». In: *IEEE International Symposium on Haptic, Audio and Visual Environments and Games*. IEEE. 2015, pp. 1–6.
- [Awatani 1955] Jobu Awatani. « Studies on Acoustic Radiation Pressure. I. (General Considerations) ». In: *The Journal of the Acoustical Society of America* **27.2** (1955), pp. 278–281.
- [Balint et al. 2018] P Balint and Kaspar Althoefer. « Medical virtual reality palpation training using ultrasound based haptics and image processing ». In: vol. 1. 2. 2018.
- [Barreiro et al. 2020] Hector Barreiro, Stephen Sinclair, and Miguel A Otaduy. « Path routing optimization for STM ultrasound rendering ». In: *IEEE Transactions on Haptics* **13.1** (2020), pp. 45–51.
- [Barreiro et al. 2021] Héctor Barreiro, Joan Torres, and Miguel A Otaduy. « Natural tactile interaction with virtual clay ». In: *IEEE World Haptics Conference*. IEEE. 2021, pp. 403–408.

- 
- [Bass et al. 1995] Henry E Bass, Louis C Sutherland, Allen J Zuckerwar, David T Blackstock, and DM Hester. « Atmospheric absorption of sound: Further developments ». In: *The Journal of the Acoustical Society of America* (1995).
- [Battista 2022] Andrew Di Battista. « Safety of high-intensity ultrasound ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 347–362.
- [Bau et al. 2010] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. « TeslaTouch: electrovibration for touch surfaces ». In: *ACM Symposium on User Interface Software and Technology*. 2010, pp. 283–292.
- [Bauer et al. 2021] Peter Bauer, Werner Lienhart, and Samuel Jost. « Accuracy investigation of the pose determination of a VR system ». In: *Sensors* **21**.5 (2021), p. 1622.
- [Beattie et al. 2020] David Beattie, William Frier, Orestis Georgiou, Benjamin Long, and Damien Ablart. « Incorporating the perception of visual roughness into the design of mid-air haptic textures ». In: *ACM Symposium on Applied Perception*. 2020, pp. 1–10.
- [Berkman et al. 2019] Mehmet Ilker Berkman and Ecehan Akan. « Presence and immersion in virtual reality ». In: *Encyclopedia of Computer Graphics and Games*. 2019, pp. 1–10.
- [Bianchi et al. 2015] Matteo Bianchi and Alessandro Serio. « Design and characterization of a fabric-based softness display ». In: *IEEE Transactions on Haptics* **8**.2 (2015), pp. 152–163.
- [Biswas et al. 2019] Shantonu Biswas and Yon Visell. « Emerging material technologies for haptics ». In: *Advanced Materials Technologies* **4**.4 (2019), p. 1900042.
- [Bolanowski Jr et al. 1988] Stanley J Bolanowski Jr, George A Gescheider, Ronald T Verrillo, and Christin M Checkosky. « Four channels mediate the mechanical aspects of touch ». In: *The Journal of the Acoustical society of America* **84**.5 (1988), pp. 1680–1694.

- 
- [Bouzbib et al. 2024] Elodie Bouzbib, Marc Teyssier, Thomas Howard, Claudio Pacchierotti, and Anatole Lécuyer. « Palmex: Adding palmar force-feedback for 3d manipulation with haptic exoskeleton gloves ». In: *IEEE Transactions on Visualization and Computer Graphics* **30**.7 (2024), pp. 3973–3980.
- [Brice et al. 2019] Daniel Brice, Thomas McRoberts, and Karen Rafferty. « A proof of concept integrated multi-systems approach for large scale tactile feedback in VR ». In: *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer. 2019, pp. 120–137.
- [Brice et al. 2021] Daniel Brice, Zara Gibson, Fintan McGuinness, and Karen Rafferty. « Using ultrasonic haptics within an immersive spider exposure environment to provide a multi-sensorial experience ». In: *Frontiers in Virtual Reality* **2** (2021), pp. 1–12.
- [Brown et al. 2020] Eddie Brown, David R. Large, Hannah Limerick, and Gary Burnett. « Ultrahapticons: “Haptifying” drivers’ mental models to transform automotive mid-air haptic gesture infotainment interfaces ». In: *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2020, pp. 54–57.
- [Brown et al. 2022a] Eddie Brown, David R Large, Hannah Limerick, William Frier, and Gary Burnett. « Augmenting automotive gesture infotainment interfaces through mid-air haptic icon design ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 119–145.
- [Brown et al. 2022b] Eddie Brown, David R Large, Hannah Limerick, William Frier, and Gary Burnett. « Validating the salience of haptic icons for automotive mid-air haptic gesture interfaces ». In: *Ergonomics & Human Factors* **2022** (2022).
- [Bruus 2012] Henrik Bruus. « Acoustofluidics 7: The acoustic radiation force on small particles ». In: *Lab on a Chip* **12**.6 (2012), pp. 1014–1021.

- 
- [Cabaret et al. 2024] Pierre-Antoine Cabaret, Claudio Pacchierotti, Marie Babel, and Maud Marchal. « Design of haptic rendering techniques for navigating with a multi-actuator vibrotactile handle ». In: *EuroHaptics*. Springer. 2024, pp. 224–236.
- [Carter et al. 2013] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. « UltraHaptics: multi-point mid-air haptic feedback for touch surfaces ». In: *ACM Symposium on User Interface Software and Technology*. 2013, pp. 505–514.
- [Cataldo et al. 2023] Antonio Cataldo, William Frier, and Patrick Haggard. « Quantifying spatial acuity of frequency resolved midair ultrasound vibrotactile stimuli ». In: *Scientific Reports* **13.1**(2023), p. 21149.
- [Cempini et al. 2015] Marco Cempini, Mario Cortese, and Nicola Vitiello. « A powered finger–thumb wearable hand exoskeleton with self-aligning joint axes ». In: *IEEE/ASME Transactions on Mechatronics* **20.2** (2015), pp. 705–716.
- [Cha et al. 2017] Hojun Cha, Hojin Lee, Junsuk Park, Hyung-Sik Kim, Soon-Cheol Chung, and Seungmoon Choi. « Mid-air tactile display using indirect laser radiation for contour-following stimulation and assessment of its spatial acuity ». In: *IEEE World Haptics Conference*. IEEE. 2017, pp. 136–141.
- [Chongyang et al. 2019] SUN Chongyang, NAI Weizhi, and SUN Xiaoying. « Tactile sensitivity in ultrasonic haptics: Do different parts of hand and different rendering methods have an impact on perceptual threshold? ». In: *Virtual Reality & Intelligent Hardware* **1.3** (2019), pp. 265–275.
- [Clarence et al. 2022] Aldrich Clarence, Jarrod Knibbe, Maxime Cordeil, and Michael Wybrow. « Investigating the effect of direction on the limits of haptic retargeting ». In: *IEEE international symposium on mixed and augmented reality*. IEEE. 2022, pp. 612–621.
- [Cornelio Martinez et al. 2017] Patricia Ivette Cornelio Martinez, Silvana De Pirro, Chi Thanh Vi, and Sriram Subramanian. « Agency in mid-air interfaces ». In: *Conference on Human Factors in Computing Systems*. 2017, pp. 2426–2439.

- 
- [Corniani et al. 2020] Giulia Corniani and Hannes P Saal. « Tactile innervation densities across the whole body ». In: *Journal of Neurophysiology* **124**.4 (2020), pp. 1229–1240.
- [Culbertson et al. 2018] Heather Culbertson, Samuel B Schorr, and Allison M Okamura. « Haptics: The present and future of artificial touch sensation ». In: *Annual Review of Control, Robotics, and Autonomous Systems* **1**.1 (2018), pp. 385–409.
- [Danilov et al. 2000] SD Danilov and MA Mironov. « Mean force on a small sphere in a sound field in a viscous fluid ». In: *The Journal of the Acoustical Society of America* **107**.1 (2000), pp. 143–153.
- [Deflorio et al. 2022] Davide Deflorio, Massimiliano Di Luca, and Alan M Wing. « Skin and mechanoreceptor contribution to tactile input for perception: a review of simulation models ». In: *Frontiers in Human Neuroscience* **16** (2022), pp. 1–19.
- [Di Luca et al. 2019] Massimiliano Di Luca and Arash Mahnan. « Perceptual limits of visual-haptic simultaneity in virtual reality interactions ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 67–72.
- [Dobrev et al. 2017] Ivo Dobrev, Jae Hoon Sim, Stefan Stenfelt, Sebastian Ihrle, Rahel Gerig, Flurin Pfiffner, Albrecht Eiber, Alexander M Huber, and Christof Röösli. « Sound wave propagation on the human skull surface with bone conduction stimulation ». In: *Hearing Research* **355** (2017), pp. 1–13.
- [Doinikov 1994] Alexander A Doinikov. « Acoustic radiation pressure on a compressible sphere in a viscous fluid ». In: *Journal of Fluid Mechanics* **267** (1994), pp. 1–22.
- [Donkov et al. 2025] Stefan Donkov, Elodie Bouzbib, Mikel Aldea, Josu Irisarri, Sonia Elizondo, Inigo Ezcurdia, and Asier Marzo. « SparkTouch: Contactless haptic spatial patterns on the Palm and fingertip using electric sparks ». In: *IEEE World Haptics Conference*. 2025.

- 
- [Drewing et al. 2005] Knut Drewing, Michael Fritschi, Regine Zopf, Marc O Ernst, and Martin Buss. « First evaluation of a novel tactile display exerting shear force via lateral displacement ». In: *ACM Transactions on Applied Perception* **2**.2 (2005), pp. 118–131.
- [Drinkwater 2022] Bruce W Drinkwater. « The physical principles of arrays for mid-air haptic applications ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 313–334.
- [Dube et al. 2023] Tafadzwa Joseph Dube and Ahmed Sabbir Arif. « Ultrasonic keyboard: A mid-air virtual QWERTY with ultrasonic feedback for virtual reality ». In: *International Conference on Tangible, Embedded, and Embodied Interaction*. 2023, pp. 1–8.
- [Dupin et al. 2015] Lucile Dupin, Vincent Hayward, and Mark Wexler. « Direct coupling of haptic signals between hands ». In: *Proceedings of the National Academy of Sciences* **112**.2 (2015), pp. 619–624.
- [Duvernoy et al. 2023] Basil Duvernoy, Emma Kindström, Anders Fridberger, and Sarah McIntyre. « Tracking submicrometer vibrations in depth in the glabrous skin ». In: *Skin* **600** (2023), p. 800.
- [Elouneg et al. 2023] Aflah Elouneg, Jérôme Chambert, Arnaud Lejeune, Quentin Lucot, Emmanuelle Jacquet, and SPA Bordas. « Anisotropic mechanical characterization of human skin by in vivo multi-axial ring suction test ». In: *Journal of the Mechanical Behavior of Biomedical Materials* **141** (2023), p. 105779.
- [Evangelou et al. 2021] George Evangelou, Hannah Limerick, and James Moore. « I feel it in my fingers! Sense of agency with mid-air haptics ». In: *IEEE World Haptics Conference*. IEEE. 2021, pp. 727–732.
- [Evangelou et al. 2023] George Evangelou, Orestis Georgiou, and James Moore. « Using virtual objects with hand-tracking: the effects of visual congruence and mid-air haptics on sense of agency ». In: *IEEE Transactions on Haptics* **16**.4 (2023), pp. 580–585.
- [Evangelou et al. 2024] George Evangelou, Orestis Georgiou, and James Moore. « Mid-air haptic congruence with virtual objects modulates the implicit sense of agency ». In: *EuroHaptics*. Springer. 2024, pp. 440–447.

- 
- [Evangelou 2025] George Evangelou. « Investigating the effects of mid-air haptics on the sense of agency ». PhD thesis. Goldsmiths, University of London, 2025.
- [Everett et al. 2013] Janine S Everett and Marilyn S Sommers. « Skin viscoelasticity: physiologic mechanisms, measurement issues, and application to nursing science ». In: *Biological Research for Nursing* 15.3 (2013), pp. 338–346.
- [Faridan et al. 2022] Mehrad Faridan, Marcus Friedel, and Ryo Suzuki. « Ultra-Bots: Large-area mid-air haptics for VR with robotically actuated ultrasound transducers ». In: *ACM Symposium on User Interface Software and Technology*. 2022, pp. 1–3.
- [Freeman et al. 2017] Euan Freeman, Ross Anderson, Julie Williamson, Graham Wilson, and Stephen A Brewster. « Textured surfaces for ultrasound haptic displays ». In: *ACM International Conference on Multimodal Interaction*. 2017, pp. 491–492.
- [Freeman et al. 2019] Euan Freeman, Dong-Bach Vo, and Stephen Brewster. « HaptiGlow: Helping users position their hands for better mid-air gestures and ultrasound haptic feedback ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 289–294.
- [Frier et al. 2018] William Frier, Damien Ablart, Jamie Chilles, Benjamin Long, Marcello Giordano, Marianna Obrist, and Sriram Subramanian. « Using spatiotemporal modulation to draw tactile patterns in mid-air ». In: *EuroHaptics*. Springer. 2018, pp. 270–281.
- [Frier et al. 2019] William Frier, Dario Pittera, Damien Ablart, Marianna Obrist, and Sriram Subramanian. « Sampling strategy for ultrasonic mid-air haptics ». In: *Conference on Human Factors in Computing Systems*. 2019, pp. 1–11.
- [Frutos-Pascual et al. 2019] Maite Frutos-Pascual, Jake Michael Harrison, Chris Creed, and Ian Williams. « Evaluation of ultrasound haptics as a supplementary feedback cue for grasping in virtual environments ». In: *ACM International Conference on Multimodal Interaction*. 2019, pp. 310–318.

- 
- [Furumoto et al. 2019] Takuro Furumoto, Keisuke Hasegawa, Yasutoshi Makino, and Hiroyuki Shinoda. « Three-dimensional manipulation of a spherical object using ultrasound plane waves ». In: *IEEE Robotics and Automation Letters* **4**.1 (2019), pp. 81–88.
- [Furumoto et al. 2020] Takuro Furumoto, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « Balloon interface for midair haptic interaction ». In: *SIGGRAPH Asia (Emerging Technologies)*. 2020, pp. 1–2.
- [Gavrilov 2008] Leonid R Gavrilov. « The possibility of generating focal regions of complex configurations in application to the problems of stimulation of human receptor structures by focused ultrasound ». In: *Acoustical Physics* **54**.2 (2008), pp. 269–278.
- [Georgiou et al. 2017] Orestis Georgiou, Valerio Biscione, Adam Harwood, Daniel Griffiths, Marcello Giordano, Ben Long, and Tom Carter. « Haptic in-vehicle gesture controls ». In: *ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*. 2017, pp. 233–238.
- [Georgiou et al. 2018] Orestis Georgiou, Craig Jeffrey, Ziyuan Chen, Bao Xiao Tong, Shing Hei Chan, Boyin Yang, Adam Harwood, and Tom Carter. « Touchless haptic feedback for VR rhythm games ». In: *IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE. 2018, pp. 553–554.
- [Georgiou et al. 2022a] Orestis Georgiou, William Frier, Euan Freeman, Claudio Pacchierotti, and Takayuki Hoshi. *Ultrasound mid-air haptics for touchless interfaces*. Vol. 3. 2022.
- [Georgiou et al. 2022b] Orestis Georgiou, William Frier, and Oliver Schneider. « User experience and mid-air haptics: Applications, methods, and challenges ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 21–69.
- [Gil et al. 2018] Hyunjae Gil, Hyungki Son, Jin Ryong Kim, and Ian Oakley. « Whiskers: Exploring the use of ultrasonic haptic cues on the face ». In: *Conference on Human Factors in Computing Systems*. 2018, pp. 1–13.

- 
- [Goldstein et al. 2022a] E. Bruce Goldstein and Laura Cacciamani. *Sensation and Perception*. 11th ed. 2022.
- [Goldstein et al. 2022b] E. Bruce Goldstein and Laura Cacciamani. « The Cutaneous Senses ». In: *Sensation and Perception*. 11th ed. 2022, pp. 357–386.
- [Gupta et al. 2013] Sidhant Gupta, Dan Morris, Shwetak N Patel, and Desney Tan. « Airwave: Non-contact haptic feedback using air vortex rings ». In: *ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2013, pp. 419–428.
- [Hajas et al. 2020] Daniel Hajas, Dario Pittera, Antony Nasce, Orestis Georgiou, and Marianna Obrist. « Mid-air haptic rendering of 2D geometric shapes with a dynamic tactile pointer ». In: *IEEE Transactions on Haptics* **13**.4 (2020), pp. 806–817.
- [Harada et al. 2022] Yuki Harada and Junji Ohyama. « Quantitative evaluation of visual guidance effects for 360-degree directions ». In: *Virtual Reality* **26**.2 (2022), pp. 759–770.
- [Hart 2006] Sandra G Hart. « NASA-task load index (NASA-TLX); 20 years later ». In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage publications Sage CA: Los Angeles, CA. 2006, pp. 904–908.
- [Hasegawa et al. 2000] Takahiro Hasegawa, Tohru Kido, Takeshi Iizuka, and Chihiro Matsuoka. « A general theory of Rayleigh and Langevin radiation pressures ». In: *Acoustical Science and Technology* **21**.3 (2000), pp. 145–152.
- [Hasegawa et al. 2013] Keisuke Hasegawa and Hiroyuki Shinoda. « Aerial display of vibrotactile sensation with high spatial-temporal resolution using large-aperture airborne ultrasound phased array ». In: *IEEE World Haptics Conference*. IEEE. 2013, pp. 31–36.
- [Hasegawa et al. 2018] Keisuke Hasegawa and Hiroyuki Shinoda. « Aerial vibrotactile display based on multiunit ultrasound phased array ». In: *IEEE Transactions on Haptics* **11**.3 (2018), pp. 367–377.

- 
- [Hasegawa et al. 2022] Keisuke Hasegawa and Hiroyuki Shinoda. « Modulation methods for ultrasound midair haptics ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 225–240.
- [Hasegawa 2023] Keisuke Hasegawa. « Two-dimensional visualization of intense ultrasound field using confined particle movement caused by radiation pressure field ». In: *IEEE International Ultrasonics Symposium*. IEEE. 2023, pp. 1–4.
- [Hashizume et al. 2017] Satoshi Hashizume, Amy Koike, Takayuki Hoshi, and Yoichi Ochiai. « Sonovortex: Rendering multi-resolution aerial haptics by aerodynamic vortex and focused ultrasound ». In: *ACM SIGGRAPH (Posters)*. 2017, pp. 1–2.
- [Hauser et al. 2019] Steven C Hauser, Sarah McIntyre, Ali Israr, Håkan Olausson, and Gregory J Gerling. « Uncovering human-to-human physical interactions that underlie emotional and affective touch communication ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 407–412.
- [Hein et al. 2022] Philipp Hein, Dursune Gönültas, Max Bernhagen, André Dettmann, and Angelika C Bullinger. « Mid-air haptic cues for safety-critical communication: (re-)exploring air-vortex rings in human-machine interaction ». In: *Mensch und Computer*. 2022, pp. 41–51.
- [Hoshi et al. 2009] Takayuki Hoshi, Takayuki Iwamoto, and Hiroyuki Shinoda. « Non-contact tactile sensation synthesized by ultrasound transducers ». In: *IEEE World Haptics Conference*. IEEE. 2009, pp. 256–260.
- [Hoshi et al. 2010] Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, and Hiroyuki Shinoda. « Noncontact tactile display based on radiation pressure of airborne ultrasound ». In: *IEEE Transactions on Haptics* 3.3 (2010), pp. 155–165.
- [Hoshi et al. 2021] Takayuki Hoshi and Yoshiki Ooka. « Experimental Verification of Nonlinear Attenuation of Airborne Ultrasound ». In: *Symposium on Ultrasonic Electronics*. Vol. 42. Institute for Ultrasonic Electronics. 2021, 3Pb4–5.

- 
- [Hoshi 2014] Takayuki Hoshi. « Noncontact tactile display using airborne ultrasound ». In: *International Display Workshops*. 2014, pp. 1529–1533.
- [Hoshi 2022a] Takayuki Hoshi. « Introduction to ultrasonic mid-air haptic effects ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 1–20.
- [Hoshi 2022b] Takayuki Hoshi. « Ultrasound exposure in mid-air haptics ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 363–383.
- [Howard et al. 2019] Thomas Howard, Gerard Gallagher, Anatole Lécuyer, Claudio Pacchierotti, and Maud Marchal. « Investigating the recognition of local shapes using mid-air ultrasound haptics ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 503–508.
- [Howard et al. 2020] Thomas Howard, Maud Marchal, Anatole Lécuyer, and Claudio Pacchierotti. « PUMAH: pan-tilt ultrasound mid-air haptics for larger interaction workspace in virtual reality ». In: *IEEE Transactions on Haptics* **13**.1 (2020), pp. 38–44.
- [Howard et al. 2022] Thomas Howard, Maud Marchal, and Claudio Pacchierotti. « Ultrasound mid-air tactile feedback for immersive virtual reality interaction ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 147–183.
- [Howard et al. 2023a] Thomas Howard, Karina Driller, William Frier, Claudio Pacchierotti, Maud Marchal, and Jessica Hartcher-O'Brien. « Gap detection in pairs of ultrasound mid-air vibrotactile stimuli ». In: *ACM Transactions on Applied Perceptions* **20**.1 (2023), pp. 1–17.
- [Howard et al. 2023b] Thomas Howard, Guillaume Gicquel, Claudio Pacchierotti, and Maud Marchal. « Can we effectively combine tangibles and ultrasound mid-air haptics? A study of acoustically transparent tangible surfaces ». In: *IEEE Transactions on Haptics* **16**.4 (2023), pp. 477–483.

- 
- [Huang et al. 1991] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, et al. « Optical coherence tomography ». In: *science* **254**.5035 (1991), pp. 1178–1181.
- [Hwang et al. 2017] Inwook Hwang, Hyungki Son, and Jin Ryong Kim. « AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback ». In: *IEEE World Haptics Conference*. IEEE. 2017, pp. 213–218.
- [Inoue et al. 2014] Seki Inoue, Koseki J Kobayashi-Kirschvink, Yasuaki Monnai, Keisuke Hasegawa, Yasutoshi Makino, and Hiroyuki Shinoda. « HORN: the hapt-optic reconstruction ». In: *ACM SIGGRAPH (Emerging Technologies)*. 2014, pp. 1–1.
- [Inoue et al. 2015] Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. « Active touch perception produced by airborne ultrasonic haptic hologram ». In: *IEEE World Haptics Conference*. IEEE. 2015, pp. 362–367.
- [Inoue et al. 2016] Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. « Scalable architecture for airborne ultrasound tactile display ». In: *International AsiaHaptics Conference*. Springer. 2016, pp. 99–103.
- [Inoue et al. 2022] Seki Inoue, Shun Suzuki, and Hiroyuki Shinoda. « Multiunit phased array system for flexible workspace ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 241–260.
- [Iodice et al. 2018] Michele Iodice, William Frier, James Wilcox, Ben Long, and Orestis Georgiou. « Pulsed schlieren imaging of ultrasonic haptics and levitation using phased arrays ». In: *arXiv preprint arXiv:1810.00258* (2018).
- [Ito et al. 2016] Mitsuru Ito, Daisuke Wakuda, Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. « High spatial resolution midair tactile display using 70 kHz ultrasound ». In: *EuroHaptics*. Springer. 2016, pp. 57–67.

- 
- [Iwabuchi et al. 2023] Sota Iwabuchi, Ryoya Onishi, Shun Suzuki, Takaaki Kamigaki, Yasutoshi Makino, and Hiroyuki Shinoda. « Performance evaluation of airborne ultrasound focus measurement using thermal imaging on the surface of a finger ». In: *IEEE World Haptics Conference*. IEEE. 2023, pp. 210–215.
- [Iwai et al. 2019] Daisuke Iwai, Mei Aoki, and Kosuke Sato. « Non-contact thermo-visual augmentation by IR-RGB projection ». In: *IEEE Transactions on Visualization and Computer Graphics* **25**.4 (2019), pp. 1707–1716.
- [Iwamoto et al. 2008a] Takayuki Iwamoto, Mari Tatezono, Takayuki Hoshi, and Hiroyuki Shinoda. « Airborne ultrasound tactile display ». In: *ACM SIGGRAPH (New Tech Demos)*. 2008, pp. 1–1.
- [Iwamoto et al. 2008b] Takayuki Iwamoto, Mari Tatezono, and Hiroyuki Shinoda. « Non-contact method for producing tactile sensation using airborne ultrasound ». In: *EuroHaptics*. Springer. 2008, pp. 504–513.
- [Jang et al. 2020] Jaehyun Jang and Jinah Park. « SPH fluid tactile rendering for ultrasonic mid-air haptics ». In: *IEEE Transactions on Haptics* **13**.1 (2020), pp. 116–122.
- [Johansson et al. 1979] Roland S Johansson and Ake B Vallbo. « Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin ». In: *The Journal of physiology* **286**.1 (1979), pp. 283–300.
- [Johansson et al. 1982] Roland S Johansson, Ulf Landstro, Ronnie Lundstro, et al. « Responses of mechanoreceptive afferent units in the glabrous skin of the human hand to sinusoidal skin displacements ». In: *Brain research* **244**.1 (1982), pp. 17–25.
- [Johansson et al. 2009] Roland S Johansson and J Randall Flanagan. « Coding and use of tactile signals from the fingertips in object manipulation tasks ». In: *Nature Reviews Neuroscience* **10**.5 (2009), pp. 345–359.

- 
- [John et al. 2024] Kevin John, Yinan Li, and Hasti Seifi. « AdapTics: A Toolkit for Creative Design and Integration of Real-Time Adaptive Mid-Air Ultrasound Tactons ». In: *Conference on Human Factors in Computing Systems*. 2024, pp. 1–15.
- [Jones et al. 2002] Lynette A Jones and Michal Berris. « The psychophysics of temperature perception and thermal-interface design ». In: *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE. 2002, pp. 137–142.
- [Jones et al. 2006] Lynette A Jones and Susan J Lederman. « Neurophysiology of hand function ». In: *Human hand function*. 2006.
- [Jones et al. 2008] Lynette A Jones and Hsin-Ni Ho. « Warm or cool, large or small? The challenge of thermal displays ». In: *IEEE Transactions on Haptics* **1**.1 (2008), pp. 53–70.
- [Jones et al. 2013] Lynette A Jones and Hong Z Tan. « Application of psychophysical techniques to haptic research ». In: *IEEE transactions on haptics* **6**.3 (2013), pp. 268–284.
- [Jones 2018] Lynette A Jones. « Tactile and Haptic Displays ». In: *Haptics*. 2018.
- [Jun et al. 2015] Jae-Hoon Jun, Jong-Rak Park, Sung-Phil Kim, Young Min Bae, Jang-Yeon Park, Hyung-Sik Kim, Seungmoon Choi, Sung Jun Jung, Seung Hwa Park, Dong-Il Yeom, et al. « Laser-induced thermoelastic effects can evoke tactile sensations ». In: *Scientific Reports* **5**.1 (2015), p. 11016.
- [Kalra et al. 2016] Anubha Kalra, Andrew Lowe, and Ahmed M Al-Jumaily. « Mechanical behaviour of skin: a review ». In: *Journal of Material Science & Engineering* **5**.4 (2016), pp. 1–7.
- [Kamigaki et al. 2017] Takaaki Kamigaki, Akihito Noda, and Hiroyuki Shinoda. « Thin and flexible airborne ultrasound phased array for tactile display ». In: *Annual Conference of the IEEE Society of Instrument and Control Engineers of Japan*. IEEE. 2017, pp. 736–739.
- [Kandel et al. 2012a] Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, Steven A. Siegelbaum, and A. J. Hudspeth. *Principles of Neural Science*. 5th Edition. 2012.

- 
- [Kandel et al. 2012b] Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, Steven A. Siegelbaum, and A. J. Hudspeth. « The Somatosensory System: Receptors and Central Pathways ». In: *Principles of Neural Science*. 5th Edition. 2012.
- [Kappers et al. 2022] Astrid ML Kappers, Max Fa Si Oen, Tessa JW Junggeburth, and Myrthe A Plaisier. « Hand-held haptic navigation devices for actual walking ». In: *IEEE Transactions on Haptics* **15**.4 (2022), pp. 655–666.
- [Kappus et al. 2018] Brian Kappus and Ben Long. « Spatiotemporal modulation for mid-air haptic feedback from an ultrasonic phased array ». In: *The Journal of the Acoustical Society of America* **143**.3\_Supplement (2018), pp. 1836–1836.
- [Karafotias et al. 2018] Georgios Karafotias, Georgios Korres, Akiko Teranishi, Wan-joo Park, and Mohamad Eid. « Mid-air tactile stimulation for pain distraction ». In: *IEEE Transactions on Haptics* **11**.2 (2018), pp. 185–191.
- [Kervegant et al. 2017] Cédric Kervegant, Félix Raymond, Delphine Graeff, and Julien Castet. « Touch hologram in mid-air ». In: *ACM SIGGRAPH (Emerging Technologies)*. 2017, pp. 1–2.
- [Kilteni et al. 2012] Konstantina Kilteni, Raphaela Grotens, and Mel Slater. « The sense of embodiment in virtual reality ». In: *Presence: Teleoperators and Virtual Environments* **21**.4 (2012), pp. 373–387.
- [Komandur et al. 2009] Sashidharan Komandur, Peter W Johnson, Richard L Storch, and Michael G Yost. « Relation between index finger width and hand width anthropometric measures ». In: *International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2009, pp. 823–826.
- [Kon et al. 2024] Shoha Kon, Eifu Narita, Izumi Mizoguchi, and Hiroyuki Kamimoto. « Variable curvature and spherically arranged ultrasound transducers for depth-adjustable focused ultrasound ». In: *EuroHaptics*. Springer. 2024, pp. 278–284.

- 
- [Korres et al. 2016] Georgios Korres and Mohamad Eid. « Haptogram: Ultrasonic point-cloud tactile stimulation ». In: *IEEE Access* **4** (2016), pp. 7758–7769.
- [Korres et al. 2017] Georgios Korres, Tamas Aujeszky, and Mohamad Eid. « Characterizing tactile rendering parameters for ultrasound-based stimulation ». In: *IEEE World Haptics Conference*. IEEE. 2017, pp. 293–298.
- [Lacôte et al. 2024] Inès Lacôte, Pierre-Antoine Cabaret, Claudio Pacchierotti, Marie Babel, David Gueorguiev, and Maud Marchal. « Do Vibrotactile Patterns on both Hands Improve Guided Navigation with a Walker? » In: *EuroHaptics*. Springer. 2024, pp. 391–404.
- [Lan et al. 2024] Ruiheng Lan, Xu Sun, Qingfeng Wang, and Bingjian Liu. « Ultrasonic mid-air haptics on the face: effects of lateral modulation frequency and amplitude on users' responses ». In: *Conference on Human Factors in Computing Systems*. 2024, pp. 1–12.
- [Larrabee Jr et al. 1986] Wayne F Larrabee Jr and Dwight Sutton. « A finite element model of skin deformation. II. An experimental model of skin deformation ». In: *The Laryngoscope* **96**.4 (1986), pp. 406–412.
- [Lederman et al. 1987] Susan J Lederman and Roberta L Klatzky. « Hand movements: A window into haptic object recognition ». In: *Cognitive Psychology* **19**.3 (1987), pp. 342–368.
- [Lee et al. 2015] Hojin Lee, Ji-Sun Kim, Seungmoon Choi, Jae-Hoon Jun, Jong-Rak Park, A-Hee Kim, Han-Byeol Oh, Hyung-Sik Kim, and Soon-Cheol Chung. « Mid-air tactile stimulation using laser-induced thermoelastic effects: The first study for indirect radiation ». In: *IEEE World Haptics Conference*. IEEE. 2015, pp. 374–380.
- [Lim et al. 2024a] Chungman Lim, Gunhyuk Park, and Hasti Seifi. « Designing distinguishable mid-air ultrasound tactons with temporal parameters ». In: *Conference on Human Factors in Computing Systems*. 2024, pp. 1–18.

- 
- [Lim et al. 2024b] Chungman Lim, Hasti Seifi, and Gunhyuk Park. « An Interactive Tool for Simulating Mid-Air Ultrasound Tactons on the Skin ». In: *Conference on Human Factors in Computing Systems (Extended Abstract)*. 2024, pp. 1–7.
- [Lin et al. 1998] Ming Lin and Stefan Gottschalk. « Collision detection between geometric models: A survey ». In: *IMA Conference on Mathematics of Surfaces*. Vol. 1. 1998, pp. 602–608.
- [Lin et al. 2008] Ming C Lin and Miguel Otaduy. *Haptic rendering: foundations, algorithms, and applications*. 2008.
- [Lobo et al. 2017] Daniel Lobo, Mine Saraç, Mickeal Verschoor, Massimiliano Solazzi, Antonio Frisoli, and Miguel A Otaduy. « Proxy-based haptic rendering for underactuated haptic devices ». In: *IEEE World Haptics Conference*. IEEE. 2017, pp. 48–53.
- [Löken et al. 2009] Line S Löken, Johan Wessberg, India Morrison, Francis McGlone, and Håkan Olausson. « Coding of pleasant touch by unmyelinated afferents in humans ». In: *Nature Neuroscience* **12**.5 (2009), pp. 547–548.
- [Long et al. 2014] Benjamin Long, Sue Ann Seah, Tom Carter, and Sriram Subramanian. « Rendering volumetric haptic shapes in mid-air using ultrasound ». In: *ACM Transactions on Graphics* **33**.6 (2014), pp. 1–10.
- [Longo et al. 2012] Matthew R Longo, Jason Jiri Musil, and Patrick Haggard. « Visuo-tactile integration in personal space ». In: *Journal of Cognitive Neuroscience* **24**.3 (2012), pp. 543–552.
- [Manfredi et al. 2012] Louise R Manfredi, Andrew T Baker, Damian O Elias, John F Dammann III, Mark C Zielinski, Vicky S Polashock, and Sliman J Bensmaia. « The effect of surface wave propagation on neural responses to vibration in primate glabrous skin ». In: *PloS One* **7**.2 (2012), pp. 1–10.
- [Marchal et al. 2020] Maud Marchal, Gerard Gallagher, Anatole Lécuyer, and Claudio Pacchierotti. « Can stiffness sensations be rendered in virtual reality using mid-air ultrasound haptic technologies? » In: *EuroHaptics*. Springer. 2020, pp. 297–306.

- 
- [Marchal et al. 2025] Maud Marchal and Claudio Pacchierotti. « Virtual reality and haptics ». In: *Robotics Goes MOOC: Interaction*. 2025, pp. 111–158.
- [Martinez et al. 2018] Jonatan Martinez, Daniel Griffiths, Valerio Biscione, Orestis Georgiou, and Tom Carter. « Touchless haptic feedback for supernatural VR experiences ». In: *IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE. 2018, pp. 629–630.
- [Martinez et al. 2019] Jonatan Martinez, Adam Harwood, Hannah Limerick, Rory Clark, and Orestis Georgiou. « Mid-air haptic algorithms for rendering 3d shapes ». In: *IEEE International Symposium on Haptic, Audio and Visual Environments and Games*. IEEE. 2019, pp. 1–6.
- [Marzo et al. 2018] Asier Marzo, Tom Corkett, and Bruce W Drinkwater. « Ultraino: An open phased-array system for narrowband airborne ultrasound transmission ». In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **65**.1(2018), pp. 102–111.
- [Marzo 2022] Asier Marzo. « Prototyping airborne ultrasonic arrays ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 335–346.
- [Matsubayashi et al. 2019] Atsushi Matsubayashi, Yasutoshi Makino, and Hiroyuki Shinoda. « Direct finger manipulation of 3D object image with ultrasound haptic feedback ». In: *Conference on Human Factors in Computing Systems*. 2019, pp. 1–11.
- [Matsubayashi et al. 2020] Atsushi Matsubayashi, Yasutoshi Makino, and Hiroyuki Shinoda. « Rendering ultrasound pressure distribution on hand surface in real-time ». In: *EuroHaptics*. Springer. 2020, pp. 407–415.
- [Matsubayashi et al. 2022] Atsushi Matsubayashi, Seki Inoue, Shun Suzuki, and Hiroyuki Shinoda. « Sound-field creation for haptic reproduction ». In: *Ultrasound Mid-Air Haptics for Touchless Interfaces*. 2022, pp. 261–279.

- 
- [Matthews et al. 2019] Brandon J Matthews, Bruce H Thomas, Stewart Von Itzstein, and Ross T Smith. « Remapped physical-virtual interfaces with bimanual haptic retargeting ». In: *IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE. 2019, pp. 19–27.
- [Maunsbach et al. 2024] Martin Maunsbach, William Frier, and Kasper Hornbæk. « MAM-MOTH: Mid-Air Mesh-based Modulation Optimization Toolkit for Haptics ». In: *Conference on Human Factors in Computing Systems (Extended Abstract)*. 2024, pp. 1–7.
- [McIntyre et al. 2019] Sarah McIntyre, Athanasia Moungou, Rebecca Boehme, Peder M Isager, Frances Lau, Ali Israr, Ellen A Lumpkin, Freddy Abnousi, and Haakan Olausson. « Affective touch communication in close adult relationships ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 175–180.
- [McIntyre et al. 2022] Sarah McIntyre, Steven C Hauser, Anikó Kusztor, Rebecca Boehme, Athanasia Moungou, Peder Mortvedt Isager, Lina Homman, Giovanni Novembre, Saad S Nagi, Ali Israr, et al. « The language of social touch is intuitive and quantifiable ». In: *Psychological Science* 33.9 (2022), pp. 1477–1494.
- [Meijer et al. 1999] R Meijer, LFA Douven, and CWJ Oomens. « Characterisation of anisotropic and non-linear behaviour of human skin in vivo ». In: *Computer Methods in Biomechanics and Biomedical Engineering* 2.1 (1999), pp. 13–27.
- [Minetto et al. 2017] Rodrigo Minetto, Neri Volpato, Jorge Stolfi, Rodrigo MMH Gregori, and Murilo VG Da Silva. « An optimal algorithm for 3D triangle mesh slicing ». In: *Computer-Aided Design* 92 (2017), pp. 1–10.
- [Montano et al. 2024] Roberto Montano, Rafael Morales, Dario Pittéra, William Frier, Orestis Georgiou, and Patricia Cornelio. « Knuckles notifications: mid-air haptic feedback on the dorsal hand for hands-on-the-wheel driving ». In: *Frontiers in Computer Science* 6 (2024), pp. 1–15.

- 
- [Moore et al. 2025] Warren Moore, Johan Nikesjö, Otmane Bouchatta, Adarsh D Makdani, Pierre Hakizimana, Mikael Rousson, Basil Duvernay, Sarah McIntyre, Laura J Pehkonen, Anders Fridberger, et al. « Robust coupling between the C-tactile afferent and the hair follicle in humans ». In: *The Journal of Physiology* **603**.16 (2025), pp. 4593–4608.
- [Moore 2007] Thomas J Moore. « A survey of the mechanical characteristics of skin and tissue in response to vibratory stimulation ». In: *IEEE Transactions on Man-Machine Systems* **11**.1 (2007), pp. 79–84.
- [Morgan et al. 2025] Zak Morgan, Youngjun Cho, and Sriram Subramanian. « Rapid in-air ultrasound holography measurement and camera-in-the-loop generation using thermography ». In: *Communications Engineering* **4**.1 (2025), p. 101.
- [Morisaki et al. 2022] Tao Morisaki, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « Non-vibratory pressure sensation produced by ultrasound focus moving laterally and repetitively with fine spatial step width ». In: *IEEE Transactions on Haptics* **15**.2 (2022), pp. 441–450.
- [Morisaki et al. 2024] Tao Morisaki, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « Noncontact haptic rendering of static contact with convex surface using circular movement of ultrasound focus on a finger pad ». In: *IEEE Transactions on Haptics* **17**.3 (2024), pp. 334–345.
- [Morzyński et al. 2024] Leszek Morzyński, Marlena Podleśna, Grzegorz Szczepański, and Anna Włudarczyk. « Ultrasonic Haptic Devices: Ultrasonic Noise Assessment ». In: *Archives of Acoustics* **49**.3 (2024), pp. 447–458.
- [Mulot et al. 2021a] Lendy Mulot, Guillaume Gicquel, William Frier, Maud Marchal, Claudio Pacchierotti, and Thomas Howard. « Curvature discrimination for dynamic ultrasound mid-air haptic stimuli ». In: *IEEE World Haptics Conference (Work-in-Progress)*. IEEE. 2021, pp. 1145–1145.

- 
- [Mulot et al. 2021b] Lendy Mulot, Guillaume Gicquel, Quentin Zanini, William Frier, Maud Marchal, Claudio Pacchierotti, and Thomas Howard. « Dolphin: A framework for the design and perceptual evaluation of ultrasound mid-air haptic stimuli ». In: *ACM Symposium on Applied Perception*. 2021, pp. 1–10.
- [Mulot et al. 2023a] Lendy Mulot, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. « Can we increase the perceived intensity of mid-air haptic shapes rendered with dynamic tactile pointers? » In: *IEEE World Haptics conference (Work-in-Progress)*. 2023.
- [Mulot et al. 2023b] Lendy Mulot, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. « Improving the perception of mid-air tactile shapes with spatio-temporally-modulated tactile pointers ». In: *ACM Transactions on Applied Perception* **20**.4 (2023), pp. 1–16.
- [Mulot et al. 2023c] Lendy Mulot, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. « Ultrasound mid-air haptics for hand guidance in virtual reality ». In: *IEEE Transactions on Haptics* **16**.4 (2023), pp. 497–503.
- [Mulot et al. 2024] Lendy Mulot, Thomas Howard, Sarah Emery, Claudio Pacchierotti, and Maud Marchal. « Designing 3D object rendering techniques for ultrasound mid-air haptics using intersection strategies ». In: *ACM Symposium on Applied Perception*. 2024, pp. 1–8.
- [Mulot et al. 2025] Lendy Mulot, Thomas Howard, Guillaume Gicquel, Claudio Pacchierotti, and Maud Marchal. « Bimanual Ultrasound mid-air haptics for virtual reality manipulation ». In: *IEEE Transactions on Visualization and Computer Graphics* **31**.9 (2025), pp. 4821–4832.
- [Nai et al. 2021] Weizhi Nai, Jianyu Liu, Chongyang Sun, Qinglong Wang, Guohong Liu, and Xiaoying Sun. « Vibrotactile feedback rendering of patterned textures using a waveform segment table method ». In: *IEEE Transactions on Haptics* **14**.4 (2021), pp. 849–861.

- 
- [Nickell et al. 2000] Stephan Nickell, Marcus Hermann, Matthias Essenpreis, Thomas J Farrell, Uwe Krämer, and Michael S Patterson. « Anisotropy of light propagation in human skin ». In: *Physics in Medicine & Biology* **45**.10 (2000), p. 2873.
- [Normand et al. 2024a] Erwan Normand, Claudio Pacchierotti, Eric Marchand, and Maud Marchal. « How Different Is the Perception of Vibrotactile Texture Roughness in Augmented versus Virtual Reality? » In: *ACM Symposium on Virtual Reality Software and Technology*. 2024, pp. 1–10.
- [Normand et al. 2024b] Erwan Normand, Claudio Pacchierotti, Eric Marchand, and Maud Marchal. « Visuo-haptic rendering of the hand during 3d manipulation in augmented reality ». In: *IEEE Transactions on Haptics* **17**.2 (2024), pp. 277–291.
- [Obrist et al. 2013] Marianna Obrist, Sue Ann Seah, and Sriram Subramanian. « Talking about tactile experiences ». In: *Conference on Human Factors in Computing Systems*. 2013, pp. 1659–1668.
- [Ochiai et al. 2016] Yoichi Ochiai, Kota Kumagai, Takayuki Hoshi, Jun Rekimoto, Satoshi Hasegawa, and Yoshio Hayasaki. « Fairy lights in femtoseconds: Aerial and volumetric graphics rendered by focused femtosecond laser combined with computational holographic fields ». In: *ACM Transactions on Graphics* **35**.2 (2016), pp. 1–14.
- [Onishi et al. 2025] Ryoya Onishi, Sota Iwabuchi, Shun Suzuki, Takaaki Kamigaki, Yasutoshi Makino, and Hiroyuki Shinoda. « Measurement of airborne ultrasound focus on skin surface using thermal imaging ». In: *IEEE Transactions on Haptics* (2025).
- [Otaduy et al. 2013] Miguel A Otaduy, Carlos Garre, and Ming C Lin. « Representations and algorithms for force-feedback display ». In: *Proceedings of the IEEE* **101**.9 (2013), pp. 2068–2080.
- [Otaduy et al. 2016] Miguel A Otaduy, Allison Okamura, and Sriram Subramanian. « Haptic technologies for direct touch in virtual reality ». In: *ACM SIGGRAPH (Courses)*. 2016, pp. 1–123.

- 
- [Pacchierotti et al. 2016] Claudio Pacchierotti, Gionata Salvietti, Irfan Hussain, Leonardo Meli, and Domenico Prattichizzo. « The hRing: A wearable haptic device to avoid occlusions in hand tracking ». In: *IEEE Haptics Symposium*. IEEE. 2016, pp. 134–139.
- [Park et al. 2025] Changhyeon Park, Yubin Lee, and Sang Ho Yoon. « Ultra-Board: Always-available Wearable Ultrasonic Mid-air Haptic Interface for Responsive and Robust VR Inputs ». In: vol. 9. 2. 2025, pp. 1–31.
- [Patchett 1979] Robin F Patchett. « Human sound frequency preferences ». In: *Perceptual and motor skills* **49**.1 (1979), pp. 324–326.
- [Peebles et al. 1998] Laura Peebles and Beverley Norris. *Adultdata: The Handbook of Adult Anthropometric and Strength Measurements : Data for Design Safety*. 1998.
- [Peirce et al. 2019] Jonathan Peirce, Jeremy R Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. « PsychoPy2: Experiments in behavior made easy ». In: *Behavior Research Methods* **51**.1 (2019), pp. 195–203.
- [Perquin et al. 2021] Marlou N Perquin, Mason Taylor, Jarred Lorusso, and James Kolasinski. « Directional biases in whole hand motion perception revealed by mid-air tactile stimulation ». In: *cortex* **142** (2021), pp. 221–236.
- [Pezent et al. 2021] Evan Pezent, Brandon Cambio, and Marcia K O’Malley. « Syn-tacts: Open-source software and hardware for audio-controlled haptics ». In: *IEEE Transactions on Haptics* **14**.1 (2021), pp. 225–233.
- [Pittera et al. 2022] Dario Pittera, Orestis Georgiou, Abdenaceur Abdouni, and William Frier. « “I can feel it coming in the hairs tonight”: Characterising mid-air haptics on the hairy parts of the skin ». In: *IEEE Transactions on Haptics* **15**.1 (2022), pp. 188–199.
- [Pittera et al. 2023] Dario Pittera, Orestis Georgiou, and William Frier. « “I See Where This is Going”: A Psychophysical Study of Directional

- 
- Mid-Air Haptics and Apparent Tactile Motion ». In: *IEEE Transactions on Haptics* **16**.2 (2023), pp. 322–333.
- [Plasencia et al. 2020] Diego Martinez Plasencia, Ryuji Hirayama, Roberto Montano-Murillo, and Sriram Subramanian. « GS-PAT: high-speed multi-point sound-fields for phased arrays of transducers ». In: *ACM Transactions on Graphics* **39**.4 (2020), pp. 138–1.
- [Price et al. 2018] Adam Price and Benjamin Long. « Fibonacci spiral arranged ultrasound phased array for mid-air haptics ». In: *IEEE International Ultrasonics Symposium*. IEEE. 2018, pp. 1–4.
- [Punpongsanon et al. 2015] Parinya Punpongsanon, Daisuke Iwai, and Kosuke Sato. « SoftAR: Visually manipulating haptic softness perception in spatial augmented reality ». In: *IEEE Transactions on Visualization and Computer Graphics* **21**.11 (2015), pp. 1279–1288.
- [Quek et al. 2014] Zhan Fan Quek, Samuel B Schorr, Ilana Nisky, William R Provancher, and Allison M Okamura. « Sensory substitution using 3-degree-of-freedom tangential and normal skin deformation feedback ». In: *IEEE Haptics Symposium*. IEEE. 2014, pp. 27–33.
- [Rahimi et al. 2019] Mehdi Rahimi, Fang Jiang, and Yantao Shen. « Non-linearity of skin properties in electrotactile applications: Identification and mitigation ». In: *IEEE Access* **7** (2019), pp. 169844–169852.
- [Rakkolainen et al. 2021] Ismo Rakkolainen, Euan Freeman, Antti Sand, Roope Raisamo, and Stephen Brewster. « A survey of mid-air ultrasound haptics and its applications ». In: *IEEE Transactions on Haptics* **14**.1 (2021), pp. 2–19.
- [Reardon et al. 2019] Gregory Reardon, Yitian Shao, Bharat Dandu, William Frier, Ben Long, Orestis Georgiou, and Yon Visell. « Cutaneous wave propagation shapes tactile motion: Evidence from air-coupled ultrasound ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 628–633.
- [Reardon et al. 2023] Gregory Reardon, Bharat Dandu, Yitian Shao, and Yon Visell. « Shear shock waves mediate haptic holography via focused ultrasound ». In: *Science Advances* **9**.9 (2023), eadf2037.

- 
- [Romanus et al. 2019] Ted Romanus, Sam Frish, Mykola Maksymenko, William Frier, Loïc Corenthy, and Orestis Georgiou. « Mid-air haptic bio-holograms in mixed reality ». In: *IEEE International Symposium on Mixed and Augmented Reality Adjunct*. IEEE. 2019, pp. 348–352.
- [Rothberg et al. 2017] Steve J Rothberg, MS Allen, Paolo Castellini, Dario Di Maio, JJ Dirckx, DJ Ewins, Ben J Halkon, P Muyshondt, Nicola Paone, T Ryan, et al. « An international review of laser Doppler vibrometry: Making light work of vibration measurement ». In: *Optics and Lasers in Engineering* **99** (2017), pp. 11–22.
- [Rutten et al. 2019] Isa Rutten, William Frier, Lawrence Van den Bogaert, and David Geerts. « Invisible touch: How identifiable are mid-air haptic shapes? » In: *Conference on Human Factors in Computing Systems (Extended Abstract)*. 2019, pp. 1–6.
- [Rutten et al. 2020] Isa Rutten, William Frier, and David Geerts. « Discriminating between intensities and velocities of mid-air haptic patterns ». In: *EuroHaptics*. Springer. 2020, pp. 78–86.
- [Saga 2015] Satoshi Saga. « HeatHapt thermal radiation-based haptic display ». In: *Haptic Interaction: Perception, Devices and Applications*. 2015, pp. 105–107.
- [Sand et al. 2015] Antti Sand, Ismo Rakkolainen, Poika Isokoski, Jari Kangas, Roope Raisamo, and Karri Palovuori. « Head-mounted display with mid-air tactile feedback ». In: *ACM symposium on virtual reality software and technology*. 2015, pp. 51–58.
- [Sarac et al. 2019] Mine Sarac, Massimiliano Solazzi, and Antonio Frisoli. « Design requirements of generic hand exoskeletons and survey of hand exoskeletons for rehabilitation, assistive, or haptic use ». In: *IEEE Transactions on Haptics* **12**.4 (2019), pp. 400–413.
- [Schott 1993] Geoffrey D Schott. « Penfield’s homunculus: a note on cerebral cartography ». In: *Journal of Neurology, Neurosurgery, and Psychiatry* **56**.4 (1993), p. 329.

- 
- [Seifi et al. 2023] Hasti Seifi, Sean Chew, Antony James Nascè, William Edward Lowther, William Frier, and Kasper Hornbæk. « Feel-illustrator: A design tool for ultrasound mid-air haptics ». In: *Conference on Human Factors in Computing Systems*. 2023, pp. 1–16.
- [Settnes et al. 2012] Mikkel Settnes and Henrik Bruus. « Forces acting on a small particle in an acoustical field in a viscous fluid ». In: *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* **85**.1 (2012), p. 016327.
- [Shakeri et al. 2018] Gözel Shakeri, John H Williamson, and Stephen Brewster. « May the force be with you: Ultrasound haptic feedback for mid-air gesture interaction in cars ». In: *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2018, pp. 1–10.
- [Shen et al. 2022] Vivian Shen, Craig Shultz, and Chris Harrison. « Mouth haptics in VR using a headset ultrasound phased array ». In: *Conference on Human Factors in Computing Systems*. 2022, pp. 1–14.
- [Shen et al. 2024a] Vivian Shen, Chris Harrison, and Craig Shultz. « Expressive, scalable, mid-air haptics with synthetic jets ». In: *ACM Transactions on Computer-Human Interaction* **31**.2 (2024), pp. 1–28.
- [Shen et al. 2024b] Zhouyang Shen, Zak Morgan, Madhan Kumar Vasudevan, Marianna Obrist, and Diego Martinez Plasencia. « Controlled-STM: a two-stage model to predict user's perceived intensity for multi-point spatiotemporal modulation in ultrasonic mid-air haptics ». In: *Conference on Human Factors in Computing Systems*. 2024, pp. 1–12.
- [Silva et al. 2009] Alejandro Jarillo Silva, Omar A Domínguez Ramirez, Vicente Parra Vega, and Jesus P Ordaz Oliver. « Phantom omni haptic device: Kinematic and manipulability ». In: *Electronics, Robotics and Automotive Mechanics Conference*. IEEE. 2009, pp. 193–198.
- [Sneddon 2018] Lynne U Sneddon. « Comparative physiology of nociception and pain ». In: *Physiology* **33**.1 (2018), pp. 63–73.

- 
- [Sodhi et al. 2013] Rajinder Sodhi, Ivan Poupyrev, Matthew Glisson, and Ali Israr. « AIREAL: interactive tactile experiences in free air ». In: *ACM Transactions on Graphics* **32**.4 (2013), pp. 1–10.
- [Sofia et al. 2013] Katherine O Sofia and Lynette Jones. « Mechanical and psychophysical studies of surface wave propagation during vibrotactile stimulation ». In: *IEEE Transactions on Haptics* **6**.3 (2013), pp. 320–329.
- [Spelmezan et al. 2017] Daniel Spelmezan, Deepak Ranjan Sahoo, and Sriram Subramanian. « Sparkle: Hover feedback with touchable electric arcs ». In: *Conference on Human Factors in Computing Systems*. 2017, pp. 3705–3717.
- [Sreng et al. 2007] Jean Sreng, Florian Bergez, Jeremie Legarrec, Anatole Lécuyer, and Claude Andriot. « Using an event-based approach to improve the multimodal rendering of 6DOF virtual contact ». In: *ACM symposium on Virtual reality software and technology*. 2007, pp. 165–173.
- [Suzuki et al. 2004] Yoiti Suzuki and Hisashi Takeshima. « Equal-loudness-level contours for pure tones ». In: *The Journal of the Acoustical Society of America* **116**.2 (2004), pp. 918–933.
- [Suzuki et al. 2005] Yuriko Suzuki and Minoru Kobayashi. « Air jet driven force feedback in virtual reality ». In: *IEEE Computer Graphics and Applications* **25**.1 (2005), pp. 44–47.
- [Suzuki et al. 2019] Shun Suzuki, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « Midair hand guidance by an ultrasound virtual handrail ». In: *IEEE World Haptics Conference*. 2019, pp. 271–276.
- [Suzuki et al. 2021] Shun Suzuki, Seki Inoue, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « AUTD3: Scalable airborne ultrasound tactile display ». In: *IEEE Transactions on Haptics* **14**.4 (2021), pp. 740–749.

- 
- [Takahashi et al. 2018] Ryoko Takahashi, Keisuke Hasegawa, and Hiroyuki Shinoda. « Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli ». In: *EuroHaptics*. Springer. 2018, pp. 276–288.
- [Takahashi et al. 2020] Ryoko Takahashi, Keisuke Hasegawa, and Hiroyuki Shinoda. « Tactile stimulation by repetitive lateral movement of midair ultrasound focus ». In: *IEEE Transactions on Haptics* **13**.2 (2020), pp. 334–342.
- [Talvas et al. 2014] Anthony Talvas, Maud Marchal, and Anatole Lécuyer. « A survey on bimanual haptic interaction ». In: *IEEE Transactions on Haptics* **7**.3 (2014), pp. 285–300.
- [Tinguy et al. 2020] Xavier de Tinguy, Thomas Howard, Claudio Pacchierotti, Maud Marchal, and Anatole Lécuyer. « Weatavix: wearable actuated tangibles for virtual reality experiences ». In: *EuroHaptics*. Springer. 2020, pp. 262–270.
- [Tinguy et al. 2021] Xavier de Tinguy, Claudio Pacchierotti, Anatole Lécuyer, and Maud Marchal. « Capacitive sensing for improving contact rendering with tangible objects in VR ». In: *IEEE Transactions on Visualization and Computer Graphics* **27**.4 (2021), pp. 2481–2487.
- [Treeby et al. 2010] Bradley E Treeby and Benjamin T Cox. « k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields ». In: *Journal of Biomedical Optics* **15**.2 (2010), pp. 021314–021314.
- [Treeby et al. 2014] Bradley E Treeby, Jiri Jaros, Daniel Rohrbach, and BT Cox. « Modelling elastic wave propagation using the k-wave matlab toolbox ». In: *IEEE International Ultrasonics Symposium*. IEEE. 2014, pp. 146–149.
- [Tsumoto et al. 2021] Kai Tsumoto, Tao Morisaki, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. « Presentation of tactile pleasantness using airborne ultrasound ». In: *IEEE World Haptics Conference*. IEEE. 2021, pp. 602–606.

- 
- [Tuthill et al. 2018] John C Tuthill and Eiman Azim. « Proprioception ». In: *Current Biology* **28**.5 (2018), R194–R203.
- [Vallbo et al. 1978] Ake B Vallbo and Roland S Johansson. *Tactile sensory innervation of the glabrous skin of the human hand*. 1978.
- [Vallbo et al. 1984] Ake B Vallbo, Roland S Johansson, et al. « Properties of cutaneous mechanoreceptors in the human hand related to touch sensation ». In: *Human Neurobiology* **3**.1 (1984), pp. 3–14.
- [Vallbo 2018] Åke Bernhard Vallbo. « Microneurography: how it started and how it works ». In: *Journal of Neurophysiology* **120**.3 (2018), pp. 1415–1427.
- [Van Neer et al. 2019] Paul Van Neer, Arno Volker, Arthur Berkhoff, Thijs Schrama, Hylke Akkerman, Albert Van Breemen, Laurens Peeters, Jan-Laurens van der Steen, and Gerwin Gelinck. « Development of a flexible large-area array based on printed polymer transducers for mid-air haptic feedback ». In: *Proceedings of Meetings on Acoustics*. Vol. 38. 1. Acoustical Society of America. 2019, p. 045008.
- [Varga et al. 2015] Ivan Varga and Boris Mravec. « Nerve fiber types ». In: *Nerves and Nerve Injuries*. 2015, pp. 107–113.
- [Verrillo 1966] Ronald T Verrillo. « Vibrotactile sensitivity and the frequency response of the Pacinian corpuscle ». In: *Psychonomic Science* **4**.1 (1966), pp. 135–136.
- [Villa et al. 2022] Steeven Villa, Sven Mayer, Jess Hartcher-O'Brien, Albrecht Schmidt, and Tonja-Katrin Machulla. « Extended mid-air ultrasound haptics for virtual reality ». In: *Proceedings of the ACM on Human-Computer Interaction* **6**.ISS (2022), pp. 500–524.
- [Vizcay et al. 2021] Sebastian Vizcay, Panagiotis Kourtesis, Ferran Argelaguet, Claudio Pacchierotti, and Maud Marchal. « Electrotactile feedback in virtual reality for precise and accurate contact rendering ». In: *TechRxiv* (2021).

- 
- [Watkins et al. 2021] Roger Holmes Watkins, Mariama Dione, Rochelle Ackerley, Helena Backlund Wasling, Johan Wessberg, and Line S Löken. « Evidence for sparse C-tactile afferent innervation of glabrous human hand skin ». In: *Journal of Neurophysiology* **125**.1 (2021), pp. 232–237.
- [Weiss et al. 2011] Malte Weiss, Chat Wacharamanotham, Simon Voelker, and Jan Borchers. « FingerFlux: Near-surface haptic feedback on tabletops ». In: *ACM Symposium on User Interface Software and Technology*. 2011, pp. 615–620.
- [Wiertlewski et al. 2012] Michael Wiertlewski and Vincent Hayward. « Mechanical behavior of the fingertip in the range of frequencies and displacements relevant to touch ». In: *Journal of biomechanics* **45**.11 (2012), pp. 1869–1874.
- [Wilson et al. 2014] Graham Wilson, Thomas Carter, Sriram Subramanian, and Stephen A Brewster. « Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion ». In: *Conference on Human Factors in Computing Systems*. 2014, pp. 1133–1142.
- [Wojna et al. 2023a] Katarzyna Wojna, Orestis Georgiou, David Beattie, William Frier, Michael Wright, and Christof Lutteroth. « An exploration of just noticeable differences in mid-air haptics ». In: *IEEE World Haptics Conference*. IEEE. 2023, pp. 410–416.
- [Wojna et al. 2023b] Katarzyna Wojna, Orestis Georgiou, David Beattie, Michael Wright, and Christof Lutteroth. « Does it par-tickle: Investigating the relationship between mid-air haptics and visual representations of surface textures ». In: *IEEE Transactions on Haptics* **16**.4 (2023), pp. 561–566.
- [Wu et al. 2004] John Z Wu, Ren G Dong, S Rakheja, AW Schopper, and WP Smutz. « A structural fingertip model for simulating of the biomechanics of tactile sensation ». In: *Medical Engineering & Physics* **26**.2 (2004), pp. 165–175.
- [Xia 2018] Pingjun Xia. « New advances for haptic rendering: state of the art ». In: *The Visual Computer* **34**.2 (2018), pp. 271–287.

- 
- [Xu et al. 2019] Jiayi Xu, Yoshihiro Kuroda, Shunsuke Yoshimoto, and Osamu Oshiro. « Non-contact cold thermal display by controlling low-temperature air flow generated with vortex tube ». In: *IEEE World Haptics Conference*. IEEE. 2019, pp. 133–138.
- [Xu et al. 2024] Shan Xu, Steven C Hauser, Saad S Nagi, James A Jablonski, Merat Rezaei, Ewa Jarocka, Andrew G Marshall, Håkan Olausson, Sarah McIntyre, and Gregory J Gerling. « Mechanoreceptive  $A\beta$  primary afferents discriminate naturalistic social touch inputs at a functionally relevant time scale ». In: *IEEE Transactions on Affective Computing* (2024).
- [Young et al. 2020] Eric M Young, David Gueorguiev, Katherine J Kuchenbecker, and Claudio Pacchierotti. « Compensating for fingertip size to render tactile cues more accurately ». In: *IEEE transactions on haptics* **13**.1 (2020), pp. 144–151.
- [Zhang et al. 2012] Dake Zhang, Yi Ding, Joanna Stegall, and Lei Mo. « The effect of visual–chunking–representation accommodation on geometry testing for students with math disabilities ». In: *Learning Disabilities Research & Practice* **27**.4 (2012), pp. 167–177.
- [Zhu et al. 2025] Qirong Zhu, Xiwen Yao, Ruoting Sun, Kotaro Hara, and Yang Jiao. « KnuckleGuide: mid-air haptic guidance system targeting dorsal hand using airborne ultrasound ». In: *IEEE World Haptics Conference*. IEEE. 2025, pp. 214–227.
- [Zilles et al. 1995] Craig B Zilles and J Kenneth Salisbury. « A constraint-based god-object method for haptic display ». In: *IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*. Vol. 3. IEEE. 1995, pp. 146–151.

## Annex A

### Résumé Long en Français

L'haptique sans contact par ultrasons (HSCU) ont pour but d'améliorer les interactions via des gestes naturels, en générant des sensations tactiles à distance via une focalisation d'ondes ultrasoniques. Cette thèse fait progresser le domaine en étudiant la perception et le rendu de stimuli UMH, ainsi que leur utilisation conjointe avec les interactions en réalité virtuelle. Ci-dessous, nous abordons plus en détail le contexte de cette thèse, ainsi que les défis scientifiques sous-jacents liés à l'HSCU. Enfin, nous présentons les contributions associées que nous proposons pour relever ces défis.

## Contexte

Le sens du toucher est un élément fondamental du système somatosensoriel humain, car il contribue à notre interaction avec l'environnement dans des contextes à la fois physiques et émotionnels. Le système somatosensoriel englobe divers sous-sens, dont :

- La proprioception [Tuthill et al. 2018] désigne la perception qu'une personne a de la position, des mouvements et des efforts de son propre corps ;
- La thermoception [Jones et al. 2002] correspond à la perception de la chaleur et des changements de température ;
- La nociception [Sneddon 2018] est responsable de la perception des stimuli douloureux ;

- 
- La pruriception [Andersen et al. 2017] utilise des mécanismes similaires pour fournir des sensations de démangeaison ;
  - Les sensations **tactiles** [Jones 2018] désignent la **perception** des stimuli mécaniques sur la peau (e.g. vibration, texture).

Parmi ceux-ci, cette thèse se concentre sur les sensations tactiles, qui constituent actuellement le sous-sens le plus étudié, offrant un large éventail de sensations extrêmement riches en réponse à une stimulation mécanique de la peau. Bien que souvent considérées comme acquises, les sensations tactiles jouent un rôle crucial dans la façon dont nous percevons notre environnement, et communiquons socialement. Par exemple, les êtres humains s'appuient généralement sur un ensemble spécifique de mouvements de la main pour sonder et explorer notre environnement à l'aide du toucher (par exemple, en appliquant une pression pour évaluer la dureté d'un objet, ou en suivant son contour pour identifier sa forme) [Lederman et al. 1987]. La stimulation tactile a également été utilisée pour compenser d'autres déficiences sensorielles (e.g. déficiences visuelles) [Lacôte et al. 2024] en allégeant la charge sur les canaux sensoriels déficients, tout en tirant avantage du canal tactile sous-utilisé. Les sensations tactiles constituent également un élément important de la communication sociale et affective, qu'il s'agisse d'une tape amicale sur l'épaule, ou de caresses affectives [Hauser et al. 2019; McIntyre et al. 2019; McIntyre et al. 2022].

Pour concevoir des sensations efficaces, il est essentiel de comprendre comment les stimuli tactiles sont perçus. La perception tactile, et la perception en général, est un processus très complexe [Goldstein et al. 2022a] : lorsqu'un stimulus, appelé **stimulus distal**, est présent ou généré dans l'environnement, il se propage à travers un milieu jusqu'à atteindre les récepteurs sensoriels, où il est alors considéré comme un **stimulus proximal**. Pour la stimulation tactile, cela correspond à la déformation de la peau vers les **mécanorécepteurs**. Ensuite, les **neurones afférents** associés peuvent générer un signal électrique en réponse au stimulus, au cours d'un processus appelé **mécanotransduction**. Les signaux individuels provenant de nombreux neurones afférents se propagent ensuite vers le cerveau, où ils sont combinés dans un processus appelé **traitement neuronal**. Enfin, les signaux traités sont transformés en réponses comportementales, comprenant une étape de prise de conscience du stimulus, qui est ensuite reconnu, compris, et interprété par le destinataire en fonction de ses connaissances et de ses expériences antérieures. Enfin, cette sensation reconnue peut amener la personne à effectuer des actions motrices (e.g., bouger sa main pour explorer une autre partie de l'environnement), ce qui lui permet de découvrir et de ressentir de nouveaux stimuli.

Pour étudier efficacement le sens du toucher, il est très utile de concevoir des technologies capables de reproduire artificiellement des sensations tactiles contrôlées et cohérentes. Ces dispositifs, appelés **interfaces haptiques**, ont été conçus pour fournir des types de stimuli spécifiques (e.g., des vibrations, une pression) [Jones 2018] à travers un processus nommé **rendu haptique**. Non seulement ces interfaces permettent-elles une évaluation cohérente des stimuli tactiles et fournissent-elles des sensations tactiles artificielles aux utilisateurs, mais elles peuvent également souvent servir d'interface utilisateur lors d'interactions 3D. Ces interactions suivent généralement un fonctionnement proche de la boucle d'interaction, illustrée en **Figure A.1**. Les actions des utilisateurs sont détectées par des périphériques d'entrée (e.g. technologie de suivi de main, ou capteurs de force), puis traduites dans un environnement physique ou virtuel selon la technique d'interaction mise en œuvre. Le retour d'information est ensuite calculé en réponse au nouvel état de l'environnement, puis envoyé aux périphériques de sortie (e.g., écran, moteur vibrant). L'information est alors rendue par les effecteurs, générant un stimulus qui peut être perçu par l'utilisateur. Ce pipeline gère la plupart des types d'interactions 3D, de la visualisation d'objets 3D. Cependant, la plupart de ces interactions pourraient bénéficier de l'ajout d'un retour tactile pour améliorer les performances, l'expérience utilisateur ou même réduire la charge cognitive. Plus précisément, les interactions conçues dans cette thèse se concentrent sur une utilisation de en réalité virtuelle (RV). La RV est une technologie interactive et souvent immersive qui permet aux utilisateurs d'interagir avec un environnement totalement virtuel grâce à un retour sensoriel généré par ordinateur. La RV immersive repose souvent sur des équipements tels que des casques de réalité virtuelle ou des interfaces haptiques afin d'offrir la meilleure expérience possible.

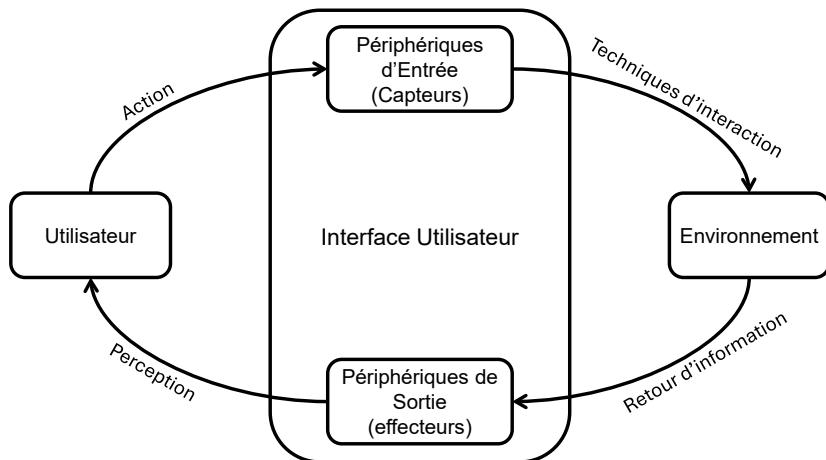


Figure A.1 – La boucle d'interaction régissant les interactions 3D.

Parmi les nombreux types d'interfaces haptiques, cette thèse se concentre sur la stimulation

haptique sans-contact par ultrasons (HSCU) [Rakkolainen et al. 2021]. Les interfaces HSCU (illustrées en [Figure A.2\(A\)](#)) reposent sur un ensemble de transducteurs à ultrasons, qui sont de petits haut-parleurs transformant l'électricité en vibrations à des fréquences ultrasoniques. En ajustant indépendamment la phase de chaque transducteur, il est alors possible de générer des interférences constructives à un endroit donné dans l'espace, appelé **point focal**, où la pression locale devient très élevée (voir [Figure A.2\(B\)](#)). Ensuite, en déplaçant rapidement et de manière répétée ce point focal le long d'une trajectoire, il est possible d'induire une vibration perceptible sur la peau à proximité (voir [Figure A.2\(C\)](#)) [Kappus et al. 2018; Frier et al. 2018]. Cependant, cette technologie reste relativement récente, ayant été introduite en 2008 [Iwamoto et al. 2008a], et il reste encore beaucoup à découvrir et à comprendre sur la perception, le rendu et l'application des stimuli HSCU.

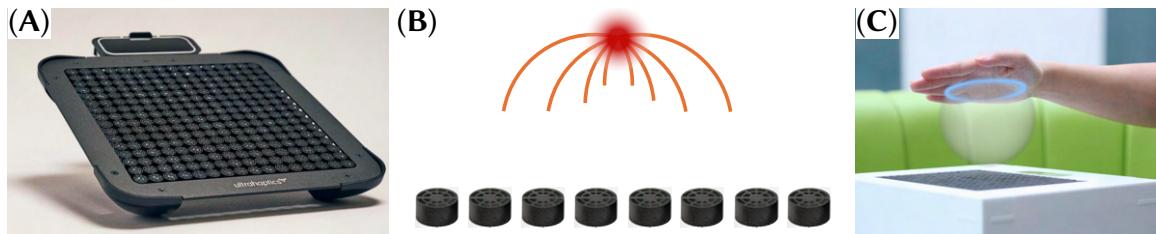


Figure A.2 – (A) Les interfaces HSCU, composées d'un ensemble de transducteurs à ultrasons, (B) focalisent les ondes ultrasonores vers des positions cibles dans l'espace. (C) En déplaçant ce point focal, elles peuvent générer une stimulation tactile le long d'une forme. Sources : (A) et (C) proviennent d'[Ultraleap](#), tandis que (B) est adaptée de [Rakkolainen et al. 2021].

C'est pourquoi cette thèse se concentre sur des *avancées en haptique sans-contact par ultrasons : études de perception, méthodes de rendu, et conception d'interactions en réalité virtuelle*. Nous proposons des avancées dans ce domaine en étudiant la perception et le rendu des stimuli HSCU, ainsi que leur utilisation conjointe avec les interactions VR. Nous étudions comment les stimuli HSCU sont **perçus** à travers des mesures physiologiques et des études psychophysiques, afin d'approfondir notre compréhension des stimuli HSCU et de permettre aux concepteurs de sensations d'optimiser leurs techniques de rendu. Nous proposons et discutons également de nouvelles primitives de **rendu** afin d'élargir la boîte à outils de conception de l'HSCU et d'offrir plus de liberté aux concepteurs HSCU, tout en repoussant les limites de ce qui peut être rendu. Nous formalisons, discutons et évaluons ensuite le pipeline de rendu pour l'HSCU afin de fournir une discussion standardisée sur le pipeline de rendu, en promouvant une plus grande reproductibilité dans ce domaine et une meilleure compréhension des mécanismes de rendu de bas-niveau liés à la conception des sensations HSCU. Enfin, nous contribuons à la conception et à l'évaluation de méthodes de retour d'information HSCU et

---

des techniques d'**interaction** associées en RV. Nous avons choisi la réalité virtuelle comme cadre pour ces interactions, car nous pensons que ces deux technologies sont très complémentaires, puisqu'elles favorisent toutes deux des interactions naturelles dans un environnement immersif.

## Défis Scientifiques

Bien que l'aspect innovant de l'HSCU est indéniable, ces technologies restent trop peu étudiées pour être utilisé de façon optimale. Plus précisément, dans le cadre de cette thèse, nous avons identifié quatre défis principaux liés à la perception (**I**), au rendu (**II**, **III**) et à l'interaction (**IV**).

**I. Comprendre la Perception Humaine des Stimuli HSCU.** Afin de fournir n'importe quel type de stimulation sensorielle efficace, il est primordial de comprendre comment le stimulus sera perçu par la personne ciblée. Cependant, la perception humaine est un processus extrêmement complexe, composé de multiples étapes [Goldstein et al. 2022a]. Mais en tant que concepteurs et conceptrices de sensations, nous ne pouvons contrôler que le stimulus distal.

D'une certaine manière, le rendu haptique (et plus généralement tout rendu sensoriel) peut être considéré comme un *encodage* d'informations, et la perception comme le *décodage*. En effet, le stimulus est une représentation d'informations que nous souhaitons perceptible et déchiffrable par la personne cible. Comme pour tout encodage d'informations, les informations décodées sont généralement plus importantes que celles qui sont encodées, car l'objectif principal de tout rendu sensoriel est d'induire des sensations cibles.

En étudiant cette relation complexe entre stimulus et sensation, nous obtiendrions une meilleure compréhension de ce processus de *déchiffrage*, ce qui permettrait aux concepteurs et conceptrices de sensations de concevoir des stimuli hautement spécifiques et contrôlés avec précision, suscitant des sensations précises, et offrant ainsi une expérience utilisateur de haute qualité. Et bien que cela soit vrai pour toute stimulation sensorielle, le caractère récent des technologies HSCU implique que très peu d'études de perception HSCU sont disponibles, ce qui conduit souvent les concepteurs et conceptrices à s'appuyer sur les résultats de perception issus d'études portant sur de la stimulation vibrotactile de contact, sans tenir compte des spécificités de l'HSCU (e.g. [Howard et al. 2023a]). C'est pourquoi la compréhension de ce processus de décodage est un objectif central de cette thèse, permettant d'élaborer des directives pour une

---

conception de stimuli plus efficaces.

**II. Améliorer les Algorithmes de Rendu et Fournir des Primitives Optimisées.** La qualité d'un algorithme de rendu dépend fortement du contexte de la tâche et de l'interaction, ce qui signifie qu'il n'existe pas d'algorithme de rendu HSCU universellement optimal. Par exemple, selon le contexte, nous pourrions vouloir rendre la même information, optimisée pour une intensité perçue élevée, un faible bruit de fonctionnement ou une netteté du signal. Et différents algorithmes de rendu seraient utilisés pour chacun de ces objectifs spécifiques.

De plus, les algorithmes de rendu HSCU peuvent généralement être considérés comme une séquence d'étapes permettant d'atteindre des objectifs généraux [Frier et al. 2019; Mulot et al. 2021b]. Disposer d'une boîte à outils regorgeant d'options (que nous appelons **primitives**) pour chaque étape est donc extrêmement avantageux pour de la conception de sensations, car cela permet de créer une gamme plus large de sensations, plus simplement. Pourtant, au moment de la rédaction de cette thèse, un nombre très limité de primitives HSCU existent ([Iwamoto et al. 2008b; Kappus et al. 2018; Takahashi et al. 2018]), et la boîte à outils de rendu HSCU manque de diversité.

**III. Fournir des Bases de Rendu Unifiées.** L'HSCU a attiré l'attention de plusieurs groupes de recherche à travers le monde, aux parcours et aux intérêts variés. Si cette diversité d'expertise et d'approches scientifiques est très bénéfique pour le domaine, elle implique également que des concepts similaires sont présentés de manières très différentes. Il est donc difficile, en particulier pour les jeunes chercheurs et chercheuses, concepteurs et conceptrices, de trouver des parallèles ou de comprendre les différences subtiles entre les techniques de rendu développées par différents groupes. Par exemple, le nom de cette technologie ne fait pas l'objet d'un consensus : alors que certains travaux, dont les nôtres, l'appellent « interface haptique sans-contact à ultrasons », ou « ultrasound mid-air haptic interface » en anglais [Mulot et al. 2021b; Seifi et al. 2023], d'autres la désignent sous le nom d'« airborne ultrasound tactile display » [Suzuki et al. 2021; Iwamoto et al. 2008a], voire « holographic haptic display » [Reardon et al. 2023]. D'autres exemples incluent la conception de techniques de rendu similaires : la modulation spatiotemporelle [Kappus et al. 2018; Frier et al. 2018] et la modulation latérale [Takahashi et al. 2018], prêtant parfois à confusion.

Cette divergence s'étend à la plupart des aspects de la présentation des travaux, avec des concepts similaires étant abordés de différentes manières. Cela rend malheureusement difficile la comparaison des travaux existants provenant de différents groupes, ou la reproduction des

---

études existantes. Afin de remédier à ce problème, nous avons besoin d'une formalisation unifiée du rendu HSCU, capable de représenter n'importe quel algorithme de rendu de manière précise et reproductible. En adoptant une telle formalisation, la communauté rendrait compte de ses résultats de manière plus cohérente, améliorant ainsi la reproductibilité dans ce domaine.

**IV. Améliorer les Interactions Naturelles en Réalité Virtuelle.** Bien que les technologies de RV aient connu d'énormes progrès techniques ces dernières années, permettant désormais le rendu visuel d'environnements très réalistes, l'absence de retour multisensoriel lors de l'interaction avec l'environnement réduit l'immersion. C'est pourquoi beaucoup de recherches portent sur le développement et l'intégration d'interfaces tactiles en RV.

Si notre objectif est d'offrir des interactions naturelles et réalistes en RV, nous devons avoir les mains libres de tout équipement et nos mouvements doivent être totalement libres afin de se rapprocher des interactions réelles. Les interfaces HSCU sont donc d'excellents candidats, car elles peuvent fournir un retour tactile, tandis que leur nature sans-contact répond à ces contraintes.

Cependant, l'HSCU présente également ses propres limites, telles que la faible intensité du retour tactile ou son espace de travail limité qui ne convient pas à tous les types d'interactions. Par exemple, lors d'interactions bimanuelles où deux mains sont positionnées l'une à côté de l'autre, les mains se retrouvent forcées à la limite de l'espace de travail des interfaces HSCU traditionnelles. Il est donc important d'étudier comment l'HSCU peut être intégré et couplé à un environnement virtuel afin de fournir des retours tactiles riches et discrets en RV. C'est ici que l'étude et la conception de schémas de couplage deviennent importantes pour représenter les relations entre les mises à jour de l'environnement virtuel simulé en réponse aux interactions virtuelles de l'utilisateur, et le retour tactile qui en résulte.

## Contributions

Alors que la section précédente met en évidence quatre défis scientifiques liés à l'étude de l'HSCU, cette thèse contribue à les relever, en élargissant notre compréhension de ces technologies et en améliorant leurs capacités. Cette thèse est organisée en quatre parties, apportant des contributions qui répondent respectivement aux quatre défis évoqués précédemment.

Chapitre 1 aborde tout d'abord les travaux connexes sur l'HSCU. Plus précisément, nous nous

---

intéressons au comportement de la peau et des afférents tactiles, avant d'aborder les solutions matérielles et logicielles permettant de stimuler ce système tactile. Nous approfondissons ensuite les spécificités de l'haptique sans-contact, et plus particulièrement de l'HSCU, en discutant de son fonctionnement, des méthodes de rendu et des outils qui ont été conçus pour l'utiliser. Enfin, nous abordons les interactions 3D et autres applications utilisant l'HSCU.

Les deux chapitres suivants, inclus dans la **Partie I**, traitent de certains aspects du **Défi I**, en se concentrant sur la perception des stimuli HSCU.

Le **Chapitre 2** vise à mesurer la réponse individuelle des afférents tactiles de premier ordre dans la peau glabre, à une large gamme de stimuli HSCU. Nous avons utilisé des techniques de microneurographie pour mesurer l'activation d'afférents tactiles individuels en réponse à des stimuli HSCU rendus avec différentes fréquences de modulation. Nous avons également mené des mesures de la déformation de la peau en profondeur, en réponse à la stimulation HSCU, à l'aide de méthodes de tomographie par cohérence optique. Cela permet de mieux comprendre comment les ondes acoustiques se propagent dans la peau et comment les afférents tactiles y répondent.

Ensuite, le **Chapitre 3** présente la conception d'une nouvelle technique de rendu HSCU, spécialement conçue pour fournir des sensations dynamiques représentant des formes. Cette technique s'inspire d'une méthode précédente [Hajas et al. 2020], mais vise à fournir une intensité perçue plus élevée avec une précision de discrimination similaire. Nous avons mené des études perceptuelles pour évaluer notre méthode et la comparer à l'état de l'art.

Ensuite, la **Partie II** se concentre sur le **Défi II** afin de fournir des primitives de rendu basées sur la physique pour l'HSCU.

Plus précisément, le **Chapitre 4** présente des primitives de rendu en forme de spirale qui peuvent être utilisées pour concentrer les ondes de choc de cisaillement résultant du mouvement rapide du point focal. Nous validons cette primitive par des simulations numériques et discutons de l'effet de ses paramètres. Enfin, nous utilisons cette primitive pour créer un algorithme complet de rendu de forme et nous évaluons son potentiel de haut niveau à travers une série d'études utilisateur.

La **Partie III** vise à fournir une formalisation générale du pipeline de rendu HSCU, adressant ainsi le **Défi III**.

Dans le **Chapitre 5**, nous formalisons les *stratégies d'échantillonnage*, qui représentent des algorithmes de rendu généraux pour les formes 2D. Nous présentons également DOLPHIN, une nouvelle infrastructure logicielle pour la conception de stimuli HSCU et leur évaluation perceptuelle. Nous avons également mené une expérience de psychophysique pour montrer com-

---

ment nous pouvions utiliser DOLPHIN en pratique pour générer des sensations, et comment nous pouvions l’interfacer avec des outils d’évaluation perceptuelle externes tels que PsychoPy [Peirce et al. 2019].

Le **Chapitre 6** étend cette formalisation au rendu de formes 3D en introduisant les *stratégies d’intersection*. Nous avons conçu différentes stratégies d’intersection pour illustrer le pipeline et mené des études utilisateurs afin de montrer comment nous pouvons utiliser les stratégies d’intersection pour modifier la perception des propriétés d’objets virtuels.

Enfin, la **Partie IV** explore les moyens d’améliorer les interactions en réalité virtuelle à l’aide du retour HSCU, contribuant ainsi au **Défi IV**.

Le **Chapitre 7** utilise le toucher comme canal permettant de fournir des informations supplémentaires à l’utilisateur ou utilisatrice, sans encombrer le canal visuel. Plus précisément, nous avons conçu des méthodes de retour HSCU pour encoder des informations de position et d’orientation 2D, qui peuvent être utilisées pour guider la main de la personne vers une pose donnée. Nous avons mené deux études utilisateurs afin d’évaluer ces méthodes, à la fois dans un environnement statique et dynamique.

À l’inverse, le **Chapitre 8** utilise le canal tactile pour fournir des informations redondantes, cohérentes avec la stimulation visuelle. Plus précisément, nous avons proposé un ensemble de retours visuels et haptiques pour représenter l’interpénétration avec un objet virtuel. Comme la stimulation tactile ne peut à elle seule empêcher l’interpénétration, et que la RV manque souvent de retour haptique pendant la manipulation, la combinaison de la stimulation visuelle et de la stimulation HSCU peut être mutuellement bénéfique. Nous avons donc mené une étude utilisateurs afin d’associer chaque retour visuel au meilleur retour HSCU associé, générant ainsi des techniques de retour visuo-haptique. Enfin, nous avons mené une deuxième étude utilisateurs afin d’évaluer chacune de ces paires visuo-haptiques, dans le but de trouver la meilleure candidate.

Ensuite, le **Chapitre 9** s’attaque aux limites de l’espace de travail de l’HSCU en proposant un système matériel dédié aux interactions bimanuelles. Nous proposons également un schéma de rendu dédié aux manipulations haptiques bimanuelles avec l’HSCU. Nous avons ensuite mené deux études utilisateurs afin de mesurer l’effet de l’HSCU lors de tâches d’exploration et de préhension en réalité virtuelle.

Enfin, le **Chapter 10** conclut cette thèse et aborde les perspectives de travaux futurs visant à approfondir ces défis scientifiques.





**Titre :** Avancées en haptique sans-contact par ultrasons : études de perception, méthodes de rendu, et conception d'interactions en réalité virtuelle

**Mot clés :** Rendu Haptique, Perception, Réalité Virtuelle

**Résumé :** Les interfaces haptiques sans-contact à ultrasons (HSCU) permettent d'améliorer les interactions via des gestes naturels, en générant des sensations tactiles à distance. L'HSCU offre un large espace de conception, et s'appuie sur des phénomènes acoustiques complexes dont les implications perceptuelles restent peu étudiées. Dans cette thèse, nous présentons des avancées dans le domaine de l'HSCU en étudiant les questions sous-jacentes selon des perspectives de perception, rendu, et interactions. Nous examinons dans un premier temps la perception des stimuli HSCU en réalisant les premiers enregistrements de l'activité nerveuse pour l'HSCU, caractérisant les réponses afférentes sur peau glabre et poilue. Nous proposons ensuite un nouvel algorithme de rendu tactile de formes, validé par études psychophysiques. Nous examinons la biomécanique de la peau sous stimula-

tion HSCU afin de développer de nouvelles primitives de rendu basées-physique, évaluées lors d'études de rendu de formes 3D. Nous présentons une infrastructure logicielle pour la conception de stimuli HSCU, définissant les composants clés des stratégies d'échantillonnage et d'intersection, permettant le rendu d'interactions avec un objet 3D. Nous démontrons également que le choix de stratégie d'intersection peut significativement altérer la perception des propriétés d'objets virtuels. Enfin, nous abordons l'HSCU dans le contexte d'interactions en réalité virtuelle en proposant et évaluant des techniques de guidage de main basées sur l'HSCU, en explorant le couplage visuo-haptique pour le rendu d'interpénétration, et en discutant des schémas de couplage HSCU bimanuels pour les scénarios de manipulation.

**Title:** Advancing Ultrasound Mid-Air Haptics: Perception Studies, Rendering Methods, and Design of Virtual Reality Interactions

**Keywords:** Haptic Rendering, Perception, Virtual Reality

**Abstract:** Ultrasound mid-air haptics (UMH) aims to improve natural gestures and interactions by rendering tactile sensations directly in mid-air. It offers a rich design space, while relying on complex acoustic phenomena whose perceptual implications remain understudied. In this thesis, we advance the field by investigating the underlying questions of UMH from the perspectives of perception, rendering, and interactions. We first study the low-level perception of UMH stimuli by conducting the first nerve recordings for UMH, characterizing afferent responses on both glabrous and hairy skin. We then propose a novel shape rendering algorithm, validated through psychophysical studies. Next, we examine skin biomechanics under UMH stimu-

lation to develop new physics-based rendering primitives, demonstrated in more complex shape rendering studies. We then introduce a formal framework for UMH rendering, defining the key components of sampling strategies and intersection strategies for rendering interactions with 3D objects. We also demonstrate through perception studies that the choice of these strategies significantly alters the perceived properties of virtual objects. Finally, we address UMH in virtual reality interactions by proposing and evaluating UMH-based hand position and orientation guidance techniques, exploring visuo-haptic coupling for interpenetration feedback, and discussing bi-manual UMH coupling schemes for manipulations scenarios.