Improving the Perception of Mid-Air Tactile Shapes With Spatio-Temporally-Modulated Tactile Pointers

Lendy Mulot, Thomas Howard, Claudio Pacchierotti, Maud Marchal

Improving the Perception of Mid-Air Tactile Shapes With Spatio-Temporally-Modulated Tactile Pointers

LENDY MULOT*, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS, France
THOMAS HOWARD*, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS, France
CLAUDIO PACCHIEROTTI, CNRS, Univ Rennes, Inria, IRISA, France
MAUD MARCHAL, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS and IUF, France

Ultrasound mid-air haptic (UMH) devices can remotely render vibrotactile shapes on the skin of unequipped users, e.g., to draw haptic icons or render virtual object shapes. Spatio-temporal modulation (STM), the state-of-the-art UMH shape rendering method, provides large freedom in shape design and produces the strongest possible stimuli for this technology. Yet, STM shapes are often reported to be blurry, complicating shape identification. Dynamic tactile pointers (DTP) were recently introduced as a technique to overcome this issue. By tracing a contour with an amplitude-modulated focal point, they significantly improve shape identification accuracy over STM, but at the cost of much lower stimulus intensity. Building upon this, we propose Spatio-temporally-modulated Tactile Pointers (STP), a novel method for rendering clearer and sharper UMH shapes while at the same time producing strong vibrotactile sensations. We ran two human participant experiments, which show that STP shapes are perceived as significantly stronger than DTP shapes, while shape identification accuracy is significantly improved over STM and on par with that obtained with DTP. Our work has implications for effective shape rendering with UMH, and provides insights which could inform future psychophysical investigation into vibrotactile shape perception in UMH.

CCS Concepts: • **Human-centered computing** → **Haptic devices**; **User studies**.

Additional Key Words and Phrases: Ultrasound Haptics; Mid-Air Haptics; Haptic Perception

ACM Reference Format:
Lendy Mulot, Thomas Howard, Claudio Pacchierotti, and Maud Marchal. 2023. Improving the Perception of Mid-Air Tactile Shapes With Spatio-Temporally-Modulated Tactile Pointers. 1, 1 (June 2023), 17 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Ultrasound mid-air haptics (UMH) [Georgiou et al. 2022; Rakkolainen et al. 2020] is a tactile feedback technology which does not require any direct physical contact between the user and device, with applications ranging from interaction with gesture-based interfaces [Freeman 2022] to interactions with virtual environments [Howard et al. 2022]. This technology stands out because of the high flexibility afforded by not needing to equip users in

*Both authors contributed equally to this research.

Authors' addresses: Lendy Mulot, lendy.mulot@irisa.fr, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS, Rennes, France; Thomas Howard, thomas.howard@irisa.fr, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS, Rennes, France; Claudio Pacchierotti, claudio.pacchierotti@irisa.fr, CNRS, Univ Rennes, Inria, IRISA, Rennes, France; Maud Marchal, maud.marchal@irisa.fr, Univ Rennes, INSA Rennes, IRISA, Inria, CNRS and IUF, Rennes, France.

XXXX-XXXX/2023/6-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

any way, and because it provides designers great freedom to manipulate the vibrotactile stimulus' spatial and temporal parameters.

Because they allow rapid motion of a point of vibration in any direction, these interfaces are able to draw tactile shapes on users' skin. This has been applied to tactile communication [Hoshi 2012], recognizing mid-air tactile user interface elements [Brown et al. 2021; Freeman 2022] or feeling the shape of virtual objects [Long et al. 2014] in augmented reality [Inoue et al. 2015; Romanus et al. 2019] and virtual reality [Howard et al. 2022; Martinez et al. 2019].

UMH shapes are mostly rendered using a method called spatio-temporal modulation (STM) [Kappus and Long 2018] or lateral modulation [Takahashi et al. 2018], both of which allow complex shapes to be drawn at high perceived intensity [Frier et al. 2018, 2019]. However, a disadvantage of these methods is that they produce rather blurry shapes [Obrist et al. 2013; Reardon et al. 2023]. As a consequence, experimental evaluations of the effectiveness of shape rendering using these approaches yield poor performances [Hajas et al. 2020; Howard et al. 2019; Rutten et al. 2019].

To overcome this issue, Hajas et al. proposed *Dynamic Tactile Pointers (DTP)* [Hajas et al. 2020], which consist in moving an amplitude-modulated (AM) focal point relatively slowly along a shape contour, optionally marking pauses at vertices to highlight them. They experimentally demonstrate that, for a small set of simple shapes, this method vastly outperforms STM in terms of shape identification accuracy. However, amplitude modulation makes DTP shapes feel weak [Frier et al. 2018; Howard et al. 2019; Takahashi et al. 2018] and requires them to be drawn over a longer minimum period of time to be recognizable [Hajas et al. 2020; Mulot et al. 2021]. This motivates our present work into Spatio-temporally-modulated Tactile Pointers (STP). Our contributions are:

- STP: A novel rendering method for overcoming the limitations of dynamic tactile pointers and STM (Sec. 3)
- Two human participant experiments assessing the benefits in terms of perceived intensity (Sec. 4) and shape identification performance (Sec. 5 and Sec. 6)

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 Sec. 7).

2 RELATED WORK

UMH devices use an array of ultrasonic transducers to emit acoustic waves whose phase is offset in such a way as to focus them at precise locations in the 3D space above the device, referred to as focal points [Carter et al. 2013; Iwamoto et al. 2008]. When they encounter the skin, these focal points generate a vibrotactile stimulus. For the stimulus to be perceivable, the high-frequency vibration generated by the array (usually 40 kHz [Carter et al. 2013] or 70 kHz [Ito et al. 2016]), which sits well outside the range of human vibrotactile perception [Bolanowski Jr et al. 1988] needs to be modulated at a lower, perceivable frequency.

The most straightforward method is to attenuate the output of the transducers cyclically over time, i.e. amplitude-modulating the transducers' outputs [Carter et al. 2013; Iwamoto et al. 2008]. This method is commonly referred to as amplitude modulation (AM).

A more recent approach consists of rapidly moving an unmodulated focal point between neighboring positions, while controlling the frequency at which the focal point passes any given position, through techniques called lateral modulation (LM) [Takahashi et al. 2018] or spatio-temporal modulation (STM) [Kappus and Long 2018]. By cyclically returning the focal point to each position along a shape contour, LM and STM end up locally amplitude-modulating the pressure signal at that location, achieving an effect as though all focal point path positions were simultaneously amplitude modulated [Kappus and Long 2018]. With the correct parameters, this can provide an impression of a singular vibrotactile shape on the skin [Frier et al. 2018]. Since the output of the

transducers is unattenuated in these methods, they render tactile shapes which are perceived as significantly stronger than AM stimuli [Frier et al. 2018; Howard et al. 2019; Kappus and Long 2018; Takahashi et al. 2018].

As previously mentioned, STM can not only make unmodulated UMH stimuli perceivable, but also "draw" complex vibrotactile shapes on a user's skin, which is one of the key advantages of UMH over other vibrotactile feedback technologies. However, several other methods based on amplitude modulation were also introduced for rendering mid-air tactile shapes.

Multi-point simultaneous AM was the first of these methods to be proposed. UMH interfaces are capable of simultaneously generating multiple focal points [Carter et al. 2013; Korres and Eid 2016] which can be arranged so as to form a shape. However, this approach produces rather weak stimuli as the full device output power is not constantly used due to amplitude modulation ([Frier et al. 2018; Kappus and Long 2018; Takahashi et al. 2018]). Furthermore, multi-point rendering divides the output power (and thus the maximum achieved intensity) across the number of rendered points, with higher point counts leading to weaker stimuli. This quickly brings the outputs below the detection threshold [Howard et al. 2019] for shapes with more than a few points [Carter et al. 2013].

Hajas et al. formalized the concept of *"Dynamic Tactile Pointers" (DTP)*, which relaxes some of the constraints posed by multi-point AM by sequentially moving a single focal point along a path [Hajas et al. 2020]. This allows the shapes to be rendered at maximum intensity (of an AM stimulus), without any limit to the number of positions used to trace the shape contour. However, achieving a recognizable AM shape requires the focal point to dwell significantly longer at each position along the shape contour than it would in STM. This yields a sensation of the shape being dynamically traced rather than static and requiring longer rendering times for larger shapes. This sensation of a dynamically traced shape can be advantageous for certain applications, but the key benefit observed by Hajas et al. was a significantly improved shape identification accuracy when compared to STM, at least for simple geometric shapes [Hajas et al. 2020].

They also proposed an improved version of *Dynamic Tactile Pointers (DTP)* in which an AM focal point not only traces the contour of a shape, but dwells longer on the shape's vertices. This highlights vertices and allows polygonal shapes to be perceptually processed in "chunks", yielding the best recorded shape identification performances to date [Hajas et al. 2020]. In the remainder of this paper, we will distinguish DTP without pauses on the vertices (*Single-Stroke DTP* or **DTP-SS**) and DTP with pauses on the vertices (*Multi-Stroke DTP* or **DTP-MS**). Despite their benefits, DTP produce rather weak stimuli because they use amplitude modulation. This is why we propose to develop an STM-based approach which attempts to reproduce the beneficial features of DTP.

3 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.

3.1 Rendering Algorithm

Here, we consider a contour to be parameterized by one variable: $s : [0, 1] \rightarrow \mathbb{R}^3$.

For a DTP rendering [Hajas et al. 2020], the shape would be represented by a set of n points $\mathbf{P_1}, ..., \mathbf{P_n}$ evenly spaced out, with $\mathbf{P_i} = s(i/n)$. Then, all of these points would be successively rendered with amplitude modulation (AM). Let f be the amplitude modulation frequency. Each point is then rendered 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.

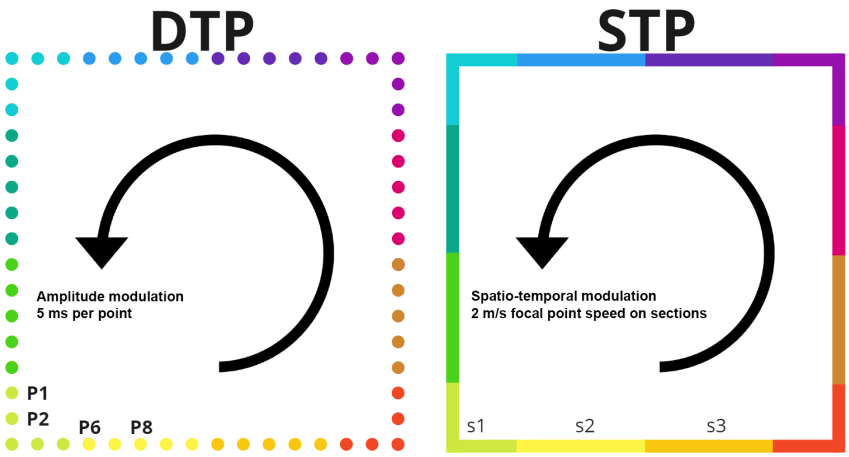

Fig. 1. (Left) DTP consists in a succession of amplitude-modulated points along the shape contour (here, a square), while (Right) STP consists in a succession of spatio-temporally-modulated sections along the contour.

The design of STP is inspired by that of DTP. We chose to cut the shape into n sections $s_1, ..., s_n$ of identical length. Each section can then be rendered with STM, using a fixed speed focal point motion. 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 Sec. 3.2. Finally, the different sections are rendered sequentially, to provide the same impression of a moving pointer as in DTP. Both techniques are illustrated in Fig. 1.

We propose two variations of our method, depending on whether the shapes to render have vertices or not. The first variation, **STP-SS**, 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**, 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].

3.2 Stimuli Parameterization for Experimental Validation

The stimuli evaluated in both subsequent user studies (Sec. 4 and Sec. 6) 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 [Komandur et al. 2009]. Their dimensions are summarized in Tab. 1.

General rendering parameters: To ensure that our results are comparable to those of Hajas et al. [Hajas et al. 2020], we rendered all shapes with a draw frequency of 0.5 Hz (2 s to fully draw 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 drawn 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 [Sun 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.'s optimum results for lateral modulation [Takahashi et al. 2018]. There are some variations on this length based on the specific shape being rendered, as we require all sections to be of exact same dimensions, 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 Tab. 1.

The shape identification experiment (see Sec. 6) also uses conventional STM rendering, with a focal point speed of 7 m/s [Frier et al. 2018, 2019]. All DTP 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 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	Octogon
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

4 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 Sec. 3.2) rendered either with STP or with DTP in both their variations.

4.1 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. [Howard et al. 2019] when investigating detection thresholds for AM and STM stimuli, we suppose that perceived intensity is independent of exploratory motion. For practical reasons, we therefore perform this experiment only in a passive configuration, where the participant's hand remains immobile throughout each trial.

4.2 Materials and Methods

4.2.1 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

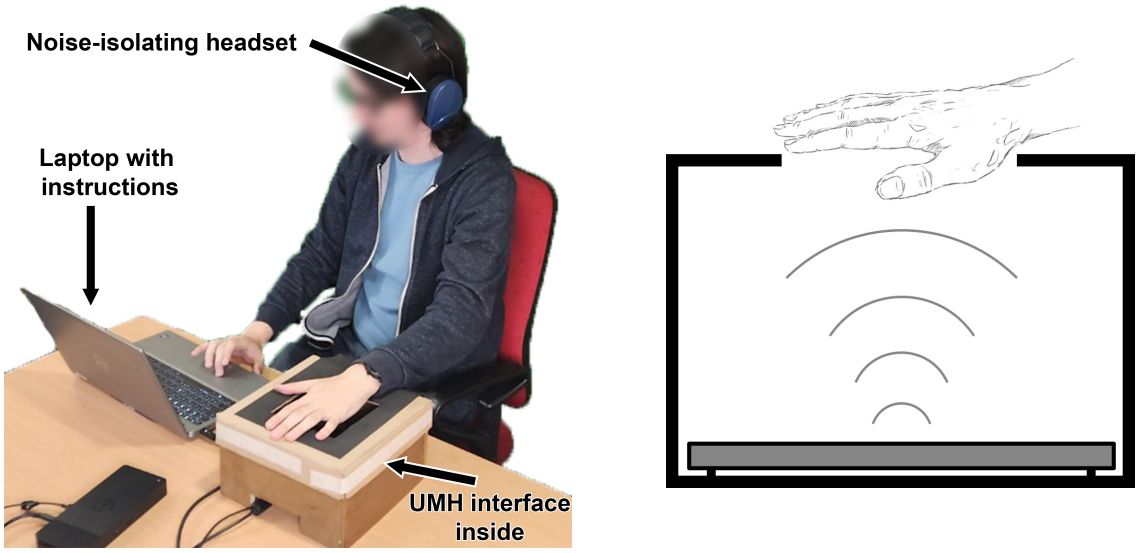

Fig. 2. Experimental setup. (Left) Participants are seated in front of the laptop running the experiment and displaying instructions, while their non-dominant hand is resting on a box containing the haptic interface (Right).

box, as illustrated in Fig. 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.

We created the experiment logic and user interface using Psychopy [Peirce et al. 2019], which triggered the playback of stimuli which were previously designed and hardcoded in C++ using Ultrealeap'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 Sec. 5 for further discussion of this point).

4.2.2 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 Sec. 3.2): 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-AFC protocol [Jones and Tan 2012], 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 Sec. 5 below.

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.

4.2.3 Participants. We recruited 18 participants (13 M, 5F), aged 18 to 52 (m: 25.8, s.d.: 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.

4.3 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" 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**.

For both the (**-SS**) and (**-MS**) blocks, we calculated the proportions P_C and P_D of "second stimulus stronger than the first" in the trials where STP and DTP were tested against themselves. Shapiro 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 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 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**) ($p < 0.001$) and (**-MS**) blocks ($p = 0.01$).

A summary of these results is shown in Tab. 2, and the distributions are illustrated in Fig. 3.

Table 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%	51.38%

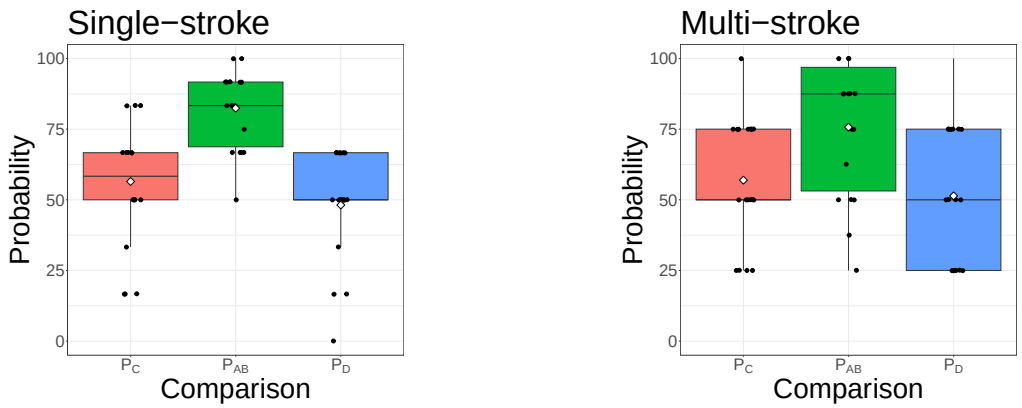

Fig. 3. Distributions of P_{AB}, P_C and P_D for the single-stroke stimulus (-**SS**) block (Left) and multi-stroke stimulus (-**MS**) block (Right). For P_C and P_D, values close to 50% indicate an absence of systemic bias in participant responses in the case of uncertainty. An average above 50% for P_{AB} shows that STP stimuli are perceived as more intense than DTP stimuli. Black horizontal lines show the median values, white diamonds represent the mean value, and black dots represent individual participant results.

4.4 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**, 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 [Howard et al. 2019; Kappus and Long 2018].

4.5 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.

5 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 (Sec. 4).

5.1 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 Sec. 4), shape (see Sec. 3.2), and block ((-**SS**) or (-**MS**)) evaluated in the stimulus intensity comparison experiment (Sec. 4). 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. Fig. 4 shows some examples of classifications of actual subject responses.

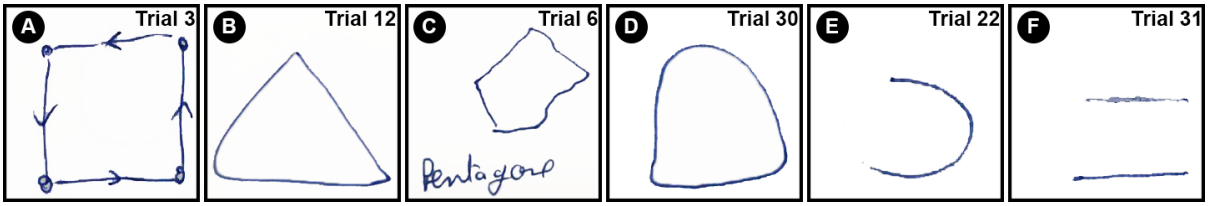

Fig. 4. Examples of (A) an unambiguous response (in this case, we are confident that the subject felt a "square"), (B) a response which could match a drawn shape but appears ambiguous (in this case, the subject probably felt a triangle, however the combination of two vertices and one rounded corner is 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").

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 Sec. 4.2.2), 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.

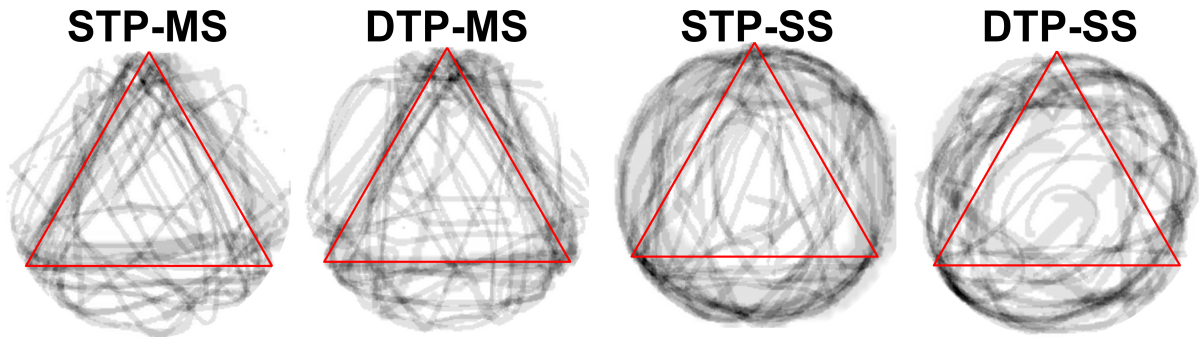

Fig. 5. Superposition of all the participant's sketches for the different rendering techniques, after a triangle was rendered on their hand. Sketches were centered and normalized. The darker the point, the more participants drew it. The red triangle corresponds to the displayed triangle. We observe more confusion with the single-stroke rendering (right) than with the multi-stroke rendering (left).

Table 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%

5.2 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 summarizes the proportions of apparently correct naive shape representations for each shape and rendering method. Fig. 5 shows global results for the triangle shape.

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

5.3 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 important 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.

6 CLARITY OF RENDERED SHAPES

In Sec. 4, 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 Sec. 5 would lead us to believe.

6.1 Hypotheses

Based on the qualitative results from Sec. 5 and prior results from Hajas et al. [Hajas et al. 2020], we hypothesize that:

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

6.2 Materials and Methods

6.2.1 Apparatus and setup. Apart from the experimental user interface, the setup for this experiment is identical to that of the first experiment (see Sec. 4.2.1 and Fig. 2).

6.2.2 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 [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.

6.2.3 Participants. We recruited a total of 20 participants (15 M, 4 F, 1 N-B), aged 21 to 32 (m: 24.85, s.d.: 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.

6.2.4 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**.

6.3 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 Fig. 6, and all average results are in Tab. 4. Recall and sharpness distributions are shown in Fig. 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, $p > 0.05$), and variances were similar between the three rendering techniques (Levene, $p > 0.05$). An 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 ($p < 0.001$). 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).

***STM** vs. **DTP-MS** vs. **STP-MS**.* Recall and sharpness distributions with **STM**, **DTP-MS** and **STP-MS** were almost normal (Shapiro, $p > 0.05$) 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 ($p < 0.001$), and post-hoc pairwise t-tests with Holm corrections revealed that both **DTP-MS** and **STP-MS** yielded significantly better performances than **STM** ($p < 0.001$). An ANOVA revealed a significant effect of rendering technique on sharpness ($p < 0.001$), and the post-hoc analysis similarly showed that **STM** shapes were felt significantly more blurry than the other two strategies.

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, $p < 0.05$), so we performed Friedman tests, which revealed a significant effect of rendering technique in both blocks. Pairwise Wilcoxon tests with Holm corrections revealed that participant were significantly more tempted to explore **STM** shapes dynamically ($p < 0.001$ for the multi-stroke block, and $p = 0.002$ for the single-stroke block).

Table 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	0.625
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

6.4 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 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 shown to use 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.

7 DISCUSSION AND PERSPECTIVES

7.1 Limitations and Future Work

While 38 people participated in the experiments, this is not a sufficiently large or varied sample 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.

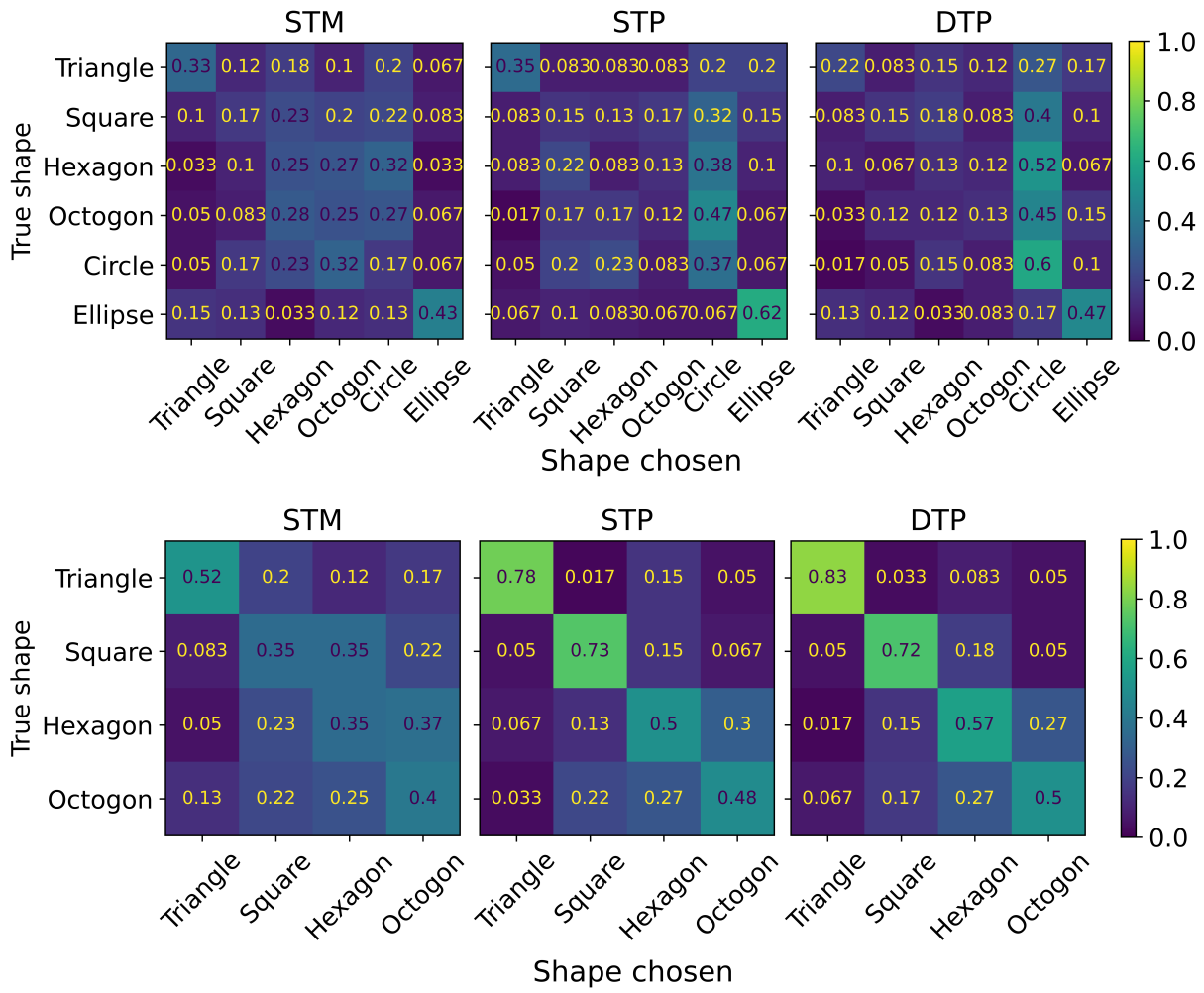

Fig. 6. Confusion matrices for the shape identification experiment, for both the (-**SS**) (Top) and (-**MS**) (Bottom) blocks.

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 Tab. 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.

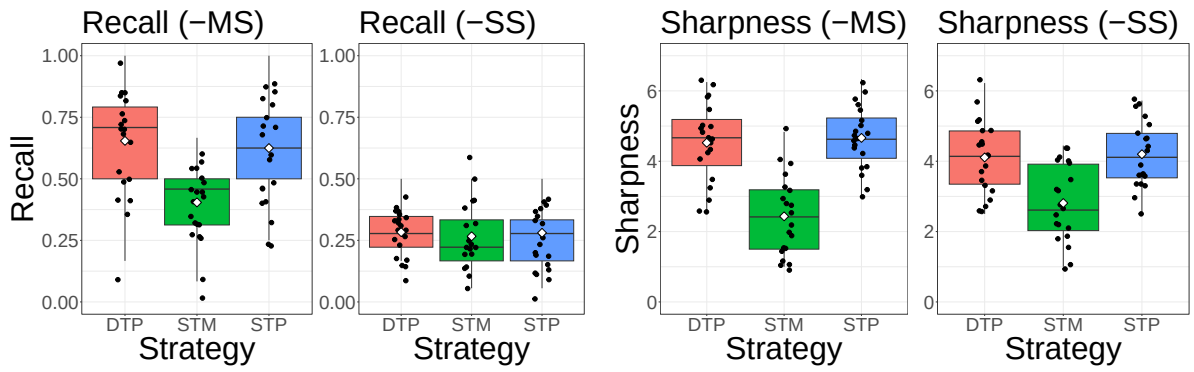

Fig. 7. Recall (Left) and Sharpness (Right) distributions for the shape identification experiment. Black horizontal lines show the median values, white diamonds represent the mean value, and black dots represent individual participant results.

7.2 General Discussions

The measured identification performances for STM and DTP are on par with the results of Hajas et al. [Hajas et al. 2020]. The slightly lower results in our case can be explained by the addition of other answer possibilities, and the slightly sub-optimal 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 confirm that the general concept of tactile pointers improves shape identification.

We also observe improved performances for the naive shape identification (Tab. 3), in the case of simple shapes. The naive and constrained results (Fig. 6) also show that users tend to report circular shapes when the amount of vertices increases. This is on par with the results of Mulot et al. [Mulot et al. 2021], who reported that small variations of curvature are mostly unperceivable with DTP.

Apart from the base shapes studied by Hajas, we added three more complex shapes which were more poorly recognized. This can be explained by a sub-optimal parametering of DTP and STP, such as the lower pause duration on vertices. This is coherent with Hajas' results for the triangle and square, showing poorer recognition rates with shorter pauses.

The improved results in a constrained answer scenario shows 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.

8 CONCLUSION

In this paper, we present *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.

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.

STP show promising results for rendering intense and recognizable mid-air haptic shapes, and could potentially improve many applications of ultrasound mid-air haptics.

ACKNOWLEDGMENTS

This project has received funding under the ANR project “MIMESIS”.

REFERENCES

Stanley J Bolanowski Jr, George A Gescheider, Ronald T Verrillo, and Christin M Checkosky. 1988. Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical society of America* 84, 5 (1988), 1680–1694.

Eddie Brown, David R Large, Hannah Limerick, William Frier, and Gary Burnett. 2021. Validating the Salience of Haptic Icons for Automotive Mid-Air Haptic Gesture Interfaces. *Contemporary Ergonomics & Human Factors 2022* (2021), 82.

Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. 2013. UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proc. of the 26th ACM Symposium on User Interface Software & Technology*. ACM, 505–514.

Euan Freeman. 2022. Ultrasound Haptic Feedback for Touchless User Interfaces: Design Patterns. In *Ultrasound Mid-Air Haptics for Touchless Interfaces*. Springer, 71–98.

William Frier, Damien Ablart, Jamie Chilles, Benjamin Long, Marcello Giordano, Marianna Obrist, and Sriram Subramanian. 2018. Using spatiotemporal modulation to draw tactile patterns in mid-air. In *Proc. of the Eurohaptics Conference*. Springer, 270–281.

William Frier, Dario Pittera, Damien Ablart, Marianna Obrist, and Sriram Subramanian. 2019. Sampling strategy for ultrasonic mid-air haptics. In *Proc. of the CHI Conference on Human Factors in Computing Systems*. 1–11.

Orestis Georgiou, William Frier, Euan Freeman, Claudio Pacchierotti, and Takayuki Hoshi. 2022. *Ultrasound Mid-Air Haptics for Touchless Interfaces*. Springer Cham. https://doi.org/10.1007/978-3-031-04043-6

Daniel Hajas, Dario Pittera, Antony Nasce, Orestis Georgiou, and Marianna Obrist. 2020. Mid-Air Haptic Rendering of 2D Geometric Shapes with a Dynamic Tactile Pointer. *IEEE Transactions on Haptics* (2020).

Takayuki Hoshi. 2012. Handwriting transmission system using noncontact tactile display. In *Proc. IEEE Haptics Symp.* 399–401.

Thomas Howard, Gerard Gallagher, Anatole Lécuyer, Claudio Pacchierotti, and Maud Marchal. 2019. Investigating the recognition of local shapes using mid-air ultrasound haptics. In *Proc. of the IEEE World Haptics Conference (WHC)*. IEEE, 503–508.

Thomas Howard, Maud Marchal, and Claudio Pacchierotti. 2022. Ultrasound Mid-Air Tactile Feedback for Immersive Virtual Reality Interaction. In *Ultrasound Mid-Air Haptics for Touchless Interfaces*. Springer, 147–183.

Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. 2015. Active touch perception produced by airborne ultrasonic haptic hologram. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, 362–367.

Mitsuru Ito, Daisuke Wakuda, Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. 2016. High spatial resolution midair tactile display using 70 kHz ultrasound. In *Proc. of the Eurohaptics Conference*. Springer, 57–67.

Takayuki Iwamoto, Mari Tatezono, and Hiroyuki Shinoda. 2008. Non-contact method for producing tactile sensation using airborne ultrasound. In *Proc. Eurohaptics Conf.* 504–513.

Lynette A Jones and Hong Z Tan. 2012. Application of psychophysical techniques to haptic research. *IEEE Transactions on Haptics* 6, 3 (2012), 268–284.

Brian Kappus and Ben Long. 2018. Spatiotemporal modulation for mid-air haptic feedback from an ultrasonic phased array. *The Journal of the Acoustical Society of America* 143, 3 (2018), 1836–1836.

Sashidharan Komandur, Peter W Johnson, Richard L Storch, and Michael G Yost. 2009. Relation between index finger width and hand width anthropometric measures. In *Proc. Ann. Int. Conf. IEEE Engineering in Medicine and Biology Soc.* 823–826.

Georgios Korres and Mohamad Eid. 2016. Haptogram: Ultrasonic point-cloud tactile stimulation. *IEEE Access* 4 (2016), 7758–7769.

Benjamin Long, Sue Ann Seah, Tom Carter, and Sriram Subramanian. 2014. Rendering volumetric haptic shapes in mid-air using ultrasound. *ACM Trans. Graph.* 33, 6 (2014), 1–10.

Jonatan Martinez, Adam Harwood, Hannah Limerick, Rory Clark, and Orestis Georgiou. 2019. Mid-air haptic algorithms for rendering 3D shapes. In *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*. IEEE, 1–6.

Lendy Mulot, Guillaume Gicquel, Quentin Zanini, William Frier, Maud Marchal, Claudio Pacchierotti, and Thomas Howard. 2021. DOLPHIN: A Framework for the Design and Perceptual Evaluation of Ultrasound Mid-Air Haptic Stimuli. In *Proc. of the ACM Symposium on Applied Perception*. 1–10.

Marianna Obrist, Sue Ann Seah, and Sriram Subramanian. 2013. Talking about tactile experiences. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. 1659–1668.

Jonathan Peirce, Jeremy R Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior research methods* 51 (2019), 195–203.

Ismo Rakkolainen, Euan Freeman, Antti Sand, Roope Raisamo, and Stephen Brewster. 2020. A Survey of Mid-Air Ultrasound Haptics and Its Applications. *IEEE Trans. Haptics* (2020).

Gregory Reardon, Bharat Dandu, Yitian Shao, and Yon Visell. 2023. Shear shock waves mediate haptic holography via focused ultrasound. *Science Advances* 9, 9 (2023), eadf2037.

Ted Romanus, Sam Frish, Mykola Maksymenko, William Frier, Loïc Corenthy, and Orestis Georgiou. 2019. Mid-air haptic bio-holograms in mixed reality. In *Proc. 2019 IEEE ISMAR-Adjunct*. 348–352.

Isa Rutten, William Frier, Lawrence Van den Bogaert, and David Geerts. 2019. Invisible touch: How identifiable are mid-air haptic shapes?. In *Extended Abstracts, CHI Conf. Hum. Fact. Comput. Syst.* 1–6.

Chongyang Sun, Weizhi Nai, and Xiaoying Sun. 2019. Tactile sensitivity in ultrasonic haptics: Do different parts of hand and different rendering methods have an impact on perceptual threshold? *Virtual Reality & Intelligent Hardware* 1, 3 (2019), 265–275.

Ryoko Takahashi, Keisuke Hasegawa, and Hiroyuki Shinoda. 2018. Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli. In *Proc. Eurohaptics Conf.* 276–288.