

It Sounds Cool: Exploring Sonification of Mid-air Haptic Textures Exploration on Texture Judgments, Body Perception, and Motor Behaviour

Roberto Montano-Murillo, *Member, IEEE*, Dario Pittera, *Member, IEEE*, William Frier, *Member, IEEE*, Orestis Georgiou, *Senior Member, IEEE*, Marianna Obrist, *Member, IEEE*, and Patricia Cornelio, *Member, IEEE*,

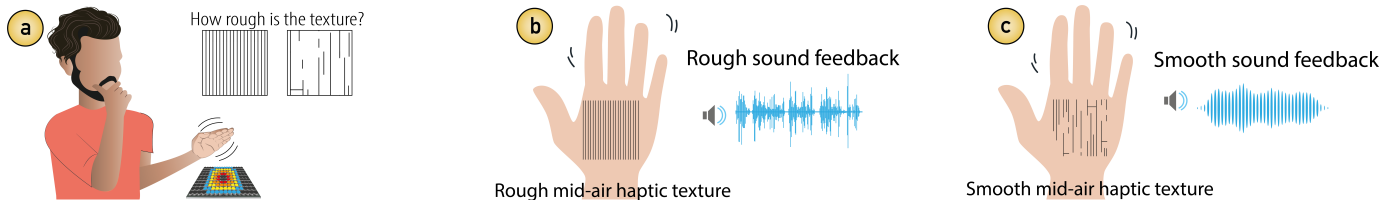


Fig. 1: (a) Perceiving different levels of roughness of mid-air haptic textures can be difficult when relying on the sense of touch alone. (b-c) We combine mid-air texture exploration with congruent audio feedback, to explore how people's judgment of the texture attributes is influenced.

Abstract—Ultrasonic mid-air haptic technology allows for the perceptual rendering of textured surfaces onto the user's hand. Unlike real textured surfaces, however, mid-air haptic feedback lacks implicit multisensory cues needed to reliably infer a texture's attributes (e.g., its roughness). In this paper, we combined mid-air haptic textures with congruent sound feedback to investigate how sonification could influence people's (1) explicit judgment of the texture attributes, (2) explicit sensations of their own hand, and (3) implicit motor behavior during haptic exploration. Our results showed that audio cues (presented solely or combined with haptics) influenced participants' judgment of the texture attributes (roughness, hardness, moisture and viscosity), produced some hand sensations (the feeling of having a hand smoother, softer, looser, more flexible, colder, wetter and more natural), and changed participants' speed (moving faster or slower) while exploring the texture. We then conducted a principal component analysis to better understand and visualize the found results and conclude with a short discussion on how audio-haptic associations can be used to create embodied experiences in emerging application scenarios in the metaverse.

Index Terms—Mid-air haptics, touchless interaction, textures, sound, body perception, Human-Computer Interaction.

I. INTRODUCTION

ADVANCES in extended reality (XR) are fast-tracking us towards a metaverse vision where one is able to go beyond 2D screens, reach out and directly interact with digital or virtual content [1]. Unlike contact-based wearable haptic devices, ultrasound-based mid-air haptic technology [2], [3] is able to provide a variety of static and dynamic tactile sensations (e.g., points, lines, shapes, textures) directly onto the user's bare hands in an unobtrusive manner. A crucial characteristic of touching virtual content in XR environments is the ability to haptically render mid-air textures [4]. While previous work has achieved rendering of mid-air textured surfaces involving complex haptic attributes (e.g., roughness [5] [6] [7], stiffness [8], softness [9], and viscosity [10]), the spatial properties of geometries intended to produce textured surfaces are still difficult to discern for human perception.

Therefore, it still remains challenging to achieve a mid-air haptic rendering that allows textures to be convincingly perceived and discriminated (see Figure 1a).

Prior work in HCI shows that sensory cues such as audio feedback, aids user perception when interacting with tactile patterns to obtain a more robust estimation of textures [11], [12]. However, most of the prior work exploiting audio effects have been focused on physical touch (e.g., vibrations or force feedback). While a few attempts combine mid-air haptics and audio feedback [13], they do not focus on textures but on simple geometric shapes or floating widgets [14]. Furthermore, it is well known that combining audio and haptic effects can modulate the perception of one's own body [15], [16]. For example, altering the sound produced when stroking people's hand (e.g., making it sound as a hammer hitting marble), can make one's hand feel stiffer/heavier [17]. However, whether this effect can be replicated with mid-air haptics is unknown.

To fill these gaps, this paper explores the sonification of mid-air haptic textures (see Figure 1b-c) and how this combination can influence people's 1) explicit judgment of the texture attributes – roughness, hardness, viscosity, temperature, moisture, and strength, 2) explicit sensations of their own hand (i.e., body perception), and 3) implicit motor behavior while exploring the mid-air texture. To do so, we perform a user study with 25 participants by delivering a mid-air texture rendering (e.g., metal texture, water texture) with a congruent associated audio effect (e.g., rubbing metal, touching water) during a haptic texture exploration task.

We found that the overall multisensory feedback (haptic, and/or audio) influenced participants' ratings of the texture attributes (only for roughness, hardness, moisture and viscosity), hand sensations (the feeling of having a hand smoother, softer, looser, more flexible, colder, wetter and more natural), and hand speed (moving faster or slower) while exploring the mid-air texture. We then conducted a principal component analysis (PCA) to better understand and visualize our results and multivariate associations between the conditions tested.

Our findings can contribute towards experience design strategies to improve mid-air haptic perception (particularly for texture judgments), using multisensory interactions. We argue that a better understanding of the capabilities and limits of human perception can provide simpler and more effective interaction techniques. For example, we can avoid complex algorithms to create precise haptic texture renderings (e.g., a real-time tactile rendering of fluids that creates sensations of viscosity [10]), but just use a congruent cross-modal association meaning given by an extra sensory modality (e.g., water sound) that will not only enhance the perception of a texture, but also influence body sensations thus amplifying the level of immersion and body ownership experienced.

II. LITERATURE REVIEW

A. A multisensory approach to mid-air haptic rendering

Focal point modulation is typically used to aid the perception of tactile sensations which is achieved through lateral, spatiotemporal, and amplitude modulation techniques [18]. Using this approach, perceptual studies provide evidence of relevant parameters to improve mid-air haptic perception. For example, determined distances between focal points to aid in discrimination tests [3], the duration of focal point presentation to produce a perception of apparent motion [19] or actual motion [20], stimulation duration and delay between subsequent points to create the perception of continuity [21], or modulation of focal point movement speed [22] and lateral modulation parameters [23] to affect perceived intensity. Other studies show perceptions of points and lines, as well as feelings of ‘bumps’ and ‘holes’ [24] and levels of ‘stiffness’ for virtual materials [8] and falling raindrops in VR [25].

While these approaches show a variety of perceptual properties informing practitioners on how to improve tactile perception, studies also acknowledge that “mid-air haptic shapes do not appear to be easily identified” due to the lack of other sensory modalities [26]. Most of these studies explore tactile perception with mid-air haptics only, without the intervention of other sensory modalities. However, haptic interactions in real life are usually accompanied by other sensory modalities [27] (e.g., visual, auditory, or olfactory).

Recent studies show that a multisensory approach can help improve mid-air haptic exploration [28]. For example, Wilson et al., [19] found that visual feedback can improve the ability to localize a focal point on the body. Hoggan and Brewster combined auditory and tactile parameters to convey icons information more effectively for mobile devices [29]. There are also studies suggesting that ultrasound-mediated touch influence other senses. For example, Ablart et al. [30] found that ultrasound tactile patterns can alter the perception of the audio and visual stimuli. Despite these efforts, the combination of multisensory cues for mid-air haptics is still scarce.

B. Combination of mid-air haptics and audio

Few studies have combined mid-air haptics with audio feedback. For example, Freeman et al. [13] added audio effects (white noise and tones) to evaluate the perceived roughness of an ultrasound haptic pattern. They found that white noise

increased perceived roughness, while pure tones did not. Ozkul et al. [14] combined auditory and mid-air haptic feedback for a holographic light-switch button and concluded that sensory combinations led to changes in the emotional responses. Maggioni et al. [31] combined audio-visual stimuli with mid-air haptic feedback and showed that the user experience resulted more pleasant, unpredictable, and creative. Thanh Vi et al. [32] presented a case study on multisensory experiences involving vision, sound, touch, smell, and taste to enhance the user experience of visual art. Their results highlight a positive effect on immersion and user experience, mainly obtained from the combination of mid-air haptic and audio. These are only a few examples of how sound can influence different mid-air haptic attributes, and that combining mid-air haptics and sound can have emotional reactions influencing perceptual qualities.

In summary, from this previous research we note that most studies have focused on understanding the perception of mid-air haptic attributes such as the sensations of motion, perception of direction, and strength. However, only a few studies have focused on texture rendering methods [5], and some of these methods have not been evaluated yet, hence, their efficacy is currently unknown [6]. With respect to multimodal feedback, while a few studies have combined mid-air haptics and audio feedback, they focused on simple geometric shapes and not actual textures. Additionally, they used neutral sounds (white noise, tones), which can be hardly associated with other sensory attributes. No study has explored the effect that the sonification of a mid-air haptic texture might produce on the perception of one’s own body sensations and behavior.

C. The effect of sound on texture and body perception

Going beyond haptic texture perception, it is known that crossmodal correspondences resulting from interactions between one’s body motion/touch and sound can produce perceived body alterations. For example, people tend to associate auditory stimuli (e.g., audio pitch, sound frequency) with haptic attributes related to one’s own body (e.g., weight, moisture, and texture), significantly influencing the actual body perception and behavior. For instance, changes in the sounds produced when rubbing one’s hands together can alter the perceived skin moisture and dryness [33], [34].

Changing the sound that one’s hand produces when tapping a surface (making it sound at a distance) results in an overestimation of the arm length [16], [35]. Modulating the volume of the sound that one’s hand produces when tapping a surface can change the perceived arm strength [36], which in turn leads to changes in motor behavior (e.g., tapping stronger or faster).

Altering the sound people produce while walking can change the perception of their own body weight [37], leading to the perception of having a lighter body [38], which in turn can affect the walking behavior (e.g., walking faster and with a straighter posture [15]).

A summary of the effects of non-veridical auditory cues on movement-related touch in relation to surface texture perception, and motor behavior can be found in Tajadura-Jiménez, et al. [39] and Stanton et al [40], respectively.

These studies suggest promising opportunities for audio-tactile integration to modify the perception of our own body

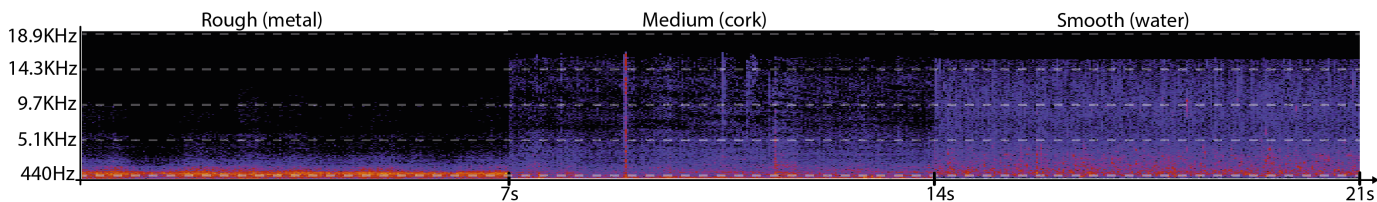


Fig. 2: Spectrogram for each of the sounds used in this work. We used the rubbing sound of metal to represent the rough texture, the rubbing sound of wood for the medium texture, and the rubbing sound of water for the smooth texture.

and leverage such effects to design particular multisensory experiences in the metaverse and beyond [41]. For example, combining a rough mid-air texture with a rough sound to produce a sensation of having dryer skin. However, this vision is hindered by the current lack of research on the sonification of mid-air haptics and the unexplored transformative potential of mid-air haptics to create embodied experiences.

III. USER STUDY – EFFECT OF SONIFICATION ON TEXTURE JUDGMENTS, BODY SENSATIONS, AND MOTOR BEHAVIOR

Motivated by the above discussion, in this study we focus on the roughness of mid-air haptic textures following evidence suggesting that the perception of roughness is altered by touch-produced sounds [42]. We therefore rendered textures with 3 levels of roughness – rough (metal roof tiles), medium (cork), and smooth (water). In order to create different haptic attributes, we then combined the haptic textures with congruent sound feedback – rubbing metal, rubbing wood, and touching water respectively, in a haptic exploration task (i.e., sonification of the texture exploration).

We only tested congruent sensory conditions as we explored audio-tactile correspondences, assuming they will yield stronger association than incongruent sensory conditions [43].

We chose our textures and sounds to emphasize audio-tactile associations that are common in our everyday environment. For instance, we usually associate a rusty metal sound with something that is rough, hard, and perhaps cold. Similarly, we more likely associate a water sound with something that is smooth, wet, and perhaps slippery. Unlike the work by Freeman et al., [13] in which the sounds used were neutral white noise and tones (hardly associated with common properties in the environment), we wanted to represent a more natural correspondence with common sensory experiences. That is, with the chosen sounds we aim not only to resemble feelings in a tactile way (e.g., soft, smooth), but also feelings that could be associated with multisensory real-life experiences, in order to enable different body sensations (e.g., heavy, big, natural). This is in line with prior work suggesting that people can associate haptic patterns (e.g., rocks, clouds) and sounds (e.g., hitting marble) with different body sensations (e.g., being stronger, being lighter, having dryer skin) [17], [34], [44].

With this approach, we included “water” texture/sound expecting to enable broader body sensations and because recent studies have made arduous efforts to render fluids textures to give sensations of viscosity with mid-air haptics (e.g., [10]), and we wanted to explore whether a simpler solution, such as adding a congruent sound (in contrast to complex rendering

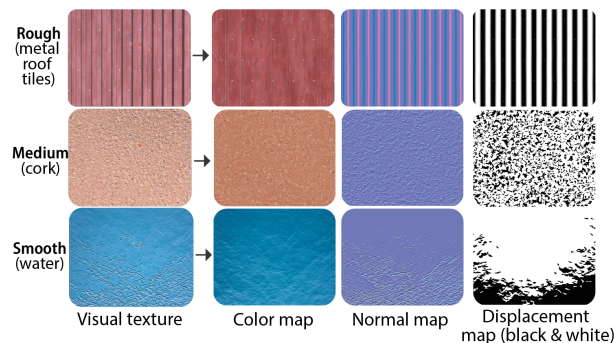


Fig. 3: Textures used for high (metal roof tiles), medium (cork) and low (water) roughness association.

algorithms and GPU processing) can produce feelings on viscosity effectively. To explore whether multisensory feedback helps participants to identify differences between textures, we compared 3 roughness levels - rough, medium and smooth within 3 sensory conditions: haptic only, audio only, and hybrid (haptic + audio). We aim to explore how the combination of mid-air haptic and auditory stimuli influence people’s perception of different roughness levels. Within our user study, we tried to answer three particular research questions:

Research question 1 (RQ1): Does the sonification of texture exploration influence participants’ judgments of a mid-air texture’s attributes – roughness, hardness, viscosity, temperature, moisture, and strength? We chose these 6 attributes following evidence on cross-modal correspondences between sound and roughness [13], hardness [45], viscosity [46], temperature [47], moisture [34], and strength [36].

Research question 2 (RQ2): Does the sonification of texture exploration influence participants’ subjective sensations of their own hand? Following the work by Tajadura et al., on altering one’s body-perception using haptic metaphors [44], we assessed 13 body sensations – speed, weight, strength, naturality, flexibility, hardness, sensitivity, size, tension, viscosity, temperature, moisture, and roughness. In other words, we explored whether some of the perceived haptic attributes of the texture are transferred to participants’ hand perception.

Research question 3 (RQ3): Does the sonification of texture exploration influence the implicit motor behavior of participants’ hand? Following the work by Tajadura et al. on influencing the motor behavior when sonifying motor actions [36], we explored hand speed changes while exploring the texture, and whether this effect remains over time.

A. Stimuli presentation

1) *Audio feedback*: It was presented via noise-canceling headphones (Sennheiser 400s) during the study. It consisted of three different sound effects – rubbing metal for a high roughness association, rubbing cork for a medium roughness association, and touching water for a smooth association.

To verify whether the sounds were properly associated to the appropriate level of roughness (rough, medium, and smooth), we conducted a pilot study with five participants within our research team. After listening to 15 sound effects, they ranked their level of roughness, resulting in the three chosen sounds for the user study. To visualize the differences in the sound waveforms chosen, we use spectrograms as seen in Figure 2. We note that the sound defined as rough, has stronger lower frequencies while the sound defined as smooth is consistently reaching higher frequencies, and the medium sound is located somewhere in between these, with a higher temporal variability. This is not to say that all rough/smooth sounds share the same spectrogram characteristics as with our small sample set. All sounds used for the pilot and main study were recorded by the authors from the environment using different materials and objects (metal, wood, marble, glass, water, etc.) and through a Zoom Am7 stereo microphone.

2) *Haptic textures rendering*: We delivered the mid-air haptic textures by using an Ultraleap STRATOS Explore Development Kit hardware platform (256-transducer array board, control board, and frame structure) which operates at 40 kHz, and an Ultraleap stereo IR 170 camera to track participants' hands. We designed a GUI for the user study sequence operated in an Alienware laptop with 16 Gb of RAM and NVIDIA RTX 2070. To render the mid-air haptic textures, we adopted the method proposed by Beattie et al. [48] which consists of an algorithm that maps the visual attributes of a texture into mid-air haptic patterns. That is, this method uses visual cues of any 2D graphical image, such as the spatial distribution of surface elements, and replicates those cues in the form of mid-air haptic attributes, thus forming a representation of how a texture should feel.

Based on this method, we first selected the visual images representing three levels of roughness. Two users explored a set of textures to get three samples that were clearly differentiable from each other in terms of visual roughness and that matched the selected sounds from the pilot study; rough (metal roof tiles), medium (cork), and smooth (water).

We used free-source images (see Figure 3) for our textures, and we then obtained their mid-air haptic representation by following the steps below:

1) *Generate Displacement and Normal Maps*: This step was done offline using a normal map generation tool [50], using a default set of parameters [contrast, strength, level, displacement, filter] as [-0.5, 2.5, 0.7, 0.3, Sobel] to generate each texture normal map and its corresponding displacement map (see Figure 3).

2) *Micro-roughness*: We then extracted the microscale roughness from the displacement map by the gradient of the power spectral density function, as described by Beattie et al. [48]. This gives us information about small texture changes over time. A larger gradient implies that changes are close to each other (high frequencies) producing smoother textures.

3) *Macro-roughness*: We then directly obtained the macroscale roughness from the displacement map values (black and white form shown in Figure 3).

4) *Haptic Synthesis*: We use a look-up table of texture roughness based on the precomputed values from the previous steps, to convert micro-roughness to Ultraleap rotation speed and waveform sampling parameters.

We used the middle finger position over the texture to convert macro-roughness of the texture points to haptic device rotation speed and focal point intensity parameters (see Figure 4b). Each texture point was rendered using a circle stimulus of 2 cm radius at a constant intensity over the circle. The intensity was set to maximum when hovering a white pixel and to minimum when hovering a black pixel.

The multisensory (haptic + audio) texture rendering workflow is shown in Figure 4a. The first stage is the extraction of local texture features (haptic mapping function block). At this stage, the system uses the current middle finger position (the intermediate-distal phalange joint) to extract the texture features of the point underneath in the form of texture positions and haptic intensities. The second stage is the sound mapping function block. Here, the output from the first stage is used to generate an intensity-to-volume dynamic mapping. Note that the intensity is a normalized value and thus the audio volume control variable receives a normalized value which enables us to directly modulate the sound volume using the image-extracted intensity values from stage one. This means that when the user feels a strong haptic sensation from the texture the sound will also be louder, on the contrary, the sound will be lower when exploring sections of the texture with lower haptic

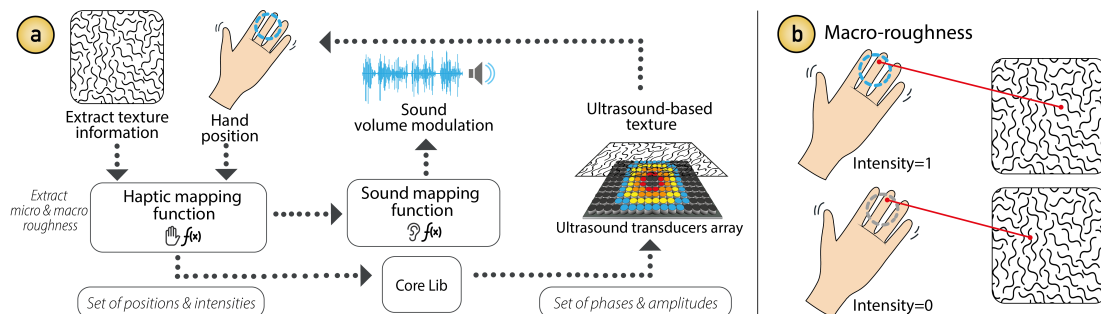


Fig. 4: (a) Multisensory (haptic + audio) texture rendering workflow. (b) Macro roughness mapping example.

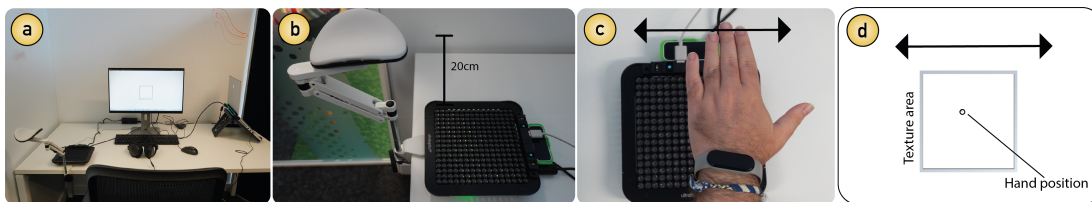


Fig. 5: Experimental setup. Participants sat in front of a computer screen (a) and rested their forearm on a support, allowing a distance of ~ 20 cm between the array of transducers and their hand (b). The texture exploration consisted of a rhythmic movement of participants' hand from side to side above the array (c). The GUI used just showed an empty square and a pointer to indicate the participants' hand position (d).

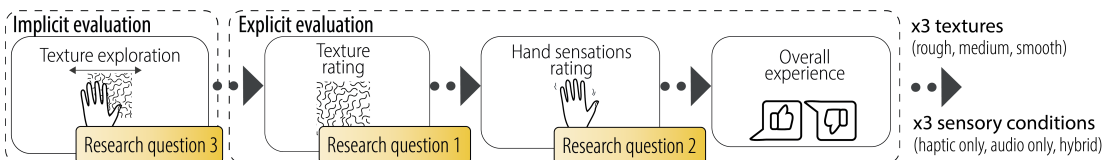


Fig. 6: The implicit evaluation consisted of recording the participants' hand behavior (hand speed) during the texture exploration task. The explicit evaluation consisted of subjective Likert scales that participants used to rate the texture attributes (research question 1), their body sensations (research question 2), and overall experience.

intensities. The third and final stage is done on the device core libraries, in which the set of positions and intensities are translated to haptic focal point phases and amplitudes for the device to operate and render the texture on demand.

Since we focused audio-haptic feedback, we reduced visual feedback as much as possible. As shown in Figure 5d, in the GUI used during the study, no actual visual texture was shown but just an empty square indicating the area where the interactive texture exploration was located. No virtual hands were used to avoid facilitating any embodied experience, thus we only used a pointer indicating the participants' hand position. Participants could see a small circle pointer with the x-y coordinates corresponding to the location of their middle finger. Neither the haptic nor the audio stimuli were presented when the participant's hand pointer was outside the square. The mid-air texture was aligned 20 cm directly above the mid-air haptic device for optimal device performance and occupied an area of approximately 15x15 cm.

B. Procedure

Participants sat on an adjustable chair in front of a computer screen with a mouse and a keyboard (see Figure 5a). They were allowed to rest their hand on a movable armrest in order to avoid joint fatigue and keep a constant distance between the array of transducers and their hand (at ~ 20 cm) as shown in 5b. Then, participants were asked to explore the three different mid-air textures (rough, medium, smooth) by moving their hand rhythmically from side to side (see Figure 5c). Participants were carefully instructed to move their hand at a speed they considered as suitable to clearly explore the mid-air texture, but while maintaining the same rhythm in their movements as much as possible in order to facilitate detecting a change in motor behaviour. We followed a similar procedure for the sonification of motor tasks as in [36].

Figure 6 shows the procedure of the study, separating the implicit evaluation (obtained through the Leap Motion

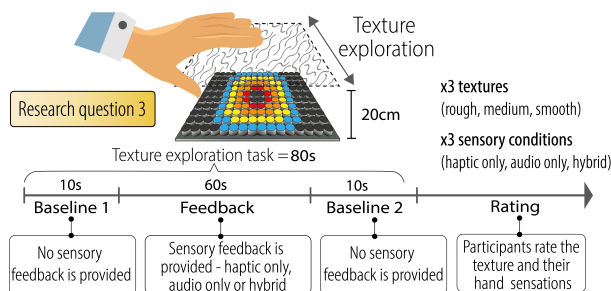


Fig. 7: The implicit evaluation consisted of the texture exploration task divided into 3 blocks: Baseline 1 and Baseline 2 for 10 seconds each (no feedback was provided), and a Feedback block for 60 seconds (sensory feedback was provided).

readings during the texture exploration task) from the explicit evaluation (obtained through texture judgement and body sensations questionnaires after the texture exploration task). At the end of the study, we collected qualitative data systematically as part of the experimental procedure. Participants were asked to provide feedback about their overall experience by writing down any thought they wished. We followed a within-subjects design, with all participants exposed to all conditions presented in blocks (one for each sensory modality) in a counterbalanced order. In particular, each participant completed nine texture explorations – 3 sensory conditions (haptic only, audio only, or hybrid), 3 texture roughness levels (rough, medium, or smooth) resulting in 9 blocks in total. The experiment lasted an average time of 33 minutes including instructions and training.

C. Explicit evaluation

To explore **RQ1**, after each texture exploration task, we asked a series of questions aimed to explore participants' judgment of the felt texture. Participants selected a score that best described the texture attributes they had just perceived during the texture exploration task using a 7-point Likert scale.

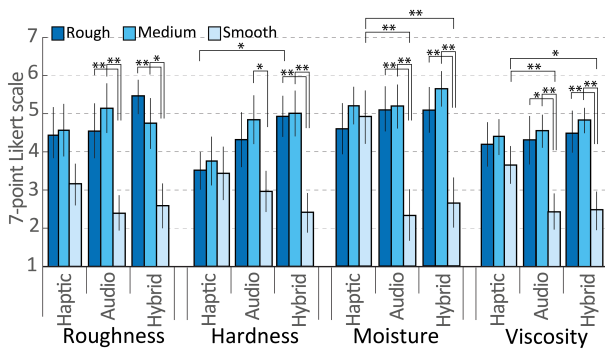


Fig. 8: Mean scores from the texture judgment questionnaire for each sensory condition and roughness levels. *Asterisks indicate significant mean differences between conditions ($p < 0.05$, $** = p < 0.01$). Error bars represent 95% CI.

Haptic attribute	Friedman test	p-value	W	Wilcoxon signed-rank tests	p-value	r	Roughness level
Hardness	$\chi^2 = 6.23$	$p < 0.05$	0.125	haptic vs hybrid	$p < 0.001$	-0.64	rough
				haptic vs audio	$p < 0.001$	-0.752	smooth
Moisture	$\chi^2 = 22.92$	$p < 0.001$	0.458	haptic vs hybrid	$p < 0.001$	-0.721	smooth
				haptic vs audio	$p < 0.005$	-0.6116	smooth
Viscosity	$\chi^2 = 14.64$	$p < 0.001$	0.293	haptic vs hybrid	$p < 0.001$	-0.663	smooth
				haptic vs audio	$p < 0.001$	-0.663	smooth

Haptic attribute	Friedman test	p-value	W	Wilcoxon signed-rank tests	p-value	r	Sensory condition
Roughness	$\chi^2 = 27.36$	$p < 0.001$	0.547	smooth vs rough	$p < 0.001$	-0.736	audio
				smooth vs medium	$p < 0.001$	-0.780	audio
				smooth vs rough	$p < 0.001$	-0.807	hybrid
Hardness	$\chi^2 = 25.26$	$p < 0.001$	0.505	smooth vs medium	$p < 0.001$	-1.103	hybrid
				smooth vs rough	$p < 0.001$	-1.103	hybrid
				smooth vs medium	$p < 0.005$	-0.617	audio
Moisture	$\chi^2 = 14.14$	$p < 0.001$	0.283	smooth vs rough	$p < 0.001$	-0.740	audio
				smooth vs medium	$p < 0.001$	-0.805	hybrid
				smooth vs rough	$p < 0.001$	-0.751	audio
Viscosity	$\chi^2 = 27.54$	$p < 0.001$	0.551	smooth vs medium	$p < 0.001$	-0.731	audio
				smooth vs rough	$p < 0.001$	-0.735	hybrid
				smooth vs medium	$p < 0.001$	-0.768	hybrid
Roughness	$\chi^2 = 23.68$	$p < 0.001$	0.474	smooth vs rough	$p < 0.001$	-0.622	audio
				smooth vs medium	$p < 0.001$	-0.684	audio
				smooth vs rough	$p < 0.001$	-0.746	hybrid
Hardness	$\chi^2 = 31.85$	$p < 0.001$	0.637	smooth vs rough	$p < 0.005$	-0.622	audio
				smooth vs medium	$p < 0.001$	-0.684	audio
				smooth vs rough	$p < 0.001$	-0.746	hybrid
Moisture	$\chi^2 = 18.60$	$p < 0.001$	0.372	smooth vs rough	$p < 0.001$	-0.746	audio
				smooth vs medium	$p < 0.001$	-0.684	audio
				smooth vs rough	$p < 0.001$	-0.746	hybrid
Viscosity	$\chi^2 = 24.73$	$p < 0.001$	0.495	smooth vs rough	$p < 0.001$	-0.800	hybrid
				smooth vs medium	$p < 0.001$	-0.800	hybrid
				smooth vs rough	$p < 0.001$	-0.800	hybrid

Fig. 9: Results of the Friedman and Wilcoxon signed-rank tests from the texture judgment questionnaire, for the comparison between sensory conditions - haptic only, audio only, and hybrid (top table), and between roughness levels - rough, medium, and smooth (bottom table).

The questionnaire comprised of the sentence “The texture was:” accompanied by six items related to the stimuli attributes, which ranged from: “smooth” to “rough” (Roughness); “soft” to “hard” (Hardness); “weak” to “strong” (Strength); “cold” to “warm” (Temperature); “wet” to “dry” (Moisture); “slippery” to “sticky” (Viscosity).

To explore **RQ2**, we used a second questionnaire aimed to explore participants’ hand sensations. Participants selected a score that best expresses their subjective hand sensations felt during the texture exploration task using a 7-point Likert scale. This questionnaire was adapted from previous studies [38], [44], by leaving the items associated to haptic attributes and adding new items to explore any possible bodily sensations elicited by mid-air haptics and the sounds used. The questionnaire comprised of the sentence “I felt my hand:” accompanied by 13 items related to the hand’s sensation, which ranged from: “smooth” to “rough” (Roughness); “soft” to “hard” (Hardness); “weak” to “strong” (Strength); “cold” to “warm” (Temperature); “wet” to “dry” (Moisture); “slippery” to “sticky” (Viscosity); “slow” to “quick” (Speed); “light” to “heavy” (Weight); “natural (as usual)” to “unnatural” (Nat-

urality); “stiff” to “flexible” (Flexibility) “small” to “large” (Size); “loose” to “tense” (Tension). This questionnaire aimed to explore whether some of the haptic attributes of the textures were transferred to the user’s hand.

D. Implicit evaluation

To explore **RQ3**, we recorded participants’ hand speed while exploring the textures. Overall, we expect that participants’ speed will be slower for rougher textures and faster for smoother textures. However, to explore any implicit motor behavior change, we divided the texture exploration task (lasted 80 seconds in total), into 3 blocks (shown in Figure 7): Baseline 1, 10 seconds (no feedback was provided), Feedback block, 60 seconds (participants received real-time sensory feedback - haptic only, audio only, or hybrid in response to their hand movements), and Baseline 2, 10 seconds (no feedback was provided). We recorded the middle finger velocity and position in all spatial dimensions during the full duration of the exploration task (80 seconds) for all the different sensory and texture roughness conditions. Questionnaires were provided after the whole trial finished (i.e., 80 seconds). With these 3 blocks, we aim to explore (1) whether participants change their speed while exploring the texture by comparing Baseline 1 and Feedback blocks, and (2) whether this effect (if any) remains after the stimulation stops by comparing Feedback and Baseline 2 blocks. We follow similar experimental conditions used in previous work to explore motor behavior changes due to audio-haptic combination [36].

E. Participants

25 participants we recruited (2 left-handed, 12 females, mean age = 30.8 years, SD = 4.4 years, range = 22 – 38 years). They gave written consent for their participation and had no injuries to their hands, sense of touch, or sense of hearing. The local ethics committee approved the study.

IV. RESULTS

All data collected were analyzed using SPSS version 29.0.1.0. We first explored the distribution of our dependent variables via a Shapiro-Wilks test that indicated a likely non-normal distribution of our data (all tests showed $p < 0.001$).

We then proceeded by applying a Friedman test for each of the six factors recorded from participants (grouping the data by sensory condition – haptic only, audio only, and hybrid and texture roughness level – rough, medium, smooth), followed by a Wilcoxon signed-rank test with Bonferroni correction to assess differences between groups (i.e., roughness levels, sensory conditions, and experimental blocks of the implicit measures). For the Friedman test, we report Kendall’s “W” as an indicator of effect size, which is used for assessing agreement among raters and in particular inter-rater reliability [49]. Then, for the Wilcoxon signed-rank, we report Pearson’s r as an indicator of effect size, the closer the value is to 0, the smaller the effect size. A value closer to -1 or 1 indicates a higher effect size.

Qualitative data from participants were analysed using an inductive approach in which the data (participants’ comments) determined our conclusion about their overall experience.

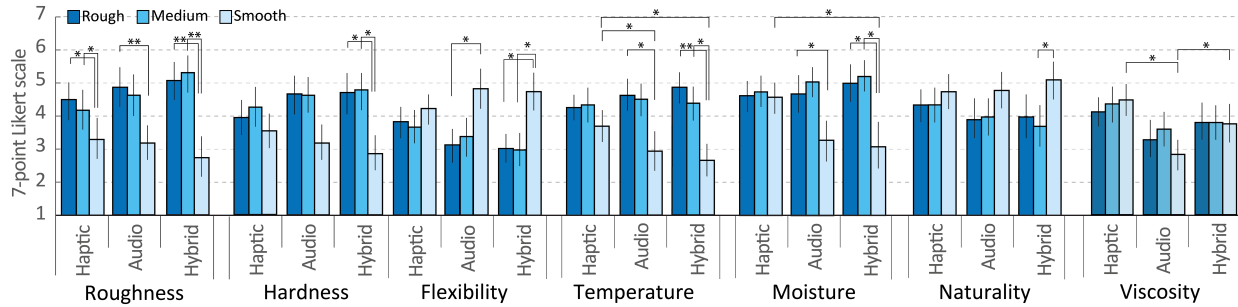


Fig. 10: Mean scores obtained from the body sensations' questionnaire for each sensory condition and roughness levels. Asterisks indicate significant mean differences between conditions ($*=p < 0.05$, $**= p < 0.01$). Error bars represent 95% CI.

Haptic attribute	Friedman test	p-value	W	Wilcoxon signed-rank tests	p-value	r	Roughness level
Viscosity	$\chi^2 = 14.58$	$p < 0.001$	0.292	haptic vs audio	$p < 0.001$	-0.671	smooth
				audio vs hybrid	$p < 0.05$	-0.514	
Temperature	$\chi^2 = 14.39$	$p < 0.001$	0.288	haptic vs audio	$p < 0.05$	-0.485	smooth
				haptic vs hybrid	$p < 0.05$	-0.622	
Moisture	$\chi^2 = 9.15$	$p < 0.05$	0.183	haptic vs hybrid	$p < 0.05$	-0.561	smooth

Haptic attribute	Friedman test	p-value	W	Wilcoxon signed-rank tests	p-value	r	Sensory condition
Roughness	$\chi^2 = 8.208$	$p < 0.05$	0.164	smooth vs rough	$p < 0.05$	-0.433	haptic
				smooth vs medium	$p < 0.05$	-0.506	
	$\chi^2 = 15.51$	$p < 0.001$	0.310	smooth vs rough	$p < 0.05$	-0.599	audio
				smooth vs medium	$p < 0.05$	-0.518	
	$\chi^2 = 27.16$	$p < 0.001$	0.543	smooth vs rough	$p < 0.001$	-0.801	hybrid
				smooth vs medium	$p < 0.001$	-0.779	
Hardness	$\chi^2 = 20.44$	$p < 0.001$	0.409	smooth vs rough	$p < 0.001$	-0.650	hybrid
				smooth vs medium	$p < 0.001$	-0.713	
Temperature	$\chi^2 = 10.48$	$p < 0.05$	0.201	smooth vs rough	$p < 0.005$	-0.581	audio
				smooth vs medium	$p < 0.001$	-0.762	
Flexibility	$\chi^2 = 25.86$	$p < 0.001$	0.517	smooth vs rough	$p < 0.001$	-0.723	hybrid
				smooth vs medium	$p < 0.001$	-0.674	
Moisture	$\chi^2 = 12.03$	$p < 0.05$	0.241	smooth vs rough	$p < 0.005$	-0.599	audio
				smooth vs medium	$p < 0.005$	-0.728	
Naturality	$\chi^2 = 21.88$	$p < 0.001$	0.438	smooth vs rough	$p < 0.001$	-0.647	hybrid
				smooth vs medium	$p < 0.001$	-0.637	
Viscosity	$\chi^2 = 17.81$	$p < 0.001$	0.365	smooth vs rough	$p < 0.001$	-0.689	audio
				smooth vs medium	$p < 0.001$	-0.704	
Temperature	$\chi^2 = 13.89$	$p < 0.001$	0.278	smooth vs medium	$p < 0.001$	-0.6408	hybrid
				smooth vs rough	$p < 0.001$	-0.6408	

Fig. 11: Results of the Friedman and Wilcoxon signed-rank tests from the body sensations questionnaire for the comparison between roughness levels - rough, medium, and smooth in each sensory condition.

A. Results (explicit) – Texture judgments and body sensations

1) *Texture judgements*: Figure 8 shows the mean scores from participants for the haptic attributes we found significant results (**roughness, hardness, moisture, and viscosity**). Figure 9 shows the results from Friedman and Wilcoxon tests.

In summary, when comparing the *sensory conditions* (haptic only, audio only, and hybrid), we found a statistically significant difference between the haptic only and audio only conditions, and between haptic only and hybrid conditions for the haptic attributes of **hardness, moisture, and viscosity**. However, no significant differences were found between audio only and hybrid conditions.

Furthermore, when comparing the *roughness levels* directly (rough, medium, and smooth) within each particular sensory condition we found that, for **roughness, hardness, moisture, and viscosity**, participants perceived the smooth texture as significantly **smoother, softer, wetter, and more slippery** compared with the rough and medium levels when they were exposed to both conditions audio only and hybrid. No effects were observed for the haptic only condition.

2) *Body sensations*: Figure 10 shows the mean scores from participants for the haptic attributes we found signifi-

cant results (**roughness, hardness, flexibility, temperature, moisture, and neutrality**). Figure 11 shows the results from Friedman and Wilcoxon tests.

In summary, when comparing the *sensory conditions* (haptic only, audio only, and hybrid), we found a statistically significant difference between the haptic only and audio only conditions, and between haptic only and hybrid conditions when participants explored the smooth texture for the haptic attributes of **viscosity, temperature, and moisture**. However, no significant differences were found between audio only and hybrid conditions.

Furthermore, when comparing the *roughness levels* directly (rough, medium, and smooth) in each particular sensory combination, we found that: for **roughness, hardness, temperature, and naturality**, participants perceived having their hand as significantly **smoother, softer, more flexible, colder, and wetter** when exploring the smooth texture compared with the rough and medium textures, when they were exposed to the hybrid condition. Moreover, participants also perceived having their hand as significantly **more flexible, and wetter** when exploring the smooth texture compared with the rough texture, when they were exposed to the audio only condition. No effects were observed for the haptic only condition.

B. Results (implicit) – Motor behavior

We carried out a series of Friedman tests with 2 DoF to explore differences between experimental blocks - baseline 1 (B1), feedback, and baseline 2 (B2) in the different sensory conditions (haptic only, audio only and hybrid) and roughness levels (rough, medium, smooth). Hence, we performed three one-factor Friedman tests – one for each sensory condition.

The participants were instructed to explore the textures by moving their hands along the x-axis (which is the appropriate exploration procedure to gain information about textures [50]). Hence, we focused on the Velocity along the x-axis obtained directly from the LeapMotion SDK.

Figure 12 shows mean velocities recorded from the experimental blocks (B1, feedback, and B2) for each sensory condition and roughness levels. Figure 13 shows the results from Friedman and Wilcoxon tests.

In summary, when comparing the *sensory conditions* (haptic only, audio only, and hybrid), we found no significant differences. However, when comparing the roughness levels directly (rough, medium, and smooth) within each particular sensory

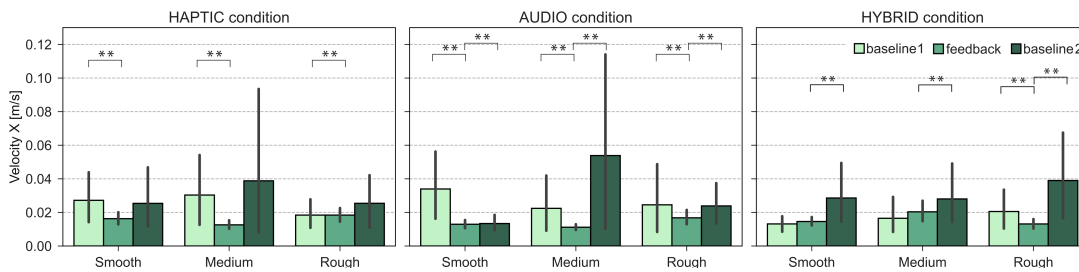


Fig. 12: Velocity change between experimental phases measured along the x-axis. These data were obtained by the LeapMotion SDK. Asterisks indicate significant differences (* = $p < 0.05$, ** = $p < 0.01$). Error bars represent 95% CI.

Sensory condition	Friedman test	p-value	W	Roughness level	Wilcoxon signed-rank tests	p-value	CLES
Haptic	$\chi^2(2) = 17.36$	$p < 0.001$	0.35	smooth	B1 vs feedback	$p < 0.01$	0.76
	$\chi^2(2) = 7.28$	$p = 0.03$	0.15	medium	B1 vs feedback	$p < 0.01$	0.71
	$\chi^2(2) = 11.12$	$p < 0.001$	0.22	rough	B1 vs feedback	$p < 0.01$	0.67
Audio	$\chi^2(2) = 17.36$	$p < 0.001$	0.35	smooth	B1 vs feedback	$p < 0.01$	0.78
					feedback vs B2	$p < 0.01$	0.78
	$\chi^2(2) = 10.32$	$p = 0.03$	0.21	medium	B1 vs feedback	$p < 0.01$	0.77
					feedback vs B2	$p < 0.01$	0.75
	$\chi^2(2) = 23.28$	$p < 0.001$	0.46	rough	B1 vs feedback	$p < 0.01$	0.74
					feedback vs B2	$p < 0.01$	0.82
Hybrid	$\chi^2(2) = 14.48$	$p < 0.001$	0.29	smooth	feedback vs B2	$p < 0.01$	0.78
	$\chi^2(2) = 21.84$	$p = 0.03$	0.44	medium	feedback vs B2	$p < 0.01$	0.78
	$\chi^2(2) = 15.68$	$p < 0.001$	0.31	rough	feedback vs B2	$p < 0.01$	0.78
					B1 vs feedback	$p < 0.01$	0.71

Fig. 13: Results of the Friedman and Wilcoxon signed-rank tests for the comparison between experimental blocks - B1, feedback, and B2 in each sensory condition.

condition, we found that in both conditions **haptic** and **audio** results showed differences for all roughness levels between B1 and feedback blocks, suggesting that participants' hand moved significantly slower while exploring the textures. However, in the audio condition we also found differences for all roughness levels between feedback and B2 blocks suggesting that they increased their speed again when the audio stimulation ended suggesting that such movement change did not remain.

For the **hybrid** condition, results showed differences for all roughness levels between feedback and B2 blocks, suggesting that participants' hand did not change its speed when receiving the hybrid stimulation. However, for the rough texture we found differences between B1 and feedback blocks as well. This suggests that when the texture was rougher participants' hand decreased their speed when receiving the hybrid stimulation and increased when such stimulation ended.

C. Discussion – Influence of sonification on texture judgments, body perception and motor behavior

Texture judgments: When comparing the different roughness levels in each particular sensory combination, results showed that audio feedback associated to the roughness of a rendered mid-air texture (e.g., water texture + touching water sound) influenced participants' ratings of the texture attributes depending on the three roughness levels (rough, medium, smooth), only for the haptic attributes of **roughness**, **hardness**, **moisture**, and **viscosity**.

Associations between audio and tactile cues were observed even without tactile stimulation (i.e., audio only conditions). In some cases, audio stimuli were enough to convey information such as texture roughness, wetness, and slipperiness. This

suggests that sound is quickly associated with haptic features that are generally attributed to texture. Audio-tactile associations were also observed in the hybrid condition (i.e., audio + haptic), whereby participants could judge haptic features such as hardness, even though our haptic rendering algorithm did not aim to render such features.

As expected, when participants received the haptic feedback only, their ratings of the texture attributes were quite similar for all the haptic attributes tested. This confirms that textural qualities are challenging to render convincingly when relying on the sense of touch only [5].

When comparing the sensory conditions directly, we found few differences between the haptic only and hybrid conditions. However, we did not find significant differences between the audio only and hybrid conditions. This suggests that in both conditions, audio only and hybrid, participants reported changes in texture attributes based on the different roughness levels. Yet, participants reported a preference for the hybrid condition when asked for feedback at the end of the study.

P1: “Having a hybrid condition actually helps in differentiating textures, and it comes with a complementary set of attributes to make the experience more complete”.

P3: “Haptics without sound most of the time felt like a different kind of air/windy intensity and trying to link it to a texture was really difficult and was a pain point.”

P13: “Sound without haptics didn't feel natural, I could guess what the sound was, but did not feel a texture. It wasn't a complete experience”.

These results address **RQ1**, suggesting that sonification of mid-air haptic textures may influence the way participants judge them. However, we cannot demonstrate such value when comparing hybrid feedback (haptics + audio) with the audio only feedback. These findings still demonstrate the value of audio to elicit explicit experiences, as suggested in the literature [37], [44], and show how such potential can be leveraged to aid in mid-air texture perception.

Body Sensations: When comparing the different roughness levels in each particular sensory combination, we found that audio only conditions were, in most cases, not sufficient for participants to establish audio-tactile associations that had influenced their own body perception.

In the hybrid condition, however, participants would transfer the haptic attributes of the texture explored to their own hand for most of the body sensations tested. That is, participants would report their hand to feel significantly different for body

sensations such as **roughness, hardness, flexibility, temperature, moisture** and **naturality** depending on the roughness level of the texture explored. This suggests that congruent audio feedback may facilitate embodiment and give rise to audio-tactile associations which extend to body sensations.

However, we did not observe such effect between sensory conditions and therefore it is unclear whether sonification influences body perception when comparing the hybrid (haptic + audio) and the single modalities (haptic only, audio only).

This could suggest that the addition of stimuli, rather than the combination of stimuli, might have produced the observed effect. Similar accounts have been discussed in the literature in which multiple sources of information help to disambiguate uncertainty compared to a single source of information [51]. We interpret these results as indicating that multisensory information might contribute to producing subjective feelings of having a hand with similar haptic attributes than a mid-air texture being explored (e.g., the sensation of having a softer and colder hand). However, we cannot demonstrate such value when comparing hybrid feedback (haptics + audio) with its single modalities, thus addressing **RQ2**.

Motor behavior: We found that for most texture roughness and sensory modality combinations, a significant difference was observed in terms of speed between B1 (no sensory feedback) and feedback blocks. This means that participants changed their motor behavior by reducing their speed when they received sensory feedback about the texture being explored. One might think that participants just reduced their speed consciously when they received sensory feedback, so that they could carefully explore the texture. However, we recall that participants were instructed to keep the same frequency/rhythm during the whole exploration task (80 seconds) i.e., during, before, and after sensory exposure, thus pointing towards an unconscious change in behavior in specific blocks of the exploration. Indeed, in the hybrid condition participants did not significantly decrease their speed while exploring the texture suggesting that there was smoother transition between not receiving feedback and feeling hybrid stimulation.

Earlier studies reported that participants stroking motion tended to be faster during exploring a smoother texture [52]. Therefore, one possible hypothesis is that hand speed in the hybrid condition would be large when exploring a smoother texture since the hybrid condition would contribute to a more realistic haptic experience.

Furthermore, since did not control how much time (from the 60 seconds given) participants took to judge the texture, we cannot exclude the possibility that they might have performed purposeless movements if they finish judging the texture before the time ran out. An earlier study reported that purposeless actions tend to produce slower movements [53], which may explain our the results on motor behaviour. Further research is needed to discard purposeless actions.

In summary, these results suggest that all feedback modalities (haptic only, audio only, and hybrid) produced a change in participants' motor behavior. We showed in previous sections that a sonified texture could significantly affect the perception of the surface viscosity. Thus, it should have followed that participants' motor system would have engaged more effort

to compensate for greater perceived viscosity and maintain constant exploration speed. We can therefore only hypothesize that this lack of change in motor behavior is explained by the intrinsic nature of mid-air haptics compared to contact-based haptic interaction. Mid-air interactions do not give rise to friction during hand movement in the first place, thus the motor control does not need to apply a mechanism to compensate for the perceived friction like in contact-based scenarios. Addressing **RQ3**, sonification of mid-air haptics does not seem to influence hand lateral motion. Future work would be required to determine whether it could influence other aspects of motor behavior or haptic exploration [50].

D. Further Analysis

The language surrounding haptic technology is difficult to define [54], despite there being well-defined taxonomies and perception dimensions, e.g., of textures [55]. For example, when we feel something, we get an immediate sense that we like or not like it, but often find it difficult to describe touch beyond that. While in previous sections we compared texture judgments and body sensations using singular descriptive attributes and metaphors (roughness, moisture, naturality, etc.), in this subsection we attempt to compare combinations of them through a principal component analysis (PCA).

PCA is often used for dimensionality reduction by projecting each data point onto only the first few principal components to obtain a lower-dimensional dataset while preserving as much of the data's variability as possible. We show in our *Supplementary Materials* that 80% of our original 19-dimensional data variability can be well described by just 2 dimensions (principal components PC1 and PC2) that are a linear combination of the original 19 attributes (6 texture attributes and 13 body sensations). Thus, we can visualize the 9 congruent conditions tested (3 sensory modalities \times 3 roughness levels) in the dimensionally reduced PCA space.

Figure 14.a) shows how the 19 different features contribute to PC1 and PC2. Weightings near zero are highlighted in grey. Indeed, we observe that features far from the axes are also the ones that were highlighted as significantly different in the previous subsections. Projecting the 9 conditions onto the dimensionally reduced principal component space by using a dot product $\mathbf{X} \cdot \mathbf{v}_i$ allows us to better visualize their differences and similarities, as shown in Figure 14.b).

We observe that PC1 sorts the stimuli conditions in terms of their smoothness (with medium/rough on the left, and smooth on the far right), while PC2 separates the modalities (with haptics-only on the top, audio-only at the bottom, and hybrid in the middle). This is quite remarkable since PCA has naturally managed to discriminate sensory conditions and also position the hybrid condition as a mix of the two unimodal conditions. Another observation from Figure 14.b) is that the hybrid and audio-only conditions form similar obtuse and quite elongated triangles in the PC1 dimension, while the haptic condition is closer to an equilateral. This picture is intuitively in agreement with results from the previous subsections that saw few differences between the haptic-only and hybrid conditions, and no difference between audio-only and hybrid conditions.

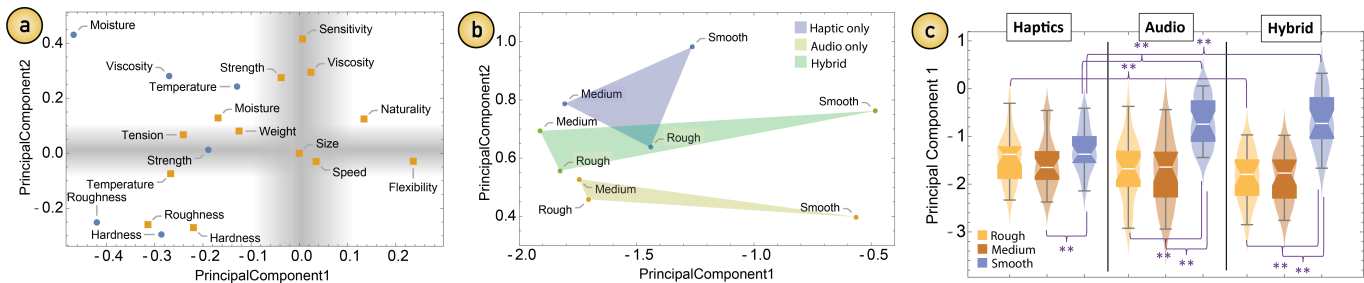


Fig. 14: a) Feature contribution for PC1 and PC2. Texture judgments are shown as blue dots, while body sensations as yellow squares. A grey cross highlights the region near the axes. b) Projection of the 9 conditions onto the 2D principal component space. Each sensory condition triplet is joined and shaded in a different color to form a triangle. c) Box & Violin plot of all 25 participants' data (6 texture attributes and 13 body sensations) projected onto PC1. *Asterisks indicate significant median differences between conditions ($** = p < 0.01$). Error bars represent 95% CI.

Finally, equipped with the PCA vectors, we can project all participant data onto the PCs and perform additional significance testing on linear combinations of attributes. Figure 14.c) shows exactly that for PC1, while also highlighting significant differences in the medians when comparing the roughness levels within and across sensory conditions. In almost all cases, we observe that participants perceived the smooth texture as significantly different from the medium and rough ones in terms of PC1 which we recall is dominated by attributes found on the left of Figure 14.a), i.e., **moisture** and **roughness**. Meanwhile, we observe that the audio-only conditions show no statistical difference from the respective hybrid conditions, in agreement with the results from the previous subsections. However, we also observe statistically different medians within the haptic only condition between the rough and medium conditions, something that we did not observe when looking at individual texture attributes and body sensations in the previous sub-sections.

V. LIMITATIONS AND FUTURE WORK

We only tested congruent conditions in our study following prior work on testing associations of haptic sensation with matching sound effects. However, multisensory research shows that naturally incongruent combinations can still lead to an integrated percept [56]. Therefore, we cannot conclude that our results could not be observed even when adding irrelevant auditory information or adding other sensory cues. Future work should be directed toward comparing congruent and incongruent conditions.

We based our results on participants' discrimination of the mid-air textures on Friedman tests followed by Wilcoxon signed-rank tests to assess differences between groups following similar studies in the literature. However, more specific tests, such as just-noticeable difference (JND) comparisons, need to be conducted to further explore whether people are able to distinguish different ultrasound-based roughness levels.

Our experimental setup used an armrest to ensure the participants' hand position was mostly moving laterally (i.e., X-axis). However, this constrained motion might have limited the influence of cross-modal correspondences to motor behavior. Future work will look at reproducing the implicit measurements part of our study in a free-exploration condition.

Our work limits cross-modal correspondences to audio and haptic sensory modalities. Mid-air interactions, especially for XR and touchless displays, are likely relying heavily on the visual components as well. Thus, cross-modal correspondences between mid-air haptics and visuals should be explored and expanded to olfactory and gustatory modalities.

The audio recordings were only volume-modulated during hand explorations. More dynamically modulated texture sounds could improve the experience of realism and hence enhance the cross-modal correspondence coupling effect. For example, the FoleyAutomatic system [57] and its modifications for real-time sound synthesis is increasingly used in contact-based immersive audio-haptic texture interactions [12].

The scores of rough and medium levels were quite similar despite them being rendered with different mid-air haptic roughness device parameters. Next work should be directed to better define medium levels of roughness.

Participants were not instructed about upper and lower limits of the sensations in the questionnaires used (e.g., criteria for the smoothest and roughest feeling). We allowed participants to introduce their own feeling limits based on the association to the sounds they heard (e.g., the smooth texture should be perceived as soft as it sounds). Therefore, future work should explore more specific criteria about the feelings limits.

VI. GENERAL DISCUSSION AND CONCLUSION

It is a challenge to replicate the natural experiences that physical touch offers using ultrasound mid-air haptics. Particularly, for texture rendering, haptic attributes (e.g., viscosity, hardness, etc.) are difficult to render convincingly due to the lack of physical effects such as friction and compliance. It is therefore essential to investigate ways to improve tactile experiences generated by ultrasound mid-air haptic displays.

To that end, this paper has explored how ultrasound-generated textures can be aided by congruent sound effects to improve roughness discrimination and induce body sensations. In summary, we found that audio cues (presented solely or combined with haptics) can influence how people judge a mid-air texture in terms of its haptic attributes and bodily perceptions, highlighting the value in designing mid-air experiences as multisensory ones, as opposed to designing mid-air haptics stimuli in isolation.

Overall, the results from our statistical analysis and further PCA, can provide practitioners and researchers with insights to consider for future investigations and implementations in a plethora of applications.

For example, in XR gaming, conveying texture attributes (e.g., roughness and moisture) and inducing hand sensations (e.g., temperature, and hardness) can be useful for dexterous object manipulation and social experiences such as a virtual handshake. Conveying texture attributes (e.g., roughness and flexibility) can be useful in the area of digital signage, e-textiles, and virtual try-on (VTO) shopping of garments. Finally, textured message notifications could be used in automotive XR applications and human-machine interfaces (HMIs) by incorporating notions of temperature during HVAC control, something that mid-air haptics alone are not designed to convey but could do so when appropriately sonified.

Our study contributes interesting results and insights that readers can consider for future research directions. With the increasing interest of haptics and digital touch in the metaverse, and the lack of research around multi-sensory perception of mid-air haptics, our contribution provides some initial insights on how to improve the perception of rendering textures.

This makes us reflect on how to further advance the field of mid-air haptics. Whether more complex haptic rendering algorithms are needed, or if research should focus on how to leverage cross-modal associations to obtain more accurate estimations or have an improved and embodied user experience.

ACKNOWLEDGMENTS

The authors would like to thank the contributions of Ana Tajadura-Jimenez for her advice on the study design.

This work was supported by UKRI Future Leaders Fellowship grant (Reference: MR/V025511/1) and the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017746 (project TOUCHLESS).

REFERENCES

- [1] A. Gallace, "Haptic interaction in virtual reality: Are we ready for the metaverse? neuroscientific and behavioral considerations," in *Handbook of Research on Implementing Digital Reality and Interactive Technologies to Achieve Society 5.0*. IGI Global, 2022, pp. 1–14.
- [2] T. Hoshi, et al., "Noncontact tactile display based on radiation pressure of airborne ultrasound," *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 155–165, 2010.
- [3] T. Carter, et al., "Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces," in *The 26th annual ACM symposium on User interface software and technology*, Conference Proceedings, pp. 505–514.
- [4] O. Georgiou, et al., "Mid-air haptic texture exploration in vr," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Conference Proceedings, pp. 964–965.
- [5] E. Freeman, et al., "Textured surfaces for ultrasound haptic displays," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, Conference Proceedings, pp. 491–492.
- [6] D. Beattie, et al., "Incorporating the Perception of Visual Roughness into the Design of Mid-Air Haptic Textures," in *Proceedings: ACM Symposium on Applied Perception*, vol. 10, no. 20. Association for Computing Machinery, sep 2020.
- [7] T. Morisaki, et al., "Midair haptic-optic display with multi-tactile texture based on presenting vibration and pressure sensation by ultrasound," in *SIGGRAPH Asia 2021 Emerging Technologies*, 2021, pp. 1–2.
- [8] M. Marchal, et al., "Can stiffness sensations be rendered in virtual reality using mid-air ultrasound haptic technologies?" in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, Conference Proceedings, pp. 297–306.
- [9] A. Matsubayashi, et al., "Rendering softness using airborne ultrasound," in *2021 IEEE World Haptics Conference*. IEEE, 2021, pp. 355–360.
- [10] J. Jang et al., "Sph fluid tactile rendering for ultrasonic mid-air haptics," *IEEE Transactions on Haptics*, vol. 13, pp. 116–122, 2020.
- [11] Y. Liu, et al., "Texture classification by audio-tactile crossmodal congruence," in *2022 IEEE Haptics Symposium (HAPTICS)*. IEEE, Conference Proceedings, pp. 1–7.
- [12] S. Chan, et al., "Hasti: Haptic and audio synthesis for texture interactions," in *2021 IEEE World Haptics Conference (WHC)*. IEEE, Conference Proceedings, pp. 733–738.
- [13] E. Freeman, "Enhancing Ultrasound Haptics with Parametric Audio Effects," in *ICMI 2021 - Proceedings of the 2021 International Conference on Multimodal Interaction*. Association for Computing Machinery, Inc, oct 2021, pp. 692–696.
- [14] C. Ozkul, et al., "Combining Auditory and Mid-Air Haptic Feedback for a Light Switch Button," in *ICMI 2020 - Proceedings of the 2020 International Conference on Multimodal Interaction*. Association for Computing Machinery, Inc, oct 2020, pp. 60–69.
- [15] G. Brianza, et al., "As light as your scent: effects of smell and sound on body image perception," in *IFIP Conference on Human-Computer Interaction*. Springer, Conference Proceedings, pp. 179–202.
- [16] A. Tajadura-Jiménez, et al., "Action sounds update the mental representation of arm dimension: contributions of kinaesthesia and agency," *Frontiers in psychology*, p. 689, 2015.
- [17] I. Senna, et al., "The marble-hand illusion," *PLoS one*, vol. 9, no. 3, p. e91688, 2014.
- [18] K. Hasegawa et al., "Modulation methods for ultrasound midair haptics," in *Ultrasound Mid-Air Haptics for Touchless Interfaces*. Springer, 2022, pp. 225–240.
- [19] G. Wilson, et al., "Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion," in *Proceedings of the SIGCHI conference on human factors in computing systems*, Conference Proceedings, pp. 1133–1142.
- [20] A. Yoshimoto, et al., "Midair haptic pursuit," *IEEE transactions on haptics*, vol. 12, no. 4, pp. 652–657, 2019.
- [21] G. Korres et al., "Haptogram: Ultrasonic point-cloud tactile stimulation," *IEEE Access*, vol. 4, pp. 7758–7769, 2016.
- [22] W. Frier, et al., "Using spatiotemporal modulation to draw tactile patterns in mid-air," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, Conference Proceedings, pp. 270–281.
- [23] R. Takahashi, et al., "Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*.
- [24] T. Howard, et al., "Investigating the recognition of local shapes using mid-air ultrasound haptics," in *2019 IEEE World Haptics Conference (WHC)*. IEEE, Conference Proceedings, pp. 503–508.
- [25] D. Pittera, et al., "I'm sensing in the rain: Spatial incongruity in visual-tactile mid-air stimulation can elicit ownership in vr users," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Conference Proceedings, pp. 1–15.
- [26] I. Rakkolainen, et al., "A survey of mid-air ultrasound haptics and its applications," *IEEE Transactions on Haptics*, no. 1, pp. 2–19, 2020.
- [27] K. E. MacLean, et al., "Multisensory haptic interactions: understanding the sense and designing for it," in *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations*, 2017, pp. 97–142.
- [28] P. Cornelio, "Opportunities for multisensory mid-air interactions featuring ultrasound haptic feedback," in *Ultrasound Mid-Air Haptics for Touchless Interfaces*. Springer, 2022, pp. 99–117.
- [29] E. Hoggan et al., "Designing audio and tactile crossmodal icons for mobile devices," in *Proceedings of the 9th International Conference on Multimodal Interfaces, ICMI'07*. New York, New York, USA: ACM Press, 2007, pp. 162–169.
- [30] D. Ablart, et al., "Using ultrasonic mid-air haptic patterns in multi-modal user experiences," in *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games*. IEEE, 2019, pp. 1–6.
- [31] E. Maggioni, et al., "Measuring the added value of haptic feedback," in *2017 ninth international conference on quality of multimedia experience (QoMEX)*. IEEE, pp. 1–6.
- [32] C. T. Vi, et al., "Not just seeing, but also feeling art: Mid-air haptic experiences integrated in a multisensory art exhibition," *International Journal of Human-Computer Studies*, vol. 108, pp. 1–14, 2017.
- [33] S. Guest, et al., "Audiotactile interactions in roughness perception," *Experimental Brain Research*, vol. 146, no. 2, pp. 161–171, 2002.
- [34] V. Jousmäki et al., "Parchment-skin illusion: sound-biased touch," *Current biology*, vol. 8, no. 6, pp. R190–R191, 1998.

- [35] A. Tajadura-Jiménez, et al., "Action sounds recalibrate perceived tactile distance," *Current Biology*, vol. 22, no. 13, pp. R516–R517, 2012.
- [36] —, "Sonification of surface tapping changes behavior, surface perception, and emotion," *IEEE MultiMedia*, vol. 22, no. 1, pp. 48–57, 2015.
- [37] A. Tajadura-Jiménez, et al., "Body weight distortions in an auditory-driven body illusion in subclinical and clinical eating disorders," *Scientific Reports*, vol. 12, no. 1, pp. 1–17, 2022.
- [38] A. Tajadura-Jiménez, et al., "As light as your footsteps: Altering walking sounds to change perceived body weight, emotional state and gait," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 2943–2952. [Online]. Available: <https://doi.org/10.1145/2702123.2702374>
- [39] A. Tajadura-Jimenez, et al., "Using sound in multi-touch interfaces to change materiality and touch behavior," in *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, ser. NordiCHI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 199–202. [Online]. Available: <https://doi.org/10.1145/2639189.2639217>
- [40] T. R. Stanton et al., "The influence of auditory cues on bodily and movement perception," *Frontiers in psychology*, vol. 10, p. 3001, 2020.
- [41] C. Velasco et al., *Multisensory experiences: Where the senses meet technology*. Oxford University Press, 2020.
- [42] N. Di Stefano et al., "Roughness perception: A multisensory/crossmodal perspective," *Attention, Perception, Psychophysics*, vol. 84, no. 7, pp. 2087–2114, 2022. [Online]. Available: <https://doi.org/10.3758/s13414-022-02550-y>
- [43] C. Spence, "Crossmodal correspondences: A tutorial review," *Attention, Perception, Psychophysics*, vol. 73, no. 4, pp. 971–995, 2011. [Online]. Available: <https://doi.org/10.3758/s13414-010-0073-7>
- [44] A. Tajadura-Jiménez, et al., "Altering one's body-perception through e-textiles and haptic metaphors," *Frontiers in Robotics and AI*, p. 7, 2020.
- [45] D. J. Freed, "Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events," *The Journal of the Acoustical Society of America*, vol. 87, no. 1, pp. 311–322, 1990.
- [46] M. Zampini et al., "Assessing the role of sound in the perception of food and drink," *Chemosensory Perception*, vol. 3, no. 1, pp. 57–67, 2010.
- [47] C. Spence, "Temperature-based crossmodal correspondences: Causes and consequences," *Multisensory Research*, vol. 33, pp. 645–682, 2020.
- [48] D. Beattie, et al., "Mid-air haptic textures," Jan. 10 2023, uS Patent 11,550,395.
- [49] M. Tomczak et al., "The need to report effect size estimates revisited. an overview of some recommended measures of effect size." *Trends in sport sciences*, vol. 21, no. 1, 2014.
- [50] S. J. Lederman et al., "Hand movements: A window into haptic object recognition," *Cognitive psychology*, vol. 19, no. 3, pp. 342–368, 1987.
- [51] L. Shams, et al., "Influences of multisensory experience on subsequent unisensory processing," *Frontiers in psychology*, vol. 2, p. 264, 2011.
- [52] T. Callier, et al., "Kinematics of unconstrained tactile texture exploration," *Journal of neurophysiology*, no. 7, pp. 3013–3020, 2015.
- [53] Y. Tanaka, et al., "Contact force and scanning velocity during active roughness perception," *PLoS One*, vol. 9, no. 3, p. e93363, 2014.
- [54] E. Kim et al., "Defining haptic experience: foundations for understanding, communicating, and evaluating hx," in *Proceedings of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–13.
- [55] S. Okamoto, et al., "Psychophysical dimensions of tactile perception of textures," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 81–93, 2012.
- [56] B. De Gelder et al., "Multisensory integration, perception and ecological validity," *Trends in cognitive sciences*, vol. 7, no. 10, pp. 460–467, 2003.
- [57] K. Van Den Doel, et al., "FoleyAutomatic: Physically-based sound effects for interactive simulation and animation," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2001*, 2001, pp. 537–544.

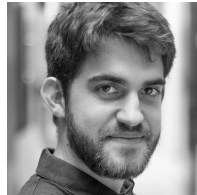


Roberto Montano is a research engineer working at Ultraleap, he worked as research fellow in Interactive Graphics at the Multi-Sensory Devices (MSD) group from University College London (UCL) and interned as a Human Computer Interaction researcher at Adobe Research. Roberto completed a PhD on computational 3D interaction techniques for immersive VR at the University of Sussex. Roberto's research work includes optimisation techniques to improve human factors related to manipulation, navigation and selection in VR, as

well as developing multimodal 3D displays using ultrasonic phased arrays of transducers to provide users with visual, auditory and haptic experiences.



Dario Pittera received his B.S. in Communication and Psychology and his M.S. in Clinical psychology, developmental psychology and neuropsychology, both from the Università degli Studi di Milano-Bicocca, Milan, Italy. He worked as a researcher at the Policlinico di Milano Hospital and at the University of Birmingham from, investigating haptic technology through psychophysical experiments. He then worked as a Lab Associate at Disney Research, Pittsburgh, PA, US before completing his PhD degree at the University of Sussex working on mid-air haptics illusions and embodiment. Currently, he is working as a behavioural research scientist at Ultraleap. His research is focused on the perception and exploration of mid-air haptic technology.



William Frier received his PhD in Human-Computer Interaction from the University of Sussex and is now the Lead Haptics Researcher at Ultraleap. Since Ultraleap Ltd early days, he has been a key investigator in studying the perception of mid-air haptics which lead to various patents, and publications in HCI and Haptic venues. He has studied Electrical and Electronical engineering at ESIGELEC (France) and Intelligent System and Robotics at University Pierre and Marie Curie (France), making him a key person in the development and integration of modular ultrasound mid-air haptic devices for robotic applications and Human-Machine Interfaces. William is the recipient of a UKRI Future Leader Fellowship and an honorary researcher at the University of Bristol.



Orestis Georgiou (Senior Member, IEEE) received a Ph.D. in Mathematics from the University of Bristol in 2012. In 2013 he was a postdoc at the Max Plack Institute of Complex Physics in Dresden. He then moved to the industry, first working for Toshiba Telecommunications Research Lab as a Senior Researcher (2013-2017), and later for Ultraleap as Head of R&D Partnerships (2017-present) where he is the PI on several publicly funded R&D projects. He is the author of 6 patents and over 100 academic papers published in leading journals and conferences of Mathematics, Physics, Computer Science, Engineering and Medicine. Orestis was awarded a Marie Curie Individual Fellowship hosted at the University of Cyprus (2019-2021), is an author of two best paper awards (ISWCS'2013 and AutoUI'2018) and is the recipient of the prestigious IEEE Heinrich Hertz Award in 2019.



Marianna Obrist is Professor of Multisensory Interfaces and, before joining UCL, she was head of the Sussex Computer Human Interaction (SCHi 'sky') Lab at the School of Engineering and Informatics at the University of Sussex. Her research ambition is to establish touch, taste, and smell as interaction modalities in human-computer interaction (HCI). As part of her research, she developed a novel scent-delivery technology that was exhibited at the World Economic Forum (WEF) 2019 and 2020 in Davos. Supported by an ERC proof-of-concept, this technology is now commercialised through OWidgets Ltd, a University start-up she co-founded in 2019. Before joining Sussex, Marianna was a Marie Curie Fellow at Newcastle University.



Patricia Cornelio is a research engineer at Ultraleap where she coordinates a 4-year R&D Future Leader Fellowship program. She also holds a honorary research position at the University College London Interaction Centre (UCLIC). Patricia obtained a PhD in Informatics at the University of Sussex under the supervision from Prof Sriram Subramanian. Before starting her PhD, Patricia obtained a MSc degree (2014) and a BSc degree (2010) both in Electronics Engineering at the National Institute of Technology of Mexico. Patricia's research is focused on multi-sensory experiences with a focus on touch and smell using mid-air haptic technology, she is passionate about the human senses and the role they play in human-machine integration.