Mouth Haptics in VR using a Headset Ultrasound Phased Array

Vivian Shen
Carnegie Mellon University
Pittsburgh, PA, USA
vhshen@cmu.edu

Craig Shultz
Carnegie Mellon University
Pittsburgh, PA, USA
cshultz@cmu.edu

Chris Harrison
Carnegie Mellon University
Pittsburgh, PA, USA
chris.harrison@cs.cmu.edu

Figure 1: Our system can deliver rich, non-contact haptic feedback to the mouth using a thin array of ultrasonic transducers integrated into the underside of a VR headset. Here a user leans forward to drink from a virtual water fountain, and can feel a sensation of a stream of water on their lips and teeth.

ABSTRACT

Today's consumer virtual reality (VR) systems offer limited haptic feedback via vibration motors in handheld controllers. Rendering haptics to other parts of the body is an open challenge, especially in a practical and consumer-friendly manner. The mouth is of particular interest, as it is a close second in tactile sensitivity to the fingertips, offering a unique opportunity to add fine-grained haptic effects. In this research, we developed a thin, compact, beamforming array of ultrasonic transducers, which can render haptic effects onto the mouth. Importantly, all components are integrated into the headset, meaning the user does not need to wear an additional accessory, or place any external infrastructure in their room. We explored several effects, including point impulses, swipes, and persistent vibrations. Our haptic sensations can be felt on the lips, teeth and tongue, which can be incorporated into new and interesting VR experiences.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Haptic devices**.

ACM Reference Format:
Vivian Shen, Craig Shultz, and Chris Harrison. 2022. Mouth Haptics in VR using a Headset Ultrasound Phased Array. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3491102.3501960

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3501960

1 INTRODUCTION

Virtual and augmented reality (VR/AR) headsets continue to make impressive strides in immersion and realism, particularly in visual and audio content. However, the delivery of rich tactile sensations continues to be a significant and open challenge. Critically, consumers want robust and integrated solutions – ones that do not require any extra devices or limit freedom of movement. For this reason, vibration motors in handheld controllers are the current consumer state of the art. While more sophisticated approaches exist (e.g., exoskeletons [1, 4, 42, 46], haptic vests [2, 16], body-cantilevered accessories [6], in-room air cannons [27, 65]), they have yet to see even modest consumer adoption.

Simultaneously, the mouth has been largely overlooked as a haptic target in VR/AR, despite being second in terms of sensitivity and density of mechanoreceptors, only behind the fingertips [14, 63, 73]. Equally important, the proximity of the mouth to the headset offers a significant opportunity to enable on- and in-mouth haptic effects, without needing to run wires or wear an extra accessory. However, consumers do not want to cover their entire face, let alone put something up against (or into) their mouth. For AR, the industry is trending towards glasses-like form factors, so as to preserve as much facial expression as possible for human-human communication. Even in VR, smaller headsets are the consumer trend, with the mouth exposed and unencumbered.

In this research, we built a thin, compact, beamforming array of ultrasonic transducers (Figure 3), which could be integrated into future headsets in a practical and consumer-friendly way. We use this hardware to focus air-borne acoustic energy onto the lips and into the mouth, creating sensations such as taps and continuous vibrations, which we can also animate along arbitrary 3D paths. In addition to the lips, our effects can be felt on the teeth and tongue. When coupled with coordinated graphical feedback, the effects are convincing, boosting realism and immersion. We built a variety of sensory demos, including raindrops, mud splatter, pushing through cobwebs, and crawling bugs. While in-air haptics using ultrasonic phased arrays is not new [24, 34], we are the first to integrate the technology into a headset for use on the mouth and explore the rich application space.

2 RELATED WORK

Our work intersects with several key areas of literature, which we now review. First and foremost, our haptic hardware is a type of ultrasonic phased array, and so we discuss the principle of operation and key related work in HCI. We then discuss tactile perception of the face, mouth and oral cavity, along with some non-VR/AR example systems. Finally, we discuss research most related to our own: VR/AR systems that achieve mouth haptic effects through a variety of mechanisms.

2.1 Ultrasonic Phased Array Haptics

Separate from contact-based ultrasonic friction effects [43, 51, 78], ultrasonic phased array devices deliver contactless haptic sensations to skin in mid-air [31, 59]. They achieve this through the focused application of high-frequency pressure waves, generally in the range of 40-70kHz. Users cannot feel the varying acoustic pressure, as the frequency is too high. Instead, these devices rely on acoustic radiation pressure, a nonlinear phenomenon, which rectifies acoustic pressure reflecting off the skin. When time-varying pressure is amplitude modulated at relevant haptic frequencies (e.g. 40-200Hz), radiation pressure can be felt by users, mediated by rapidly adapting Meissner and Pacinian corpuscles. Importantly, this effect can be spatially localized [22], though localization is limited by the spatial wavelength of sound in air (around 8.5mm at 40kHz). Localization is achieved by adjusting the timing and amplitude information of a phased transducer array to create superpositions of waves in space that constructively or destructively interfere.

Iwamoto et al. was the first to describe and position this effect [33] using a 2D array of transducers, and suggested possible applications in VR [31]. Further improvements in focal point rendering and position sensing were introduced by Carter et al. [9], leading to an eventual commercialization of the technology [71]. Due to the popularity of this technique, open source systems [48] also exist. A range of cutaneous tactile sensations can be rendered, from moving trajectories [77] to general shape of the focal region [45] or even sensations of texture [20, 50]. Ultrasonic phased arrays have been used in prior VR systems for rendering haptics onto the hands (but not the face), including Kervegant et al. [40], Martinez et al. [47], Georgiou et al. [23] and Sand et al. [62]. Conversely, LipNotif [35] and Whiskers [24] systems utilize ultrasonic phased arrays for actuating the mouth, but do not consider VR or integration into a worn apparatus. We cover additional mouth psychophysical related work in Section 4.

2.2 Haptics on the Mouth and Face

Compared to the hands, little has been studied about the tactile mechanoreceptors and internervation of the face. We report some of the known tactile capabilities of the face and mouth, and review interfaces that have explored this area of the body.

The face and mouth can be broken down into three basic regions, the non-glabrous (hairy) regions (e.g. cheeks, brow, nose, forehead, etc), the glabrous skin around the outer lips and mouth, and the inside of the oral cavity, including the tongue [14]. Both slowly and rapidly adapting afferents can be found in these regions, with a slightly higher percentage of slowly adapting afferents than rapidly adapting ones. The highest percentage of receptors is around the lips and mouth. Tactile innervation and acuity is similar to that of the hand, but there is no 200Hz peak in sensitivity associated with rapidly adapting type II afferents (Pacinian corpuscles), although tactile response does extend into this frequency range. Exterior facial afferent receptive fields size (4-6mm^2) and density (48-84 units/cm^2) compares favorably to the phalanges and palm of the hand (11-101mm^2 and 10-100 units/cm^2, respectively) [14, 36].

The oral cavity also contains a wealth of mechanoreceptors, such as extremely force sensitive periodontal mechanoreceptors in the teeth [69], receptors on the inner lip [5], and other soft tissues in the mouth, tongue, and gums. Surfaces of the mouth are in constant interaction with each other (as the mouth is not easily observed by the eyes or ears) and thus they generate rich patterns of somatosensory input which lead to vivid sensations during eating, drinking, and speaking [28]. In particular, the oral cavity appears sensitive to air pressure (most likely to aid in speech), which makes it a unique part of the body for ultrasonic mid-air stimulation. The skin of the face and inner mouth, like the rest of the body, is also sensitive to temperature and pain, though we do not explore those aspects here.

Innumerable interactive systems have taken advantage of the face's broad array of tactile sensors. In the VR/AR domain specifically, systems have been presented that use methods such as vibrotactile motors (Figure 2A) [15, 39, 41, 55, 56, 58, 75, 79], Peltier elements (thermal) (Figure 2B) [11, 12, 57, 68], skin suction actuators (Figure 2C) [37, 38], skin stretch motors/pulleys [10, 74], and flywheels for applying inertial forces to the head [26]. Most often these techniques are added or integrated into the headset itself, especially the liner that contacts the face. However, more closely related to our work are systems that generate haptics specifically for the mouth region, which we discuss next.

2.3 VR/AR Mouth Haptics

There are a limited number of systems that have explored haptics on the mouth. To avoid instrumenting the user (with extra weight, wires, etc.), some approaches have utilized external equipment installed in the user's environment. For instance, Hapmosphere [29] and Haptic Around [30] are large, ceiling mounted pieces of equipment, while [32] is a fully instrumented CAVE. All three of these systems can generate wind and weather effects (i.e. heat, cold), which can be felt on the mouth (and body in general), to enhance VR experiences. However, the cost, complexity, aesthetics and size of these systems make consumer adoption unlikely.

A more practical approach is to have the haptic hardware be integrated with the headset. In this area, most systems have focused on simulating wind and airflow around the head and mouth. For instance, Head Mounted Wind [8] uses a series of fans externally attached to the headset to blow air onto the mouth and face. VaIR [61] is similar, but uses a pneumatic system with adjustable nozzles

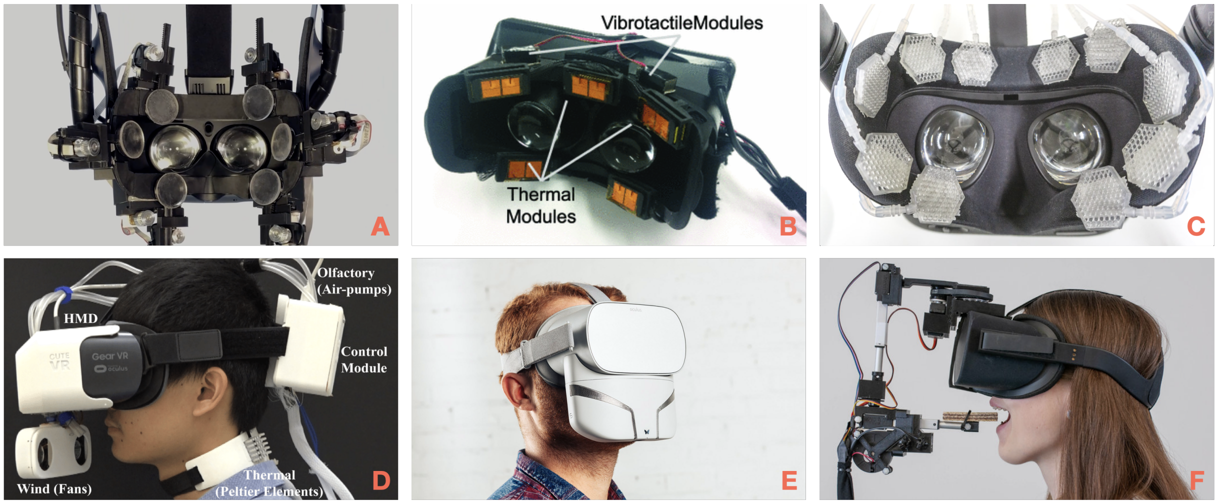


Figure 2: Top three images are prior VR/AR systems that can deliver haptic effects to the face (but not the mouth): A) Masque [74], B) Liquid Reality [56], and C) Haptopus [37]. Bottom three images are VR/AR systems with mouth haptic effects: D) Season Traveller [60], E) FeelReal [19], and F) FaceHaptics [76].

and tubes. Other systems combine blown air with other effects, for example, Ambiotherm incorporates Peltier Elements for thermal control. Season Traveller (Figure 2D) [60] adds thermal and olfactory elements. Finally, FeelReal (Figure 2E) [19] offers thermal, olfactory, and vibration feedback, and further incorporates an ultrasonic ionizing system to create water mist.

Perhaps the most related previous work is FaceHaptics [76], which consists of a robotic arm attached to a headset (Figure 2F). On this arm are several modules, including a spray nozzle, a heating wire, a fan, and an interchangeable soft rubber tip for tactile feedback. While this system does enable a wide range of localizable haptic effects, the form factor is less than ideal for consumer use. Latency can also be high for some effects, like multiple impulses, as the robot arm must physically translate to new positions. Our work, in many ways, aims to achieve a similar pallet of tactile effects, but in a more practical form factor.

3 IMPLEMENTATION

We now describe the hardware and software elements that comprise our system. We note that the success of haptic effects is particularly difficult to convey in a research paper. Thus, to encourage replication and exploration by researchers and practitioners, we have open sourced our software and PCB designs at http://github.com/FIGLAB/mouthhaptics.

3.1 Hardware

As a proof-of-concept platform, we selected an Oculus Quest 2 [53] VR headset (though our technique is broadly applicable to other VR headsets and even AR headsets such as the Hololens). To the bottom of this device, we fitted a custom PCB populated with 64 Murata MA40S4S 40kHz ultrasonic transducers [52]. These transducers have an 80° emission cone, and so we oriented them towards the mouth at a 45° angle. The transducer array measures 17.9x10.6x1.5cm and weighs 107g (although already thin, there are opportunities for further miniaturization that we discuss in Future Work Section 7). Ultrasonic arrays for in-air haptics have been found to be safe in typical use [7, 17]. Indeed, such devices are commercially available today (e.g., UltraHaptics), though the included documentation notes the devices should not be focused into the ears – a recommendation we endorse. We measured the sound pressure level (SPL) of a focal point centered 3cm in front of the array (141 dB SPL), and also at the user's ears when wearing our headset (100 dB SPL). This is well below the maximum threshold recommended by Battista [17].

Our transducer array connects via ribbon cables to an Ultraino driver board [48], which we upgraded with IX4427N IC chips [13] to be able to drive our transducer array from 0-30V using an external power supply (power consumption is discussed later). The driver board is mated to an Arduino Mega [3] flashed with the Ultraino DriverMEGA firmware [48, 70]. This setup allows for independent control of all 64 transducers with a time step resolution of 2.5us, which equates to $\pi/5$ phase resolution at 40kHz, and a simulated nodal position accuracy of 0.9mm (see [48] for an extended discussion of the mathematics and mechanics of ultrasound beamforming). The Arduino Mega connects over USB to a laptop running our Unity-based software (described next), which also powers the graphics of the Oculus Quest 2 over OculusLink. Note that the inclusion of a laptop and an Arduino was purely to ease development; the Oculus Quest 2 hardware is more than capable of running the haptic elements of our system, and we envision future commercial-grade systems being tightly integrated with a skin-like array of transducers (Figure 22).

3.2 Software

We developed our software stack in Unity [72]. Using the known locations of the virtual object and headset (and thus also the geometry of the ultrasonic phased array; Figure 3), our software calculates the phase array timings to create a haptic effect at the requisite 3D

Figure 3: Our proof-of-concept hardware features 64 ultrasonic transducers located on the underside of the headset.

position. This "firing solution" is calculated (Figure 4) and transmitted to the Arduino at 10Hz. Although this data is transmitted over USB to the Arduino, it still "beats" the visuals rendered on the Oculus Quest 2 by 2ms on average (which is close enough in time to be perceived as synchronous). The update rate of 10Hz is sufficiently fast such that virtual objects feel bound to their haptic output.

There are several ways to create an acoustic node. One option is to transmit two alternating firing patterns to the Arduino – one pattern that creates a beamformed node, while the other simply turns the phased array off (i.e., amplitude modulation). However, a stronger haptic sensation can be created with two alternating patterns that create adjacent nodes (we use a displacement of 1mm in the X-axis), which is called Lateral Modulation (LM) in the literature [21, 66, 67]. This means the transducer array is continuously outputting acoustic energy (vs. being 50% off). We adopted this second method for all of our demos.

At present, our software only generates haptics for the closest haptic object in front of the mouth (true multi-node haptics is a challenge we leave to future work). Fortunately, it is most typical for people to bring a single haptic object to their mouth at one time (e.g., drinking from a water fountain). In the case of particle systems, our software simply continuously animates the particles in the stream. This creates a sensation of particles hitting the mouth in quick succession, though not more than one particle can hit at the same instant, and thus it works better for lower frequency events (i.e., <10Hz). For high-frequency effects, it is better to utilize lower-level haptic parameters (described below).

Finally, before use, a quick calibration procedure is required to ascertain the user's mouth position (X/Y/Z) with respect to the headset. For this, we generate a persistent vibration node which can be moved using keyboard arrow keys. Once the node is centered between the lips, the offset value is saved for future use.

3.2.1 World-Bound Haptic Effects.

As previously noted, haptic effects can be attached to any objects in Unity. Haptic objects located in the world scene can only be felt by moving one's mouth to that location (e.g., a fixed asset, such as a water fountain), or by moving an object to the mouth (e.g., an insect jumping onto the face). We call these world-bound haptic effects. We emphasize this is not a fixed location on the mouth, but is rendered into the air in the object's 3D location. This means that users can stick out their tongue to feel an effect, or even move forward to translate the haptic node into their mouth and onto their teeth or tongue. To save power, we only generate haptic output when objects are within 10cm of the face.

Would-bound haptic effects create two distinct sensations. One is when the mouth is held in the haptic region, and the node is translated left/right and up/down (either by moving the mouth or object). This manifests as a swipe-like sensation, with the node position tracking on the surface of the face (Figures 5A, 5B). Alternatively, the node can translate along the Z axis (again, either by the mouth or by the object moving), causing the haptics to enter/exit the mouth. When the lips are closed, this feels like a momentary impulse – the moment the acoustic focus is on the lips (Figure 5C). If the mouth is open, the effect can enter the mouth and be sensed by the teeth and tongue.

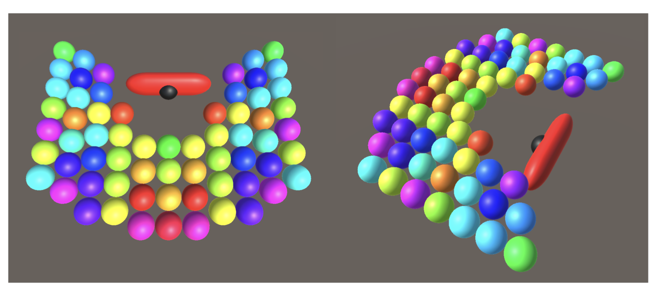


Figure 4: Screenshot of our Unity interface. In order to beamform an acoustic node (black sphere) onto the lips (red capsule), our pipeline must calculate the firing phase offset for each of the 64 ultrasonic transducers (colored spheres; hue denotes phase).

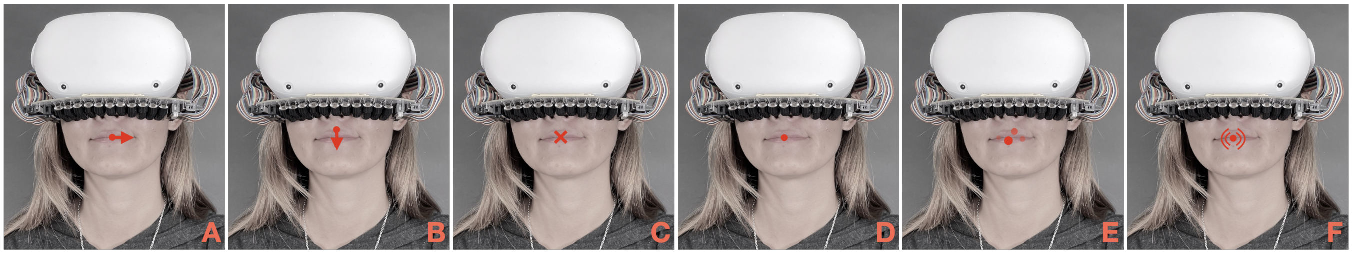


Figure 5: The library of animations we utilize in our applications: A) X-swipe, B) Y-swipe, C) Z-swipe, D) Single Impulse, E) Impulse Train, and F) Persistent Vibration.

3.2.2 Mouth-Bound Haptic Effects. Haptic effects can also be attached to mouth-bound objects, which use the local coordinate system of the mouth. This is useful for creating haptic effects that track with the head, such as sweat or tears, or pre-programmed animations such as swipes across the lips. To achieve this effect, one must attach a HapticEffect actuation script to a mouth-bound object in Unity, which can animate its position programmatically. We note that it is sometimes advantageous to use a mouth-bound haptic effect even when the event is triggered by a world object. For example, in our haunted forest demo (described later), the user must walk through cobwebs. While we could have attached a haptic effect to the cobweb object, a user moving through it would only experience a momentary impulse (Z-axis translation; see previous section). We found the experience was more visceral when a swiping effect was used, as if breaking through the cobweb and it being drawn off the face as one walks forward.

In this mouth-bound mode, we use two haptic designs. First, and most simple of all, is to create a momentary tap or impact (Figure 5D). This is achieved by creating a transient haptic object at a desired position, after which the object is deleted or disabled. These haptic objects can be rapidly generated, creating a structured or random stream of impulses if desired (Figure 5E). If this haptic object persists on the lips, it becomes a persistent vibration effect (Figure 5F). Second, it is also possible to create animated paths simply by moving objects programmatically in Unity. For instance, a swipe can be created by creating and translating a haptic object across the mouth (Figures 5A, 5B and 5C). Irregular paths are equally possible, such as an insect walking randomly on one's lips (Figure 10). One can also translate a mouth-bound object along the Z-axis to create an almost peristalsis-like effect (which we use in our coffee mug demo to simulate liquid entering the mouth).

3.2.3 Phased Array Parameters. The haptic effects described above are all spatial manipulations: creating, deleting and moving haptic objects within Unity. While this moves haptic effects around in space, it does not modify the underlying actuation signal. For this, we expose a series of parameters that alters the waveform emitted by our transducer array.

Modulation Frequency – The most significant tunable parameter to add haptic expressivity is the modulation frequency of the array. We found the best results between 20Hz and 200Hz (which matches prior work [24, 34]; see also our Perception Study). The sensation is what one would expect – below ~50Hz, it is possible to feel the individual impulses. Above this threshold, the sensation feels like a continuous vibration or buzzing.

Node Separation – As noted previously, when our software places a haptic node in 3D space, it is actually two adjacent ultrasonic nodes separated by 1mm, which toggle back and forth rapidly at the modulation frequency (called Lateral Modulation [66, 67]). This separation distance can be altered to create different sensations.

Amplitude – The amplitude of the haptic effect can be controlled in hardware by varying the duty cycle of the 40kHz signal. Our prototype supports amplitudes from 0% to 100% in steps of 20%. Alternatively, the amplitude can be controlled by varying the input voltage to the driver board. We note, however, that in all of our demos we used 28V, opting to vary other parameters instead.

3.3 Power Consumption and Battery Life

As determined in our subsequent perception study, we need to drive our phased array at a minimum of 20.1V to ensure reliable perception of our haptic effects. At this level, the total power consumption of the driver board and phased array is 2.71W when generating a persistent node. Of course, our duty cycle is very low, as we only emit when there is a haptic event to render. When in standby and not emitting, the driver board consumes 0.30W. If we assume a one-second haptic event every 30 seconds, we find a power consumption of 0.38Wh. Our prototype platform, the Oculus Quest 2, contains a 14Wh battery [54], which means one hour of haptically-enhanced gameplay would consume 2.71% of the device's battery life. With further hardware optimization, especially in regards to leakage current when the driver board is off, power consumption should drop below 0.1Wh, and thus have a negligible impact on runtime.

4 PERCEPTION STUDY

Although tactile perception on the face, lips, mouth, and oral cavity has been studied in general (see e.g., [5, 14, 28, 69]), there is limited work looking specifically at haptics delivered to these locations using mid-air ultrasonics. Fortunately, the few papers that do exist provide an excellent characterization on which we grounded much of our work. For instance, in Whiskers [24], the authors examined the sensitivity of the face at three locations: the cheek, center of forehead, and eyebrow. Using static cues and different durations (0.5, 1, and 1.5s) of 40Hz amplitude modulation, they found high detection rates with little variation between locations. Their dynamic

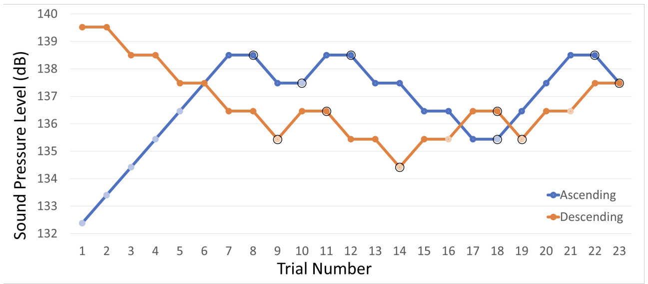


Figure 6: An example (P2) of two staircases (ascending and descending) in the transformed two-down, one-up method. Reversals are denoted with black circles, and undetected stimuli are denoted by the lighter colors.

tactile apparent motion stimuli, discretely changing the modulation location along a line in rapid succession, was found to be limited in its ability to render motion. The center of the forehead achieved a localization error of 3.77mm, and was the highest performance in terms of duration and movement perception. Mizutani et al. [49] also studied absolute thresholds of mid-air ultrasonic tactile perception on the face, using both lateral modulation (LM) and amplitude modulation methods. Across six locations (on the forehead, cheeks, and lower jaw) they found no significant difference between stimulus locations using the same type of modulation method, and that LM was perhaps the most effective method for stimulation (presumably because there is no reduction in output power with lateral modulation).

More recent and relevant to this work, Jingu et al. [34] psychophysically characterized perception of ultrasound stimulation on the lips. Trends from other areas of the face and body were also found in this work, mainly that lateral modulation (specifically circular LM) has a lower detection threshold than amplitude modulation. They also found the valley-shaped area in the middle of the lips to be the most sensitive spot out of those tested, and saw no evidence of additional 200Hz sensitivity (as expected from other studies on facial sensitivity). Indeed, 40Hz was found to be the most sensitive frequency, suggesting strong response from RAI or Meissner corpuscles. It was even found that circular LM at 40Hz on the lips was as much as 3dB more sensitive than circular LM at 200Hz on the palm, more evidence of the high sensitivity of the lips.

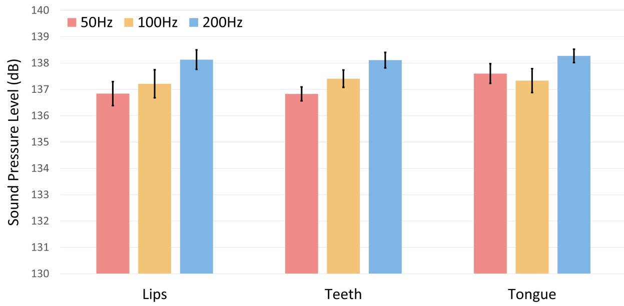
While this prior work was comprehensive for the face and lips, a few significant questions remained. Foremost, it was not immediately obvious that having an ultrasonic phased array placed so close to the mouth would permit accurate beamforming, let alone the fact the transducer geometry had to be rather irregular (previous work was performed with large rectangular arrays directly in-front and in-plane to a user's face). Second, prior work did not explore sensations in the oral cavity – specifically the teeth and tongue – which we believe opens new and interesting interactive avenues in VR/AR. For this reason, we ran our own supplemental perception study to help inform our later explorations.

4.1 Procedure

The study was conducted with 11 participants, mean age of 24. Participants remained seated in the center of a large room for the duration of the study, which lasted around 45 minutes and paid $20. After the headset was worn, participants were shown the default Unity VR scene consisting of a blue sky and a brown ground without other visualizations, in order to not distract from the haptic effect. We performed the same procedure at each of our three locations of interest – lips, teeth, tongue – with randomized presentation order. To start, users first calibrated the position of their mouth (see Software Section). Participants then familiarized themselves with the haptic sensation until they felt comfortable making judgments. Of note, three participants could not consistently feel the node in all three location conditions and so they were dismissed from the study. In two cases, the participants could not feel the sensation at all (potentially due to poor node calibration), and in the other case, the participant could feel it on the teeth and tongue, but interestingly, not on the lips. Thus our analysis below uses data from the eight remaining participants.

For our amplitude testing, we varied the voltage supplied to the board using a variable power supply. We used a transformed two-down/one-up double staircase method, targeting a 70.7% detection response threshold on the psychometric curve (see Levitt [44] for more details on this method). Two staircases were run simultaneously – one ascending and one descending – which were interleaved to reduce user bias. An example staircase session can be seen in Figure 6. The ascending staircase started at 7V, while the descending staircase started at 28V. The haptic effect was played for 1000ms and the participant was verbally asked if they could detect the stimulus. The stimulus amplitude was increased (by 3V) after a single report of no detection and lowered (by 3V) after two consecutive reports of positive detection (following a two-down/one-up paradigm). The staircases were individually stopped after six reversals, meaning that the entire test was over after both the ascending and descending staircases had reached six reversals. This double staircase procedure was performed three times, at modulation frequencies of 50, 100, and 200Hz in a randomized order. The above procedure was repeated for all three mouth locations.

We also performed a spatial acuity test. For this, we generated two haptic nodes at varying separation distances, with one node always centered on the mouth. The first haptic node was active for 500 ms, followed by a 500ms break, and then followed by the second haptic node for 500 ms. Participants could request the stimuli be replayed as many times as they wished. We also used a double staircase method, with one staircase starting with both nodes centered on the mouth at the same location, while the other staircase had one node start 1.8 cm to the right of center. If the participant was able to distinguish the two different node locations twice in a row, the judgement of location was made "harder" and the outer node was moved towards the center by a step size of 1 mm. If the participant reported the two nodes to be in the same location, then the staircase was made "easier" by moving away from the center by the same step size of 1 mm. To minimize variations in sensation caused by other factors, this test was held at a constant modulation frequency and amplitude of 50Hz at 28V.



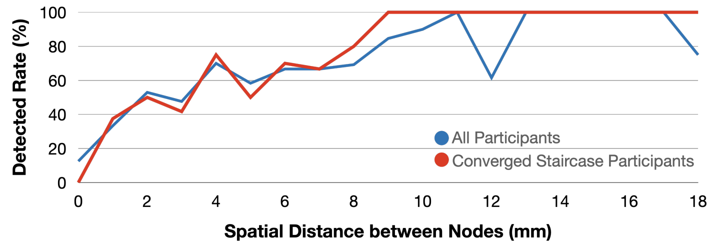


Figure 7: Results from our perception study. Left: Mean minimum detection thresholds across three mouth locations and three modulation frequencies. Right: Percentage of correct detections at varying separation distances of two nodes. Blue is the average of all our participants, while red is the average of our participants who had converging staircases.

4.2 Results

We took the average of the last six peaks and valleys of each of the interleaved staircases for analysis. For both the lips and teeth, the 50Hz modulation had the lowest detection threshold, while 200Hz had the highest. For the tongue, 100Hz actually had a lower detection threshold than 50Hz. This increased sensitivity to lower frequencies is in line with prior work [24, 34, 49], and we used this to inform the creation of our demo interactions, mostly keeping our modulation rates between 50 and 100Hz. We also found varying detection thresholds by location, with the lips having a threshold slightly lower than the teeth, with the tongue being the least sensitive. Results can be seen in Figure 7, left.

Figure 7, right, provides a summary of our spatial resolution results for the lips. In general, by 9mm of separation, most participants were confidently reporting they could feel two nodes on the lips. This almost exactly matches the wavelength of 40kHz ultrasound in air: 8.5mm. Below 8.5mm of separation, two nodes will have overlapping waveforms and thus be harder to separate. We did find that five participants did not have converging staircases, and were inconsistent in some of their answers (i.e., they reported feeling two separate nodes in one trial, but upon returning to that trial, reported they could not). For this reason, their staircases did not converge in the traditional manner. For our other three participants, their staircases converged in a typical manner, and using their reversals we found a mean separation distance of 5.6mm. This suggests that although a single node's spatial extent may be around 8.5mm, node location steps smaller than this may be perceivable by some users (the lips are known to have some of the smallest receptive fields, even smaller than the fingertips [14]). For completeness, we plot all participants' data along with these three more reliable participants' data in Figure 7, right.

In the case of our spatial tongue trials, none of our participants had staircases that converged and responses were very inconsistent. This suggests it was not possible for our participants to make

Scenario	Interaction	Animation Type	Duration	Position & Distribution	Generation Rate	Mod. Freq.
Haunted forest	Spiderwebs	X swipe	1000ms	Center start, animates to side of collision.	1 total	90 hz
Haunted forest	Spider on face	Impulse train	1250ms	Random impulses inside 1x1cm area, animates to right of mouth.	5/sec	40 hz
Haunted forest	Spider splatter	Impulse train	1000ms	Random impulses inside 3x1.5 cm area, centered.	10/sec	70 hz
Haunted forest	Dripping Goo	Y swipes	Continuous upon collision	Swipes with random starting X position.	2/sec	50 hz
School Simulator	Water fountain	Impulse train	Continuous upon collision	Random impulses in 1x1 cm area	5/sec	70 hz
School Simulator	Brushing teeth	X swipe	Continuous upon collision	Continuous back and forth swipe, centered at object collision spot.	1 total	80 hz
School Simulator	Cigar	Persistent vibration	Continuous upon collision	At object collision.	1 total	90 hz
School Simulator	Cup of liquid	Z swipes	Continuous upon collision	At object collision.	1/sec	100 hz
Racetrack	Wind	X swipes	Continuous	Center start, two alternating nodes animate to sides of mouth.	1.5/sec	50 hz
Racetrack	Crate impact	Single impulse	1000ms	At object collision.	1 total	90 hz
Racetrack	Rain	Impulse train	Continuous	Random impulses inside 2x1 cm area, centered.	3.33/sec	40 hz
Racetrack	Puddle splash	Impulse train	500ms	Random impulses inside 3x1.5 cm area, centered.	20/sec	80 hz

Figure 8: Haptic parameters used in our example scenarios.

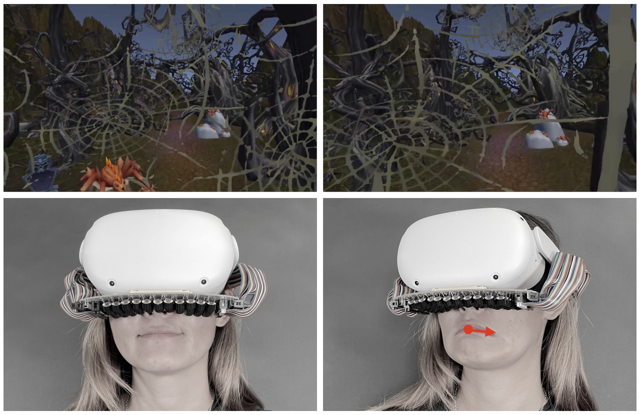


Figure 9: Walking through a spiderweb, which triggers an X-swipe.

Figure 10: Spider jumping onto the user's face, which triggers a series of haptic impulses.

consistent spatial judgements on the tongue, implying the effect was not particularly localized. Further, only one participant had a converging staircase on the teeth, indicating that the teeth also do not have very fine spatial resolution either.

Anecdotally, we note that our participants described the sensations at each location as feeling different: the lips felt more like a localized point of vibration, while haptics on the teeth were more often described as being a general sensation across the teeth and surrounding gums. For the tongue, participants would often experience a general vibration within the mouth that was not localized, sometimes saying that it felt like "the back of [my] throat was vibrating," which may be due to reflections in the oral cavity. The latter could allow for haptics on areas of the mouth without line-of-sight to the transducer array, and perhaps unlock other interesting haptic sensations in future work (e.g., vibrating the uvula).

In summary, our study characterized the perception thresholds for the lips, teeth and tongue. However, only the lips were sensitive enough to assess spatial acuity. These results directly informed the design of our haptic effects and animations, utilized in a series of example scenarios that we built, discussed next.

5 EXAMPLE SCENARIOS

To help convey the potential of our approach, we created three demo VR scenarios, each containing four interaction examples. We used these applications in our user experience study to gather feedback on the strength, expressivity, and overall success of our technique. Figure 8 provides a more detailed breakdown of each sensory examples' type, duration, positioning, generation rate, and modulation frequency (all demos used 100% amplitude). Please also see our Video Figure.

5.1 Haunted Forest

In this scenario, users must walk through spiderwebs to get to a small clearing. Upon touching a spiderweb, an X swipe is animated, simulating the spiderweb being drawn off the face (Figure 9). In the clearing, a spider jumps onto the user's mouth, which generates random impulses in an 1cm area that drifts to the side, simulating insect feet scurrying around and running off the side of the face (Figure 10). Later, the user must shoot a flare gun at a spider, which explodes into goo and splashes the user. This too utilizes random

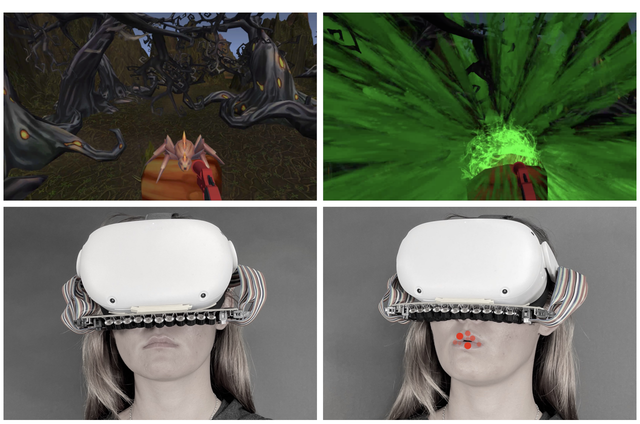


Figure 11: Shooting a spider, which explodes with goo, triggering as series of haptic impulses.

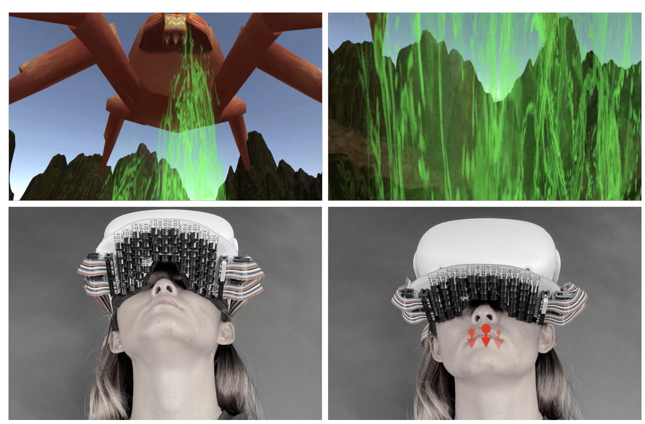


Figure 12: Walking into dripping spider venom, which triggers a series of Y-swipes.

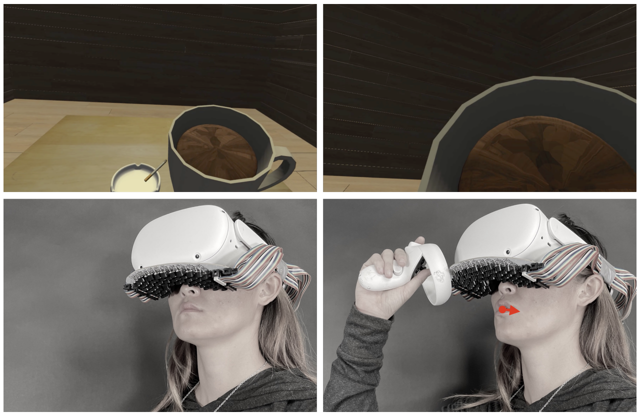


Figure 13: Drinking from a mug, which triggers continuous Z-swipes.

impulses, centered on the mouth and at a higher frequency of 70 hz (Figure 11). Finally, there is a large boss spider, dripping with venom that rains down onto the user, manifesting as random Y swipes on the mouth (Figure 12).

5.2 School Simulator

In this scenario, we created a school scene with various real-world objects. Users start off at a water fountain; if they can lean down and have their lips meet the water stream, they feel a series of rapidly generated impulses (Figure 1). Next, they find themselves in the break room where they sit down to drink some coffee, utilizing animated Z swipes to simulate liquid entering the mouth (Figure 13). They can also take a puff of a cigarette, which manifests as a high-frequency persistent vibration node (Figure 14) where the cigarette "sits" on the lips. Finally, they brush their teeth before returning to work, and our system animates a haptic node moving back and forth on the teeth when the toothbrush collides with the mouth (Figure 15).

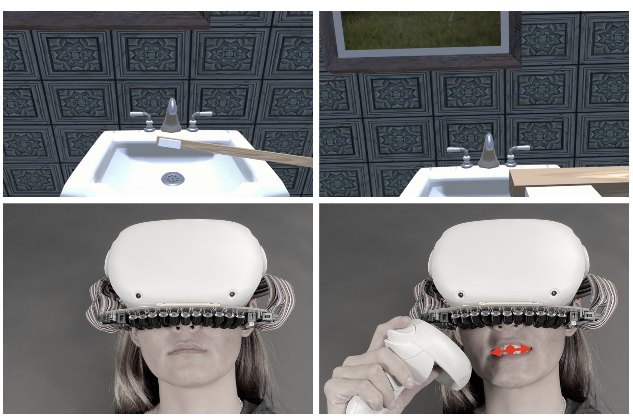


Figure 15: Brushing teeth uses back-and-forth X-swipes.

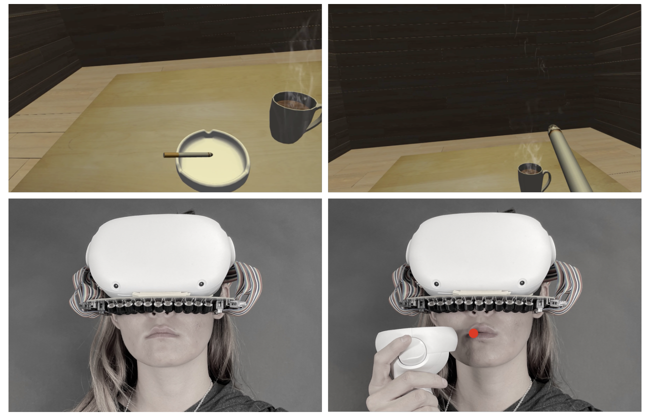


Figure 14: Taking a puff of a cigarette, which manifests as a continuous node.

5.3 Racing Game

Our final scenario is a motorcycle racing game. As the user rides down an open road, our system animates a "wind" effect by having impulses start at the center of the mouth and move outwards to the corners of the lips in an alternating pattern (Figure 16, left). On the track, there are boxes that break apart upon collision. When the user rides into them, they feel a single impulse on their lips for 1000ms (Figure 17). Later, it starts to rain, which manifests as a series of slow impulses that hit the user's mouth in random locations (Figure 16, right). After the storm clears there are puddles on the track, which can splash the user, manifesting as a series of rapid, random-location impulses on the lips (Figure 18).

6 USER EXPERIENCE STUDY

To gather feedback on the haptic effects we devised, as well as to evaluate the immersion and realism that comes from integrating our mouth haptic feedback, we ran a second user study focusing on subjective experience.

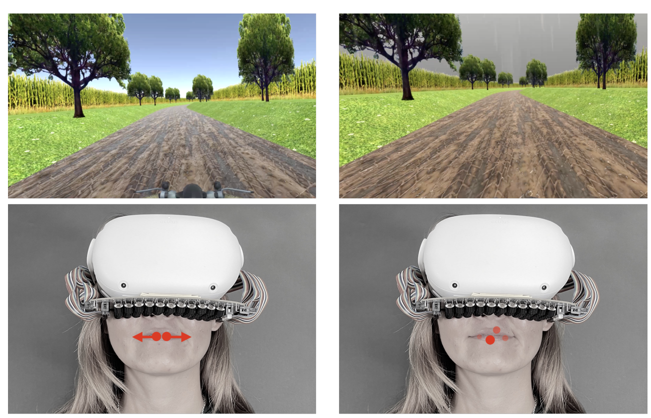


Figure 16: Left: Wind is animated onto the mouth using two alternating X-swipes. Right: Rain hits the mouth as a series of random impulses.

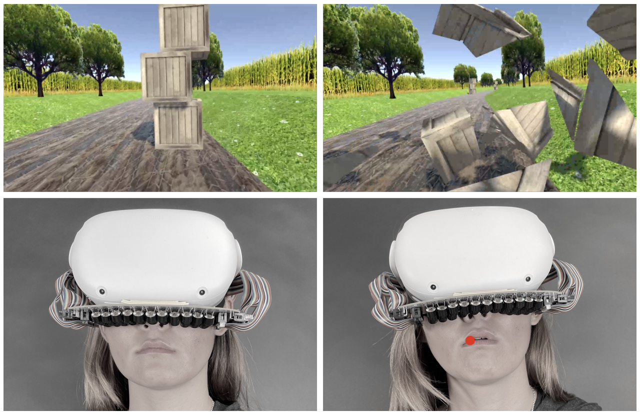


Figure 17: Colliding with a crate, which triggers a 500ms impulse.

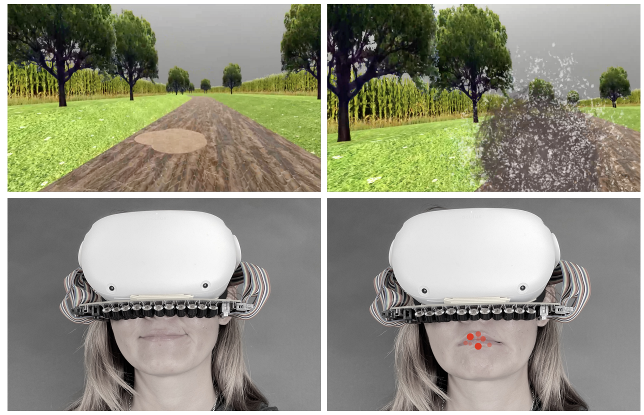


Figure 18: Splashing into a puddle, which triggers a series of short random impulses.

6.1 Procedure

We segmented each of the demo scenarios above into four different interactions, as seen in Figure 8. For each interaction, we ran two conditions: no-mouth haptics and with-mouth haptics, leading to 3 scenarios x 4 interactions x 2 conditions = 24 trials per participant. The only difference for the no haptics version of each trial was that our haptics array was not turned on, and did not generate any output. The order of the conditions, as well as the order of the example scenarios, was randomized across participants.

We recruited 8 participants (mean age of 25), who were compensated $20 for the 40-60 minute study. All participants had some prior VR experience, and two participants had previously taken our perception study and were familiar with the sensation of a persistent node. Most of the scenarios were done while seated in a chair, but three interactions (water fountain, spiderwebs, dripping goo) required the participant to stand up and move around. Participants were encouraged to think aloud during the study. Participants tried each scenario in two conditions (no feedback and mouth feedback; presentation order counterbalanced). Upon entering the scene, they were given a brief verbal description about the interaction, which they could repeat as many times as they wished. After each interaction, participants completed a six-question questionnaire verbally (to avoid having to remove the headset) with a seven-point Likert scale. At the end of the study, they filled out a 29 question questionnaire. Both questionnaires are described next.

6.2 Feedback Questionnaires

We included questionnaires to investigate two key questions: 1) Does having haptic feedback on the mouth increase immersion and realism? 2) Does our system enhance the overall experience? We utilized two different questionnaires as part of our user experience study: a six-question survey between trials, as well as a 29 question post-study survey. The inter-trial survey included four questions from the Embodiment Questionnaire [25], three of which evaluate tactile sensations (Q10, Q12, and Q13), with a fourth capturing reaction to external stimuli (Q23). In our subsequent results, we refer to these questions by the following shorthand: realism, localization, tactility, and reaction, respectively. We also included one question from the HX model [64], which asks about immersion. We included one final custom question: "Does the sensation/experience match the graphics?" All questions used a seven-point Likert scale ranging from strongly disagree (-3) to strongly agree (3), matching the format in prior haptics work [18, 25, 64]. This survey was given every time a participant completed a trial (i.e., 24 times in total).

The post-study survey (i.e., after all scenes were completed) consisted of 29 questions that also utilized a seven-point Likert scale (-3 to 3). The first seven questions were modeled on Fang and Harrison [18], and consisted of 1) The *with-mouth feedback* examples felt realistic; 2) The *no-mouth feedback* examples felt realistic; 3) The *no-mouth feedback* examples made me feel more immersed in the scene; 4) The *with-mouth feedback* examples made me feel more immersed in the scene; 5) The *no-mouth feedback* examples were fun; 6) The *with-mouth feedback* examples were fun; and 7) I preferred the *no-mouth feedback* over *with-mouth feedback* examples. To mitigate order effects, we counterbalanced the ordering of the *no-mouth* and *with-mouth* conditions in the questions between participants. The other twenty-two questions were taken directly from Sathiyamurthy et al. [64], which are used to assess the five haptic experience dimensions: autotelic, expressivity, immersion, realism, and harmony. We note that validation of this scale was ongoing at the time of writing.

6.3 Results

Figure 19 provides the mean ratings of our inter-trial survey questions, on which our haptic condition uniformly outperformed experiences without haptics. In short, and perhaps unsurprisingly, haptics boosts realism and immersion, and heightens user reactions. Participants rated tactility and localization highly, underscoring the fine-grained nature of our technique and animations. In particular, haptic sensations helped with VR events that fell outside the user's field of view: "without haptics it was difficult to tell when things were supposed to be touching the face [...] I don't even know when the cigarette is in my mouth" (P4), and "haptic effects made those virtual actions really cool" (P6). Our system rated well in visuo-haptic match, though some participants remarked on scenarios that

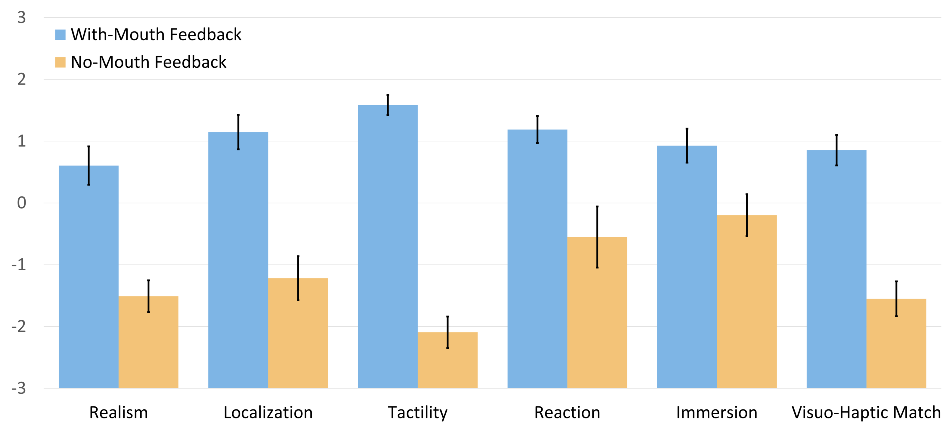


Figure 19: Results from our inter-trial questionnaire. Error bars are standard error.

were less convincing: "even though the haptic effect was interesting and well placed, the sensations themselves didn't necessarily match [my] expectations" (P8) (e.g., wetness from splashes cannot be realistically simulated with vibrotactile actuation alone).

The first seven questions of our post-study questionnaire directly compared our haptic feedback to no haptic feedback (Figure 20). As before, our system rated highly in realism and immersion. Participants rated both haptic and no haptic versions of our demos as fun. In our forced choice question, every single participant preferred our system, averaging 2.0 ("agree") on the Likert scale. In the 22 questions from Sathiyamurthy et al. [64], our system performed best in the autotelic and harmony categories (21). The worst performing category was expressivity, which pertains to how different the feedback felt across the presented scenarios. As already noted above, our system is fundamentally limited to delivering vibrotactile actuation to the mouth, which cannot fully realize the sensation of many experiences. Nonetheless, our user feedback results clearly show that even this reduced set of haptic effects can still enhance VR/AR experiences.

7 FUTURE WORK AND LIMITATIONS

While our approach and proof-of-concept hardware demonstrate promise, challenges and future work remain. While relatively compact, our phased array was still 15mm thick, adding about 16.5% to the total height of the headset. This also had the negative effect of bringing the transducers closer to the mouth – with an 80° emission cone, it meant that many transducers could not participate in the beamforming pattern, reducing the haptic effect strength. Weight is also a significant issue for VR headsets, especially with respect to rotational inertia. Our transducer array weighs 107g, adding 15.8% to the Oculus Quest 2 with Elite Strap's mass of 676g (we do not include the mass of the Arduino and driver board, as these would be rendered moot or significantly optimized in a commercial implementation).

Fortunately, there are several avenues open for size and weight miniaturization. Most promising is moving to smaller and lighter transducers. Our most recent prototype uses surface mount Murata MA40H1S-R transducers, measuring just 5.2 x 5.2 x 1.15mm each and weighing just 0.8g. These are less powerful than the open-can-type transducers we used. However, in tests, we found that

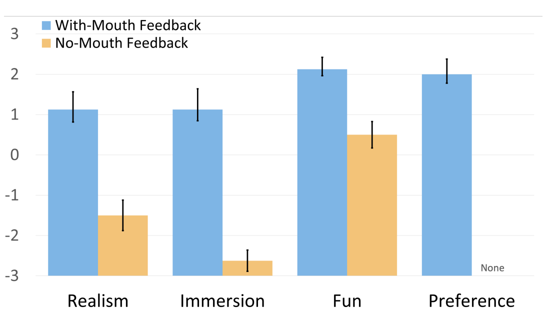


Figure 20: Mean Likert scores of the seven-question study adapted from Fang and Harrison [18]. Error bars are standard error.

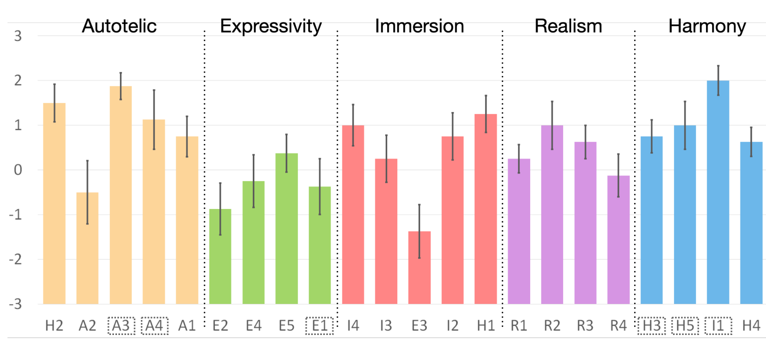


Figure 21: Mean Likert scores of the 22 questions from Sathiyamurthy et al. [64]; see this work for key to X-axis categories. Negatively weighted questions are denoted with square boxes, and their scores have been flipped. Error bars are standard error.

Figure 22: Our latest hardware using thin form factor ultrasonic transducers in a similar array pattern.

clustering four together (acting as a single emitter) running at 10 VRMS can achieve comparable sound pressure levels (122.7 dB SPL) to one can-type transducer of our original array (118.7 dB SPL). Thus our latest design, seen in Figure 22, consists of 64 groups of transducer quartets. The only downside of this new arrangement is the planar configuration of the surface mount transducers, making it hard to focus energy towards the mouth. Our next design will include angled daughter boards to achieve a similar geometry to the can-type array used in our earlier studies. Another option to increase haptic strength is to use a domed transducer array, though this would increase thickness. In regards to impact on battery life, our technique is already performing well. As reported earlier, one hour of haptically-enhanced gameplay will consume roughly 2.71% of an Oculus Quest 2's battery, and further optimization should get this down to under 1%.

One of the difficulties we encountered with our system was the manual calibration required to find the relative position of the phased array and user's lips. For this, we simply moved a persistent node in 1mm steps on every axis until the user reported the sensation was well centered. For most of our participants, this only required a small tweak of our starting values. However, for a quarter of our participants, a more extensive search was required due to significantly different face morphologies. In the future, this calibration could be automated by having a programmatic raster-scan-like process (i.e., where the user presses a button when they feel the sensation), or perhaps by putting a small multi-zone range-finding sensor onto our PCB (e.g., STMicroelectronics' VL53L1).

Finally, although we demonstrated a range of haptic effects, the vocabulary of sensations is still limited compared to that of the real world (a high bar). That said, it is roughly comparable to that of vibration motors in handheld controllers and many mobile phones, but with more spatial expressivity. Nonetheless, we believe VR/AR systems should strive for greater realism, and future work is necessary to expand upon our work. We note that our search was not exhaustive and that more advanced systems could experiment with e.g., fluctuating modulation frequencies as a dynamic parameter. Similarly, multi-node haptic effects are entirely possible with additional computation. We also note that our method can complement other VR/AR haptic approaches, such as thermal and vibratory actuators built into headsets. We also reiterate that 3 out of 19 participants could not feel the haptic effect, and so work remains to study if this is truly a technology accessible to all.

8 CONCLUSION

In this paper, we presented a novel implementation of mouth based haptics in VR using an ultrasonic phased array, developed as a thin, compact component that can be attached to a headset. In a perception study, we quantified the parameters at which our haptic system works best across the lips, teeth, and tongue. We defined a library of haptic animations that our system is capable of rendering onto the mouth, including swipes, impulses, and persistent vibrations, and we illustrated potential uses across several interaction scenarios. After running participants through these scenarios, our user experience study results indicate that our system enhances realism and immersion in virtual reality. Additionally, participants uniformly preferred using our system to having no haptic feedback at all, signalling that mouth haptics could be an engaging addition to consumer VR systems.

ACKNOWLEDGMENTS

We would like to thank Asier Marzo and UpnaLab, Universidad Pública de Navarra, for open sourcing and providing guidance for their driver board and firmware.

REFERENCES

[1] Ahmed Al Maimani and Anne Roudaut. 2017. Frozen suit: designing a changeable stiffness suit and its application to haptic games. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2440–2448.

[2] Mohammed Al-Sada, Keren Jiang, Shubhankar Ranade, Mohammed Kalkattawi, and Tatsuo Nakajima. 2020. HapticSnakes: multi-haptic feedback wearable robots for immersive virtual reality. *virtual reality* 24, 2 (2020), 191–209.

[3] Arduino. 2021. Mega 2560 Rev3. https://store-usa.arduino.cc/products/arduino-mega-2560-rev3

[4] bHaptics. 2021. https://www.bhaptics.com

[5] M. Bukowska, G. K. Essick, and Mats Trulsson. 2010. Functional properties of low-threshold mechanoreceptive afferents in the human labial mucosa. *Experimental Brain Research* 201, 1 (Feb. 2010), 59–64. https://doi.org/10.1007/s00221-009-2005-0 Number: 1.

[6] Andres Andres Alejandre Calvo. 2017. *A body-grounded kinesthetic haptic device for virtual reality.* Ph.D. Dissertation. Massachusetts Institute of Technology.

[7] Samuele Carcagno, Andrew Di Battista, and Christopher J Plack. 2019. Effects of high-intensity airborne ultrasound exposure on behavioural and electrophysiological measures of auditory function. *Acta Acustica united with Acustica* 105, 6 (2019), 1183–1197.

[8] Sylvain Cardin, Daniel Thalmann, and Frederic Vexo (Eds.). 2007. Head Mounted Wind. *Proceeding of the 20th annual conference on Computer Animation and Social Agents (CASA2007)* (2007).

[9] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. 2013. UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST '13).* Association for Computing Machinery, New York, NY, USA, 505–514. https://doi.org/10.1145/2501988.2502018

[10] Hong-Yu Chang, Wen-Jie Tseng, Chia-En Tsai, Hsin-Yu Chen, Roshan Lalintha Peiris, and Liwei Chan. 2018. FacePush: Introducing Normal Force on Face with Head-Mounted Displays. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18).* Association for Computing Machinery, New York, NY, USA, 927–935. https://doi.org/10.1145/3242587.3242588

[11] Zikun Chen, Roshan Lalintha Peiris, and Kouta Minamizawa. 2017. A Thermal Pattern Design for Providing Dynamic Thermal Feedback on the Face with Head Mounted Displays. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction (TEI '17).* Association for Computing Machinery, New York, NY, USA, 381–388. https://doi.org/10.1145/3024969.3025060

[12] Zikim Chen, Wei Peng, Roshan Peiris, Liwei Chan, and Kouta Minamizawa. 2017. Thermoreality: thermally enriched head mounted displays. In *SIGGRAPH Asia 2017 Emerging Technologies.* ACM, Bangkok Thailand, 1–2. https://doi.org/10.1145/3132818.3132823

[13] Integrated Circuits. 2021. IX4427N. https://www.digikey.com/en/products/detail/ixys-integrated-circuits-division/IX4427N/4009906

[14] Giulia Corniani and Hannes P Saal. 2020. Tactile innervation densities across the whole body. *Journal of Neurophysiology* 124, 4 (2020), 1229–1240.

[15] Victor Adriel de Jesus Oliveira, Luca Brayda, Luciana Nedel, and Anderson Maciel. 2017. Designing a Vibrotactile Head-Mounted Display for Spatial Awareness in 3D Spaces. *IEEE Transactions on Visualization and Computer Graphics* 23, 4 (April 2017), 1409–1417. https://doi.org/10.1109/TVCG.2017.2657238 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[16] Alexandra Delazio, Ken Nakagaki, Roberta L Klatzky, Scott E Hudson, Jill Fain Lehman, and Alanson P Sample. 2018. Force jacket: Pneumatically-actuated jacket for embodied haptic experiences. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–12.

[17] Andrew Di Battista. 2019. *The effect of 40 kHz ultrasonic noise exposure on human hearing.* Universitätsbibliothek der RWTH Aachen.

[18] Cathy Fang and Chris Harrison. 2021. Retargeted Self-Haptics for Increased Immersion in VR without Instrumentation.

[19] FEELREAL. 2021. Multisensory VR Mask. https://feelreal.com

[20] Euan Freeman, Ross Anderson, Julie Williamson, Graham Wilson, and Stephen A. Brewster. 2017. Textured surfaces for ultrasound haptic displays. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI '17).* Association for Computing Machinery, New York, NY, USA, 491–492. https://doi.org/10.1145/3136755.3143020

[21] Euan Freeman and Graham Wilson. 2021. Perception of ultrasound haptic focal point motion. In *Proceedings of the 2021 International Conference on Multimodal Interaction.* 697–701.

[22] L. R. Gavrilov. 2008. The possibility of generating focal regions of complex configurations in application to the problems of stimulation of human receptor structures by focused ultrasound. *Acoustical Physics* 54, 2 (March 2008), 269–278. https://doi.org/10.1134/S1063771008020152 Number: 2.

[23] Orestis Georgiou, Craig Jeffrey, Ziyuan Chen, Bao Xiao Tong, Shing Hei Chan, Boyin Yang, Adam Harwood, and Tom Carter. 2018. Touchless haptic feedback for VR rhythm games. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, 553–554.

[24] Hyunjae Gil, Hyungki Son, Jin Ryong Kim, and Ian Oakley. 2018. Whiskers: Exploring the use of ultrasonic haptic cues on the face. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–13.

[25] Mar Gonzalez-Franco and Tabitha C Peck. 2018. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI* 5 (2018), 74.

[26] Jan Gugenheimer, Dennis Wolf, Eythor R. Eiriksson, Pattie Maes, and Enrico Rukzio. 2016. GyroVR: Simulating Inertia in Virtual Reality using Head Worn Flywheels. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16).* Association for Computing Machinery, New York, NY, USA, 227–232. https://doi.org/10.1145/2984511.2984535

[27] Sidhant Gupta, Dan Morris, Shwetak N Patel, and Desney Tan. 2013. Airwave: Non-contact haptic feedback using air vortex rings. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing.* 419–428.

[28] Patrick Haggard and Lieke de Boer. 2014. Oral somatosensory awareness. *Neuroscience and Biobehavioral Reviews* 47 (Nov. 2014), 469–484. https://doi.org/10.1016/j.neubiorev.2014.09.015

[29] Ping-Hsuan Han, Yang-Sheng Chen, Chiao-En Hsieh, Hao-Cheng Wang, and Y. Hung. 2019. Hapmosphere: Simulating the Weathers for Walking Around in Immersive Environment with Haptics Feedback. *2019 IEEE World Haptics Conference (WHC)* (2019). https://doi.org/10.1109/WHC.2019.8816140

[30] Ping-Hsuan Han, Yang-Sheng Chen, Kong-Chang Lee, Hao-Cheng Wang, Chiao-En Hsieh, Jui-Chun Hsiao, C. Chou, and Y. Hung. 2018. Haptic around: multiple tactile sensations for immersive environment and interaction in virtual reality. *VRST* (2018). https://doi.org/10.1145/3281505.3281507

[31] Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, and Hiroyuki Shinoda. 2010. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Transactions on Haptics* 3, 3 (July 2010), 155–165. https://doi.org/10.1109/TOH.2010.4 Number: 3 Conference Name: IEEE Transactions on Haptics.

[32] F. Hülsmann, N. Mattar, Julia Fröhlich, and I. Wachsmuth. 2014. Simulating Wind and Warmth in Virtual Reality: Conception, Realization and Evaluation for a CAVE Environment. *J. Virtual Real. Broadcast.* (2014).

[33] Takayuki Iwamoto, Mari Tatezono, and Hiroyuki Shinoda. 2008. Non-contact Method for Producing Tactile Sensation Using Airborne Ultrasound. In *Haptics: Perception, Devices and Scenarios (Lecture Notes in Computer Science)*, Manuel Ferre (Ed.). Springer, Berlin, Heidelberg, 504–513. https://doi.org/10.1007/978-3-540-69057-3_64

[34] Arata Jingu, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. 2021. Tactile Perception Characteristics of Lips Stimulated by Airborne Ultrasound. In *2021 IEEE World Haptics Conference (WHC).* IEEE, 607–612.

[35] Arata Jingu, Takaaki Kamigaki, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. 2021. LipNotif: Use of Lips as a Non-Contact Tactile Notification Interface Based on Ultrasonic Tactile Presentation. In *The 34th Annual ACM Symposium on User Interface Software and Technology.* 13–23.

[36] Roland S Johansson, Mats Trulsson, KÅ Olsson, and K-G Westberg. 1988. Mechanoreceptor activity from the human face and oral mucosa. *Experimental brain research* 72, 1 (1988), 204–208.

[37] Takayuki Kameoka and Hiroyuki Kajimoto. 2021. Tactile Transfer of Finger Information through Suction Tactile Sensation in HMDs. In *2021 IEEE World Haptics Conference (WHC).* IEEE, 949–954.

[38] Takayuki Kameoka, Yuki Kon, Takuto Nakamura, and Hiroyuki Kajimoto. 2018. Haptopus: Haptic VR Experience Using Suction Mechanism Embedded in Head-mounted Display. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings (UIST '18 Adjunct).* Association for Computing Machinery, New York, NY, USA, 154–156. https://doi.org/10.1145/3266037.3271634

[39] Oliver Beren Kaul and Michael Rohs. 2017. HapticHead: A Spherical Vibrotactile Grid around the Head for 3D Guidance in Virtual and Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 3729–3740. https://doi.org/10.1145/3025453.3025684

[40] Cédric Kervegant, Félix Raymond, Delphine Graeff, and Julien Castet. 2017. Touch hologram in mid-air. In *ACM SIGGRAPH 2017 Emerging Technologies.* 1–2.

[41] KnoxLabs. 2021. Tactal Haptic Face Cushion for HMDs. https://www.knoxlabs.com/products/tactal-haptic-cushion-for-face

[42] Yukari Konishi, Nobuhisa Hanamitsu, Benjamin Outram, Kouta Minamizawa, Tetsuya Mizuguchi, and Ayahiko Sato. 2016. Synesthesia suit: the full body immersive experience. In *ACM SIGGRAPH 2016 VR Village.* 1–1.

[43] Vincent Levesque, Louise Oram, Karon MacLean, Andy Cockburn, Nicholas D. Marchuk, Dan Johnson, J. Edward Colgate, and Michael A. Peshkin. 2011. Enhancing physicality in touch interaction with programmable friction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 2481–2490. https://doi.org/10.1145/1978942.1979306

[44] H. Levitt. 1971. Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America* 49, 2B (Feb. 1971), 467–477. https://doi.org/10.1121/1.1912375 Number: 2B Publisher: Acoustical Society of America.

[45] Benjamin Long, Sue Ann Seah, Tom Carter, and Sriram Subramanian. 2014. Rendering volumetric haptic shapes in mid-air using ultrasound. *ACM Transactions on Graphics* 33, 6 (Nov. 2014), 181:1–181:10. https://doi.org/10.1145/2661229.2661257 Number: 6.

[46] VR Electronics Ltd. [n.d.]. TeslaSuit.

[47] Jonatan Martinez, Daniel Griffiths, Valerio Biscione, Orestis Georgiou, and Tom Carter. 2018. Touchless haptic feedback for supernatural VR experiences. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, 629–630.

[48] Asier Marzo, Tom Corkett, and Bruce W. Drinkwater. 2018. Ultraino: An Open Phased-Array System for Narrowband Airborne Ultrasound Transmission. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 65, 1 (Jan. 2018), 102–111. https://doi.org/10.1109/TUFFC.2017.2769399 Number: 1 Conference Name: IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control.

[49] Saya Mizutani, Masahiro Fujiwara, Yasutoshi Makino, and Hiroyuki Shinoda. 2019. Thresholds of Haptic and Auditory Perception in Midair Facial Stimulation. In *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*. 1–6. https://doi.org/10.1109/HAVE.2019.8920999

[50] Yasuaki Monnai, Keisuke Hasegawa, Masahiro Fujiwara, Kazuma Yoshino, Seki Inoue, and Hiroyuki Shinoda. 2015. Adding Texture to Aerial Images Using Ultrasounds. In *Haptic Interaction: Perception, Devices and Applications*, Hiroyuki Kajimoto, Hideyuki Ando, and Ki-Uk Kyung (Eds.). Springer Japan, Tokyo, 59–61. https://doi.org/10.1007/978-4-431-55690-9_11

[51] Joe Mullenbach, Craig Shultz, J. Edward Colgate, and Anne Marie Piper. 2014. Exploring affective communication through variable-friction surface haptics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3963–3972. https://doi.org/10.1145/2556288.2557343

[52] Murata. 2021. MA40S4S Ultrasonic Sensors. https://www.murata.com/en-us/products/productdetail?partno=MA40S4S

[53] Oculus. 2021. Oculus Quest 2. https://www.oculus.com/quest-2/

[54] Oculus. 2021. Quest 2 Product Description. https://www.oculus.com/legal/product-information-sheet-oculus-quest-2

[55] Victor Adriel de Jesus Oliveira, Luca Brayda, Luciana Nedel, and Anderson Maciel. 2017. Experiencing guidance in 3D spaces with a vibrotactile head-mounted display. IEEE Computer Society, 453–454. https://doi.org/10.1109/VR.2017.7892375 ISSN: 2375-5334.

[56] Roshan Lalintha Peiris, Liwei Chan, and Kouta Minamizawa. 2018. LiquidReality: Wetness Sensations on the Face for Virtual Reality. In *Haptics: Science, Technology, and Applications (Lecture Notes in Computer Science)*, Domenico Prattichizzo, Hiroyuki Shinoda, Hong Z. Tan, Emanuele Ruffaldi, and Antonio Frisoli (Eds.). Springer International Publishing, Cham, 366–378. https://doi.org/10.1007/978-3-319-93399-3_32

[57] Roshan Lalintha Peiris, Wei Peng, Zikun Chen, Liwei Chan, and Kouta Minamizawa. 2017. ThermoVR: Exploring Integrated Thermal Haptic Feedback with Head Mounted Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5452–5456. https://doi.org/10.1145/3025453.3025824

[58] Yi-Hao Peng, Carolyn Yu, Shi-Hong Liu, Chung-Wei Wang, Paul Taele, Neng-Hao Yu, and Mike Y. Chen. 2020. WalkingVibe: Reducing Virtual Reality Sickness and Improving Realism while Walking in VR using Unobtrusive Head-mounted Vibrotactile Feedback. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376847

[59] Ismo Rakkolainen, Euan Freeman, Antti Sand, Roope Raisamo, and Stephen Brewster. 2021. A Survey of Mid-Air Ultrasound Haptics and Its Applications. *IEEE Transactions on Haptics* 14, 1 (Jan. 2021), 2–19. https://doi.org/10.1109/TOH.2020.3018754 Number: 1 Conference Name: IEEE Transactions on Haptics.

[60] Nimesha Ranasinghe, Pravar Jain, Nguyen Thi Ngoc Tram, Koon Chuan Raymond Koh, David Tolley, Shienny Karwita, Lin Lien-Ya, Yan Liangkun, Kala Shamaiah, Chow Eason Wai Tung, Ching Chiuan Yen, and Ellen Yi-Luen Do. 2018. Season Traveller: Multisensory Narration for Enhancing the Virtual Reality Experience. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3174151

[61] Michael Rietzler, Katrin Plaumann, Taras Kränzle, Marcel Erath, Alexander Stahl, and E. Rukzio. 2017. VaiR: Simulating 3D Airflows in Virtual Reality. *CHI* (2017). https://doi.org/10.1145/3025453.3026009

[62] Antti Sand, Ismo Rakkolainen, Poika Isokoski, Jari Kangas, Roope Raisamo, and Karri Palovuori. 2015. Head-mounted display with mid-air tactile feedback. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*. 51–58.

[63] K Sathian and A Zangaladze. 1996. Tactile spatial acuity at the human fingertip and lip: bilateral symmetry and interdigit variability. *Neurology* 46, 5 (1996), 1464–1464.

[64] Suji Sathiyamurthy, Melody Lui, Erin Kim, and Oliver Schneider. 2021. Measuring Haptic Experience: Elaborating the HX model with scale development. In *2021 IEEE World Haptics Conference (WHC)*. 979–984. https://doi.org/10.1109/WHC49131.2021.9517220

[65] Rajinder Sodhi, Ivan Poupyrev, Matthew Glisson, and Ali Israr. 2013. AIREAL: interactive tactile experiences in free air. *ACM Transactions on Graphics (TOG)* 32, 4 (2013), 1–10.

[66] Ryoko Takahashi, Keisuke Hasegawa, and Hiroyuki Shinoda. 2018. Lateral modulation of midair ultrasound focus for intensified vibrotactile stimuli. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 276–288.

[67] Ryoko Takahashi, Keisuke Hasegawa, and Hiroyuki Shinoda. 2020. Tactile Stimulation by Repetitive Lateral Movement of Midair Ultrasound Focus. *IEEE transactions on haptics* 13, 2 (June 2020), 334–342. https://doi.org/10.1109/TOH.2019.2946136

[68] ThermoReal. 2021. http://thermoreal.com

[69] M. Trulsson. 2006. Sensory-motor function of human periodontal mechanoreceptors. *Journal of Oral Rehabilitation* 33, 4 (April 2006), 262–273. https://doi.org/10.1111/j.1365-2842.2006.01629.x Number: 4.

[70] Ultraino. 2018. ArduinoMega64 Firmware. https://github.com/asiermarzo/Ultraino

[71] Ultraleap. 2021. https://www.ultraleap.com

[72] Unity. 2021. Unity Development Platform. https://unity.com

[73] Robert W Van Boven and Kenneth O Johnson. 1994. The limit of tactile spatial resolution in humans: grating orientation discrimination at the lip, tongue, and finger. *Neurology* 44, 12 (1994), 2361–2361.

[74] Chi Wang, Da-Yuan Huang, Shuo-wen Hsu, Chu-En Hou, Yeu-Luen Chiu, Ruei-Che Chang, Jo-Yu Lo, and Bing-Yu Chen. 2019. Masque: Exploring Lateral Skin Stretch Feedback on the Face with Head-Mounted Displays. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 439–451. https://doi.org/10.1145/3332165.3347898

[75] Ke Wang, Yi-Hsuan Li, Chun-Chen Hsu, Jiabei Jiang, Yan Liu, Zirui Zhao, Wei Yue, and Lu Yao. 2021. A Research on Sensing Localization and Orientation of Objects in VR with Facial Vibrotactile Display. In *Virtual, Augmented and Mixed Reality (Lecture Notes in Computer Science)*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Springer International Publishing, Cham, 223–241. https://doi.org/10.1007/978-3-030-77599-5_17

[76] Alexander Wilberz, Dominik Leschtschow, Christina Trepkowksi, Jens Maiero, Ernst Kruijff, and Bernhard Riecke. 2020. FaceHaptics: Robot Arm based Versatile Facial Haptics for Immersive Environments. (2020), 14.

[77] Graham Wilson, Thomas Carter, Sriram Subramanian, and Stephen A. Brewster. 2014. Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1133–1142. https://doi.org/10.1145/2556288.2557033

[78] Laura Winfield, John Glassmire, J. Edward Colgate, and Michael Peshkin. 2007. T-PaD: Tactile Pattern Display through Variable Friction Reduction. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*. 421–426. https://doi.org/10.1109/WHC.2007.105

[79] Dennis Wolf, Leo Hnatek, and Enrico Rukzio. 2018. Face/On: Actuating the Facial Contact Area of a Head-Mounted Display for Increased Immersion. https://doi.org/10.1145/3266037.3271631 Pages: 146-148.