
There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/225807/

Deposited on 30 October 2020

Enlighten – Research publications by members of the University of Glasgow

http://eprints.gla.ac.uk
A Survey of Mid-Air Ultrasound Haptics and Its Applications

Ismo Rakkolainen1, Euan Freeman2, Antti Sand3, Roope Raisamo1, Stephen Brewster2

Abstract— Ultrasound haptics is a contactless haptic technology that enables novel mid-air interactions with rich multisensory feedback. This paper surveys recent advances in ultrasound haptic technology. We discuss the fundamentals of this haptic technology, how a variety of perceptible sensations are rendered, and how it is currently being used to enable novel interaction techniques. We summarize its strengths, weaknesses, and potential applications across various domains. We conclude with our perspective on key directions for this promising haptic technology.

Index Terms— Ultrasound haptics, Haptics, Human computer interaction, Mid-air haptics.

I. INTRODUCTION

Haptic devices enable human-computer interfaces to create rich and visceral digital sensations pertaining to the sense of touch. The ability to recreate real physical sensations – or create entirely new ones – can enrich communication between computer and human and improve the way we interact. Most haptic devices require physical contact between an actuator or end effector and the skin, however there are several emerging alternative technologies that allow haptic sensations to be perceived without physical contact.

Contactless haptic devices are compelling because they offer the potential to create tactile sensations without direct physical contact or the need to wear a device that may disrupt feelings of immersion and presence. They are well suited to mid-air user interfaces, where users interact with digital content and user interface controls through hand and finger movements in air. A notable usability issue with such interfaces is that they lack implicit tactile cues experienced when interacting through touch [1], [2]. The ability to ‘feel’ content in mid-air is therefore desirable, as it can address fundamental usability challenges with gesture interfaces [2]–[4] and improve user engagement [5], amongst other benefits.

Ultrasound haptic feedback [6], [7] is one example of a contactless haptic technology that has attracted considerable attention in the human-computer interaction (HCI) and haptics literature. This technology has also received great interest from industry, particularly in the areas of digital advertising [5], [8], automotive user interfaces [9]–[12] and Virtual and Augmented Reality (VR and AR) [13]–[16]. In the past decade, ultrasound haptic feedback devices have become more accessible and available through commercial (e.g., Ultradev [17]) and open source initiatives [18], allowing this technology to expand into new application areas.

As well as enabling non-contact haptic sensations, ultrasound haptic devices have unique advantages that expand the range of sensations which can be created. Haptic sensations are transmitted to the skin through precisely focused sound waves, so multiple points of stimulation can be created and moved without the constraints imposed by a physical end effector. These sensations can cover a wide range, stimulating all of the hand, for example, within a relatively large workspace. Finally, high sample rates and the speed of sound enable a high degree of temporal precision, which can lead to novel tactile sensations that are not possible with some alternative technologies.

Early research on ultrasound haptics focused on fundamental technical and hardware aspects of this new haptic technology. As the technology progressed and became more accessible to a wider audience, newer research has started to investigate the perceptual aspects of these novel haptic sensations and applied research in specific domains has become increasingly common. As a new approach for creating tactile sensations, however, it is not yet clear, which are its most compelling use scenarios, user needs and potential applications.

There are many surveys on haptics in general (e.g., [19]–[22]), and recent surveys on haptics for VR [23], [24]. Whilst they all briefly mention mid-air ultrasound haptics, there is a real need for a comprehensive survey on this topic, since there is a rapidly growing body of research, with commercially available devices targeted at consumers and industry alike. This article addresses this gap, with a detailed survey of mid-air ultrasound haptic feedback. There are existing state-of-the-art surveys on phased array ultrasound that we are aware of. A review from 2012 [25] discusses mostly experiments in the MHz range for the purposes of physiology and medicine. A short HCI-oriented survey from 2015 [26] and a survey on non-contact haptics from 2013 [27] are outdated as technology and rendering methods are advancing rapidly. A recent survey [28] gives a succinct overview of the topic, but lacks depth and breadth of discussion, which we aim to give here.

Our aims with this survey are to introduce ultrasound haptic feedback in an accessible way, provide an overview of current research into its perception and use in HCI, discuss practical issues around its usage and deployment, and reflect on its...
strengths and weaknesses, drawing on our own expertise with this technology. We finish with a look to the future for ultrasound haptics, giving our view on key research directions and on the areas where it can have a significant impact on how we interact with digital information and services.

II. CONTACTLESS HAPTIC TECHNOLOGIES

Cutaneous tactile feedback is often provided in direct contact to skin through physical end effectors. These may be dedicated haptic devices such as tactile gloves or exoskeletons, or other devices (e.g., smartphones, video game controllers) that contain embedded actuators for presenting contact vibration to the hands. Another class of haptic device provides kinesthetic force feedback, again requiring physical contact between the user and end effector. End effectors must be touched by the user, moved against the skin (e.g., with actuated arms), or worn on the body. Some of these approaches may be obtrusive or inconvenient, or poorly suited to certain interaction contexts. Novel approaches, such as wireless, foldable, lightweight haptic epidermal patches [29] are less cumbersome, but contactless haptics without the need to wear anything is a compelling challenge.

Several technologies have emerged for presenting cutaneous haptic information without direct physical contact with an end effector. Unlike contact-based haptic devices, there are no mechanically constrained end effectors – making it possible to create complex tactile sensations with a fine-grained spatial resolution and a high temporal resolution. Our survey begins with a brief overview of contactless haptic approaches, before turning our attention to ultrasound haptics – the predominant contactless haptic technology in use today.

By necessity, a contactless haptic device must transfer tactile sensations through air to the skin, resulting in a perceptible tactile experience. A simple approach is to use controlled air streams to exert pressure against the skin: e.g., from fans or pressurized air jets [30], [31]. Whilst this method has limited control over the spatial resolution of the tactile sensations, it can produce relatively strong and continuous forces.

A more nuanced alternative used subwoofers in an enclosed space to compress air through a narrow aperture, yielding compressed air vortex rings that maintain pressure over distances of up to 3m [32]–[34]. Upon contact, the pressure is transferred to the skin or clothes. There is an inherent latency as the vortex travels through the air, and continuous sensations are not possible. Spatial resolution is also relatively coarse (e.g., AIREAL [33] made vortices with a diameter of 8.5 cm).

Contactless tactile sensations can also be presented using lasers [35]–[38] and electric arcs [39]. These induce tactile and thermal effects with a very high spatial resolution, although range is often limited (e.g., the electric arcs in Sparkle were only 6 mm in length [39]). Distal thermal effects have been created, leveraging the skin’s ability to feel heat from a distance, e.g., HeatHapt [40] directed the heat generated from a lamp towards the hand. Electromagnetic fields have also been used to transfer haptic sensations to the body. These typically attract or repel magnetic implements worn on the hand [41], [42]. Body hair has also been treated with magnetic gel or wax so that it becomes responsive to magnetic fields, resulting in perceptible sensations from the hair follicles [43].

The methods discussed in this section allow the delivery of haptic sensations without physical contact with a device or actuator. In most cases, additional hardware does not need to be worn by the user, resulting in a truly contactless haptic display. However, these methods suffer from a range of limitations. An alternative contactless haptic technology, which we focus on in the rest of this paper, uses ultrasound to exert pressure on the skin [6], [7]. Ultrasound haptic devices have recently received a lot of academic and commercial interest and it is an active area of research in the HCI and haptics communities. The main advantages of this technology compared are that it allows multiple points of stimulation, a high degree of spatial and temporal resolution, and almost instantaneous and continuous presentation over a larger workspace.

III. ULTRASOUND HAPTIC FUNDAMENTALS

A. Phased Ultrasound Arrays for Mid-Air Haptics

Ultrasound is vibration of air that propagates as a pressure wave with frequencies higher than the upper limit of human hearing (~20 kHz). Focusing dozens or hundreds of waves from an array of emitters towards a single ‘focal point’ increases the achievable amplitude significantly. When an obstacle (e.g., a person’s hand) comes into contact with the focal point, acoustic radiation pressure arises as a non-linear phenomenon of in-air ultrasound [6]; i.e., the energy from the sound waves results in positive pressure as they are reflected off the skin.

Phased-array focusing techniques are used to produce focal points. By independently controlling the phase of each emitted wave, their amplitude peaks are timed to arrive synchronously at a given location, where they constructively interfere to create a focal point with cumulative amplitude. Fig. 1 illustrates this principle; note that waves are emitted from outermost elements first, such that they arrive at the target point in synchrony with the waves from the innermost elements of the array.

Fig. 1. This image sequence demonstrates the principle of focusing ultrasound from a phased array of eight transducers. The phase of each wave is offset such that they arrive at the desired point at the same time.

The first mid-air ultrasound haptics device was presented by Iwamoto et al. [6]. They produced a hexagonal array with four rings of emitters on separate phase lines, which created a focal point 25 cm above the middle of the array, which could be moved perpendicularly. Hoshi et al. [44] later used dynamic phase control to allow the focal point to be moved in three dimensions, permitting complex haptic patterns to be produced in air for the first time. In work that would eventually lead to the founding of Ultraleap [17] (formerly Ultrahaptics) as a commercial producer of this technology, Carter et al. [7] described a method for producing multiple focal points at the same time. Fig. 2 shows an example of an ultrasound haptics device with a typical rectangular array layout, as used by most state-of-the-art devices.
Variations on these fundamental focusing methods are still used in state-of-the-art devices, with focal points used as the building blocks in most haptic rendering techniques. The ability to focus ultrasound through phase manipulation allows tactile sensations to be rendered in a wide volume above the emitter array. Whilst there are limitations inherent in the use of ultrasound (which we discuss later), there are no mechanical constraints that restrict focal point movement.

If a single stationary focal point is all that is necessary, it is possible to focus ultrasound through the physical orientation of the emitters in the array. Simplicity, although output is limited because focal points cannot be moved.

In most phased-array systems, ultrasound is focused without considering how the waves are scattered by the skin. However, recent work has started to investigate adaptive focusing methods, where reflections from a single finger are integrated into the sound field synthesis algorithms to create better defined focal points [48]. Such methods show the potential for more precise focusing from a phased array, although there is a great degree of complexity in considering reflections from a whole hand in real time.

Ultrasound haptic arrays typically use piezo style emitters to produce waves, although other emitter technologies are being explored for this purpose, e.g., printed polymer transducers [49], [50]. Most devices use 40 kHz ultrasound emitters. These are readily available (due to their primary use in the automotive industry), offer reasonable attenuation over the range typically used for mid-air human-computer interaction, and have a good balance between power consumption, size and performance. Other frequencies have also been used (e.g., 70 kHz [51]), although this is less common.

The number of elements in a transducer array affects the maximum amplitude of a focal point and the range in which focal points can be produced. Increasing the number of emitters can increase the amplitude, although there are diminishing returns. Due to the directionality of sound waves, simply adding more emitters may not necessarily contribute to stronger amplitude in certain locations. However, large-scale arrays have increased range, enabling contactless haptic feedback over a wider interaction area [52], [53], for multiple hand orientations [54], [55], or for many users [56].

B. Focal Point Modulation

In their initial evaluation of their first device, Iwamoto et al. [6] observed that only the onset and offset of pressure could be perceived. Humans cannot typically feel vibration at ultrasonic frequencies (i.e., cannot perceive the pressure variation in a focal point). However, Iwamoto noted that if the amplitude was modulated at a frequency within the range of vibrotactile perception (around 5-1000 Hz [57]), then the localized rapid change in pressure of the focal point is perceived as a vibration-like tactile sensation.

Amplitude modulation was originally the predominant way of improving perception of focal points. One of its key limitations, however, was the necessity of modulating the amplitude of emitted sound waves. Averaged over time, this means reducing the power output by as much as 50%, potentially limiting the strength of tactile sensations [58]. In recent years, alternative modulation techniques have emerged to address this.

Takahashi et al. [58] described lateral modulation, a method whereby a focal point is continuously moved back and forth across a target position (e.g., along a 5 mm trajectory). Since the focal point repeatedly moves across the target position, repeated onset and offset of a sensation is achieved, similar to amplitude modulation. This makes the sensation perceptible, as before, but with the added bonus of allowing full power to be used. As a result, the strength of the sensation increases. The authors hypothesized that this was not just caused by the increase in total power, but the continuous motion of the focal point likely also contributed to the stronger sensation.

Frier et al. [59] described a similar technique called spatio-temporal modulation, where a focal point continuously moves along an arbitrary trajectory, of any shape and size. The aim of this method is to induce tactile sensations across the entire path of the focal point; contrast this to lateral modulation [58], which produced a singular point. In this work, Frier et al. note that continuously moving a focal point across the skin causes measurable skin displacement, which likely contributes to its perception. A user study found that perceived intensity is highest when the focal point moves at a rate similar to the rate of wave propagation on the skin of the hand, adding further weight to this hypothesis. Takahashi et al. [60] similarly applied lateral modulation to circular trajectories and reported stronger sensations than a sequence of amplitude modulated points.

It is possible to perceive an unmodulated ultrasound field, as demonstrated by Inoue et al. [54] who utilized standing waves to create a perceptible pressure pattern. However, most works in the literature have employed a temporal modulation pattern to aid perception. At the time of writing, these three modulation methods (amplitude, lateral and spatio-temporal) are the main mechanisms for turning an ultrasound focal point into a reliably perceptible sensation. We discuss perception of these methods in more detail in a later section of this article.

![Fig. 2. An Ultralap UHEV1 device, consisting of a 16 × 16 array of 40 kHz transducers (10 mm diameter) and a Leap Motion optical finger tracker (top edge). Its dimensions are 167 mm × 167 mm × 25 mm.](image-url)
C. Ultrasound Haptic Rendering

We now discuss how focal points can be used as the building blocks of more complex tactile sensations, like 2D/3D shapes or textures. These are not recommendations or a how-to guide, as more research is needed to investigate alternative methods and to verify if existing ones succeed in creating their desired effects. We present an overview of key rendering approaches and discuss what is currently known about their success.

As discussed previously, focal points are the basic unit from which more complex tactile sensations can be created. Focal points can be rapidly moved (e.g., Ultraleap devices have sampling rates of up to 16 kHz) and multiple focal points can be presented at the same time, if necessary. It is important to remember that ultrasound haptic sensations have both a spatial and temporal component, since a single focal point likely needs to traverse several positions in sequence.

1) Sensations of Motion

One of the earliest dynamic tactile sensations created using an ultrasound haptic display was the sensation of motion across the palm. Hoshi et al. [44], [61] achieved this by presenting amplitude-modulated focal points sequentially (as in Fig. 3). Wilson et al. [62] conducted psychophysical experiments into the illusion of apparent tactile motion experienced using this approach, finding it successful in inducing the feeling of motion. However, the convincingness varied with factors such as number of points and the time of presentation for each point.

![Fig. 3. Sensation of motion can be induced using the illusion of apparent tactile motion, where a sequence of discrete focal points are perceived as smooth and continuous movement along the trajectory.](image)

With the recent emergence of spatial modulation [58]–[60], sensation of motion can be achieved through the continuous movement of a focal point. For example, Georgiou et al. [11] used continuous motion of a single focal point along a circular trajectory, switching between clockwise and anti-clockwise direction to show the increase or decrease of a value, respectively. This approach uses true focal point motion, rather than apparent tactile motion. In earlier work, the illusion of apparent tactile motion was necessary because amplitude modulated focal points needed to be presented in a static position for a brief duration, in order to be clearly perceived.

2) Shapes

One of the key advantages of ultrasound haptics compared to other tactile technologies is that it allows designers to manipulate spatial properties of its output (i.e., focal point position), in addition to its temporal properties. This presents the opportunity to create haptic shapes in mid-air.

Long et al. [63] investigated the feasibility of presenting volumetric shapes, by rendering the cross-section intersected by the hand in mid-air. The hand intersection was computed using hand tracking data, then the 2D cross-section was computed as a set of focal points distributed around the outline of the shape. As an example, moving the hand through a sphere resulted in a sequence of circular cross-sections of varying diameter. A large circle may have had 16 individual focal points on its circumference, for example. A user study showed reasonably successful shape identification (ranging from 61–94%). Due to the focal point size, however, similar shapes were frequently confused (e.g., a cone and pyramid were most often mistaken for each other)

Korres et al. [64], [65] used a similar method to create in-air shapes, distributing focal points along their outlines. Amplitude-modulated points were presented in sequence, at a speed intended to create the perception of simultaneous presentation. They investigated the effect of presentation parameters, such as the duration of each individual point and the number of points along the outline, resulting in guidelines for rendering convincing shapes.

An alternative shape rendering approach was proposed by Inoue et al. [54], using standing waves (stationary sound waves) to create the desired shape. This allowed the lines and edges of a shape to be presented, rather than unconnected focal points along its outline. Standing waves are a static interference pattern, resulting from waves travelling in different directions. This requires multiple transducer arrays (in this case, eight arrays were placed in an octagonal prism), thus is not usable with standard ultrasound haptic device form factors.

By necessity, spatio-temporal and lateral modulation both require focal points to move continuously along a trajectory. This simplifies the creation of haptic polygons by repeatedly rendering their outlines (e.g., Frier et al. [59] created circles in this way). The simplest way of creating shapes using spatio-temporal or lateral modulation is to traverse the outline. Circles are the simplest shape to present in this way. Fig. 4 shows an example of how amplitude and spatio-temporal modulation can be used to render the outline of a circle.

![Fig. 4. Shapes can be created using a sequence of amplitude-modulated points (left, e.g., [63]) or through continuous point movement (right, e.g., [59]).](image)

Rutten et al. [66] investigated shape identification using spatio-temporal modulation. They found poor accuracy (below 50%), even after a learning phase. Their results suggest straight lines are more reliably identified than connected polygons (e.g., circles and squares).

Shapes with multiple edges (e.g., rectangles, triangles) are
more complicated, as changes in direction between edges may be difficult to perceive. Hajas et al. [67] investigated the use of a “dynamic tactile pointer”, intended to emphasize the discrete edges that constitute a polygon by pausing at the vertices (e.g., the four corners of a rectangle). Their method improved perception, leading to a recommendation that polygons are rendered as discrete edges rather than a continuous outline with a uniform movement speed.

3) Textured Surfaces

There have recently been formative attempts to render textural qualities like ‘roughness’ using ultrasound haptics. Freeman et al. [68] described a method for rendering macro-scale textures (e.g., gratings). They defined a surface as a tessellation of geometric features. When the hand intersected a surface, a focal point moved to recreate the features at each point of intersection. The spatial properties of these geometric surfaces were difficult to discern, so temporal properties were also manipulated to create a wider range of perceptions.

Beattie et al. [69] presented a method for recreating textural features in images. Images were analyzed to identify degrees of macro- and micro-roughness, which were then mapped to haptic features. Unlike Freeman’s naïve method, Beattie considered dynamic aspects of texture exploration, by integrating hand movement speed in the rendering algorithm.

These approaches have yet to be evaluated, so their efficacy is currently unknown. Textural qualities are challenging to render convincingly, since fine-grained features that typically constitute material texture are orders of magnitude smaller than ultrasound focal points (e.g., 40 kHz ultrasound has a wavelength of 8.6 mm, yielding similarly sized focal points). An alternative is to modulate temporal properties of a focal point to recreate the sensation of roughness, similar to methods used to simulate roughness using piezoelectric vibrotactile actuators [70]. For example, amplitude modulation frequency can yield different sensations [71].

4) Abstract Dynamic Patterns

The basic rendering elements discussed so far can be combined to create abstract haptic patterns, which do not necessarily correlate with real physical objects or sensations. The lack of mechanical constraints on focal point movement means designers have a great degree of creative freedom to render novel haptic patterns.

Ultraleap have created a set of abstract ‘sensations’ using such haptic patterns. For example, an erratic motion trajectory that creates the sensation of an electric spark, or a random presentation of points that mimics running water [72]. They have similarly developed patterns aiming to recreate sensations such as a heart beating and static electricity [5]. Ablart et al. [73] and Vi et al. [74] similarly developed abstract patterns to accompany visual artforms, with the aim of enhancing the experience. Such patterns intend to evoke an affective response from users and succeed in doing so [75]. Sand et al. [76] used abstract patterns for delivering notifications, similar to Tactons [77]. There are currently no guidelines for crafting such sensations, reflecting their often arbitrary nature.

D. Ultrasound Haptic Perception

1) Perceptual Fundamentals

Focal point modulation, using one of the methods discussed earlier, is typically used to aid the perception of tactile sensations. Most of the literature has focused on stimulating the glabrous (non-hairy) skin, such as the palm-side of the hand. Therefore, we focus on this aspect of perception. Not much is known about perception of ultrasound haptics in non-glabrous skin, although formative results show it is indeed possible [60] and that with sufficient emitters (over 4000, in this case), enough pressure is created to stimulate through clothing [78].

There are four mechanoreceptors in the glabrous skin, contributing to a wide range of tactile percept: Pacinian corpuscles, Ruffini endings, Meissner corpuscles, and Merkel’s discs [79]. The primary mechanoreceptors stimulated by ultrasound haptic sensations are Pacinian corpuscles (PCs). There are approx. 2500 PCs on each hand, most densely located near the fingertips [79]. PCs respond to the onset of pressure (i.e., the repetitive onset of pressure arising from vibration), enabling them to sense vibration across a wide range of frequencies (5-1000 Hz, with peak sensitivity around 200 Hz [57]).

The predominance of PCs in vibrotactile perception meant early systems typically used amplitude modulation around 200-250 Hz, close to peak sensitivity of these mechanoreceptors (e.g., systems by Iwamoto [6], Carter [7], Palovuori [80], Korres [81]). However, there is also evidence that amplitude modulation can stimulate Meissner corpuscles (MCs), which are sensitive to lower frequency vibrations [57]. Obrist et al. [71] compared 16 Hz and 250 Hz modulation, resulting in different tactile experiences. Gil et al. [82] also targeted MCs, this time on the face. These results suggest both MCs and PCs can be stimulated by ultrasound haptic focal points in areas of glabrous skin. Slow adapting receptors may also be sensitive to ultrasound haptic patterns. Suzuki et al. [78] delivered ultrasound haptics to the upper body in a workspace larger than a 1 m × 1 m × 1 m cube. They targeted Merkel disks and Ruffini endings and used modulation frequencies of 25, 50 and 100 Hz. They could deliver ultrasound haptics even through thin clothes. Inoue et al. [54] used standing waves to create complex patterns that could be perceived without temporal modulation, providing evidence that slow adapting mechanoreceptors may also be sensitive to acoustic radiation pressure.

Contemporary modulation methods (i.e., lateral and spatio-temporal) achieve repeated onset of pressure through cyclic focal point movement, allowing stimulation of PCs and MCs. However, skin deformation is also thought to contribute to the perception of these sensations. Frier et al. [59] found improved perception when patterns were rendered to coincide with wave propagation across the skin. Reardon et al. [83] similarly investigated surface wave propagation across the skin, using optical vibrometry. They found that skin waves behaved differently, depending on the speed of focal point motion, and that these waves contribute to the perception of tactile sensations. To better understand and utilize such effects, Chilles et al. [84] also conducted laser vibrometry measurements and used simulation results to inform future developments.

By targeting PCs and MCs (the rapidly adapting receptors in
glabrous skin), ultrasound haptic devices are essentially creating vibrotactile stimuli with a high degree of spatial control. However, other aspects of tactile perception can be stimulated using ultrasound. For example, Kamigaki et al. [85] presented a method to create thermal sensations to hands with static ultrasound pressure. Gloves were worn to absorb energy from the ultrasound and turn it to heat. Ochiai et al. [35] also demonstrated a method for creating multimodal tactile sensations, combining ultrasound haptics with highly precise laser-based tactile stimulation, although this required two haptic display technologies.

2) Perceptual Studies

Psychophysical and perceptual studies have been performed to inform the design of ultrasound haptic sensations. The aim of these studies is typically to identify parameters that maximized perceived intensity, or support identification and discrimination of shapes, patterns, etc. Carter et al. [7] described a method for generating multiple focal points. They investigated the effect of focal point separation and amplitude-modulation frequency in a two-point discrimination test. They found improved discrimination as points moved further apart (reaching acceptable levels around 5 cm). However, modulating the points at different frequencies improved discrimination at closer distances (~3 cm).

Wilson et al. [62] investigated localization of a static point and perception of apparent movement. The average localization error was 8.5 mm (approximately the same size as the 40 kHz focal points), although this varied depending on location on the palm. For convincing illusions of apparent motion, they found that longer focal point presentation durations (50-200 ms) and longer distances (>3 cm) can improve recognition of motion. Yoshimoto et al. [86] investigated perception of actual motion, finding that a moving point could be followed up to 10 cm/s.

Vo et al. [87] investigated how accurately users could localize a focal point in an active exploration task. Average error was 14.1 mm, almost double the passive localization error reported by Wilson et al. [62]. They reported that users located the focal points significantly faster when visual feedback was given as well, suggesting some guidance can be beneficial in reducing ambiguity over where the hand should be placed for optimal perception.

Korres et al. [64] investigated parameters for rendering perceptually smooth and continuous sensations when presenting a sequence of amplitude-modulated points. Their experiment measured the minimum stimulation duration and minimum delay between subsequent points that are perceived as continuous. Results suggest a perceptible stimulation can occur for a duration of 5-50 ms and the minimum acceptable delay between points is below 40 ms.

Frier et al. [59] investigated the perceived intensity of spatio-temporal modulated circles, using a magnitude scaling experiment. Their results reveal a relationship between circle diameter, focal point movement speed, and perceived intensity. A later study [88] investigated the relation between device sample rate and the perceived intensity of the circular patterns. These results suggest sample rate does indeed affect perceived magnitude, leading to the suggestion that sample rate should be optimized for the intended haptic patterns.

Takahashi et al. [58] compared amplitude and lateral modulation for a static focal point. Their studies characterized the effects of modulation method, lateral modulation parameters, and perceived intensity. Lateral modulation typically resulted in lower sensitivity thresholds (i.e., more intense stimuli) and improved perception on non-glabrous skin.

Howard et al. [89] investigated the perception of points and lines, as well as feelings of ‘bumps’ and ‘holes’. The 50% detection threshold for a single amplitude-modulated focal point was approx. 560 Pa, significantly less than the pressures typically achieved with a state-of-the-art device. For polygons rendered using spatio-temporal modulation, they found an even lower threshold (consistent with other work [58], [60]).

Raza et al. [90] presented a perceptually correct haptic rendering algorithm. They measured the output with several different input values (intensity, frequency, distance, position of focus) for an ultrasound haptic device, and further measured the perceived sensations in a user test. They then created a correcting algorithm, which produces perceptually correct and constant values across a volumetric object. The system produces better perceived quality and only the measurements of the technical features are needed for a new device.

Marchal et al. [91] created various levels of ‘stiffness’ for virtual materials, rendered using ultrasound haptics. Their study showed that the percept of stiffness can be successfully created, even though an ultrasound haptic device is unable to produce sufficient force to resist movement.

A limited body of work have investigated if haptic illusions can be induced using ultrasound haptic stimuli. The previously discussed work by Wilson et al. [62] confirms that apparent tactile motion can be convincingly conveyed. Pittera et al. [92] investigated an illusion involving the perception of falling raindrops in VR (similar to the ‘rubber hand’ illusion). Users perceived an incongruent visual-tactile stimulation as being congruent, such that the users felt a virtual hand as their own.

These studies show a variety of perceptual properties that have been investigated in the literature. Their findings have resulted in guidelines about how to create simple sensations, like multiple points that can be reliably discriminated and polygons that can be reliably identified. There is a great need for more research into perception, however, e.g., to establish psychometric curves for aspects of perception. Recent work by Frier et al. [59], [88] and Takahashi et al. [58], [60] take steps in the right direction, giving empirical evidence that can be used to inform pattern design. In this section, we gave an overview of how ultrasound haptic devices work, how they can be used to render tactile sensations and create haptic representations of objects, and how these sensations are perceived by users.

IV. ULTRASOUND HAPTIC INTERACTION TECHNIQUES

In this section of this paper, we discuss how the fundamental principles of ultrasound haptics have been used to create novel haptic interaction techniques. In the following section, we will discuss the main contexts where these are being applied.
A. Mid-air User Interface Components

The simplest use of ultrasound haptic feedback is to create focal points to evoke the sensation of touching something in mid-air. In this case, the presence of a focal point against the fingertips is more important than what the focal point feels like, because it creates a sense of presence and assures users that the system is responding to their actions [3]. This approach has been used to create the experience of touching a screen in mid-air (e.g., [3, 7, 16, 93, 94]), with focal points mapped to one or more fingertips to recreate the implicit tactile cues experienced when pressing a touchscreen. The binary sensation of feeling feedback, or not, is sufficient to inform users about system state and improve usability [3].

Others have adapted this method to create the sensation of touching individual user interface controls (e.g., buttons), rather than the ‘screen’ that contains them. The simplest method for creating this feedback is to present a focal point when a finger is situated within the boundaries of a button [16], helping users identify when they are able to make a selection. Others have rendered the edges of buttons [9] to try help users locate them.

An alternative button feedback design is to try create the sensation of activating a button. For example, Rümelin et al. [12] presented a short burst of feedback after a ‘tap’ gesture, informing the user that a button was successfully activated. Cornelio-Martinez et al. [2] also investigated this style of activation feedback, finding that it improves the sense of agency over button activation (i.e., made users feel more in control of the user interface). Hwang et al. [14] and Ito et al. [95] used two layers of feedback, with stronger feedback in the lower layer, to mimic the experience of pressing a physical button and experiencing an increase in mechanical resistance.

Screens and buttons are the most commonly represented user interface widgets in mid-air haptic systems, but others have found ultrasound haptic feedback to be just as effective for other types of control. Harrington et al. [9] created a horizontal slider with haptic feedback; as users moved the slider (by moving their hand left or right), pulses of feedback were given to the palm each time the slider moved to the next ‘notch’. Georgiou et al. [11] and Shakeri et al. [10] created dials, with focal points moving along a circular trajectory to indicate direction of change. For example, increasing a dial by turning it clockwise resulted in a clockwise focal point trajectory, increasing in speed as the underlying value increased.

B. Mid-Air Gesture Feedback

Ultrasound haptic devices have been used to give feedback and guidance about mid-air gesture interactions. For successful mid-air interaction, users need to know where to perform gestures so they can be sensed reliably [1]. Freeman et al. [4] used haptic feedback to inform users about how close they were to the ‘optimal’ area for interaction, using a circle that grew in diameter as users approached the target position. Suzuki et al. [53] also investigated haptic feedback for guidance, creating a mid-air haptic ‘hand rail’ to guide hand movements.

Haptic patterns have been developed for gesture feedback as well, informing users about how a system is responding to their actions. For example, Shakeri et al. [10] described static and dynamic patterns to confirm gesture recognition. Static patterns were short bursts of feedback, used to indicate that a hand pose gesture was recognized. Dynamic patterns used focal point movement to indicate the effects of gestures, e.g., clockwise motion around the palm to show an increasing value. Georgiou et al. [13] described a set of patterns used to give feedback about gestures, e.g., points tapping the palm for a ‘tap’ gesture, and a line sweeping across the palm for a ‘swipe’ gesture. Similarly, Dzidek et al. [96] also described a set of bespoke feedback patterns for gesture interactions, e.g., a forcefield sensation when approaching virtual objects.

C. Virtual Object Representations

A third category of common interaction techniques have used ultrasound haptics to create representations of virtual objects in mid-air. We omit details about the haptic rendering of these object representations, since this was discussed previously.

Ultrasound haptics paired with “holographic” displays (e.g., [97]) and VR headsets (e.g., [14–16]) have been presented (examples in Fig. 5). As both the virtual ‘floating’ image and tactile feedback are in mid-air, they can be spatially linked and may be more effective and engaging than each alone. In addition to increasing the immersion of interacting with a virtual object, this use of haptic feedback also has interaction benefits. For example, experiencing spatially congruent feedback can help users grasp a virtual object between their fingers [98], as depicted in Fig. 6.

Fig. 5. Left: desktop-sized mid-air fogscreen with mid-air haptics [80] ©2015 IEEE. Right: mid-air ultrasound haptic feedback in front of a VR HMD [16].

Fig. 6. An example of spatially-congruent feedback for virtual objects. As a finger contacts an object, feedback is presented in an appropriate location [98].

V. APPLICATIONS OF ULTRASOUND HAPTICS

Ultrasound haptic devices have been explored in new use cases across a variety of application areas. They are often used where contactless interaction is desirable, e.g., for spontaneous and unencumbered interaction, or where hygiene or sterility are concerns. We now provide a brief overview of four application areas that have been explored in the literature and in new product concepts. These give insight into how this technology is currently being applied.
A. Sterile Medical Interfaces

Contactless user interfaces have many potential applications in medical and healthcare settings, where sterility is a concern that impedes interaction [99]. In sterile conditions, interaction often happens by proxy [100], leading to inefficiency. Mid-air interaction can potentially alleviate such barriers to interaction, with haptic feedback used to improve usability. Mid-air user interfaces can also improve interaction with 3D data, e.g., contactless interaction with medical images could allow health professionals to visualize, analyze, and collaboratively interact with patient data.

Mid-air tactile sensations can also be used to present health information in new ways. Hung et al. [46], [47] developed a mid-air haptic system to train cardiologists to search for a pulse. Users placed their hand between a display and ultrasound array; an augmented reality version of their hand was rendered on the screen, and haptic patterns were used to simulate the sensation of feeling for a pulse. In this case, the attraction of ultrasound haptic feedback was its temporal resolution, giving control over the pulse sensation. Others have also explored the use of ultrasound haptics for palpation simulation [101].

Ochiai et al. [35] proposed the combination of ultrasound haptic feedback with augmented reality displays, with one use case being medical visualizations. Romanus et al. [102] similarly presented a mid-air haptic bio-hologram, where the user can see, touch and feel an augmented version of their own heart beating. Their system uses AR glasses, a wearable heart rate sensor and an ultrasound array placed on a desk (Fig. 7).

Fig. 7. Ultrasound haptic feedback for an AR medical visualization [102] (©2019 IEEE).

These works have begun to investigate the potential of ultrasound haptics for training and visualization in the medical domain. A trend seen in work so far is that haptic feedback is being used to present temporal information (i.e., pulse) [46], [47], [101], [102]. More generally, the presence of feedback for input commands may improve usability. Limited spatial resolution, however, means ultrasound haptic feedback is unlikely to be used to encode spatial information, e.g., in medical images.

B. Automotive Applications

Contactless user interfaces have been adopted by automotive companies as a potentially less distracting interaction modality for drivers, allowing imprecise eyes-free hand gestures as an alternative to visually demanding touchscreens. Ultrasound haptic feedback has been investigated for touchless in-car user interfaces, with the aim being to deliver feedback whilst allowing drivers to focus their visual attention on the road. This can reduce visual demand, shorten interaction times, improve input accuracy, minimize eyes-off-the-road time (EORT) on displays, and thus improve safety [9], [10], [103], [104].

Automotive user interfaces typically use ultrasound haptics to deliver feedback for interface controls. Rümelin et al. [12] created tactile feedback for mid-air buttons, so that users could feel the ‘click’ and know a control was activated without having to divert attention from the road. Harrington et al. [9] also investigated feedback for buttons in an automotive interface, comparing them with horizontal sliders (illustrated in Fig. 8). They found that sliders, in particular, benefit from the mid-air haptic feedback, but button pressing was still better suited to touchscreens. Georgiou et al. [11] and Shakeri et al. [10] created ultrasound haptic feedback for mid-air gestures. Shakeri’s findings suggest that for maximum efficacy (and therefore maximum safety), ultrasound haptic feedback should be combined with at least one other modality, to reduce distraction as a result of uncertainty over interaction.

Fig. 8. Ultrasound haptic feedback for in-vehicle interactions [9].

These works have employed ultrasound haptics as a means of presenting eyes-free interaction feedback, helping drivers determine the success of their actions when manipulating widgets like buttons, dials and sliders. Evaluation using driving simulators suggests good efficacy, helping drivers focus on the road instead of the dashboard interface [9], [10].

C. Digital Advertising, Retail and Signage

Digital advertising aims to catch attention and reach potential customers in new and engaging ways. Such advertising displays
are typically limited to visual and auditory modalities, but the introduction of tactile sensations could allow more informative or entertaining experiences with the marketed products. Mid-air haptics is a particularly appealing way of creating these tactile experiences because of hygiene concerns.

Limerick et al. [5] investigated the effect of adding mid-air haptic feedback to interactive digital posters. A lab-based evaluation found that mid-air haptic feedback increased user engagement and also improved usability and aesthetic appeal. Haptic feedback also led to greater feelings of immersion with the advertised content. Their results suggest that mid-air haptic feedback can enrich interactive digital posters. Other work has explored this potential as well [105], [106].

Others have explored how ultrasound haptic feedback can be used to remotely convey tactile qualities of products to users. Touch-enabling technologies can provide utilitarian and hedonic value to consumers [107], encouraging and informing product purchases. For example, Kim et al. [108] presented a demonstration combining ultrasound haptics with a holographic display, for presenting a catalogue of bathroom appliances. Users could, for example, view a virtual image of a tap or shower and experience the flow of water against their hand.

In these examples, ultrasound haptics was used to entertain and attract interest, with the goal of increasing engagement with advertised content. Formative studies suggest this can be effective [5], with novelty likely being part of its initial success.

D. Mixed Reality

Mixed reality applications could benefit from mid-air ultrasound haptics, as it is unobtrusive, maintains freedom of movement, and can be used to render a wide variety of sensations. A close coupling between visual and haptic sensations is likely to increase the efficacy of the haptics in this context, overcoming issues relating to its limited ability to present spatial information.

Augmented haptics have been developed for synthetic worlds since 1967 [109]. Haptics can greatly enhance feelings of immersion and improve input performance and interaction with mixed reality systems, although many tactile technologies are rudimentary compared to the high fidelity sensations achievable with auditory and visual displays. The advantages of ultrasound haptic devices, discussed at length in this paper, create new opportunities for rich tactile sensations for mixed reality experiences, across a variety of application domains.

Ultrasound haptics has been used to add tactile sensations to augmented and virtual reality images. For example, the medical visualization examples we discussed previously combined AR images with a haptic heart pulse. The tactile sensations were used to inform about data associated with the visual content.

Similar systems attempt to recreate physical objects using a combination of visual and haptic stimuli. For example, Makino et al. [55] and Kervegant et al. [110] used different augmented reality display methods, paired with ultrasound haptic devices, with the goal of creating a convincing multisensory virtual representation for a user to explore; Fig. 9 shows an example of the latter work.

Mixed reality systems have also been used for novel playful and entertaining experiences; for example, Hoshi et al. [61] describe a mixed reality experience where users feel raindrops falling onto their hand, or the footsteps of a small elephant that walks across the palm. Similar concepts have been widely explored (e.g., [55], [61], [94], [98], [111]–[113]), with spatially congruent visual and haptic content, an example of which is shown in Fig. 10.

In similar fashion, Ultraleap have created novel experiences where haptic feedback is used to create “supernatural” feelings [15]. In this work, a user wearing a VR headset sees their hands in front of them, above a book of spells that is in the same location as a haptics array (see Fig. 11). As users cast new spells, they experience haptic effects, like the flames from a ball of fire or the sparks from a ball of electricity. They have also created similar installations, with multisensory experiences not possible with visual displays alone [114], [115].
Recent innovations attempting to use haptic feedback in specific application areas. Some of these limitations influence the practicality and haptic output quality of the feedback. Ultrasound haptics has some limitations that may affect its use in mixed reality interfaces. An early example was presented by Sand et al. [16], who presented tactile feedback corresponding to virtual reality buttons. The visual and haptic feedback was spatially congruent, such that users could feel the buttons that were visible in front of their face. Yoshino et al. [56] and Palovuori et al. [80] likewise presented tactile feedback for buttons on projected mid-air displays, mimicking the sensation of touching a screen.

Dzidek et al. [96] expanded this concept, discussing a fully augmented reality workspace for productivity environments. Their vision was of a workspace mixing conventional graphical user interface elements with virtual content (e.g., product models), with corresponding haptic feedback. In their work, they discussed a wide design space, suggesting how the spatial, temporal, and acoustic properties of focused ultrasound can be mapped onto the wide variety of feedback needs in this type of augmented environment.

VI. LIMITATIONS OF ULTRASOUND HAPTICS

Ultrasound haptics has some limitations that may affect its use in specific application areas. Some of these limitations impose trade-offs between practicality and haptic output quality. In this section, we discuss some important limitations and recent innovations attempting to overcome them.

A. Precision

Ultrasound haptic patterns consist of focal points whose size varies with the wavelength, such that smaller wavelengths yield smaller focal points. For 40 kHz ultrasound, the wavelength is 8.6 mm in air. The term ‘focal point’ is perhaps misleading, since the region of high sound pressure is elongated along the direction of wave propagation (Fig. 12). This region may be several centimeters in length and its orientation tilts as it moves further from the center of the transducer array. An implication of this is that users may feel a tactile sensation at the furthest extent of the high-pressure region, leading them to not find the strongest part of the focal point. This could explain why Freeman’s localization study [4] found that users consistently placed their hands too high when attempting to localize a point.

Another use of ultrasound haptics is providing interaction feedback for mixed reality interfaces. An early example was presented by Sand et al. [16], who presented tactile feedback corresponding to virtual reality buttons. The visual and haptic feedback was spatially congruent, such that users could feel the buttons that were visible in front of their face. Yoshino et al. [56] and Palovuori et al. [80] likewise presented tactile feedback for buttons on projected mid-air displays, mimicking the sensation of touching a screen.

Dzidek et al. [96] expanded this concept, discussing a fully augmented reality workspace for productivity environments. Their vision was of a workspace mixing conventional graphical user interface elements with virtual content (e.g., product models), with corresponding haptic feedback. In their work, they discussed a wide design space, suggesting how the spatial, temporal, and acoustic properties of focused ultrasound can be mapped onto the wide variety of feedback needs in this type of augmented environment.

VI. LIMITATIONS OF ULTRASOUND HAPTICS

Ultrasound haptics has some limitations that may affect its use in specific application areas. Some of these limitations impose trade-offs between practicality and haptic output quality. In this section, we discuss some important limitations and recent innovations attempting to overcome them.

A. Precision

Ultrasound haptic patterns consist of focal points whose size varies with the wavelength, such that smaller wavelengths yield smaller focal points. For 40 kHz ultrasound, the wavelength is 8.6 mm in air. The term ‘focal point’ is perhaps misleading, since the region of high sound pressure is elongated along the direction of wave propagation (Fig. 12). This region may be several centimeters in length and its orientation tilts as it moves further from the center of the transducer array. An implication of this is that users may feel a tactile sensation at the furthest extent of the high-pressure region, leading them to not find the strongest part of the focal point. This could explain why Freeman’s localization study [4] found that users consistently placed their hands too high when attempting to localize a point.

Another use of ultrasound haptics is providing interaction feedback for mixed reality interfaces. An early example was presented by Sand et al. [16], who presented tactile feedback corresponding to virtual reality buttons. The visual and haptic feedback was spatially congruent, such that users could feel the buttons that were visible in front of their face. Yoshino et al. [56] and Palovuori et al. [80] likewise presented tactile feedback for buttons on projected mid-air displays, mimicking the sensation of touching a screen.

Dzidek et al. [96] expanded this concept, discussing a fully augmented reality workspace for productivity environments. Their vision was of a workspace mixing conventional graphical user interface elements with virtual content (e.g., product models), with corresponding haptic feedback. In their work, they discussed a wide design space, suggesting how the spatial, temporal, and acoustic properties of focused ultrasound can be mapped onto the wide variety of feedback needs in this type of augmented environment.

VI. LIMITATIONS OF ULTRASOUND HAPTICS

Ultrasound haptics has some limitations that may affect its use in specific application areas. Some of these limitations impose trade-offs between practicality and haptic output quality. In this section, we discuss some important limitations and recent innovations attempting to overcome them.

A. Precision

Ultrasound haptic patterns consist of focal points whose size varies with the wavelength, such that smaller wavelengths yield smaller focal points. For 40 kHz ultrasound, the wavelength is 8.6 mm in air. The term ‘focal point’ is perhaps misleading, since the region of high sound pressure is elongated along the direction of wave propagation (Fig. 12). This region may be several centimeters in length and its orientation tilts as it moves further from the center of the transducer array. An implication of this is that users may feel a tactile sensation at the furthest extent of the high-pressure region, leading them to not find the strongest part of the focal point. This could explain why Freeman’s localization study [4] found that users consistently placed their hands too high when attempting to localize a point.

Another use of ultrasound haptics is providing interaction feedback for mixed reality interfaces. An early example was presented by Sand et al. [16], who presented tactile feedback corresponding to virtual reality buttons. The visual and haptic feedback was spatially congruent, such that users could feel the buttons that were visible in front of their face. Yoshino et al. [56] and Palovuori et al. [80] likewise presented tactile feedback for buttons on projected mid-air displays, mimicking the sensation of touching a screen.

Dzidek et al. [96] expanded this concept, discussing a fully augmented reality workspace for productivity environments. Their vision was of a workspace mixing conventional graphical user interface elements with virtual content (e.g., product models), with corresponding haptic feedback. In their work, they discussed a wide design space, suggesting how the spatial, temporal, and acoustic properties of focused ultrasound can be mapped onto the wide variety of feedback needs in this type of augmented environment.
lobes. Another alternative which has been commercially produced uses a Fibonacci spiral arrangement [118], which can remove most of the side lobes (visible in Fig. 12), improving the overall quality of device output.

Focal point size is another potential limitation. Since 40 kHz ultrasound has an 8.6 mm wavelength in air, these devices cannot produce more fine-grained spatial patterns. Higher frequencies have a shorter wavelength (e.g., 70 kHz has a wavelength of 4.9 mm and has been used for mid-air haptics [51]). However, 40 kHz is the predominant frequency because of hardware (mass-produced 40 kHz transducers offer a good balance between size, power and cost) and attenuation (higher frequency ultrasound attenuates more quickly); these practical issues mean improved spatial resolution is unlikely for now. The spatial resolution of an ultrasound haptics device is far from resolution and capabilities of the human tactile system [22]; nevertheless, these can be successfully used to elicit a variety of tactile sensations.

Ultrasound mid-air haptics has relatively high spatial and temporal resolution, but a precise tracking system is crucial so that the ultrasound stimulation does not go to a wrong location. The Leap Motion sensor is relatively accurate and has a similar functional range as an ultrasound haptics device, so it is typically the preferred device for mapping haptic patterns to hand position in air. For mid-air haptics to work effectively in a VR or AR setting, the haptic feedback needs to be targeted in a precise manner, relative to the user’s body position. An implication of this is that room-scale tracking will need to be able to precisely determine the user’s hand position, as well as the location of the user’s head. A dedicated hand-tracker on the VR/AR headset could be used to overcome this problem.

Temporal precision also influences the perception of ultrasound haptic feedback. Fier et al. [88] noted that the hardware sample rate has an effect on the user’s perception of high-speed mid-air tactile patterns. A faster sample rate is not necessarily better; instead, the sample rate should scale with the size of the haptic pattern (i.e., because a smaller pattern results in a smaller focal point displacement at each update).

Rendering method (i.e., how focal points are modulated to construct a pattern) may also influence spatial and temporal perception. Sun et al. [121] studied if different ultrasound haptic rendering methods have an impact on the perceptual threshold. Their results show that the center of the palm is more sensitive to ultrasound haptics than the fingertip, the palm is more sensitive to a stationary or slow-moving than a fast-moving focal point. When the modulating wave has a DC offset, the palm is sensitive to a much smaller modulation amplitude.

B. Strength and Perception

The perceptible strength of an ultrasound haptic pattern is limited: the total tactile force (~0.016 N) [61] achieved with one phased array is only a small fraction of the force thresholds of physical hardware buttons (e.g., 1.5 N). Increasing the number of transducers can increase the maximum amplitude to an extent. However, this does not increase the producible force linearly, due to the limited directivity of the transducers (since the waves from additional transducers cannot contribute to focal points that are too far away). Likewise, there are also size, power and cost implications of using larger arrays. Furthermore, ultrasound haptics cannot generate a strong static force or torque (kinesthetic force), so it is limited to cutaneous tactile sensations [70]. Dedicated force feedback devices have various force ranges, typically about 3-8 N, e.g., 3D Systems Touch [122], orders of magnitude in excess of the resistive force that can be produced using ultrasound.

Attempts have been made to strengthen the haptic effects by using novel array setups. In a recent experiment, a hybrid focus from 40 kHz and 70 kHz arrays produced a marginally stronger sensation [123] than either individual frequency, but still insufficient for resistive force.

Most ultrasound arrays are planar grids of transducers. Non-planar arrays may improve strength and precision (or alternatively, reduce the number of transmitters). Hung et al. [46], [47] employed a parabolic hexagonal array and Ciglar [45] used a spherically shaped array. These arrangements use physical orientation to create a fixed focal point, equidistant from each transducer. Since the main lobes are aligned towards the focal point, they can use the maximum intensity of all transducers, without loss due to angular dissipation. An obvious limitation, however, is the fixed focal point position, since the transducers are physically oriented to focus the waves. Due to the weak pressure involved, ultrasound haptic patterns require modulation to be perceived. The rendering methods discussed earlier each have their own trade-offs. Amplitude modulation renders perceptible focal points in a static position. Points must remain static for a brief duration, so that the modulation results in perceptible and recognizable signals. For more complex shapes, a discrete sequence of focal points must be presented [63], creating indistinct shapes.

Lateral modulation [58] oscillates the focus point quickly and subtly (within a few millimeters) while maintaining maximum amplitude. A localized change in pressure is obtained from movement and is rated as stronger than a comparable amplitude modulated point. Complex shapes can be rendered using a set of multiple focal points [60].

Spatio-temporal modulation [59] draws shapes continuously and rapidly while also maintaining maximum amplitude. For example, a circle is rendered by moving a focal point along a continuous circular trajectory. As the amplitude does not change, the pattern has the potential to feel stronger. An advantage of this method is that it is not limited to discrete focal points; however, a single focus point would require an amplitude modulated or lateral modulated point.

Inoue et al. [124] built a large octagonal-prism type phase array, where eight planar arrays faced each other. They formed a cavity where the users could insert their hand. It did not employ hand tracking or interactivity, as the resulting acoustic radiation field was a standing wave pattern, with waves representing the edges in a polygon. The method does not require any temporal modulation due to the large number of transducers. It has no limitations on hand orientation (as the palm is always facing an emitting array), but the prototype had reported limitations as well as impractical power requirements.
C. Range

The range of effective perception is limited because ultrasound attenuates as distance increases from the emitter. Focal points also cannot be clearly defined within the order of a few wavelengths from the emitter array (up to the Fresnel distance). Very close to the array, it will therefore be difficult to create precisely defined pressure patterns. Likewise, far from the emitter (typically beyond 40 cm, depending on array size), stimuli become increasingly weak and difficult to precisely perceive. For an array of ~256 transducers (e.g., a 16×16 grid), focal point intensity peaks around 10 cm above the middle of the array and will be difficult to perceive after approximately 40-50 cm. The base frequency of the transducers also has an effect, with 40 kHz attenuating less than higher frequencies (e.g., 70 kHz [51], [123]).

Phased ultrasound arrays are currently limited by transducer technology. For example, 40 kHz piezoelectric transducers typically have a diameter (1 cm) larger than the wavelength (8.6 mm in air), which limits the range and precision of the sound field. Acoustic metamaterials are physical structures whose shape affects wave propagation [125], [126]. These may extend the performance and range of transducer arrays by allowing analogue manipulation of the sound field. These are materials that are engineered to have specific acoustic properties, e.g., adding a desired phase delay (e.g., using meandering paths to delay a sound wave) and can be used for acoustic focusing, steering, etc. They can simplify the design of acoustic devices because the underlying transducer(s) do not need to digitally manipulate phase. Metamaterials can be smaller than transducers, allowing a greater degree of precision (e.g., smaller focal points).

The available output range can be a limitation for some applications, thus the placement of the array needs to be considered carefully for each application and environment. Ultrasound tactile feedback is not suitable for large distances of several meters from the transducers, as that would require massive ultrasound arrays or much lower frequencies, which may have safety concerns. As an example of a large scale array, Hasegawa and Shinoda [127] presented a multi-unit ultrasound phased array with an order of magnitude more transducers distributed over a wider space; this produced feedback over a distance of 1 m, improving on the range of a single device.

Transducer arrays are typically fixed in place, offering a range of output that is strongest in the space directly above it (although focal points can still be created outside of the device footprint). As a result, users need to keep their hands above an ultrasound array to experience feedback. A small phased array mounted on the user is very mobile and such systems increase the range of interaction by forming an effective range of output near the user. Wilson et al. [62] suggested an array mounted on the wrist would offer portable mid-air feedback for in-air finger gestures, so long as the hand was oriented towards the wrist.

Sand et al. [16] mounted an array onto a head-mounted display (HMD), allowing feedback in the space in front of the user’s head, regardless of orientation. This allowed tactile sensations in front of the user, so long as the palms of their hands were facing back towards the headset. Whilst this allowed the haptics device to move with the user, it had its own limitations; for example, users could not reach forward and explore a virtual object in front of them, as their hands had to face the array. Alternatively, a head-mounted array could be oriented to stimulate the face rather than the hand. Gil et al. [82] investigated if ultrasound haptic cues are viable for AR glasses. They presented stimuli to the cheek, center of the forehead, and above the eyebrows. Their results show the viability of ultrasound tactile cues on the face.

Others have considered mechanisms for actuating an array, to allow contactless haptic feedback to be presented over a wider area. The “Pan-tilt Ultrasound Mid-Air Haptics” device (PUMAH) [128] is an actuated array with 2 degrees-of-freedom, which the authors claim increases the haptic interaction space by a factor of 14. Whilst the array is still grounded in front of the user, they can experience tactile sensations over a larger area than achievable with a fixed array alone. Brice et al. [129] actuated an ultrasound array attached to the arm of a robot, greatly expanding the range of haptic output to 1.5 m². In that work, the robot remained stationary, with its arm providing range of motion. In future, a mobile robot could further expand that range, allowing a single array to provide room-scale coverage of haptic feedback.

D. Physiological Limitations

One physiological limitation of this technology is that the signals can mostly only be perceived by the palm of the hand. The dorsal (back) side of hand and most other areas of the human body are unable to perceive it and can’t feel the modulated signals. This is largely because perception relies on Pacinian corpuscles, which are densely distributed in the glabrous skin of the palm. Thus, the orientation of the hand (and array location) has an impact on the ability to perceive ultrasound tactile feedback.

The use of gloves during cold weather or during surgical operations would block the effect, even though the signal can be perceived through thin clothes [78]. Lower frequency modulation can be felt in other body parts (through Meissner corpuscles), as discussed before. Research suggests the face is also sensitive to ultrasound [82], although this may have safety concerns due to the very high pressure close to the ear (in excess of 140 dB [130], [131]).

E. Size, Weight, Cost, Power Consumption

A current limitation of this technology is the size and weight of the arrays, because a large number of transducers are needed to increase feedback intensity to easily perceptible levels. There is a trade-off between portability and feedback intensity. At least 100 transducers are needed for a strong effect, but this depends on the desired range, intensity, arrangement and tilt of the transducer array, etc.

The large number of transducers affects the cost. As the transducer price is about 1 USD/piece in large quantities, a 100-piece array would cost around 100 USD only for the transducers, not including the other needed electronics. If individual transducers need to be independently amplified (e.g., supported by UltraLeap arrays) then the complexity of the driving circuitry is increased. Current commercial ultrasound arrays cost several
thousand USD due to this complexity. State-of-the-art devices use piezoelectric transducers, but they are relatively large (typically 10 mm), considering the large numbers required. Transducer manufacturers have smaller transducer versions (e.g., 5 mm × 5 mm Murata MA40H1S) and are developing even smaller transducers, although these do not yet have power levels comparable to the transducers used in state of the art ultrasound haptic devices.

Microelectromechanical systems-based and piezoelectric-polymer-based approaches have been proposed to make the transducers thinner, but they may be fragile, costly or have smaller piezoelectric coefficients compared with ceramics. Polymer-based transducers have made some progress [132]. Flexible printed circuit technology could make very thin (e.g., 1.3 mm / 0.25 mm) and flexible ultrasound emitters [49], [50]. These show the potential for smaller form factors, although currently have impractical power requirements and a lower order of magnitude signal strength, so are not yet ready for market.

New types of transducers could possibly also use transparent electrodes, enabling the mid-air tactile elements to be pasted onto a visual display. The small size and weight would match well with the needs of mobile computing, and they could even be integrated to a phone cover or screen surface. The printing technology would also potentially bring down the cost, and it would disrupt the ultrasound tactile feedback technology and its applications significantly. It would for example enable built-in, light-weight mid-air tactile feedback for HMDs, as in [16]. However, the proof-of-concept prototypes are not yet ready for commercial exploitation.

Current ultrasound arrays require a lot of power to drive their output. For example, the prototype of 128 transducers reported in [80] consumes max. 700 mA / 24 V + 200 mA / 5 V for the FPGA. UltraLeap’s Stratos Inspire device with 256 transducers consumes 80 W of electrical power at its peak. The power consumption and heat dissipation may be a problem, especially for mobile and portable form factors (e.g., for VR HMDs).

F. Noise

The ultrasound signals of 40 kHz are inaudible to humans. However, there may be some faint but audible noise from the hardware. Modulating the position or amplitude of an ultrasound focal point may create some inherent noise when the modulation occurs at an audible frequency, but overall ultrasound haptic devices are relatively silent – comparable to the audible vibrations from a smartphone or game controller. In noisy environments such as a car, this goes unnoticed. There are also recent attempts to reduce the noise [133].

A potential issue for other devices is that microphones may sense in the low frequency ultrasound range, which may cause disturbances, e.g., in some nearby mobile phones or ultrasound-based position trackers. We tested some phones near an array while making a call, and many of them picked up the noise even from the sides of the array and resulted in audible artifacts. The noise was audible in phones such as LG G3 and Samsung S7, but e.g., Huawei B160 fixed wireless terminal had good filtering and worked well.

G. Safety

Safety is an important consideration. Ultrasound cannot be heard, so there is a possibility that dangerous levels of exposure could go unheard and unnoticed. Unlike audible sound, extreme levels of ultrasound would not cause noticeable discomfort that warns the user that their hearing is at risk. The highest pressure in the sound field is within the focal point(s), which may be in excess of 140 dB [130], [131], [134]. This pressure is highly localized, attenuating away from the focal point(s).

Occupational long-term workplace exposure to ultrasound in excess of 110 dB may lead to hearing loss. Exposure limit recommendations for continuous 40 kHz ultrasound are 110 dB and peak 140 dB [130], [135]. Exposure in excess of 155 dB may produce heating effects that are harmful to the human body [136]. A report by the UK’s Health Protection Agency recommends an exposure limit to airborne ultrasound sound pressure levels (SPL) of 100 dB (at 25 kHz and above) [137].

According to Lenhardt [138], “Current exposure standards are based on the concept that detectability and the potential for damage to hearing are related.” Apart from the barely audible by-product low-frequency buzz of phased arrays, its operation is not detectable by hearing or other human senses outside the focus area. Lenhardt recommended 145 dB maximum exposure at 40 kHz, which is comparable to the pressure from a commercially available device.

An investigation into the effects of an ultrasound haptics device on human hearing [131] found that the human ear is typically exposed to 110 to 120 dB of sound pressure during normal use of an ultrasound haptics device. This is significantly lower than the 140 to 150 dB of sound pressure found within a typical focal point [130]. To investigate the potential for hearing loss, they also studied if ultrasound exposure caused a shift in hearing thresholds. Audible sounds (up to a maximum of 8 kHz) were emitted in the presence of the 40 kHz ultrasound. Their results suggest that no threshold shift occurred, leading to the conclusion that an ultrasound haptics device is unlikely to contribute to hearing loss.

A further study into the effects of 40 kHz ultrasound on human hearing [134] found no evidence of a threshold shift indicative of potential hearing loss. They also found no evidence of brain activity (through EEG) in response to the ultrasound frequency or its lower harmonics. Based on other ultrasound safety studies (e.g., [135]), it seems that extreme airborne ultrasound is well tolerated in industrial settings.

Another recent study shows that even far away from the focal region, the limit value of 110 dB will be exceeded near users’ ears [130]. Further studies are necessary to clarify whether the long-term workplace guiding limits apply also to short-term exposure.

These studies suggest that the inaudible sound pressure levels from an ultrasound haptics device may pose little risk to hearing, even though more research is needed. The extreme levels of sound pressure are highly localized at the focal point and the ultrasound carrier does not appear to be perceived by the user. Because ultrasound in the mid-air haptics systems is highly directional, the high sound pressure is largely confined to the space immediately in front of the device, in the direction that
ultrasound travels. However, the side lobes and other issues may cause some problems.

For VR/AR use, an ultrasound array on an HMD directs the exposure on the user’s body (like in Sand et al. [16]). The array can also be grounded in front of the users, facing up towards their hands. With this in mind, we are satisfied that ultrasound haptic feedback likely poses minimal safety risk in most desktop and VR/AR applications.

VII. Conclusions

Contactless haptic technologies can create new opportunities for human-computer interaction. Ultrasound haptic devices, in particular, have captured the imagination of designers and researchers alike, inspiring a large and multi-disciplined research community, as well as making waves in industry. In this paper, we took a broad and detailed look at ultrasound haptics, from its fundamental principles, to its perception and use in state-of-the-art interactive systems. While much research has been done in this area, there is still a great need for more work to fully take advantage of these mid-air haptic capabilities. We conclude our discussion with some perspectives on the key directions for future work in this area.

Our first call to action is for more research to investigate the perception of ultrasound haptic sensations. A limited body of work has used psychophysical methods to gain insight into how these are perceived (Section III.D.2). A better understanding of perception would provide clearer recommendations for designing more effective mid-air haptic feedback. A better understanding of perception is important for reasons other than the design of good feedback and sensations, however. This knowledge can also inform the development of hardware and software tools. For example, recent work by Frier et al. [88] shows how perceptual results can be used to optimize something as easily taken for granted as the hardware sample rate, to create better haptic output. We see work in this area as being beneficial to the entire community, from those developing new hardware and software, to those creating new ‘supernatural’ [15] experiences for users.

Closely related to perception is the development of new methods for rendering ultrasound haptic sensations. The design vocabulary is constrained by the rendering techniques available at the time the work was undertaken. For example, several groups have developed techniques for rendering 2D shapes using a discrete set of amplitude-modulated focal points [54], [63], [64]. Newer rendering methods, such as spatio-temporal modulation, would do this in a different way and may give different results. For example, work by Hajas et al. [67] into ‘dynamic tactile pointers’ suggest considerable nuance is required to render easily identifiable 2D polygons using spatio-temporal modulation. It is, therefore, important for researchers to consider the capabilities of the technologies they are using when looking at past results.

A greater challenge is finding novel ways to render material properties of virtual objects, rather than just spatial properties like their size and shape. Qualities like texture and roughness have received some attention [68], [69], but the work is in its infancy. A critical challenge here is overcoming the difference in size between the coarse-grained focal points used to render ultrasound haptic sensations and the fine-grained physical elements we associate with ‘real’ sensations of texture. An alternative approach is necessary, for example breaking texture down to macro- and micro-elements, which are better suited to the technology [69]; however, this challenge may be insurmountable. Others have had success in creating more experiential sensations, like ‘electricity’ and ‘fire’ [5], [15], but there are no systematic guidelines about how to create such patterns and no knowledge about why some of them have been successful. More work is therefore needed to develop our understanding about what range of tactile experiences can be created using this technology, and how these can be replicated.

Almost all algorithms currently used to synthesize acoustic fields are open-loop, which do not consider the disruptive effect of a person’s hand on the desired sound field. An exception is work by Inoue et al. [48], which adapted the synthesized sound field by considering scattering from a single finger, to create better defined focal points. Closed-loop methods like this could improve the quality of ultrasound haptic feedback. It will of course be necessary to investigate if users can perceive the difference between an open- and closed-loop rendered stimulus, to determine if increased computational complexity is justified.

We discussed four application areas where ultrasound haptics has been used to enhance, or enable the creation of, novel user interfaces. We focused on augmented and virtual reality as this is a particularly promising topic. New application areas will emerge, however, and a particularly timely domain is retail. At the present time, much of the world is investigating how to emerge safely from limitations imposed by the response to COVID-19 pandemic. Many of the interactive systems used in retail contexts involve touchscreens or physical buttons, where hygiene and sterility are key concerns. Touchless user interfaces have potential to address these, with mid-air haptic feedback helping to improve usability and create new ways of supporting the retail experience. Recent work by Paneva et al. [139] investigated an ATM with an ultrasound haptic braille display, as one example of how an existing utility interface could be adapted for touchless haptic interaction.

To conclude, ultrasound haptics is still in its infancy, but has excited and inspired a large body of research to investigate its potential. As this technology becomes more widely accessible to users, researchers and designers, we hope this review serves as a useful reference for understanding its basic principles and the myriad ways in which it can be used, as well as inspiring new research directions and application areas.

Acknowledgment

We thank the reviewers and our colleagues for their very helpful comments, which have greatly improved this paper.

References


To appear in IEEE Transactions on Haptics (DOI: 10.1109/TOH.2020.3018754)


To appear in IEEE Transactions on Haptics (DOI: 10.1109/TOH.2020.3018754) 16


Ismo Rakkolainen received the Ph.D. degree in signal processing from the Tampere University of Technology, Finland, in 2002. He is currently a University Research Fellow at the Tampere University, Finland. He has also been the CTO of FogScreen Inc. He has published over 80 journal and conference papers, 5 book chapters, and he holds 4 patents. His primary research interests include interaction techniques, novel user interfaces, displays and mixed reality. Dr. Rakkolainen has received the 2005 European IST Prize and the national Inno-Suomi Innovation Award.

Euan Freeman is a Lecturer at the University of Glasgow, UK. He received a Ph.D. in computing science from the University of Glasgow in 2016. He has published over 30 journal and conference papers. His primary research interests include mid-air interaction, ultrasound haptics, acoustic levitation and novel mobile interaction techniques.

Antti Sand received the M.Sc. degree in computing science in 2013 from the University of Tampere, Finland, and he will receive the Ph.D. degree in 2020. He is a Researcher at the Tampere University, Finland. He has published 13 journal and conference papers. His research interests include novel spatial user interfaces, immaterial particle screens, ultrasound haptics and novel embodied interaction.

Roope Raisamo received the Ph.D. degree in computing science from the University of Tampere, Finland, in 1999. He is currently a Professor of computer science at the Faculty of Information Technology and Communication Sciences, Tampere University, Finland. He is the head of TAUCHI Research Center (Tampere Unit for Computer-Human Interaction), leading the Multimodal Interaction Research Group. He has published over 300 journal and conference papers. His 25-year research experience in human-technology interaction is focused on multimodal interaction, XR, haptics, gaze, gestures, interaction techniques, and software architectures.

Stephen Brewster received a Ph.D. degree in computing science from the University of York, UK, in 1994. He is currently a Professor of computer science at the University of Glasgow, UK. He has published over 500 journal and conference papers or book chapters. His research interests include multimodal human-computer interaction, sound, haptics and gestures.