Abstract - Mid-air haptic technology works by focusing ultrasonic radiation pressures onto the user’s bare hands, thereby delivering a localized vibrotactile effect. Advanced spatial and temporal modulation of the focused pressures expands the haptic capabilities of this technology and can enable the rendering of volumetric shapes. To that end, we explore ten different algorithms for the haptic rendering of virtual 3D objects and describe their respective implementations, weight their pros and cons, and discuss their applicability in different settings and paradigms. The two best performing algorithms are then user-tested to assess their applicability. We conclude that salient features, and especially object corners, should be emphasized by the mid-air haptic rendering algorithms when shape information is important and provide easy to follow guidelines. We also note that the visual representations of haptified objects can pre-condition our haptic expectations. We thus propose alternative graphic representations in order to mitigate such discordance.

I. INTRODUCTION

Real-time, robust, and accurate 3D skeletal tracking of a user's hands using depth cameras has been the subject of many studies and commercial ventures [1] and has returned truly impressive results [2]. This progress has translated into an increasingly high number of synthetic 3D interfaces [3] not only in virtual reality (VR), but also in other mixed modes and gesture-based interfaces [4]. At the same time, rendering haptic feedback [5] and therefore the act of instilling the sense of touch to dynamic object manipulation has primarily advanced in a slightly orthogonal direction to that of optical tracking. One reason for this has been that most wearable haptic apparatus [6] has relied mostly on accelerometer sensor data or similar rather than optical tracking. As such, the two technologies, optical tracking and haptic feedback have had little interaction, with two notable exceptions: 1) pseudo-haptics (also referred to as visuo-haptics) where the user’s virtual hand is dynamically displaced according to the physics engine of the simulator, thereby creating a physical illusion [7], and 2) mid-air haptics where a phased array of ultrasound speakers is used to create a tactile effect on the user’s bare hands [8] [9]. Both these approaches can improve user interaction and experience, and can amplify the overall immersion of the experience in many different ways [10] [11].

To that end, this paper focuses on mid-air haptics and describes different algorithms for rendering 3D shapes, an emerging part of mixed reality and gestural interfaces and controller-free games. Our main contributions are: 1) We propose 3D shape haptic rendering methods based on hand-object intersections. The collisions between the outer shell of the virtual object and the hand is rendered in a time-efficient manner. Some variations to the algorithm are also explored to change the user perception. 2) We explore different feature rendering methods. Specifically, we adopt the notion that identification of 3D objects is mostly driven by salient features [12], and therefore describe effective methods to haptically render edges and corners in mid-air. 3) We propose accompanying visuo-haptic particle rendering methods as an approach to evoke mental models that are closer to the rendered mid-air haptic sensations.

II. BACKGROUND AND MOTIVATION

We are used to receiving a rich stream of tactile information when interacting with real-world objects. During our daily routines we use our hands to identify these with almost 100% accuracy [13] [14]. Therefore, the role of haptic cues and perception has been the focus of many research studies and in particular in the way that we explore the world with our hands [15]. Similarly, holding and manipulating objects requires “knowledge” of what the object is, its shape and size, its inertial tensor, the material it is made of, its mass, and so on. Without touch information, manipulation becomes nearly impossible or incredibly clumsy. Think for example of trying to put a key into a keyhole when your fingers are cold. Therefore, haptic feedback is crucial for physical interactions in the real world. By extension, haptic feedback is crucial for physical interactions with virtual simulations and controller-free gaming [16].

A large number of studies have haptically rendered 3D shapes before including the use of force feedback devices [13], wearables [6], and morphable devices [17]. A key performance indicator (KPI) has traditionally been the range of achievable shapes and the haptic accuracy thereof. Other studies have focused on rendering a specific feature, for instance the simulation of friction and texture through vibrotactile feedback to create the illusion of touching an edge on a flat or curved display [18]. In general, however, a consistent challenge faced in many of these approaches is that of translating the rich and dynamic stream of tactile information into the often small number of hardware-limited degrees of freedom or the equally small number of touch actuated contact points. In other words, how does one adequately display rich haptics using the available haptic feedback device? On the other hand, with increasing degrees of freedom and touch actuated points, the complexity of the haptic rendering algorithms grows significantly, thus making the task of mapping out optimal techniques and procedures ever so difficult. In the case of mid-air haptics using ultrasonic phased arrays, this complexity is yet to be fully understood since the full palmar region of the hand can be stimulated with high spatial and temporal resolutions.

Mid-Air Haptic Algorithms for Rendering 3D Shapes

Jonatan Martinez, Adam Harwood, Hannah Limerick, Rory Clark and Orestis Georgiou

Ultrahaptics Ltd., The West Wing, Glass Wharf, Bristol, BS2 0EL, UK
To that end, we revisit the challenge of rendering volumetric 3D shapes using mid-air haptics [19], now that new modulation techniques have been proposed [20] and study different haptic rendering methods for 3D shapes. Pros and cons are discussed in each case and implementation details are given. User testing is applied to a subset of these methods providing us with useful qualitative insights and best practices for the 3D haptic rendering algorithms. Finally, we discuss how the visual appearance of the 3D objects precondition our haptic expectations and propose alternative graphic representations in order to reduce discordance and improve user experience.

III. ULTRASONIC MODULATION TECHNIQUES

We use a STRATOS Explore (USX) development kit from Ultrahaptics, a device consisting of an array of 16x16 ultrasonic transducers driven with a carrier frequency of 40 kHz. The transducers are individually electronically controlled to change their phase and amplitudes so that the resulting interference pattern in the acoustic field creates focused points in mid-air. The diameter of the focal points is determined by the wavelength of the ultrasound (about 8.6 mm) thus defining the smallest possible haptic pixel we can render. The focusing range defines an interaction region from 10 to 60 cm, and an operating angle of 60º, thus extending the interaction zone beyond the edge of the board.

The focal points are moved along the surface of the palm, distributing the acoustic radiation force in an optimized way and creating tactile patterns. For this the hand of the user needs to be accurately tracked in space and at a high framerate which is accomplished using a Leap Motion [1] infrared optical tracking controller.

To induce a perceptible tactile effect, an ultrasonic focal point needs to be modulated at a lower frequency than the carrier frequency and within the perceptual peak range of our various mechanoreceptors (5 - 500 Hz). As discussed in [19], two modulation techniques can be used:

Amplitude Modulation (AM): This scheme modulates the intensity of the focal point using a fixed sine waveform thus creating a localised vibrotactile effect. The frequency of modulation is usually fixed at 200 Hz as this has been found to result in maximal response. This scheme is good for static or slowly moving focal points. However, rendering volumetric shapes requires the creation of multiple focal points, thus dividing the power available to the device. For example, in order to create a square, one needs a minimum of four focal points to represent the four corners. Multiple optimisation algorithms have been proposed in [19] that efficiently make use of device available resources.

Spatio-Temporal Modulation (STM): This scheme changes the position of the focal point rapidly and repeatedly along a trajectory while maintaining the ultrasound focal point intensity at its maximum. When the focal point completes the trajectory fast enough it is perceived as a continuous and complete shape such as a circle or square. This technique requires very fast refresh rates in order to maintain resolution as the shapes and the length of the haptic trajectories grows. The USX development kit has a haptic refresh rate of 40 kHz. STM can further adjust the frequency (number of times the trajectory is being traversed per second) or speed (meters per second) of the focal points as to improve haptic perception [20]. Current rule of thumb is to aim for focal point speeds of about 7 m/s.

Within the Ultrahaptics API, the position and intensity of the focal point on a given instant of time is defined using Control Points (CP). On a higher abstraction level, we define a polyline as a sequence of continuous points in space that we want to haptically draw using STM. Polylines are interpolated by taking into account the desired modulation frequency, thus obtaining the paths that are going to be traversed by the focal point. A multiline is defined as a set of joint polylines that are traversed by the CP at approx. 7 m/s.

IV. HAPTIC RENDERING METHODS

The objective we have set out to meet is to devise rendering algorithms that create the perceptible illusion that a user is touching a virtual 3D object. Specifically, in the present study we have focused on creating the perception of volumetric shapes, but other properties such as texture or weight are the subjects of future work. In this section we expose 10 algorithms from A to J that we have developed and tested on the same 3D shapes (cube, sphere and cone), in order to generate the tactile patterns used to modulate the ultrasonic focal points in mid-air and comment on their subjective efficacy in meeting the objective.

The discussion that follows is geared towards providing
trial-and-error insights into what works and what doesn’t, why, and also what side-effects one may expect. See Figure 2 for a graphical representation of the various algorithms. The output of each algorithm (A to J) is the path (a sequence of points) that a focal point will traverse. The green lines are basically one or more groups of discrete 3D positions in space that are being traversed by the focal point in that specific instance of time using STM. The red point in Figure 2C is the discrete position in 3D space where an AM focal point is located. In the following subsections a brief description of each algorithm is presented, followed by a summary of the subjective evaluation of four haptic designers from our in-house team. Note that the palm of the hand is stimulated in the same way as the fingers.

A. Intersection

When we touch an object, it is the areas in contact that are primarily felt through the skin, even though recent evidence supports that tactile vibrations at the contact points diffuse and are excite mechanoreceptors far away from the source [21]. We will therefore limit our study to different ways of rendering the intersection between the virtual object and the tracked hand of a user. Further, since the ultrasonic haptic device cannot prevent a user’s hand from penetrating the object (there is no force feedback effect), the algorithms we describe next calculate and haptically render the intersection between the object outer shell and the palmar surface of the hand in the following sequence: 1) Discretize hand surface, calculate position of the points in 3D space using tracking information and calculate depth of the object at that positions to get a 2D depth map (see Fig. 1A). 2) Binarize and perform edge detection on the 2D depth map (see Fig. 1B). 3) Vectorize raster data to get sequences of connected points (2D polylines) (see Figure 1C). 4) Transform 2D polylines to 3D space using the hand tracking information (see Figure 1D).

Subjective evaluation: The algorithm provides good and strong feedback. It effectively renders the outer shell of the objects, so the subjective sensation is fairly accurate in terms of the object shape, but the objects feel hollow.

B. Features (corners and edges)

Detecting a salient feature of an object is key to recognizing the shape and the object it belongs to [12]. Therefore, the objective here is to render only selected salient features so that the user can more clearly identify the shape of the object. For this algorithm we have selected two types of features, corners and edges. We consider edges as the sharp boundaries between two faces, which are haptically rendered as polylines. Corners are the points where two or more edges meet. We haptically represent them by one of the following 2D STM curves: Circle, Pulsing circle, Rotating line, Triangle, or Random points.

Algorithm Method: 1) Edges and corners are manually defined for the haptic object. 2) Calculate intersection between the edges and the bounding box of phalanges and palm and generate a multiline. Optionally project them to the hand surface. 3) Calculate the corners that collide with the bounding box of phalanges and palm. Optionally project them to the hand surface. Generate a multiline parallel to the surface of the skin using a 2D drawer. 4) Combine multilines from edges and corners.

Subjective evaluation: The method works well for some objects such as the cube and the pyramid. Others such as the sphere do not have any of the selected salient features and thus needs to be rendered using a different method. Edges perpendicular to the hand are barely noticeable (and especially those that are also perpendicular to the USX). From all the different 2D drawers we have tested, a rotating line is perceived as the strongest sensation for corners.

C. Intersection + AM deepest point

In an attempt to add a depth sensation while touching a virtual object, the point of maximum penetration into the
object is also rendered. This has been implemented using two simultaneous focal points. One CP renders the results of the intersection method A using STM while the other focal point uses the AM scheme with the CP at the location of maximum penetration depth.

Subjective evaluation: The point rendered using AM is barely noticeable. Adding a second focal point adds a significant side-effect of amount of audible noise.

D. Intersection solid

This is a different approach from than in A that aims to avoid feeling the objects as if they are an empty shell. This algorithm renders one multiline following the contour of the object (like the intersection algorithm A) and one more near the point of maximum penetration into the object.

It has been implemented using a modified version of the intersection algorithm. First an intersection is calculated performing the threshold step at penetration depth as usual. Then, a second intersection is calculated with the threshold step configured at the maximum penetration depth found in the first step. The two multilines obtained have been tested using one and two simultaneous focal points.

Subjective evaluation: With this algorithm the sensation of a hollow shell is reduced at the cost of a weaker overall strength. This is due to the increased length of the rendered polylines, which spreads the available energy output of the USX. Using two focal points simultaneously provides a similar haptic sensation but produces more audible noise.

E. Solid random

This algorithm renders small random polylines on the regions of the hand that are inside a virtual object.

Subjective evaluation: Diffuse sensation, the intensity is inversely proportional to the area stimulated which is as one would expect, due to the spread of device power.

F. Solid texture

This algorithm renders small random polylines on the regions of the hand that are inside a virtual object. Unlike the Solid random algorithm E, the randomness is generated as a function of space, so that the polylines don’t change unless the user moves their hand.

Subjective evaluation: Diffuse sensation, but more intense than the one before. Increased sensation of granular texture inside the object, with more spatial coherence.

G. Blobs

This algorithm draws circular polylines on the fingertips and at the palm with a radius proportional to the penetration into the object.

Subjective evaluation: The overall intensity is weak. The length of the polylines tends to be higher than with other methods, so the energy is more spread.

H. Features + Intersection (per finger / gap)

Both the salient features and the intersection are rendered combining both multilines. Features are prioritized using two different approaches.
- Finger based: Any finger that is touching a salient feature (corner or edge) does not receive intersection haptics.
- Gap based. A minimum distance or gap between features and the intersection is required to render the intersection, thus making them more discrete.

Subjective evaluation: The approximation based on “per finger” produces confusing results, as any collision of a finger with a salient feature disables the intersection haptics on that finger, producing haptic discontinuities. The “gap” solution seems to be more stable in its implementation, thus creating a good combination of the two algorithms (A & B).

I. Features + Intersection Solid

This is the combination of rendering salient features and the intersection solid algorithm. The prioritization algorithm based on distance has been used.

Subjective evaluation: Weak sensation due to excessive length of the polylines.

J. Features + Blobs

This is the combination of rendering salient features preferentially and blobs otherwise with prioritization based on distance.

Subjective evaluation: Weak sensation due to excessive length of the polylines.

V. USER STUDY

After the subjective evaluation of the ten previously described rendering algorithms, we selected the two best performing ones to proceed with a more in-depth user study; A. Intersection and H. Feature + Intersection (gap). The first one provides strong and accurate perception of the collision hand-object, and hence a good baseline for comparison. The second method is of interest to determine the relevance of adding salient features to the haptic rendering output. Both these algorithms produce stronger subjective haptic sensations which is primary KPIs we seek to optimize. The secondary KPI is the accuracy of the perceived 3D shape. The objective of the user study was hence to develop insights on how people expect to feel a virtual object and to determine what strategies users applied to compare the two different haptic rendering methods. These would help us to optimise the haptic experience of 3D object interactions using the appropriate rendering methods.

A. Procedure

A qualitative study was conducted, which involved two tasks. The first task was designed to determine what people
expect to feel when visually presented with a virtual 3D object. The second task was designed to uncover what features are important when comparing two different mid-air haptic rendering methods. Three regular geometric objects were used for the experiment: cube, cone, and sphere.  

1st Task: During the first task, people were presented with a visual representation of the objects but with no haptic rendering. Participants were asked to reach out and move their virtual hand towards the virtual object. We used a speak out loud methodology meaning that participants were asked to describe what they would expect to feel if they could feel something. Here, we captured their expectations about what the haptic rendering should represent.

2nd Task: For the second task, participants were presented with two instances of the same virtual objects displayed on two separate haptic USX devices, each using a different ultrasonic rendering method. All the objects (cube, cone, sphere) were contained within a 10 cm³ volume at 20 cm height centred above the USX. Participants were asked to explore the two objects with their hands and answer the following questions while also giving a confidence ranking on a 5-point Likert scale, 5 being the most confident:

Q1. Which of the two sensations is the most similar to the shape of the object seen on the screen?  
Q2. Which sensation feels the strongest?  
Q3. Which sensation has the most obvious surface?

Tasks were split into blocks and trials and counterbalanced to reduce any possible bias. The answers to the questions were recorded and a short post-task interview was performed to discuss their experience. The experimental setup can be seen in Figure 3.

B. Results

Sixteen participants took part in the user study experiment. They were aged between 23 and 42 (mean 35) with six of them female. All participants had experience in mid-air haptics. Their anonymous responses to the questions were recorded by the experimenter.  

1st Task: The qualitative results from task one were reviewed after the session and statements were classified into three common themes: 1) visual representation, 2) hand exploration movements, and 3) haptics.

1) Visual representation: Participants had expectations about the weight of the object, material properties and behaviour in response to the hand. The visual representation used in the experiment suggested a heavy and solid object.  
2) Hand exploration: The following movements in relation to the shape were commonly exhibited:  
   - Moving the fingertip to the corner of the cube.  
   - Moving fingertips over the surface of the shape.  
   - Holding finger and thumb to opposite sides of the shape.  
   - Holding the shape in the centre of the palm and expecting to feel the shape outline.  
   - Grasping shape in the centre of the hand and wrapping fingers around it.  
   - Not expecting to be able to penetrate the object.  
3) Haptic representation: Participants had expectations around what they would feel given the shape. Salient points were noted frequently, such as the shape corners, the edges, and the slope of the cone. A consistent haptic feeling across areas of the hand were expected, with some surface texture.  

2nd Task: The selections made in task two in response to the 3 questions were fairly equal across the two haptic rendering methods with a slight preference for algorithm H in all questions: 56%, 63%, 63% voted for H instead of A. Taking into consideration the confidence scores (weighting each vote by the reported confidence scores) gave further weight to algorithm H with the revised results being: 61%, 67%, 70%. Due to the small sample size we were not expecting any significant quantitative response, rather to understand how people assess virtual objects and how they use the haptic cues to discriminate.

The following conclusions were reached. Firstly, the haptic perception of the shapes did not match the users’ expectations they had from visuals in terms of weight or solidity. Secondly, salient features are extremely important in our case for using mid-air haptics to discriminate between two virtual objects. Objects with corners (such as a cube) are better rendered with the feature rendering method, however edges were not so clearly perceived. Moreover, one of the limiting factors we discovered was that edges that are nearly perpendicular to the haptic device are barely perceived. If the hand is parallel to the device then the edge projection on the skin results in a very short polyline. On the contrary, if the hand is perpendicular then the acoustic radiation force is not efficiently transmitted to the surface of the skin.

![Particle-based visualization proposed for 3D objects.](image)

VI. LIMITATIONS AND DISCUSSION

The main limitations towards attaining our objective of devising haptic rendering algorithms of virtual 3D objects derive from the nature of the system. Ultrasound mid-air haptics offer primarily a vibrotactile feedback effect. Further, the small working volume offered by a 16x16 transducer array limits exploratory hand movements and therefore the number and size of 3D object that can be rendered within the interaction zone. This is further restricted since a user’s palm must be facing the device in order to receive mid-air haptics.

The limited amount of power available from the acoustic radiation force highlights the importance of optimizing the energy distribution between focal points. The longer the polylines we render, the weaker the sensation will be. It also makes it impractical to render polylines with different degrees of intensity, which is the main challenge faced when trying to highlight features (such as edges and corners) from the rest of the haptic shape.
Hand-object intersections are well resolved by the algorithms we have described, however rendering haptics for situations where large areas of the hand are touching the object are still challenging, due to the limited amount of power available to the USX device. Adding more devices arranged in different planes, e.g., at 45-90 degrees to each other would solve the limitation in exploration movements thus making the sensations stronger.

After performing the user study, it was clear that the standard solid appearance of the 3D objects evoked mental models to the participants that did not match the penetrability of the visual effect of the avatar hand going through the object. Thus, we propose a particle-based soft appearance of 3D objects as an alternative shown in Figure 4. A future user study will analyse the perceptual differences when using such accompanying graphics.

Finally, we while the salient feature rendering showed promising results for algorithm $H$, further research is needed to explore more efficient rendering methods as well as different kinds of features such as curvature and texture.

VII. CONCLUSION

This paper investigated different algorithms to render 3D shapes based on hand-object intersection information and object salient features using an ultrasonic phased array. The algorithms proposed herein utilise a combination of state-of-the-art modulation techniques (AM and STM) and explore the pros and cons of haptically rendering different features during virtual hand-object interactions. After testing these algorithms and giving our subjective opinions about their performance, the paper describes a user study that compared the best performing two algorithms in order to extract further insights and guide future implementations.

We concluded that intersection algorithms successfully allowed participants to feel the outer shape of the objects. Additional rendering of salient features can be very beneficial towards the transmission of shape information, particularly for corners, but no so much for edges, thus further helping in identifying some of the shape properties. Finally, a particle-based soft graphical appearance of 3D holographic objects was proposed to better manage pre-haptic expectations, thus minimising the difference between what one expects to feel, and what haptic feedback they actually experience from ultrasonic mid-air haptic interfaces.

ACKNOWLEDGMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 801413, H-Reality.

REFERENCES

[1] Leap Motion Inc, 2019. http://leapmotion.com
[2] Melax, S., Keselman, L. and Orsten, S., Dynamics based 3D skeletal hand tracking. In Proceedings of Graphics Interface 2013 (pp. 63-70). Canadian Information Processing Society.
[3] Dachselt, R. and Hübner, A., 2007. Three-dimensional menus: A survey and taxonomy. Computers & Graphics, 31(1), pp.53-65.