User-Defined Mid-Air Haptic Sensations for Interacting with an AR Menu Environment

Lawrence Van den Bogaert[(✉)] and David Geerts

Meaningful Interactions Lab (Mintlab), KU Leuven, Leuven, Belgium
{lawrence.vandenbogaert,david.geerts}@kuleuven.be

Abstract. Interfaces that allow users to interact with a computing system by using free-hand mid-air gestures are becoming increasingly prevalent. A typical shortcoming of such gesture-based interfaces, however, is their lack of a haptic component. One technology with the potential to address this issue, is ultrasound mid-air haptic feedback. At the moment, haptic sensations are typically designed by system engineers and experts. In the case of gestural interfaces, researchers started involving non-expert users to define suitable *gestures* for specific interactions. To our knowledge, no studies have involved – from a similar participatory design perspective – laymen to generate mid-air haptic sensations. We present the results of an end-user elicitation study yielding a user-defined set of mid-air haptic sensations to match gestures used for interacting with an Augmented Reality menu environment. In addition, we discuss the suitability of the end-user elicitation method to that end.

Keywords: Mid-air haptics · End-user elicitation study · Participatory design

1 Introduction

While advancements in gesture recognition technologies (e.g., Microsoft Kinect, Leap Motion) open up exciting new ways of human-computer interaction, one important constraint they face is their lack of a haptic component. One technology that carries the potential to address this shortcoming, without impeding the user with wearables or hand-held devices, is ultrasound mid-air haptic feedback [2,3]. By generating pressure fields in mid-air via an array of ultrasound transducers, this technology allows the user to experience a sense of touch on the palm and fingers of the unencumbered hand. The ultrasound waves that create these pressure fields can be meticulously altered, opening up a quasi-infinite set of sensations, shapes and patterns. Typically, it is experts and specialists who are tasked with their design. For gestures, however, it has already been argued that even though designer-designed gestures might be expertly crafted, end-user elicited gestures are more intuitive, guessable and preferred [4]. Accordingly, we believe this might be true for mid-air haptic sensations too. To our knowledge,

so far only one other study has involved end-users to create mid-air haptic sensations [5]. However, whereas their goal was for participants to create mid-air haptic sensations to convey and mediate emotions to others, we wanted to test whether novice users can ideate haptic sensations to match gestures used for operating a menu in Augmented Reality (AR). To this end, we turned to Wobbrock et al.'s *end-user elicitation methodology* [12]. Typically, such elicitation studies present participants with one or more *referents* for which they are asked to ideate a suitable *symbol*. A referent is the effect of an interaction with an interface (e.g., an increase in volume on a music player). The symbol is the corresponding action (input) that is required to invoke the referent (e.g., turning the volume knob clockwise). In principle, a symbol can be anything that computing systems are able to register as input. This method has been used to design symbols (input) to actuate, among others, touchscreens [13], virtual & augmented reality [6], smart glasses [9], home entertainment systems [10] and public displays [7]. Whereas up until now participants of end-user elicitation studies were tasked with the ideation of the symbol, we assigned ours with the elicitation of what we suggest to call *intermediary referents*. 'Intermediary' means it concerns not the main system output (referent), but rather a form of feedback that accompanies it. Common examples might be the vibrotactile feedback we feel when typing on our smartphone, or the beeping tones we hear when adjusting the volume.

With this work, we present a set of user-defined intermediary referents in the form of mid-air haptic sensations to match a set of gestures used for interacting with an AR interface. We discuss the eligibility of the end-user elicitation method for their design.

2   Study Setup

Twenty-four non-specialist participants were invited individually to our lab in Leuven, Belgium to elicit a set of five mid-air haptic sensations to match five gestures actuating an AR menu. Given the novelty of mid-air haptics, participants were recruited from a list of people who had previously taken part in a study involving ultrasound mid-air haptics. They were still non-experts, but had at least a basic understanding of what the technology does and how it feels. Ages ranged from 19 to 56 ($\mu = 26, 2$). Ten participants were male, fourteen were female.

2.1   Apparatus and Gesture Selection

For our participants to interact with an AR menu, we chose Microsoft's HoloLens as it is by default gesture-controlled and free of hand-held controllers or wearables. Three proprietary gestures allow for its actuation: 1) 'bloom' (used to evoke the home menu at any time to easily navigate in and out of applications); 2) 'air tap' (used to select things, equivalent to a mouse click on a desktop); 3) 'tap and hold' (equivalent of keeping a mouse-button pressed, e.g., for scrolling

on a page). 'Air tap' was left out because it is similar to 'tap and hold' in every aspect except for its duration ('air tap' being more brief and thus less suitable for mid-air haptic actuation). In addition to the default HoloLens commands, we turned to Piumsomboon et al.'s user-defined AR gesture set [6] to supplement the HoloLens gestures. From it, we selected three additional gestures for our study: a) 'swipe' (used to rotate a carousel); b) 'stop' (used to pause, e.g., a song or video); and c) 'spin [index finger]' (used to increase or decrease the volume). We chose these three gestures because they were directly applicable to interactions that are possible in the existing HoloLens menu environment (e.g., by using the multimedia player) and because they leave at least part of the palm and fingers unadorned and thus free to be actuated by ultrasound waves (whereas, e.g., a clenched fist would occlude each operable area of the hand).

2.2 Procedure

The goal of the study was explained to each participant as follows: while wearing a Microsoft HoloLens, they would execute five basic menu-related tasks in AR by using bare-hand gestures in mid-air. For each gesture, they would ideate a mid-air haptic sensation that suits this gesture best and feels most 'logical' for that interaction. At the start of each session, participants were presented a set of ten distinct mid-air haptic sensations to (re)acquaint themselves with the different adjustable properties (i.e., location on the hand, shape of the pattern, duration, dynamics and single-point vs. multi-point feedback).

Using the five gestures described above, participants then completed a set of tasks in AR. For each gesture, participants were encouraged to think out loud, depict and describe what mid-air haptic sensation they would want it to be accompanied by. Sessions had a duration of approximately 45 min. Because wearing an AR headset for a prolonged time becomes cumbersome and considering the time required to elicit and explain in detail each mid-air haptic sensation, we deliberately kept the set of tasks relatively small in comparison to other elicitation studies. For each participant, the entire process was video and audio recorded. In addition, the study moderator took notes; inquired for clarification, and reminded participants to go over all the adjustable properties that constitute a mid-air haptic pattern (cfr. supra).

Naturally, the HoloLens is by default unable to detect Piumsomboon's gestures and respond to them. This is why during each session, the proprietary HoloLens gestures were always visited first. After having performed the two 'functional' HoloLens gestures and having ideated a matching mid-air haptic sensation for them, participants were able to more easily perform the 'non-functional' gestures from Piumsomboon's user-defined set and imagine them having the targeted effect. No participants indicated having difficulties with this.

As to not curb creativity, we emphasized that technical limitations and feasibility did not have to be considered. Participants were told to simply imagine what they would want to feel and not worry about how the mid-air haptic sensation would be emitted on the hand. For example, we explained that the mid-air

haptics could be emitted from the HoloLens itself (as in a similar setup presented by [8]) rather than from a unit standing on the table.

3   Analysis

All elicited sensations were analyzed and categorized based on the researcher transcripts and audiovisual recordings made during the elicitation phase. Because we allowed participants a high degree of freedom, often minor variations on similar ideas and patterns were proposed. Rather than discriminating sensations from each other for each identifiable objective inconsistency, we assessed and categorized them on a conceptual level. For example, to summon the 'home menu tile' in AR by making a 'bloom' gesture, 11/24 participants suggested a single short mid-air haptic sensation felt on the palm of the hand right after finishing the movement. Whether this sensation had the shape of, e.g., a square or a circle was not deemed imperative to the conceptualization and thus all 11 designs in which 'a short sensation on the palm to confirm that the gesture was well registered' was described, were placed in the same conceptual group.

   As such, all 120 ideated sensations were classified based on their conceptual similarity. This resulted in 26 different conceptual groups, each representing a distinct conceptual model of mid-air haptic feedback. To understand the degree of consensus on the conceptual groups among participants, the revised agreement rate AR by Vatavu & Wobbrock [11] was calculated for each gesture:

$$AR(r) = \frac{|P|}{|P| - 1} \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|} \right)^2 - \frac{1}{|P| - 1}$$

where P is the total amount of elicited sensations per task and Pi is the subset of sensations with a similar conceptual model for that task. To revisit 'bloom' as an example; the conceptual groups contained respectively 11, 8, 2, 2 and 1 user-elicited mid-air haptic design(s). The agreement score for this gesture is then calculated as follows:
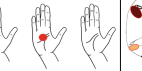
$$AR_{bloom} = \left(\frac{|24|}{|23|} \right) \times \left(\frac{|11|}{|24|} \right)^2 + \left(\frac{|8|}{|24|} \right)^2 + \left(\frac{|2|}{|24|} \right)^2 + \left(\frac{|2|}{|24|} \right)^2 + \left(\frac{|1|}{|24|} \right)^2 - \left(\frac{|1|}{|23|} \right) = 0.308$$

4   Results

Each participant ideated one mid-air haptic sensation for each of the five gestures, resulting in a total of 120 sensations. Table 1 shows the agreement rate for each gesture/task.

Bloom. The bloom gesture, used to summon the home menu tile at any given moment in the HoloLens environment requires all 5 finger tips to be pressed together and then opened into a flat hand, palm facing up, with space between

Table 1. Consensus set of user-defined mid-air haptic sensations

Gesture	Bloom	Tap and hold	Swipe	Stop	Spin
Visual depiction (darker = stronger)					
Agreement rate (AR)	0.308	0.308	0.228	0.576	0.297

each finger. Out of 24 participants, 11 suggested to receive a simple, short one-time feedback on the hand palm that immediately disappears again and confirms that the gesture has been registered well (AR = 0.308). The other conceptual groups were a) 'opening circle on the palm of the hand to mimic the menu being *spawned* from the hand (n = 8); b) 'constant sensation on the entire palm that remains present as long as the menu is opened' (n=1); c) 'horizontal line rolling from the bottom of the palm up to the fingertips' (n = 2); and d) 'mid-air haptic sensation on each fingertip that becomes sensible starting from the moment the fingertips are pressed together until the moment the gesture is completed (n = 2).

Tap and Hold. The 'tap and hold' gesture is used to scroll through pages and requires the user to pinch together their stretched-out index finger and thumb and maintain this posture in order for the scroll modus to remain activated. Users can then lift or lower the hand to move the page up and down. Scroll speed can be adjusted by moving the hand further up and down in respect to the original location. The conceptual sensation that was elicited most often here (n = 11, AR = 0.308) was 'a constant sensation on the tip of the thumb and index finger that remains sensible as long as the scroll functionality is active, with the intensity depending on the speed at which the user scrolls'. An almost similar design proposed by 8 other participants also suggested a constant sensation on the tips of the index finger and thumb. However, they made no notion of the intensity corresponding to the scrolling speed.

Swipe. To rotate through a carousel, HoloLens users normally have to point and click on elements in the far left or far right of that carousel. We, however, asked our participants to imagine that instead, they could use a full-hand 'swipe' to rotate the elements in the carousel. We clarified that this 'swipe' gesture was appropriated from Piumsomboon et al.'s user-defined gesture set for AR, and that it would not actually work with the HoloLens. Nonetheless, participants were asked to perform the gesture multiple times and imagine the carousel actually rotating. None of them indicated having trouble imagining this work and multiple participants spontaneously commented that this would indeed increase the ease of use. With an agreement rate of 0.228, the conceptual group on which there was most consensus contained the design of 10 participants who suggested to feel a 'vertical line that "rolls" over the entire length of the hand (i.e., from

the bottom of the palm over to the fingertips) while making the gesture'. The two (out of five) other most popular conceptual groups were 'sensation (static) that decreased in intensity while moving the hand' (n = 4) and a concept similar to the consensus sensation but with the line being fixed to one location instead of 'rolling' over the hand (n = 5).

Stop. Using the HoloLens media player, participants watched a short movie and were asked to imagine they could pause it by briefly holding out a flat hand ('stop gesture') in front of them. This was the second non-functional gesture, yet again, none of the participants indicated having trouble to imagine it work. The consensus on elicited sensations for this gesture was 'very high' (AR = 0.576) [11]. Eighteen participants proposed 'a single short sensation to confirm that the gesture has been registered', similar to the consensus sensation for 'bloom'. The other 6 participants (n = 3, respectively) wanted to feel a) 'a sensation on the hand as long as the movie remained paused (as opposed to one short confirmation)'; and b) 'a sensation that enlarged and/or intensified while the hand was brought forward, to only be emitted at full strength when the arm was completely stretched out and the screen paused.

Spin. We asked participants to imagine that they could adjust the volume of the AR environment by spinning their stretched index finger clockwise (increase) or counter-clockwise (decrease). Seven different conceptual groups were needed to classify the heterogeneous ideas that were elicited for this gesture. The design on which there was most consensus (n = 12, AR = 0.297) was described as 'a constant sensation on the index fingertip or on the entire index finger, with intensity increasing or decreasing according to the volume'

5   Discussion

The presented sensations constitute a user-defined consensus set of mid-air haptic sensations that match gestures used to interact with an Augmented Reality menu environment. Despite the quasi-infinite range of possible mid-air haptic sensations and the idiosyncrasy of its features (timing, dynamics, location on the hand, intensity, ...), the agreement scores of our final five gestures (between 0.228 and 0.576) can be regarded as medium (0.100–0.300) to very high (>0.500) according to [11]. As the consensus set shows, the majority of participants seemed to prefer relatively simple and straightforward sensations to amplify their gestures. In addition to an inclination towards non-complex sensations, also noteworthy was how participants used mid-air haptic *intensity* in their elicited sensations. Depending on the functionality of the associated gesture, in some cases intensity was indeed a key feature. When the gesture actuated a discrete function (e.g., bloom and stop), a short monotone sensation was usually preferred, whereas more continuous menu-actions (e.g., scrolling through a page or adjusting volume) came with more continuous (i.e., changing intensity) sensations.

Administering the end-user elicitation method to generate what we propose to label as *intermediary referents* instead of *symbols*, shows promise but requires some remarks too. Our study suggests this method to be useful for the ideation of user-defined *mid-air haptic* sensations, however, for other sensory modalities it might not be as easy to have non-experts elicit novel designs and forms. Novices may not be familiar enough with the adjustable variables and properties of a sensorial modality in order to ideate variants of it. This contrasts with the end-user elicitation of, e.g., gestures, as people are apprehensive of their own physical possibilities (and limitations) and therefore inherently more capable of expressing themselves physically. Asking untrained participants to ideate, e.g., auditory intermediary referents (sound design), would assumedly require more training and/or facilitating tools. The think-aloud protocol that was used for our study, in combination with the available UltraHaptics kit on which simulations of elicitations could be depicted, was in our case proficient. However, when having end-users elicit other types of intermediary referents, we advise to assess well up front whether a modality is at all suitable for end-user elicitation and to think of the means or tools necessary to allow participants to elicit new variants of it.

Finally, one way to further validate the results of an elicitation study is what Ali et al. [1] describe as end-user *identification* studies. They are the conceptual inverse of elicitation studies: they present participants with a symbol and ask which referent would be invoked by it. In the case of intermediary referents, then, a control group would be asked to identify from a set of mid-air haptic sensations, which one suits a specific symbol (gesture) best. This would be a beneficial topic for follow-up research.

References

1. Ali, A.X., Morris, M.R., Wobbrock, J.O.: Crowdlicit: a system for conducting distributed end-user elicitation and identification studies. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, pp. 255:1–255:12. ACM, New York (2019). https://doi.org/10.1145/3290605.3300485, event-place: Glasgow, Scotland Uk
2. Carter, T., Seah, S.A., Long, B., Drinkwater, B., Subramanian, S.: UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In: Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST 2013, pp. 505–514. ACM, New York (2013). https://doi.org/10.1145/2501988.2502018, event-place: St. Andrews, Scotland, United Kingdom
3. Iwamoto, T., Tatezono, M., Hoshi, T., Shinoda, H.: Airborne ultrasound tactile display. In: ACM SIGGRAPH 2008 New Tech Demos. Los. ACM (2008)
4. Morris, M.R., Wobbrock, J.O., Wilson, A.D.: Understanding users' preferences for surface gestures. In: Proceedings of Graphics Interface 2010, GI 2010, pp. 261–268. Canadian Information Processing Society, Toronto, Ont., Canada (2010). http://dl.acm.org/citation.cfm?id=1839214.1839260, event-place: Ottawa, Ontario, Canada
5. Obrist, M., Subramanian, S., Gatti, E., Long, B., Carter, T.: Emotions mediated through mid-air haptics. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 2053–2062. ACM, New York (2015). https://doi.org/10.1145/2702123.2702361, event-place: Seoul, Republic of Korea

6. Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2013, pp. 955–960. ACM, New York (2013). https://doi.org/10.1145/2468356.2468527, event-place: Paris, France
7. Rodriguez, I.B., Marquardt, N.: Gesture elicitation study on how to opt-in & opt-out from interactions with public displays. In: Proceedings of the Interactive Surfaces and Spaces on ZZZ - ISS 2017, pp. 32–41. ACM Press, Brighton (2017). https://doi.org/10.1145/3132272.3134118, http://dl.acm.org/citation.cfm?doid=3132272.3134118
8. Sand, A., Rakkolainen, I., Isokoski, P., Kangas, J., Raisamo, R., Palovuori, K.: Head-mounted display with mid-air tactile feedback. In: Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology, VRST 2015, pp. 51–58. ACM, New York (2015). https://doi.org/10.1145/2821592.2821593, event-place: Beijing, China
9. Tung, Y.C., et al.: User-defined game input for smart glasses in public space. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 3327–3336. ACM, New York (2015). https://doi.org/10.1145/2702123.2702214, event-place: Seoul, Republic of Korea
10. Vatavu, R.D.: A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. J. Ambient Intell. Smart Environ. **5**(2), 187–211 (2013). http://dl.acm.org/citation.cfm?id=2594684.2594688
11. Vatavu, R.D., Wobbrock, J.O.: Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 1325–1334. ACM, New York (2015). https://doi.org/10.1145/2702123.2702223, event-place: Seoul, Republic of Korea
12. Wobbrock, J.O., Aung, H.H., Rothrock, B., Myers, B.A.: Maximizing the guessability of symbolic input. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2005, pp. 1869–1872. ACM, New York (2005). https://doi.org/10.1145/1056808.1057043, event-place: Portland, OR, USA
13. Wobbrock, J.O., Morris, M.R., Wilson, A.D.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2009, pp. 1083–1092. ACM, New York (2009). https://doi.org/10.1145/1518701.1518866, event-place: Boston, MA, USA