HapToes: Vibrotactile Numeric Information Delivery via Tactile Toe Display

Preeti Vyas, *Member, IEEE*, Feras Al Taha, Jeffrey R. Blum, and Jeremy R. Cooperstock, *Member, IEEE*

***Abstract*— Tactile rendering of numeric information via a single actuator has been considered for such purposes as fitness progress tracking. However, multi-actuator designs, leveraging spatial mapping, may offer superior performance. Motivated to explore this approach without requiring hardware on the fingers or wrist, we designed HapToes, a novel ten-digit spatial mapping of numeric information to the toes, which overcomes inter-toe discrimination ambiguity. Compared to ActiVibe, a single-actuator wrist-based numeric rendering technique, under similar distraction conditions, HapToes demonstrates equivalent performance for single-value identification, and improved accuracy, response time, and cognitive load when conveying three values sequentially in a single message.**

I. Introduction

Wearable devices often use haptic actuators to convey information in the form of tactons [1]. Applications such as navigation (http://lechal.com) [2], may benefit from the ability to convey numeric information such as distance and time. For fitness and health monitoring systems, it is helpful to display numeric information such as progress rate [3], or physiological parameters [4]. In systems employed in control room environments [5], it may prove valuable to convey monitoring parameters mapped to numeric information. Certain applications can also benefit from haptic number rendering to provide feedback on user input [6].

A single actuator can be a convenient and efficient mechanism for rendering numeric information haptically as demonstrated by Cauchard *et al.* in their ActiVibe study [3]. However, single-actuator based renderings can be restrictive due to actuator constraints and to a limited number of distinguishable patterns that can be achieved when varying parameters of frequency, amplitude and duration [7]. Multi-actuator systems add a spatial aspect to the haptic rendering, allowing for more distinguishable patterns, and ultimately providing a larger design space for hapticians [7]. These systems benefit from having rendering locations with better two-point discrimination. Researchers have used multi-actuators displays in wearables to design spatio-temporal haptic patterns [8], [9] demonstrating promising results. However, these tactons have been used almost exclusively for categorical or abstract representations. The semantic associations of such abstract patterns to concrete numerical information might not be obvious. Indeed, abstract representations for ordinal information would seem likely to impose an additional layer of cognitive load.

Since there is evidence for a spatially defined mental "number line" [10], it is natural to consider the fingers as a particularly effective locus for spatial tactile display to represent values of 1–10. However, rendering haptic patterns on fingers may interfere with day-to-day interactions with the world, as we hold or manipulate objects, or perceive the environment through touch. As such, toes may serve as a more practical ten-unit tactile display alternative. They may equally facilitate location-based semantic associations [11] since, as fingers they are physically separated from each other.

Unfortunately, the differentiability of haptic stimuli applied to individual toes, especially for the middle toes, is worse than that of fingers [12], [13]. This limitation prompted our investigation of an improved tactile toe encoding method [14], for which we achieved perceptual accuracy approaching that of the fingers. As a result, we find that the toes are a viable location for delivery of tactile numeric information.

In practice, foot-based haptic displays suffer from imperfect coupling of actuators and dampening of vibration in mobile conditions such as during walking and running activity [15]. Demonstrating the ecological validity of such a rendering requires the development of compact hardware for shoes as well as field testing. As such, the results presented here can be seen as an initial exploration of the possibilities of tactile communication through the toes, which can be used for seated applications such as rendering alert messages in control room environments [5] or conveying information to users in smart wheelchairs. However, significant further work is required before the solution can be considered for mobile real-world deployment.

Our contributions include (1) the design of *HapToes*, a novel haptic numeric rendering on toes based on our fundamental toe perception study [14], and (2) the findings of a laboratory user study comparing our method with the existing ActiVibe haptic numeric rendering proposed by Cauchard *et al.* [3], but under distraction conditions, and for rendering both single and three values. We will use the term *values* throughout to refer to the discrete information presented to users in our study, which are the numerals 1–9, and 10.

P. Vyas https://orcid.org/0000-0003-4608-2542,
J. R. Blum https://orcid.org/0000-0002-0835-4343, and
J. R. Cooperstock https://orcid.org/0000-0002-3392-2541
are with the Centre for Interdisciplinary Research in Music Media and Technology and the Department of Electrical and Computer Engineering, McGill University, Montréal, Canada.

F. Al Taha https://orcid.org/0000-0001-5588-9431
is with the Department of Electrical and Computer Engineering, McGill University, Montréal, Canada.

{preeti.vyas, feras.altaha, jeffrey.blum}@mail.mcgill.ca, jer@cim.mcgill.ca

Manuscript received September 27, 2019; revised *blank*.

II. Background

Researchers have explored numerosity, the ability to count sequential tactile, visual or audible pulses, in terms of how it is represented in the human brain [10] and compared numerosity discrimination across these modalities [16]. Most haptic numeric rendering applications, including progress tracking [3], reporting time [17], and feedback for menu selection [6] used single-actuator temporal patterns rendered on the wrist using a smartwatch, or to the hand via a smartphone. Other works have used multi-actuator rendering on locations such as the back [4], arm [5], or fingers [18] to convey alpha-numeric characters using spatio-temporal renderings.

Cauchard *et al.* [3] created ActiVibe, a temporal vibrotactile rendering on a commercial smartwatch, to convey numbers from 1 through 10. ActiVibe achieved a recognition rate of up to 96 % in a laboratory study and 88.7 % when used in the wild for conveying single values. The authors suggested that incorporating a pre-signal vibration before the actual haptic pattern may prove useful to cue users about incoming vibrations. Blum *et al.* [19] performed a follow up study using ActiVibe with such a pre-signal vibration, and compared it against other duration-based methods under distraction conditions for rendering single and three values. The results suggested that ActiVibe maintains a performance advantage in terms of accuracy and subjective preference against other methods for rendering single values but not for three values. Participants not only demonstrated decreased performance with ActiVibe when multiple values were rendered at a time, but also in conjunction with the added load of the distractor task. This raises questions about the robustness of ActiVibe's rendering strategy for real life scenarios. We hypothesized that as an alternative, a spatial, multi-point approach may prove effective for number rendering since it may reduce the complexity and rendering duration of vibrotactile patterns.

We note that ActiVibe is a strictly temporal, rather than spatio-temporal, haptic numeric rendering technique. Nevertheless, given the dearth of alternatives, and in particular, the lack of multi-actuator systems that have been evaluated for numeric information delivery, we chose this technique as a benchmark against which to evaluate the performance of our method. We are interested in this comparison because prior research suggests that the mental "number line" is not only intimately related to space, but also to time [10].

Prior research has explored haptic rendering on the feet for use cases where the visual and audio channels are overloaded [2], and when the hands are occupied with other tasks [20]. A haptic rendering on the feet offers advantages of being discreet, and its hardware assembly can be embedded in footwear. This approach has been explored for conveying information in various applications including navigation (`http://lechal.com`) [2], physical training [20], and dance training [21]. In all these applications, a small set of semantic messages, such as directions, are conveyed by vibrotactile signals at different locations beneath the feet. However, to the best of our knowledge, the rendering of broader sets of more complex information has yet to be explored for feet. This may require either designing a complex spatio-temporal rendering or increasing the number of spatial rendering locations.

Haptic rendering on the toes has been explored in prior work. Panarese *et al.* [22] reported that humans can incorporate spatial force feedback to the toes into their sensorimotor loop during robotic teleoperation tasks. Iijima *et al.* [23] used toes as a location for creating the sensation of a haptic illusion on the sole. Cicmil *et al.* [12] performed a study comparing perception of toes and fingers, finding high error rates for the trials on the middle toes, with confusion between adjacent toes. Manser *et al.* [13] reported a similar trend from their follow-up study, which compared glabrous and hairy surfaces. A large number of errors involved a specific directional bias. Counting from the little toe as first, the second and third toes were biased toward the little toe, and the fourth toe was biased toward the big toe. In our initial study [14], we designed a haptic rendering strategy that accounted for this directional bias and successfully mitigated the specific confusion around the middle toes, improving the recognition of stimulus on toes by 17%. However, there was still a possibility of further improvements, with approximately 88 % of the errors due to misidentification of the toe directly adjacent to the targeted toe. This motivated the work described in the following sections, intended specifically to reduce confusion between adjacent toes.

III. Haptic number rendering

In this section, we discuss the ActiVibe Final rendering (AVF) described by Cauchard *et al.* [3] and HapToes (TOE), our proposed method of haptic numeric rendering on toes.

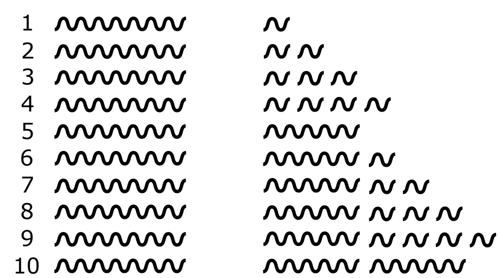

Fig. 1: AVF rendering based on Cauchard *et al.*'s study [3]

A. ActiVibe

ActiVibe is a temporal haptic number rendering in which each number from 1 through 10 is represented by one or more short (150 ms) or long (600 ms) vibrations, separated by pauses, rendered on the wrist by a single vibrotactile actuator, as illustrated in Figure 1. A pre-signal vibration of 750 ms followed by a pause of 1200 ms is used as a cue before rendering the actual values. To represent three consecutive values (AVF3), the pre-signal vibration was rendered only once at the beginning of the sequence, followed by the three values, rendered sequentially and each separated by a pause (800 ms). The total duration of vibration and

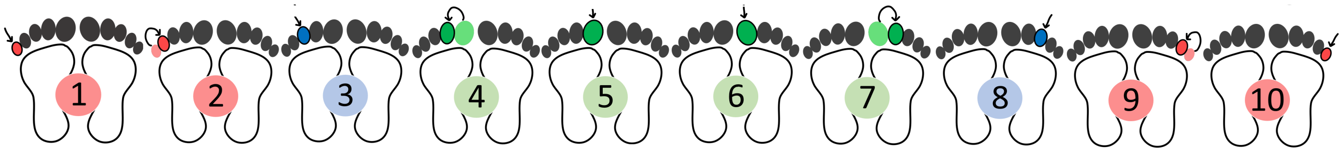

Fig. 2: Final design of HapToes for both feet. For numbers 1, 3, 5, 6, 8 and 10, a long vibration of 1200 ms is rendered on the targeted toe (darker color). For the remaining numbers 2, 4, 7 and 9, a short vibration of 400 ms is rendered on the closest edge toe, followed by a longer vibration of 800 ms on the targeted toes (darker color).

pauses varies with the values being rendered: for single-value renderings, this was in the range of 2050 ms to 3900 ms, while for three-value renderings it was 3950 ms to 9500 ms.

Due to a coding implementation error on the smartwatch when replicating the rendering, there was an additional 200 ms inter-value pause after each 5 or 10 value in the AVF3 condition (e.g., the AVF3 value sequence 1-5-3 would have a normal 800 ms pause between the values 1-5, but a 1000 ms pause between the values 5-3). This extra 200 ms pause was also added to the end of each presentation of AVF or AVF3 encoded values that ended in a 5 or 10, such that the gap between presentations is extended by 200 ms. We believe these discrepancies are immaterial, and if the extra 200 ms gap between values has any effect at all, it is likely to make it easier to distinguish which vibration pulses are associated with each value. This will give AVF3 a slight advantage, albeit with extended total rendering time in these cases.

B. *HapToes*

HapToes is a haptic numeric rendering based on our fundamental perception study [14], designed to map values from 1 to 10 to the toes across both feet—proceeding from left to right, the leftmost toe of the left foot represents 1, and the rightmost toe of the right foot represents 10. This spatial association is reminiscent of a keyboard number layout with the toes as keys, and the zero replaced by ten. Unlike our basic design [14], with all the actuators placed below the toes, here, we alternate the placement of actuators directly below (for toes 1, 3, 5, 6, 8, 10) and directly above the toe, at the nail, to better differentiate adjacent toes.

As shown in Figure 2, for toes corresponding to values 1, 3, 5, 6, 8 and 10, termed as directional toes, a continuous long vibration of 1200 ms is rendered; for toes 2, 4, 7 and 9, a short vibration of 400 ms on the directional toe is immediately followed by a long vibration of 800 ms on the targeted toe. For example, to render 7, a short vibration was rendered on the big toe of the right foot followed immediately by a longer vibration on the second toe of the right foot. We posit that this difference between a continuous long vibration on one toe vs. two vibrations over separate toes provides a further distinction between adjacent toes, facilitating discrimination.

To represent three sequential values (TOE3), each value was separated from the next by a pause of 800 ms. The total duration of vibration was 1200 ms for single-value rendering and 5200 ms for three-value rendering.

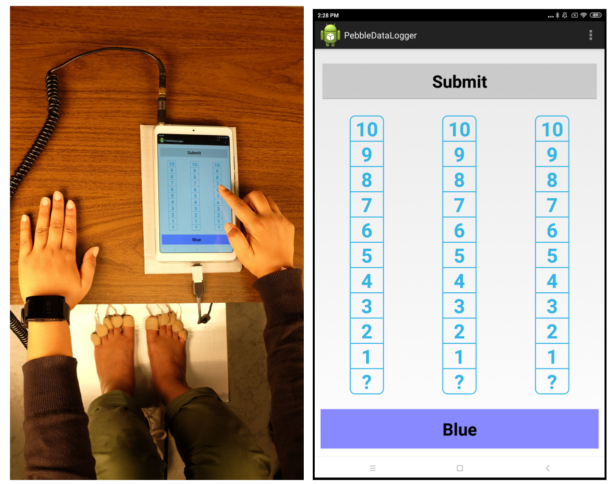

Fig. 3: Experiment setup and tablet UI for entering responses. In the experiment, the feet are obscured by the table, included here for illustration purposes.

IV. User Study

A user study was conducted to investigate the research questions and to assess the performance of the proposed rendering.

A. *Haptic Number Communication Task*

Participants were seated at a table with a smartwatch strapped comfortably tight to their non-dominant wrist and the toe rendering device strapped to their toes. They used a tablet interface to enter their responses to the values they perceived. Participants were instructed not to look at their toes during the experiment. They wore headphones playing pink noise to mask the audible vibration from the rendering devices.

In each trial, participants received either a single value or three values, rendered by vibration, for which they were asked to make a best-guess interpretation to submit using the tablet. The UI of the tablet, shown in Figure 3, consisted of one or three columns for entry of a single value or three values respectively. Columns contained numbers from one to ten as well as a question mark "?" (selected by default) for participants to use when they were completely uncertain of the rendered value. The responses in the UI could be modified until the participant pressed the "Submit" button to finalize their entry. Trials continued automatically upon submission of an answer or after eight seconds with no answer. The UI columns disappeared when the "Submit" button was pressed and reappeared at the end of the next trial's rendering, immediately after the vibrations stopped.

B. Simultaneous Distractor Task

Participants performed an audio task [19], [24] in parallel with the numeric perception task to simulate a real-world distracted use of the rendering device. Throughout the experiment round, they wore headphones through which they heard color names being read along with the pink noise. The Android text-to-speech engine was used to generate the audio. Participants were instructed that their primary task was to report the color blue, every time it was spoken, using the "Blue" button in the UI. The color blue was spoken four times out of 20 in a randomized list along with 16 other color names, one color per second, after which the list was re-randomized. Thus, four blue stimuli occurred every 20 seconds, or 20% of the time.

C. Apparatus

The tactile toe display consisted of ten vibrotactile actuators encased in a rubber piece and attached to the participant's toes with adhesive medical tape (AUPCON Self Adherent Bandage). ERM actuators (2 mm Mini Vibrating Disk Motor, RB-See-403, Seeed Studio) were used to deliver the vibrations, driven by a microcontroller (Teensy 3.2), which was serially connected to an Android tablet (RedMI 4) sending commands driving the trials. For AVF rendering, a smartwatch (Pebble model 301) was used. Participants reported their answers on a UI displayed on the tablet.

D. Methodology

The experiment used a within-subjects design to compare the performance of HapToes (TOE) against ActiVibe (AVF). We recruited 25 participants (11 male, 13 female, 1 gender-neutral, ages 18-31, median = 24) from the McGill University community and compensated them CAD$10 for their participation, which lasted approximately one hour. None of these participants took part in our previous study [14]. Data from one participant (P2) was excluded from the analysis for non-compliance with the instructions.

After participants signed the consent form and completed a pre-test questionnaire, the experimenter explained the Blue audio task and the simultaneous task of identifying numbers rendered by vibrations on the toes or wrist. The tablet was kept on the table, which obscured the participants' feet. The experiment was divided into training and testing phases. The following training was given to the participants:

1) *UI Training:* Participants briefly practiced using the tablet UI, responding to the Blue audio task while entering numerical values on the tablet based on the number of fingers held up by the experimenter near the tablet screen.
2) *TOE/AVF Familiarization:* Participants received a scripted verbal description of the rendering supported by visual aids. Following this, they were exposed to the values from one to ten in the form of vibrations. The value was displayed graphically on the smartwatch or the tablet during the rendering. Participants were asked to pay attention to the rendering.
3) *TOE/AVF Training:* Twenty randomized single-value trials were presented to simulate the real experiment round. Participants were asked to report the perceived value via the tablet UI while also responding to the Blue audio task in parallel. After they submitted their responses, the correct values were displayed on the tablet.
4) *TOE3/AVF3 Familiarization:* The same process as TOE/AVF Familiarization was run for ten randomized three-value trials.

Following training, we carried out separate testing rounds for single-value rendering and three-value rendering, lasting approximately 30 minutes in total. In each round, two sets of trials were run for each of TOE and AVF conditions, presented in reverse counterbalance order (ABBA or BAAB) across participants. The single-value rendering round (AVF, TOE) was followed by the three-value rendering round (AVF3, TOE3). For single-value rendering, within each set, the numbers were shuffled in blocks of ten using Fisher–Yates shuffle and two shuffled blocks were appended together, resulting in 20 trials per set. For three-value rendering, within each set, ten trials were presented, with each possible value included once in each of the three positions in random order. Although offered, none of the participants took a break between the rounds. Participants then completed a post-test questionnaire and were compensated for their time.

E. Measures

We measured missed rate (MR), error rate (ER), absolute Difference between Input and Answered value (DIA), response time (RT), render+response time (RRT) and performance on the distractor task. MR is the percentage of missed trials, i.e., trials where the participant either failed to enter a response within eight seconds or entered "?". ER is the sum of MR and the percentage of trials for which the submitted value differs from the correct rendered value. The DIA is the absolute difference between the rendered and perceived values, which measures the magnitude of the error. Missed values are assigned a DIA of 10. MR, ER and DIA are recorded per value for both single-value and three-value rendering.

Response time (RT) is defined as the time from the end of the haptic rendering to the user's submission of their perceived response. This represents the time required to interpret the rendering as one or more numerical values. The render+response time (RRT) is defined as the time from the start of the haptic rendering to the participant's submission of a response, equivalent to the rendering time plus the response time (RT). This metric is relevant to practical systems since it considers the overall time to convey numeric values to a user, including the time to communicate and interpret. Performance on the distractor task is the mean percentage of acknowledged blue stimuli. Tapping the "Blue" button within 3.5 s of a blue stimulus onset was considered acknowledged, and otherwise missed. We then analyzed the percentage of blue stimuli whose onset occurred during the one second before through one second after the vibrations in each trial. This provides a measure of cognitive effort required for the perception and interpretation of the haptically rendered values.

F. Hypotheses

We expected participants to exhibit more errors in the temporal rendering (AVF) condition than in the spatio-temporal rendering condition (TOE). The former requires participants to remember a sequence count, whereas the latter requires only memory of the stimulated site. We also anticipated that the spatial association of numbers [10] will help the participants to remember single and three values in their memory after it is rendered. Hence, we hypothesized that our spatio-temporal tactile rendering method (TOE) will perform better than single-site temporal rendering (AVF) method delivered on the wrist, both in terms of accuracy (ER, MR and DIA) and response time (RRT and RT). For both conditions, we expected accuracy to drop, and time response and DIA to rise in the three-value trials as opposed to the single-value trials. Participants' performance in the distractor task was also expected to be superior with TOE than AVF for both single and three-value trials.

V. RESULTS

We selected non-parametric statistical tests since the experimental data did not follow a normal distribution. The experiment used a repeated measures design for comparing two conditions, hence, we used the McNemar Test for nominal data and the Wilcoxon signed-rank test for interval data. We implemented these tests via Pingouin 0.2.8 package [25] in Python. The effect size (r) reported is the RBC- rank-biserial correlation.

A. Pre-Questionnaire Results

Twenty participants reported their right hand and foot as dominant, two left hand and foot, one left hand and right foot, and one right and left foot. No participant reported reduced tactile sensation in their hands, arms, feet or toes. Five participants had already experienced vibration from a smartwatch prior to the experiment. Participants' foot width measured at the toes were in the range of 8.5 cm to 10.5 cm (median = 9.5 cm) and shoe sizes were in the range of 22.8 cm to 28.3 cm (median = 26.0 cm).

B. Missed Rate (MR)

As shown in Table I, for single-value trials, mean MR for both TOE and AVF was 0.1%. Participants were able to recognize the correct value for almost all the trials. For three-value trials, MR increased for both the conditions, as expected (Figure 4). As per Wilcoxon signed-rank test results, we did not find a statistically significant difference in MR between AVF3 and TOE3 ($p > 0.05$).

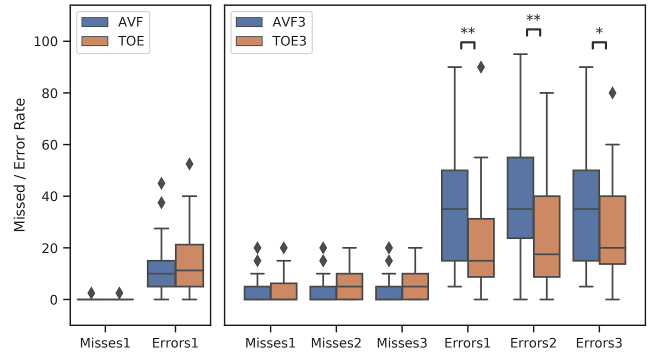

Fig. 4: Boxplot of missed rate and error rate for each condition. Error rate is inclusive of missed rate. Lines in the box centers represent medians.

TABLE I: Mean missed rate and error rate (%) across all conditions

Condition	Missed Rate (%)			Error Rate (%)		
	Val 1	Val 2	Val 3	Val 1	Val 2	Val 3
AVF	0.1			11.9		
TOE	0.1			14.9		
AVF3	3.3	3.8	4.0	35.8	40.4	38.5
TOE3	4.2	5.2	5.2	21.7	25.6	26.2

C. Error Rate (ER)

As also shown in Table I, for single-value trials, mean ER for TOE was 3% higher than AVF. However, we failed to find any statistically significant difference between these conditions as per a Wilcoxon signed-rank test ($p > 0.05$). For three-value trials, ER increased for both the conditions (Figure 4). Table II provides the results of a Wilcoxon signed-rank test, which indicated a statistically significant difference between AVF3 and TOE3.

TABLE II: Wilcoxon signed-rank test for ER in three-value trials

AVF3 vs TOE3	W-val	z-val	p-val	r-val
Val 1	39.5	-2.59	<0.005	0.68
Val 2	48.5	-2.66	<0.005	0.68
Val 3	54	-2.08	<0.025	0.57

D. Difference between Input and Answered value (DIA)

For both single-value and three-value answered trials, DIA for AVF was mostly concentrated in the range of one to five but for TOE, the maximum proportion of DIA was between one and two (Figure 5). A Wilcoxon signed-rank test indicated a statistically significant difference between conditions as shown in Table III.

TABLE III: Wilcoxon signed-rank test values for DIA

Condition	Values	W-val	z-val	p-val	r-val
AVF vs TOE	Val 1	10557	-1.83	<0.05	0.16
AVF3 vs TOE3	Val 1	7617.5	-4.53	<0.001	0.36
AVF3 vs TOE3	Val 2	10615.5	-3.59	<0.001	0.28
AVF3 vs TOE3	Val 3	10275	-3.07	<0.005	0.25

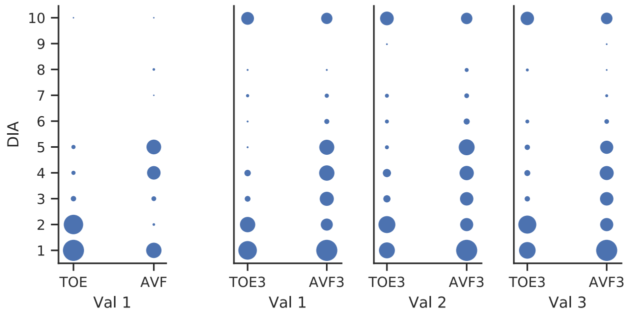

Fig. 5: Distribution of trials in which DIA>0, with misses assigned DIA=10. Dot sizes are proportional to the percentage of trials at each DIA level.

E. Response Time and Render+Response Time

As seen in Figure 6, the median value of render+response time (RRT) was higher for AVF than TOE for both single-value and three-value trials. Similarly, the median response time (RT), i.e., time to interpret the value once it has been rendered, was higher for AVF than TOE both for single- and three-value trials. However, the effect size is insignificant for providing meaningful insight in practical applications.

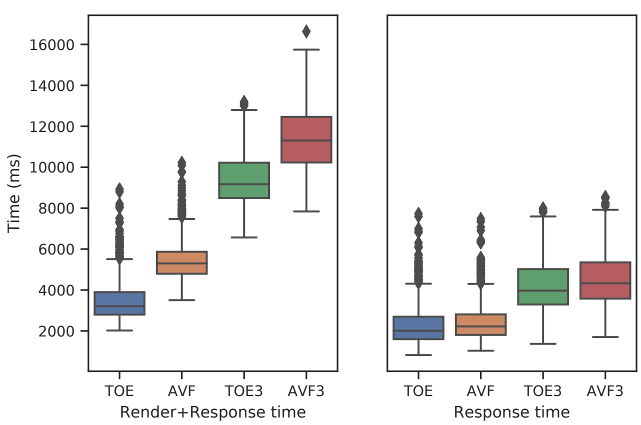

Fig. 6: Box plots of render+response times (RRT) and response times (RT). Lines in the box centers represent medians.

F. Distractor Task Performance

For single-value trials, the mean percentage of acknowledged blue stimuli was $67\,\%$ for AVF and $69.5\,\%$ for TOE. For three-value trials, this was $43\,\%$ for AVF and $52\,\%$ for TOE. This difference in performance was significant only for three-value trials as determined by a Wilcoxon signed-rank test ($W = 56$, $z = -2.43$, $p < 0.01$, $r = 0.63$). As we hypothesized, this indicates that AVF requires more cognitive effort than TOE for three-number rendering.

G. Post-test Questionnaire Results

In the post-test questionnaire, we asked participants to select between AVF and TOE for their overall preference, accuracy and effort required. Since these data were nominal, we used the McNemar test to calculate statistical parameters. Out of 24 participants, 15 selected TOE over AVF in terms of overall preference (no significant difference), 19 participants selected TOE over AVF in terms of accuracy ($\chi^2 = 7$, $p < 0.01$), and 19 participants reported AVF to require more effort than TOE ($\chi^2 = 7$, $p < 0.01$). These subjective responses are consistent with our aforementioned objective measurements obtained during the study.

VI. Discussion

Our proposed method, HapToes (TOE), performed on par with ActiVibe (AVF) for single-value rendering and exhibited a lower error rate than AVF for three-value rendering. This could be attributed to the spatial association of numbers in cognition, for example using a mental "number line". Although haptic signals can leverage duration to represent numbers, it has been reported that temporal tasks are easily disrupted by secondary tasks [10].

Another significant advantage of TOE/TOE3 over AVF/AVF3 is its smaller average DIA. Although TOE often exhibits DIAs of one, i.e., the perceived value is one unit off the rendered value, AVF suffers from DIAs in the range of one to five. By the nature of the AVF rendering, if a user confuses the long vibration representing five with a pre-signal vibration, and thus ignores it, or confuses the long vibration with a short one, this can result in a DIA of five or four, respectively. DIAs of two and three may also result from counting the short vibrations incorrectly.

Following termination of the haptic rendering, the time required for participants to interpret the rendered values differs by a negligible amount of $200\,\mathrm{ms}$ on average, with TOE obtaining faster responses than AVF. Additionally, the reduced rendering time of TOE results in a shorter overall time required to convey a value and have it interpreted, i.e., time from the start of the haptic rendering to the participant's confirmation on the tablet of the perceived value. This is, naturally, desirable for improved efficiency of information communication.

Furthermore, the results related to distractor task performance suggest that TOE rendering imposes lower cognitive demands than AVF. Indeed, the spatial mapping in TOE does not require the same sustained concentration effort to keep track of a series of vibrations, as does AVF. As one participant described, they "associated vibration to toes quite early so there was a visualisation component". Once one identifies which toes were targeted and in what order, this information can be retained in short-term memory while attending to another (distractor) task. As soon as the distractor task is no longer occupying attention, the memorized toe sequence can then be interpreted as the intented values. In contrast, this strategy does not apply with a temporal mapping, requiring counting, as does AVF.

We note that the comparison of TOE to AVF is not intended to demonstrate a general superiority of one method over another. These are inherently very different techniques, with the former requiring instrumentation over multiple toes, and employing spatio-temporal tactons, while the latter uses a single consumer device worn on the wrist, and renders temporal tactons. These two options are therefore suited for different scenarios, one in which simplicity of hardware is preferable, and the other for which accurate delivery of multiple numeric values is required, or where the wrist is otherwise unavailable as a locus for information delivery.

Although, the performance of TOE is favorable when used for seated applications, its suitability to tasks involving a user in motion remains to be investigated. The current hardware only serves as a proof of concept prototype for the encoded rendering. An improved design would be required to ascertain its use as a wearable, and overcome challenges of inconsistent haptic coupling and interference from motion, which create haptic noise. While we expect that accuracy would decrease when the user is standing, it would be important to characterize the performance drop while walking or running, compared to the effects on wrist-based rendering. Moreover, improvements in foot-based haptic interfaces could help enhance the performance, leading to a more acceptable and practical device suitable for everyday use.

VII. Conclusion

We presented HapToes, a novel numeric haptic rendering method for toes. HapToes outperformed ActiVibe, a previously reported method for communicating numeric values, establishing its possibility of serving as a ten-unit tactile display. Although it might not be possible to convey phone numbers haptically to distracted users using this method, the results of our study suggest that reasonably accurate identification of large numbers under modest cognitive load is feasible. Increased accuracy could be achieved with further improvements to the rendering device. For example, more capable vibrotactile actuators such as voice-coil actuators would enable delivery of a different frequency of vibration to each toe, to enhance the discriminability of adjacent toes, which remained the source of most erroneous trials. It would also be interesting to investigate how the rendering would perform using an entirely different haptic stimulus such as electrotactile actuation.

Acknowledgment

We thank members of the Shared Reality Lab for their suggestions and help in the development of the prototype. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), RGPIN-2017-05013, and the MITACS Globalink Graduate Fellowship Award. McGill REB #83-0814.

References

[1] S. Brewster and L. M. Brown, "Tactons: structured tactile messages for non-visual information display," in *Proc. 5th conf. Australasian user interface-Vol. 28*. Australian Computer Society, Inc., 2004, pp. 15–23.

[2] M. Schirmer, J. Hartmann, S. Bertel, and F. Echtler, "Shoe me the way: a shoe-based tactile interface for eyes-free urban navigation," in *Proc. 17th Int. Conf. Human-Computer Interaction with Mobile Devices and Services*, 2015, pp. 327–336.

[3] J. R. Cauchard, J. L. Cheng, T. Pietrzak, and J. A. Landay, "Activibe: Design and evaluation of vibrations for progress monitoring," in *Proc. Conf. Human Factors Comput. Syst.* New York, NY, USA: ACM, 2016, pp. 3261–3271.

[4] T. K. Ferris and N. Sarter, "Continuously informing vibrotactile displays in support of attention management and multitasking in anesthesiology," *Human Factors*, vol. 53, no. 6, pp. 600–611, 2011.

[5] E. Pescara, V. Diener, and M. Beigl, "Vibraid: comparing temporal and spatial tactile cues in control room environments," in *Proc. 12th ACM Int. Conf. Pervasive Technol. Assistive Env.* ACM, 2019, pp. 138–145.

[6] Y.-C. Liao, Y.-C. Chen, L. Chan, and B.-Y. Chen, "Dwell+: Multi-level mode selection using vibrotactile cues," in *Proc. 30th Annu. Symp. User Interface Softw. Technol.* ACM, 2017, pp. 5–16.

[7] S. D. Novich and D. M. Eagleman, "Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput," *Exp. Brain Res.*, vol. 233, no. 10, pp. 2777–2788, 2015.

[8] S. C. Lee and T. Starner, "Buzzwear: alert perception in wearable tactile displays on the wrist," in *Proc. Conf. Human Factors Comput. Sys.* ACM, 2010, pp. 433–442.

[9] Y.-C. Liao, Y.-L. Chen, J.-Y. Lo, R.-H. Liang, L. Chan, and B.-Y. Chen, "Edgevib: effective alphanumeric character output using a wrist-worn tactile display," in *Proc 29th Annu. Symp. User Interface Softw. and Technol.* ACM, 2016, pp. 595–601.

[10] V. Walsh, "A theory of magnitude: common cortical metrics of time, space and quantity," *Trends in cognitive sciences*, vol. 7, no. 11, pp. 483–488, 2003.

[11] H. Dempsey-Jones, D. B. Wesselink, J. Friedman, and T. R. Makin, "Organized toe maps in extreme foot users," *Cell reports*, vol. 28, no. 11, pp. 2748–2756, 2019.

[12] N. Cicmil, A. P. Meyer, and J. F. Stein, "Tactile toe agnosia and percept of a "missing toe" in healthy humans," *Perception*, vol. 45, no. 3, pp. 265–280, 2016, pMID: 26562866.

[13] K. Manser-Smith, L. Tamè, and M. Longo, "Tactile confusions of the fingers and toes," *J. Experimental Psychology: Human Perception and Performance*, 2018.

[14] P. Vyas, F. Al Taha, J. Blum, A. Weill-Duflos, and J. Cooperstock, "Ten little fingers, ten little toes: Can toes match fingers for haptic discrimination?" *IEEE Transactions on Haptics*, 2020.

[15] J. Anlauff, J. Fung, and J. R. Cooperstock, "Vibewalk: foot-based tactons during walking and quiet stance," in *World Haptics Conf.* IEEE, 2017, pp. 647–652.

[16] M. Tokita and A. Ishiguchi, "Precision and bias in approximate numerical judgment in auditory, tactile, and cross-modal presentation," *Perception*, vol. 45, no. 1-2, pp. 56–70, 2016.

[17] S. Töyssy, J. Raisamo, and R. Raisamo, "Telling time by vibration," in *Int. Conf. Human Haptic Sensing Touch Enabled Computer Appl.* Springer, 2008, pp. 924–929.

[18] H. Nicolau, J. Guerreiro, T. Guerreiro, and L. Carriço, "Ubibraille: designing and evaluating a vibrotactile braille-reading device," in *Proc. 15th Int. SIGACCESS Conf. on Computers and Accessibility*. ACM, 2013, p. 23.

[19] J. R. Blum and J. R. Cooperstock, "Single-actuator vibrotactile numeric information delivery in the face of distraction," in *World Haptics Conf.* IEEE, 2019, pp. 461–466.

[20] D. S. Elvitigala, D. J. Matthies, L. David, C. Weerasinghe, and S. Nanayakkara, "Gymsoles: Improving squats and dead-lifts by visualizing the user's center of pressure," in *Proc. Conf. Human Factors Comput. Syst.* ACM, 2019, p. 174.

[21] M. Narazani, "Designing foot-based interaction for wearable skill transfer systems," in *Proc. ACM Int. Joint Conf. and Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput.* ACM, 2018, pp. 1787–1788.

[22] A. Panarese, B. B. Edin, F. Vecchi, M. C. Carrozza, and R. S. Johansson, "Humans can integrate force feedback to toes in their sensorimotor control of a robotic hand," *Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 6, pp. 560–567, 2009.

[23] Y. Iijima, M. Uchida, T. Hachisu, and Y. Hashimoto, "Enhancement of range of creation of foot sole tactile illusion by vibration stimulation of the foot instep," in *World Haptics Conf.* IEEE, 2019, pp. 31–36.

[24] A. Chan, K. MacLean, and J. McGrenere, "Learning and identifying haptic icons under workload," in *1st Joint Eurohaptics Conf. Symp. Haptic Interfaces Virtual Env. Teleoperator Syst. World Haptics Conf.* IEEE, 2005, pp. 432–439.

[25] R. Vallat, "Pingouin: statistics in python." *J. Open Source Software*, vol. 3, no. 31, p. 1026, 2018.