Investigating Interactions for Text Recognition using a Vibrotactile Wearable Display

Conference Paper · January 2018

DOI: 10.1145/3172944.3172951

Granit Luzhnica
Graz University of Technology

Eduardo E. Veas
Graz University of Technology

Investigating Interactions for Text Recognition using a Vibrotactile Wearable Display

Granit Luzhnica
Know-Center
Graz, Austria
gluzhnica@know-center.at

Eduardo Veas
Know-Center
Graz, Austria
eveas@know-center.at

ABSTRACT

Vibrotactile skin-reading uses wearable vibrotactile displays to convey dynamically generated textual information. Such wearable displays have potential to be used in a broad range of applications. Nevertheless, the reading process is passive, and users have no control over the reading flow. To compensate for such drawback, this paper investigates what kind of interactions are necessary for vibrotactile skin reading and the modalities of such interactions. An interaction concept for skin reading was designed by taking into account the reading as a process. We performed a formative study with 22 participants to assess reading behaviour in word and sentence reading using a six-channel wearable vibrotactile display. Our study shows that word based interactions in sentence reading are more often used and preferred by users compared to character-based interactions and that users prefer gesture-based interaction for skin reading. Finally, we discuss how such wearable vibrotactile displays could be extended with sensors that would enable recognition of such gesture-based interaction. This paper contributes a set of guidelines for the design of wearable haptic displays for text communication.

Author Keywords

vibrotactile feedback; skin reading; stimulation; haptic display; wearable; user study; interaction, interaction design; gesture interaction; gesture recognition; HCI

INTRODUCTION

Reading is a fundamental skill in human development, as a way to acquire knowledge from written text, and also a major component in communication (e.g., letters, telegram, chat, email). To most people, reading is a visual task that recruits the most developed sense of vision to decode messages encoded in graphical symbols. People with visual impairments resort to the tactile sense for reading by using Braille. Besides them, non-impaired individuals can benefit from a means to perceive messages that do not recruit the visual or auditory senses. The primary feedback modalities of mobiles and wearables are visual and auditory. As such, they compete for visual and auditory attention and distract the user from important tasks.

IUI'18, March 7–11, 2018, Tokyo, Japan

ISBN 978-1-4503-4945-1/18/03...$15.00

DOI: https://doi.org/10.1145/3172944.3172951

Alternative display modalities, such as tactile displays, can reduce demands on the predominant visual display, but are largely under-utilised [4].

With the proliferation of wearable devices, devices with vibrotactile capabilities are accessible to a substantial number of end users. With them, it becomes possible to transmit generic symbols with abstract meanings [14, 29, 24] that can be combined in arbitrary messages (words, sentences) [14, 29]. Luzhnica et al. refer to this as vibrotactile *skin-reading*. While encoding alphanumeric symbols in tactile patterns to transmit words is not new [14], recent developments in vibrotactile technology and stimulation methods overcome issues in effective training, throughput, accuracy as well as portability [28]. Vibrotactile skin reading uses a vibrotactile display to convey dynamically generated textual information similar to refreshable Braille displays [43, 41]. However, it is much more portable, and mobile as the entire vibrotactile display is packed within a wearable glove or sleeve [29]. Moreover, it requires substantially less effort to learn. Recent research has demonstrated that sighted individuals are able to learn the entire English Alphabet (26 letters) within three hours of training [29] as opposed to six months of training required for learning Polish Alphabet (32 letters) using Braille [3]. While being able to efficiently transmit an alphabet is a necessary step, reading is inherently a much more complex task [38]. Learning to read is learning how to use the conventional forms of printed language to obtain meaning. For skin-reading, the print is transformed into vibrotactile patterns, so that reading entails obtaining meaning from them.

A common belief that reading is a sequential task, where *glide smoothly across the page*, is merely an illusion [37]. At the word level, well-established research postulated that words are recognized as units [22, 11, 40, 5] and they are even recognized before individual letters [5]. Reading depends on the mechanics of the visual system to stop at fixed spots in the text (fixations) and jump quickly to other spots (saccades, covering about 8 letter spaces) [37]. Skilled readers fixate on about $2/3$ of the words in a text. Beside forward movements to advance in reading, they reread nearby material backwards in the text about 10 to 15% of the time, occasionally driven by breakdowns on comprehension. Conversely, beginning readers fixate every word (often more than once), perform shorter saccades, and up to 50% of their eye movements are regressions, as they rely more on context to identify words [37]. Obtaining meaning from printed words is not sequential; it depends on processing words as units and uses backward jumps at word level to aid understanding.

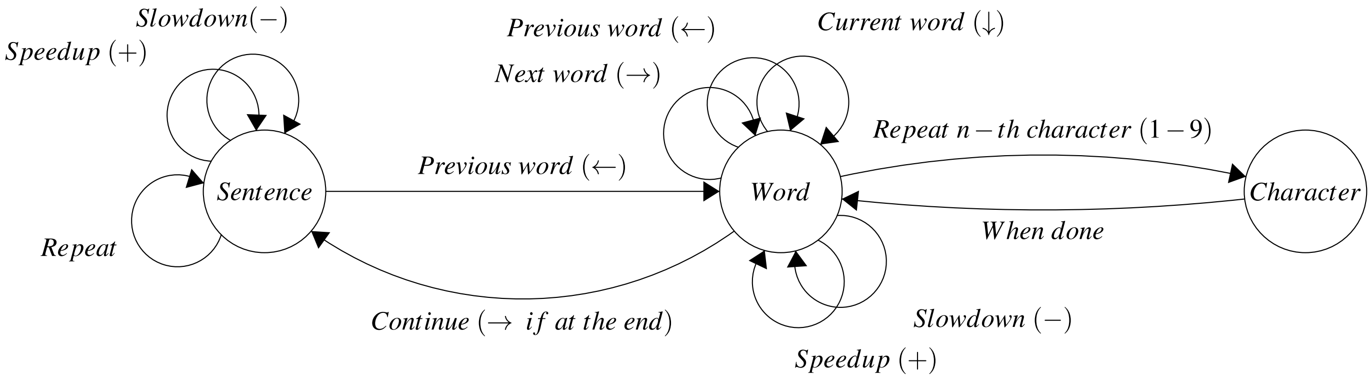

Figure 1. Interaction concept during sentence transmission. States (Sentence, Word and Characters) represent what the system is transmitting to user.

In Braille, it is not possible to form a global shape recognition of the entire word, so the text has to be processed character by character [7, 31, 30]. The perception and flow of information in Braille are controlled by moving the hand forward and occasionally backwards to revisit information [30, 20]. Thereby, Braille readers control reading speed, focus on particular letters or re-scan entire words. On the other hand, vibrotactile skin reading is passive: a pattern of vibrations is stimulated by the device from start to end while users have no control over the transmission. Yet, vibrotactile displays can evoke the perception of words as units, by means of tactile animations [21, 48]. The question that drives this research is what interactions are needed for efficient skin-reading? Users may not understand parts of the text due to lack of concentration or training. They need ways to pause, resume and jump to previous units of meaning or change the speed of transmission to account for progress in their reading skills.

In this paper, we propose using navigation interactions for skin reading to equip users with means of controlling the reading process. An interaction concept was created to enable reading interactions for vibrotactile displays. A formative study trained novice users to recognise letters and words, and tested their behaviour while skin-reading sentences. We analyse participants' interaction behaviours and a questionnaire to determine what interactions are useful and what are the preferred means of interactions. Finally, we map our interactions to gestures and discuss the wearable design choices that could allow the used wearable vibrotactile display to be extended for supporting such gesture-based interaction concept.

RELATED WORK

Starting with Braille's invention of the Braille coding in 1824, haptic displays have long been widely used by people with visual impairments. Research on tactile displays equipped with actuators has been ongoing since at least 1924 [13], where Gault [13] used a piezoelectric unit to convert entire recorded speech to touch. Similarly, Kirman [21] used a 15×15 vibrator matrix on the palm to teach six participants to differentiate between the patterns of 15 different words. Other researchers attempted to utilise a visually oriented approach, where a low-resolution image of the object is projected to an array of stimulators. For instance, White [52] transformed images captured from a video feed to a 20×20 vibrotactile display placed on the back. After training, participants were able to distinguish simple shapes like circle, square and triangle. Bliss [2] developed the first commercial device capable of capturing text from the video feed and then imprinting each letter on the finger with a 6×24 matrix of vibrators.

A more successful approach of transmitting information through haptics was provided by Geldard [14]. The device was named Vibratese and used five vibromotors on the chest to encode 45 symbols (letters, numbers and most frequent short words). The author reported that after 65 hours of training one participant was able to understand 38 wpm (words per minute). More recently, Luzhnica et al. followed a different encoding scheme using only the location of vibromotors to encode 26 letters of English alphabet [29]. The authors used six vibromotors on the back of the hand and were able to train users to perceive letters, words and phrases through skin within only five hours, although they needed repetition of stimuli.

A crucial aspect of tactile displays is how they encode information. It requires providing patterns that are discriminative while delivering them as fast as possible. Typically a combination of variations in amplitude [47, 49, 54], frequency [47, 49, 54], duration [16, 14] and body locations [14, 54, 34, 44] have been used. For instance, Geldard [14] in his Vibratese used five locations, a variation of three durations and three intensities to encode the desired symbols. Recently, Novich [35] showed that spatiotemporal encoding, where vibromotors in a pattern are turned on and off sequentially one after the other, results in significantly better discrimination than the spatially encoded patterns where all vibromotors in a pattern onset simultaneously. Liao [24] utilised such a spatiotemporal encoding to encode symbols on the wrist. Although such encoding works well [24, 35] in terms of being identified by participants, it is many times slower than the spatial encoding. Luzhnica [29] used a prioritised overlapping spatiotemporal encoding where vibromotors are activated in sequence after a gap, and they stay on until the pattern is finished. This method resulted in better recognition accuracy than spatial encoding, and it is faster than spatiotemporal encoding, as vibromotors share most of the activated time. Recent research leverages spatial acuity (sensitivity of locations of actuators) for achieving a better perception of encoded information [28, 27]. We employ the same method as Luzhnica [29] for encoding char-

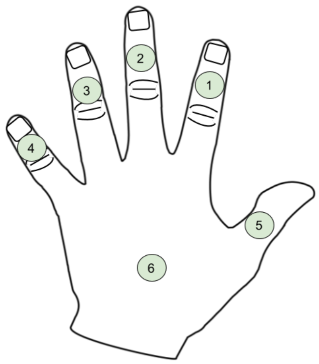

Figure 2. The vibrotactile display investigated in this paper containing six vibromotors. The design has been borrowed from [29]. The number of the vibromotor indicates the priority of activation.

acters, words and sentences namely for skin reading. Our contribution and novel aspect of this paper lie on proposing to equip users with interactions for skin reading.

INTERACTION CONCEPT

We design an interaction concept for textual vibrotactile skin reading, illustrated in Figure 1. The concept is based on a virtual fixation point metaphor. The fixation point represents the word that is currently being transmitted to the user. We will refer to it as the current word. While perceiving text/sentences, at whatever point in time, users have the possibility to request re-transmission of the current word. In this case, the system transmits the current word and transfers itself into the pausing state (corresponds to Word state in Figure 1) where no further text is transmitted until resumed. Let us assume that the user paused on the n-th word of a text. While on the pausing state, the user can repeat the current (n) word or navigate to the previous ($n-1$) word in the text. In this case, the fixation point shifts to the left in the text and the ($n-1$) word is transmitted and becomes the current word. At this point, the user can repeat the current word ($n-1$), regress to the previous one ($n-2$), or go to the next one (n). Hereby, the user navigates back an forth and scans the text. If the fixation point is at position n and the user navigates to the next word (beyond the point where it was paused), the system resumes and starts transmitting the rest of the words. Additionally, when the system is in the pause mode, the user can repeat particular characters of the fixated word. Furthermore, users can also change the speed of transmission which would proportionally change gaps between characters and words; and the activation time of each vibration motor (see Figure 4).

USER STUDY

To investigate our interaction concept and determine which of the interactions are useful for the user, we conducted a user study that combines participant training and testing of characters, words and sentences. An additional goal of the study was to investigate the word recognition process. However, the topic is out of the scope of this paper, and thus we will not discuss the results and findings concerning this investigation.

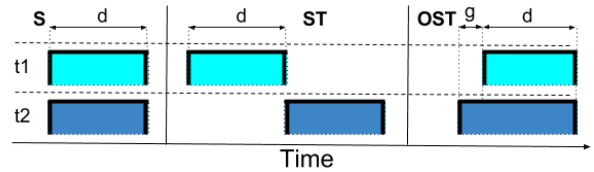

Figure 3. Pattern types composed of two vibromotors/locations: spatiotemporal (ST), overlapping spatiotemporal (OST), spatial (S). Base duration (d) represents the activation time of a vibromotor (t1 and t2). The gap between the activation of vibromotors is denoted by g.

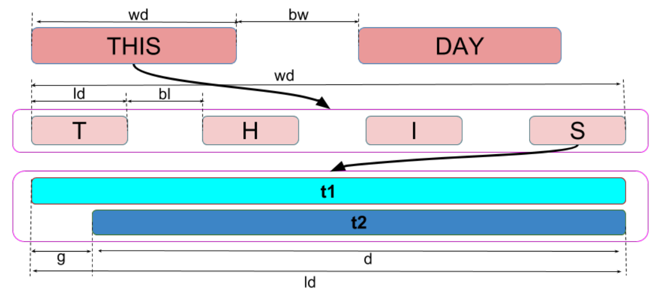

Figure 4. Stimulation process of characters, words and sentences. Base duration (d) represents the activation time of a vibromotor. Words in a sentence are transmitted in series separated by a gap (bw = 600 ms). Within words, characters are transmitted in series with a gap in between (bl = 200 ms). The characters are encoded using OST patterns where vibromotors are activated in sequence with a gap in between (g = 10ms).

Wearable Haptic Display Design

We use the glove-based design of Luzhnica et al [29] for the vibrotactile display. It contains six vibromotors placed on the back of the hand (see Figure 2). The vibromotors can be fitted in a fingerless glove, leaving the fingers free for interaction.

Vibrotactile Patterns and Encoding

Each character is encoded with one or two vibromotors using an OST (overlapped spatiotemporal) stimulation as introduced by Luzhnica and Veas. [28]. Figure 3 illustrates the details of OST and the differences with other stimulation forms. The activation of vibromotors is done in sequence, but they share most of the activation time. Moreover, the activation is prioritised based on the sensitivity of the finger as such prioritisation yields a higher accuracy of identification of locus [28]. Sensitivity order is assumed as reported on studies on the subject [9, 51, 19], suggesting that sensitivity decreases from the index finger towards the little finger: the index finger is more sensitive than the middle, ring, and pinky finger. The thumb has the lowest sensitivity [46]. For example, for a character encoded in index and pinky finger, the vibromotor on index finger would be activated first, and then after a gap, the vibromotor placed on the pinky finger. Figure 5 illustrates which vibromotors are used to encode each of the characters used in the study.

Figure 4 illustrates the technical details of the stimulation process of characters, words and sentences. Character encoding uses a base duration (d) of 200 ms and a 10 ms gap (g) between the activation of vibromotors. This means that the duration of a character (ld) is 200 ms for one vibromotor and 210 ms for two-vibromotor letters. When constructing words, a between letter gap (bl) of 200 ms is used to separate sequential letters. With such encoding, a word containing

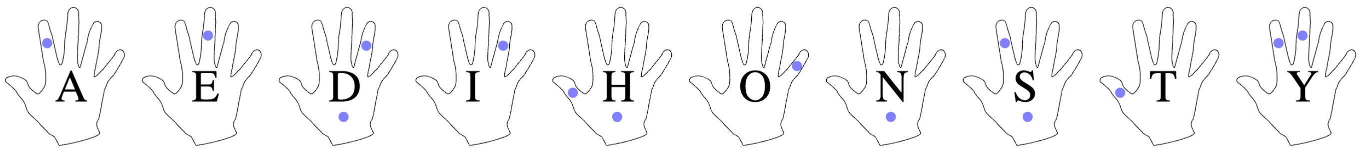

Figure 5. The encoding scheme of each character used during the study.

2-char	3-char	4-char	5-char
is	tea	easy	shiny
he	say	does	stand
it	hot	this	notes

Table 1. The list of words that are used during the experiment

four characters can be transmitted within 1400-1440 ms. Note that, users can be trained to recognise letters and words with shorter duration when exposed to longer training periods [29]. However, we aimed at having training and testing in a single session. Hence, we decided for longer durations. Additionally, sentence encoding uses a between word gap of 600 ms.

Characters

Similar patterns have been used to train users to recognise the entire English alphabet [29], and a wearable glove based layout capable of encoding 36 characters with only one or two vibromotors have been proposed [28], which demonstrate the capabilities of such wearable device for text reading. Nevertheless, in this study, we use only ten characters: A, E, I, O, T, N, S, H, D and Y. A small alphabet ensures a shorter period of training for characters and still enables us to create words and sentences for investigating interactions.

Words

With the chosen characters, we composed a list of 12 words, containing two to five characters. The words have been selected from the list of basic English words[1] and they will be used to train the user with words.

Sentences

We created a list of 29 sentences (see Table 2). But only 15 of them were stimulated during testing, whereas the rest are there just to create more choices. The goal of the sentences testing was to observe how participants interact with the system while reading a sentence and how well they perform.

He is the one, **She is the one**, **The tea is hot**, The sea is hot, **It is hot**, **He is hot**, She is hot, **Today is hot**, He says no, **She says no**, I say yes, **I say no**, **He says yes**, She says yes, I say yes, **This is easy**, It was easy, It is easy, **I hate this**, He hates this, **Hide this idea**, **It does stand**, This does stand, **It is done**, This is done, **This is shiny**, It is shiny, **It is noisy**, This is noisy.

Table 2. Sentences used during the experiment. All sentences (29) were presented to the user in a list to select from. Only the bold marked ones (15) are used to test the user. The rest of the sentences (14) are used as decoys to make the process more challenging.

Procedure

The entire study was organised in several blocks which we will refer to as rounds, each serving different purposes:

[1] `https://simple.wikipedia.org/wiki/Wikipedia:List_of_1000_basic_words`

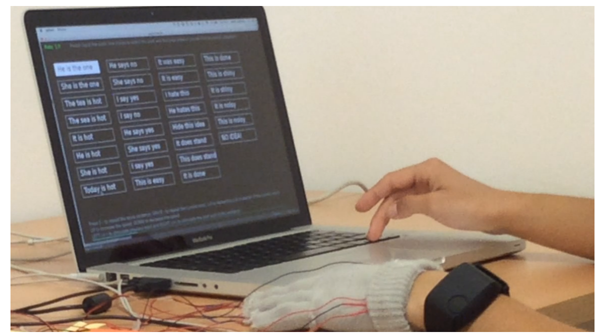

Figure 6. A participant performing a round of sentence reinforcement

Character Training trained users to associate a symbol with a vibrotactile pattern. During this process, participants were stimulated with patterns representing a character, the character was displayed on the screen, and an audio spelling of the character is simultaneously played as shown in Figure 7. Such a simultaneous technique of tactile, auditory and visual stimulation has been demonstrated to be efficient [29].

Character Reinforcement. Participants were stimulated with a pattern and asked to input the character associated with it. After entering the answer, they were notified whether their input was correct and saw the correct answer (see Figure 7). This way they would learn from their mistakes. Participants were allowed to repeat the stimuli before answering.

Word Training exposed users to simultaneous stimulation of vibrotactile, auditory and visual stimuli all representing a word. The process was similar to character training, but instead of characters words were used. Words were transmitted as a series of characters (see Figure 4).

Word Reinforcement. This is similar to character reinforcement, but words are used instead. Participants are allowed repeat the stimuli. After stimulation, they were asked to select the answer from a list constructed with all the words shown in Table 1, plus 21 other words words including "No idea!". Participants were instructed to choose "No idea!" if they do not know what is stimulated. We also pointed out that for every stimulated word there are other similar words in the list and thus they should avoid guessing based on few characters. Upon entering the answer, participants were informed on the display what would have been the correct answer.

Words Testing: was similar to word reinforcement, but participants were not allowed to repeat the stimuli, and they were not notified the correct answer.

Sentence Reinforcement. The process was similar to the word reinforcement but used sentences instead of words. Participants were stimulated with a sentence and asked to select an answer from a list of sentences. As shown in Table 2, the list of choices contained 29 sentences plus the "No idea" op-

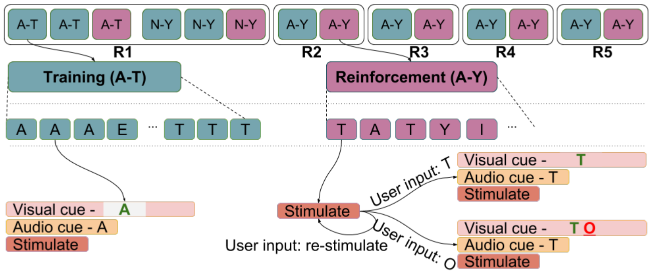

Figure 7. Character training and reinforcement process. Initially only five characters were introduced (A-T) then the next five letters (N-Y). For the the rest of letter training, all ten characters were used (A-Y). Colour coding: ◯ - train round, ◯ - reinforcement round.

tion. However, users were tested only in 15 sentences and the rest of them were used to make the process more challenging. Each sentence was composed of three to four words (see Table 2). The main purpose of sentence reinforcement was to study interactions during skin reading.

Participants went through five rounds of character training and reinforcement as shown in Figure 7. The first round was split into two short rounds, where only half of the characters are used in each (A-T and N-Y). This way participants were introduced first to five characters and then to the next five. The following four rounds used all ten characters. In every training round, a character was displayed three times whereas in the reinforcement rounds each character was tested twice.

Thereafter, participants went through a round of word testing followed by a round of reinforcement using words from Table 1. Then, users went through four rounds of word training each followed by a round of word reinforcement. Finally, users were subject to one round of word testing followed by one round of word reinforcement where they were exposed to a combination of words they trained on (from Table 1) and an equal number (balanced by word length) of words they did not trained on. To finish the study, participants were subject of one round of sentence reinforcement where each of 15 sentences (see Table 2) was tested once.

Interaction

During the sentence reinforcement rounds, participants could use the interaction concept presented in Figure 1. They, could repeat the current word by using the keyboard space key, navigate words using left and right arrow keys, change transmission speed by using up and down arrow keys. They could choose to re-stimulate only a particular letter by pressing a number from 1-9, which would re-stimulate the n-th letter of the current word. Additionally, participants were able to repeat the entire sentence by pressing S. During word reinforcement rounds, participants were also able to repeat the entire word or particular characters of it. Similarly, during the character reinforcement, participants could repeat the character.

Data Collection

First, all user responses and interactions during testing and reinforcement rounds were logged. Additionally, at the end of the session, users filled a questionnaire asking questions about how they would use such a wearable device on a daily basis. Initially, they were asked to rate whether they find the interactions for skin reading as: (i) Necessary, (ii) Optional or (iii) Not useful. Also, for each available interaction, users were asked to rate how often they think they would use it by selecting one of the available choices:

1. Continuously - every couple of seconds or minutes
2. Often - every couple of hours
3. Not very often - every couple of days or weeks
4. Rarely - few times only, and
5. Never.

Furthermore, we proposed three possible modes of interaction:

1. Gesture-based: the user uses hand gestures to interact,
2. Smartphone-based: an application in the user's smartphone is used to interact, and
3. Physical buttons based: physical buttons would be added to the vibrotactile glove for interaction.

Participants were asked to rate (0-10) how suitable each of the proposed modality would be for interactions with the wearable display. Furthermore, they were asked to choose one modality they would use for interactions/commands that they would use more often and one for the interactions they would use rarely.

Apparatus

Our device consisted of an Arduino Due board which controls 3.4mm vibrotactile motors of type ROB-08449 (Voltage range: $2.3V \sim 3.6V$; Amplitude vibration: $0.8G$).

Participants

Twenty-two (22) individuals (12 male and 10 female) aged between 17 and 38 (M=26.7, STD=5.5) years old participated in this study. The overall study took approximately 90 minutes. Only one of them was left-handed. All of them used the left hand for stimulation and the right to interact with the computer as depicted in Figure 6. One participant, for personal time constraint reasons, completed the character and word rounds but did not continue with sentence reinforcement round.

Results

Let us define four common variables: accuracy, repetition, total duration. Accuracy will be defined as a binary variable set to be one if the user's answer is correct. Repetition describes how many times a user repeated the stimulation (character, word or sentence) within a reinforcement round. The total duration represents the entire duration from the time stimulation was first initiated by the system until the user responded, including repetitions. Additionally, we define an interaction to be a repetition of any kind. E.g. during sentence reinforcement round, each repetition such as: current word, previous word, next word, a particular letter of the current word or the entire sentence is considered to be an interaction. Although, we will provide a brief overview of performance to give an impression of users training level prior to sentence reading, we will mostly focus on interactions during sentence reinforcement round as the reading performance is out of the scope of this paper.

Performance

Each character reinforcement round collected 20 probes (2 $\times$ 10 characters) for each user. Table 3 presents the results

Round	Accuracy	Total Duration (s)	Repetitions
1a	0.98 (0.15)	3.29 (4.03)	0.22 (0.60)
1b	0.82 (0.38)	5.06 (8.41)	0.57 (1.28)
2	0.86 (0.25)	5.57 (9.58)	0.72 (1.52)
3	0.95 (0.23)	5.58 (23.8)	0.72 (1.81)
4	0.95 (0.22)	4.39 (6.56)	0.63 (1.37)
5	0.95 (0.22)	3.33 (5.58)	0.39 (0.90)

Table 3. Results of character reinforcement rounds. Note that we consider the first two rounds (1a and 1b) as two parts of round one as each of them contained only half the characters. This table shows the correct recognition rate (accuracy), average total duration, average duration and average repetition rate.

Round Type	Accuracy	Total Duration (s)	Repetitions
With Repetition	0.81 (0.39)	12.41 (12.27)	2.32 (4.28)
No Repetition	0.55 (0.50)	9.34 (8.45)	0.0 (0.0)

Table 4. Results of word recognition in the last reinforcement (with repetitions) and testing (no repetitions) rounds.

of character recognition, including the average accuracy, repetition, duration and total duration. By the third round, participants could already recall characters with a high accuracy (M=95%). While on the next two rounds the accuracy does not improve, there is an improvement in repetition and duration which could be interpreted as them being more confident.

For the word recognition, we focus on the last round of reinforcement and testing as we consider them to be the end result of the word training process. Note that, in the reinforcement round users are allowed to repeat word or letters whereas in the testing repetitions are not allowed. In each of the two rounds, 24 probes were collected for each user. The recognition accuracy, total duration and repetition rate are presented in Table 4. Additionally, the user recognition accuracy (averaged per user) is shown in the Figure 8. Both Table 4 and Figure 8 reveal that when repetitions are allowed, participants achieve a higher accuracy. Furthermore, a chi squared test reveals that participants achieve a significant higher accuracy ($M = 0.81, STD = 0.39$) in the round where they can perform repetitions compared to the round where they are not allowed to repeat ($M = 0.55, STD = 0.5$); $\chi^2(2, 1056) = 81.67, p = 0.0$.

The sentence recognition round collected 15 probes for each user. The average accuracy, number of interactions, duration and total duration for sentence recognition are presented in Table 5. On average, users needed 37.29 seconds to recognise sentences with an accuracy of 82%.

Interactions

First, we analyse the interactions in the last round of word reinforcement. On average participants performed 2.32 (SD=4.28) interactions for each word. From the interactions, participants repeated 67% of words completely (at least once) whereas they repeated one or more single characters only in 6.8% of the words. Table 5 shows that participants needed on average needed 7.14 interactions to recognise sentences. Moreover, the histogram in Figure 11 shows that the vast majority of users needed a relatively low number of interactions. Eleven users needed on average five or fewer interactions per sentence but some users performed even over 20 interactions per sentence.

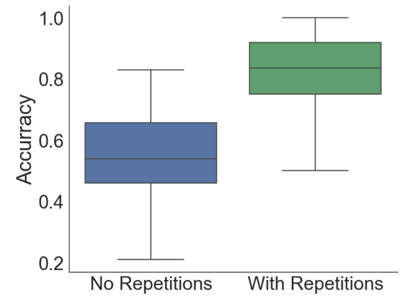

Figure 8. Averaged user accuracy for the last reinforcement (with repetitions) and testing rounds (without repetitions).

Generally, word repetition was used (at least once) in 82.8% of the sentences, character repetition in 6.3% of them and sentences were repeated entirely in 21.2% cases. Within word interactions, 62.8% of them were repetition of current word, 25.6% of them were repetition of the previous word and only 11.6% were a repetition of next word in the sentence.

Figure 9 shows the overall state transitions probabilities between interactions during sentence reading. The chart is constructed from the interaction data of the sentence reinforcement rounds. The transition plot in Figure 10 shows the state transitions probabilities for the first ten iterations of interactions. The start state represents the time point when the sentence is fully transmitted the first time. The finish state represents the user providing the answer for sentence recognition.

The most likely interaction at the start is repeating the current word (probability = 0.57). Participants were also likely to start with previous word interaction (0.23) which could be interpreted as they already understood the current word. Precipitants were also likely to start with repeating the entire sentence (0.17). After a current word repetition, participants were most likely to continue with another current word repetition (0.6), meaning that they did not understand the word from the last repetition. They were also fairly likely to continue with previous word (0.16) or next word repetitions (0.1). They were quite likely to provide the answer (0.10). After a previous word repetition, participants were most likely to continue with another previous word repetition (0.38), meaning that they understood the word that was repeated and they were scanning the sentence backwards. They were also highly likely to continue with current word repetition (0.36), in cases where they did not understand the repeated word. On the other hand, they were relatively less likely to continue with a next word interaction (0.1). They were quite likely to provide the answer (0.16), meaning that they finished backwards scanning. After the next word repetition, participants were most likely to use a current word repetition (0.48) in the case where the did not understand the repeated word. They were also likely to use the next word interaction (0.18) again; scanning forward the sentence, or use the previous word interaction (0.1). They were quite likely to provide the answer (0.22), meaning that they finished forward scanning.

After the entire sentence repetition, users were most likely to continue with another sentence repetition (0.47), which could be interpreted as some users were simply repeating the

Accuracy	Interactions	Total Duration (s)
0.82 (0.38)	7.14 (7.59)	37.29 (27.04)

Table 5. Averaged sentence recognition results (M, STD).

sentence over and over until they were able to understand it completely. Such an interaction was relatively less but still likely followed by current word repetition (0.13) or previous word repetition (0.09). Users also were likely (0.3) to provide an answer. A similar behaviour pattern occurs after character repetition. Users were most likely (0.9) to repeat character again as users who used this interaction were repeating different letters of the current word.

Only five users adjusted the transmission speed during sentence recognition. Two of them set it to a higher than the default speed whereas three of them did slow it down.

Additionally, we explored the relationship between sentence recognition accuracy and the number of interactions. Figure 11 shows that there is a positive Pearson correlation between the average recognition accuracy and the average number of interactions, meaning that users that interacted more, also recognised sentences more accurately; $r = 0.47, p = 0.03$.

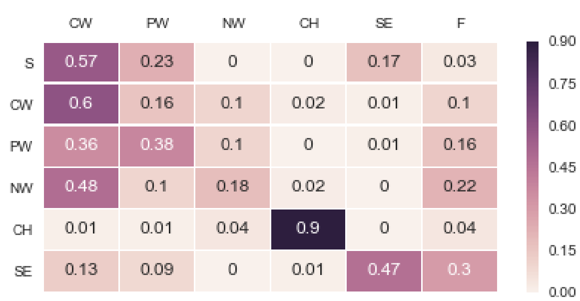

Figure 9. State transitions probabilities between interactions for sentence reading constructed from the sentence reinforcement round. Interaction states: S -start, CW - current word, PW -previous word, NW - next word, C - character, SE - sentence, F- finish.

Questionnaire

On the question of how useful they found the interaction concept when performing sentence reading, 76% (16) of users rated it as "Necessary", 24% (5) of them rated it as "Optional" and no user rated it as "Not useful". When asked why users thought such an interaction was optional, they all argued that with a proper amount of training they would get proficient and there would not be a need for such interaction.

User ratings on how frequent they think each of the commands/interactions they would use are presented in Figure 12. For all word repetition interactions: repeat the current word, previous word, and next word, the vast majority of participants thought that they would use them continuously (every couple of seconds or minutes). Quite contrary, for adjusting the speed the majority if users think that they would rarely use. For repeating the n-th character of the current word, there is some divergence. While most of the users think that they would never use it, two users think they would use it continuously, and another four think they would use it often.

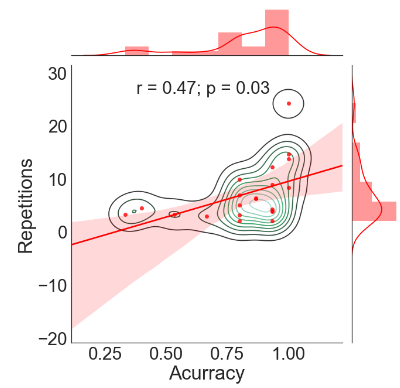

Figure 11. The relation between average sentence recognition accuracy and the average number of interaction. The bar plots on the top and on the side represent histograms and calculated the univariate distribution of the variable in the given axis. The contours represent the multivariate distribution of both variables. The straight line and the shades around it represent the fitted regression and its confidence. The Pearson correlation index and the confidence value are annotated as r and p.

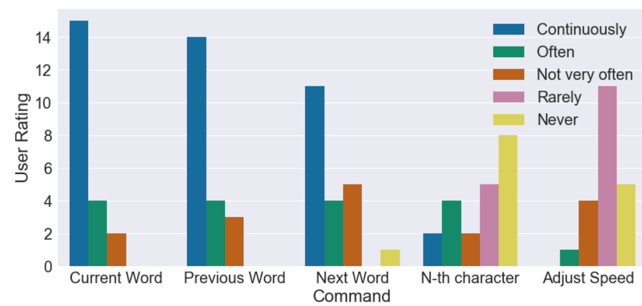

Figure 12. User ratings on how frequent they think they would use each of the interactions for text reading through a vibrotactile display.

User ratings on how suitable the proposed modalities of interaction would be for skin reading application are presented in Figure 13. Gesture interaction received the highest rating $(M = 7.9, STD = 1.7)$. But, a paired t-test reveals that the difference with physical buttons interaction $(M = 6.57, STD = 2.99)$ is not significant ; $t(42) = 1.72, p = 0.101$. On the other hand users rated gesture interaction $(M = 7.9, STD = 1.7)$ significantly higher than interaction using a smart phone $(M = 4.76, STD = 2.64)$; $t(42) = 4.1, p = 0.001$[2]. Additionally, when users were asked to choose one preferred modality of interaction for commands they would regularly and for ones they would rarely use, users mainly prefer gesture-based interaction for regular interactions and smartphone-based interaction for rarely used interactions (see Figure 13).

Discussion

Our study was designed to investigate and identify useful interactions for skin reading with a wearable vibrotactile display. The evidence from all sources such as user performance, interaction behaviour and questionnaire point out that when performing skin reading, users benefit from means of interactions with the vibrotactile display. First, the majority (76%) of the users explicitly expressed in the questionnaire that they

[2]The significance level is considered $\alpha = 0.025$ according to Bonferroni correction for two comparisons

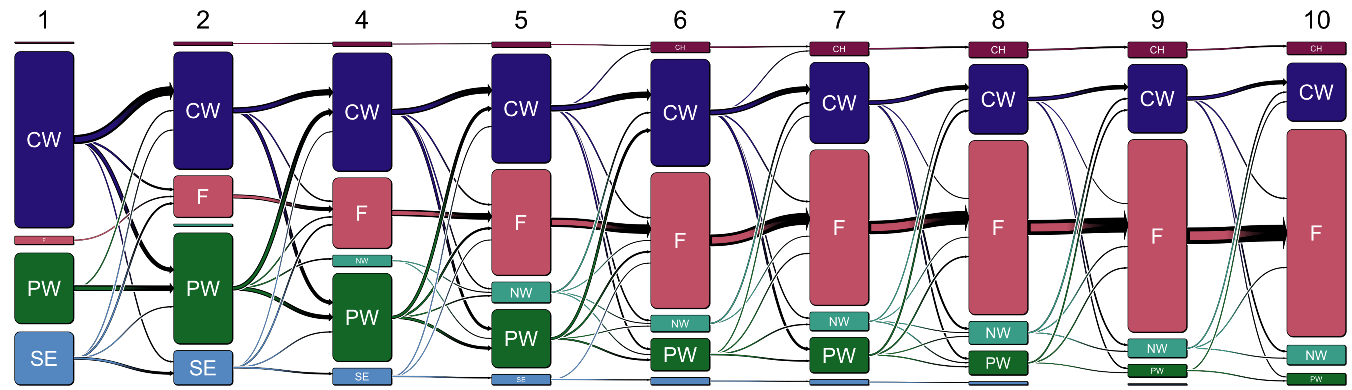

Figure 10. State transitions diagram between interactions for the first ten interaction during the sentence reading. Interaction states: S -start, CW - current word, PW -previous word, NW - next word, C - character, SE - sentence, F- finish. The size of the bar represents the probability of being in that state for the given interaction whereas the width of the arrow represents the probability of the state transition.

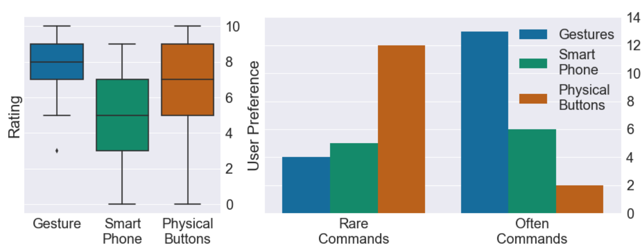

Figure 13. The box plot on the left visualises user ratings (0-10) on how suitable different modalities would be for interactions with a vibrotactile display during skin reading. The bar plot on the right visualises user preferences choices on which modality would be more useful for interactions/commands that they would use very often (at least every couple of hours) versus the interactions they would use rarely.

think having interactions similar to our experiment is necessary for skin reading. While some users (24%) expressed that when proficient, they would not need interactions, none of the believed that such interactions were not useful at all.

The necessity for interactions is also expressed in users' performance during sentence and words recognition. Interactions had a positive effect on the recognition accuracy. Participants performed significantly better (see Figure 8) in the word recognition rounds where repetition is allowed. Additionally, participants who on average performed more interactions in sentence recognition, achieved a better accuracy (see Figure 11).

Interaction usage also demonstrates that the interactions were necessary. Participants, on the last round of word recognition with repetition, on average needed 2.32 (SD=4.28) interactions for each word. Furthermore, they used on average 7.14 interactions for each sentence during sentence recognition.

As participants had no prior experience in skin reading, our study shows that at least for novice users, interactions are crucial. Users with minimal training can start perceiving words and phrases and would be able to understand them if the navigation interactions are at their disposal. Thereby, users do not need to become proficient before they can start using such a wearable device. On the other hand, as users' skin reading skills increase, they would presumably need fewer interactions. They might learn to recognise entire words as units similar to visual reading [53, 45, 11, 40, 5]. Nevertheless, that does not invalidate the need for interactions. First, users might need to repeat certain words from time to time as result of attention breaks or simply due to misperception. In both visual and Braille reading such interactions occur very often even if readers are not aware of it as it occurs unconsciously and naturally. Readers jump backwards to revisit already visited letters and words [22, 37, 38, 39]. This phenomenon is known as back regression, and skilled readers make regressions back in $10-15\%$ of the reading time [37, 38, 39]. Such regression is common practice also in Braille reading [30, 20].

Besides emphasising the importance of interactions in skin reading, our study can be used to derive details of which interactions are most important. Both behaviour analysis and questionnaire analysis confirm that word repetition, and navigation interactions are critical. Participants used word interaction in 82% of sentences and the vast majority of users were convinced that such interactions would constantly be used. Repeating the n-th character of a word was rated as unnecessary by users and was mostly irrelevant to sentence recognition. A closer look at the usage of this interaction reveals that it was only used by seven participants. One participant used it in every sentence; one user used it only in two sentences whereas five users used it only in one sentence. Thus, only one user used character repetitions regularly in sentence recognition, whereas the rest did try to use them in one or two sentences to explore how well that would work but did not continue afterwards with the rest of the sentences. Adjusting the speed seems not to be frequently used. It was used only by five participants. Participants set the speed they were comfortable with and used it for the rest of the sentence recognition. This was also mirrored in the questionnaire where most of the participants expressed that they would only use it a few times during the lifetime of such device. Most of the participants pointed out that they would use it in accordance with the progress of their skin reading skill.

Interestingly and contrary to expectations, the repetition of the entire sentence was used by some users. Although, they were a small number, yet we did not expect any user to rely on it.

Our prior expectation was that at first, users might be tempted to try it, but they would realise that it is challenging with their level of training to perceive the entire sentence at once. Such a scenario did occur with five users, where they used it only in four sentences or less. As shown in the Figure 10, at the first, second or third interactions some users switch to word-based interactions. Initially, they were curious and tried it for one or two iterations (see Figure 10), but then realised it was difficult and switched to other interactions. However, there were users who persisted using sentence repeats. Three users relied on this interaction almost entirely (used in more than 12 sentences), and two others used it moderately (in 7-9 sentences). Figure 10 shows that some participants repeatedly used this interaction until they provided an answer. Although it was interesting to provide such interaction in our study and explore user behaviour, only a small number of users used them. Moreover, in a real-world scenario, where text contains multiple sentences, and they are much longer, repeating the entire text might not very be useful as it does not scale.

The questionnaire reveals that the preferred modality for constantly used interaction would be gesture-based, whereas participants would prefer a settings smartphone application for rarely used interactions (adjusting the speed). For the constantly used interaction, one would need to provide a gesture-based system that would support the basic interactions: repeat the current word, go to previous word (and transmit it) and go to next word. Such a system would be a simplified version of the interaction system we initially designed (see Figure 1). Stopping and resuming to transmit the rest of the text could be automatically achieved using the current word and next word interaction as explained in Section Interaction Concept.

GESTURE BASED INTERACTION

Besides user preferences, from the engineering, designing and manufacturing perspective, a gesture-based interaction would be a perfect fit for a glove-based vibrotactile wearable display as the same glove could be equipped with sensing capabilities to enable gesture recognition. In this section, we map our interactions to hand gesture interactions and explore what sensors would we need to recognise them.

Gesture Mapping

Our study concluded that only the interactions related to word navigation are essential for skin reading. Here we map each of them to a hand-based gesture. The used gestures should be easy to remember, fast to perform, and contain simple movements so that they can be recognised with a minimal set of sensors. Thus, we map the previous word interaction to swiping left gesture and next word interaction to swiping right. The mapping is natural, as it corresponds to movement of focus point within the sentence. Additionally, the interactions are easy to perform and considered in the research community as natural hand gestures [42, 15, 26] meaning that most users would be familiar with them. As for current word interaction, we map it to swipe up gesture as it shares the simplicity and popularity with the other selected gestures.

Gesture Recognition

Gesture recognition has been a subject of many researchers. At a core abstract level, the main approaches for gesture recognition has been based on either environmental sensors such as: camera [8, 6, 12, 8, 1, 18], radar signals [10, 25] wifi [23, 36], etc... or hand/body-worn sensors such as : motion sensors, flexion and pressure sensors [26, 32, 56, 33, 57]. Even though each of them has its advantages and disadvantages, for our specific use case, wearable based sensors approach is a more suitable solution. First, it is not bound to the location (where sensors are placed) and second, adding the sensors to the same wearable glove, makes the setup much more convenient.

Many gesture recognition systems using wearable sensor have been proposed over the years, most of them rely on either hand worn sensors [26, 32, 56, 33, 57] or wrist-based sensors [17, 50, 55, 58]. For recognizing gestures, typically statistical and machine learning approaches such as neural networks [32, 56], support vector machines [26], linear discriminant analysis (LDA) [26, 15], logistic regression [26, 58], decision trees [58], etc.. have been employed.

Our gesture recognition system uses an existing framework and dataset of Luzhnica et al. [26]. The authors, collected data from 18 participants performing 31 gestures using a custom made data glove, where each participant performed each gesture 5-10 times. The data were annotated manually. Their data glove was equipped with seven inertial measurement units (IMU), one on each finger, one on the back of the palm and one on the wrist. Additionally, the glove was equipped with 13 bend sensors to cover main finger joint and wrist movement; and also five pressure sensors on each fingertip. The recording frequency of data glove was 85Hz (85 frames per second). For their recognition system, the authors used a sliding window approach upon which they extracted time domain features such as minimum, maximum, range, average, standard deviation and signal energy and time domain features such as Fast Fourier Transform for every sliding window. The authors evaluated different parameters for sliding window and different machine learning algorithms. They concluded that using LDA for dimensional reduction and then logistic regression for classification yielded the best results, where they achieved a f_1 score accuracy of 98.5%.

For our use case, we will use the same approach in data processing, algorithm, training and evaluation procedure and thus we will skip some of the extensive details (c.f., [26]). From the dataset provided by the authors [26], we use only three (swipe left, right and up) gestures and thus also explore what sensors we would need to correctly identify the given gestures. More precisely, we explore two particular sensors from their set: IMU on the back of the and the IMU on the wrist. Such motion sensors should be sufficient to capture motion characteristics of our gestures.

Considering that we reduce the number of gesture classes to three, the number of windows with rest class is very imbalanced. The rest class represents the data where the user is not performing any gesture such as: not moving or performing arbitrary movements. Thus we reduce the number of windows with rest class by randomly sampling a portion (5%) of them.

	Wrist				Palm			
	I	**L**	**R**	**U**	**I**	**L**	**R**	**U**
I	**385**	4	2	2	**386**	3	1	2
L	5	**61**	2	0	1	**68**	0	0
R	2	3	**54**	0	2	0	**57**	0
U	0	0	0	**60**	0	0	0	**60**

Table 6. Confusion matrix for classification in the test set using IMU on the wrist (left) and IMU on back of the hand (right). Classes: I - rest, L - swipe left, R - swipe right, U - swipe up.

In total, the entire resulting dataset (both training and test) contains 2959 windows. We use LDA for dimensional reduction and logistic regression for classification trained on the training set which represents the 80% the data. The test set (the rest 20% the data) will be used to report on performance.

With only the IMU on back of the hand, our classifier achieves a f_1 score accuracy of 98.4% on the test set. Using only the IMU on the wrist results in an accuracy of 96.5%.

Lessons Learnt

Overall using a single motion sensor (IMU) one would be able to recognise the necessary gesture-based interactions with a very high accuracy. The gestures can be better recognised by placing the sensor on the back of the hand (98.4%) as opposed to placing on the wrist (96.5%). The confusion matrix presented on Table 6 reveals that when using only the IMU on the wrist, there are some more misclassifications for classes left, right and the rest. Such misclassifications are less evident when using the IMU on the back of the hand. This could be explained by the fact that such gestures involve physical flexion and extension of the wrist which can easily be captured by the sensor on the hand but not on the wrist.

However, besides the accuracy, there are design and practical implications that might influence the decision of sensor location. First, there is a vibromotor located on the back of the hand of our vibrotactile wearable glove (see Table 2), which is approximately located nearby the IMU on the data glove used to record the data [26]. Having both IMU and a vibromotor nearby might introduce noise in measurements when the vibromotor is active. A possible overcome could be achieved by shifting in opposite direction (left and right) to maximise the space in between. Alternatively, one could move the IMU on the palm side of the hand. On the other hand, a wrist-worn device makes it impossible to wear it along watches or wristbands. However, considering that a lot of users might already possess smart watches or wristbands equipped with a motion sensor, the motion data from their existing watch or wristband could be used to classify the required hand gestures. This would reduce both costs and power consumption of the wearable vibrotactile display.

LIMITATIONS AND FUTURE WORK

One limitation of our interaction system is that it deals only with text reading and not text comprehension, and thus it does not offer means of navigation beyond neighbouring words. For instance, while reading, users might want to revisit text 3-4 sentences backwards to better comprehend the text. Although, jumping larger portions of the text such as sentences could be provided analogously to our current interactions. Such interaction could be mapped to rotation-based gestures like hand pronation and supination which could be easily be recognised using the motion sensors we proposed. For that, we would need to train users for longer periods so that they would be able to perceive and understand larger messages in the first place. Furthermore, we do not evaluate whether hand motion while performing the gesture could affect the ability of the user to perceive information during skin reading. Such effects need to be studied. If that were the case, a less convenient solution would be to use one hand for skin reading and the other one for interaction. We will consider both limitations mentioned above for conducting additional studies in the future.

CONCLUSION

This paper investigates interactions for skin reading using a wearable vibrotactile display. Initially, we designed an interaction concept for skin reading. We conducted a formative study with 22 users to evaluate our concept during word and sentence reading with a six-channel wearable vibrotactile display. Participants were trained to recognise ten characters, trained on words and then tested on word and sentence recognition, during which they were able to use the designed interactions. Furthermore, participants filled a questionnaire expressing their opinion about the interaction concept in general, different interactions within it and preferred modality of interaction.

The results of our study and analysis of questionnaire indicated that interactions are beneficial for skin reading. Furthermore, this study shaped our interaction concept by characterising interactions like character repetition as not necessary and transmission speed adjustment as less important. As a result, our end concept contains three main interactions all of them providing word repetition and navigation within the sentence. For such interactions, users preferred gesture-based interaction as an interaction modality. Following such a preference, we mapped our interactions to swipe-based hand gestures. Furthermore, we used motion sensors on the back of the hand and wrist to examine how well such gesture-based interactions would be recognisable using machine learning algorithms. Our results showed that a single motion sensor either on the wrist or hand is sufficient to recognise our gesture-based interactions the with a high accuracy. Also, hand (98.4%) is a better choice for locating the sensor compared to wrist (96.5%) in terms of gesture recognition accuracy. Such a sensor could be incorporated into the same wearable glove providing one single solution for both skin reading and interaction. Thus, our contribution could serve as a guideline for designers and manufactures of such wearable vibrotactile displays.

ACKNOWLEDGMENTS

This work was funded by MoreGrasp (H2020-ICT-2014-1 643955). The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

REFERENCES

1. Henrik Birk, Thomas B. Moeslund, and Claus B. Madsen. 1997. Real-Time Recognition of Hand Alphabet Gestures Using Principal Component Analysis. In *In 10th Scandinavian Conference on Image Analysis*.

2. J. C. Bliss, M. H. Katcher, C. H. Rogers, and R. P. Shepard. 1970. Optical-to-Tactile Image Conversion for the Blind. *IEEE Transactions on Man-Machine Systems* 11, 1 (March 1970), 58–65. DOI: http://dx.doi.org/10.1109/TMMS.1970.299963

3. Łukasz Bola, Katarzyna Siuda-Krzywicka, Małgorzata Paplińska, Ewa Sumera, Paweł Hańczur, and Marcin Szwed Szwed. 2016. Braille in the sighted: Teaching tactile reading to sighted adults. *PLoS ONE* (2016). DOI: http://dx.doi.org/10.1371/journal.pone.0155394

4. Stephen Brewster and Lorna M. Brown. 2004. Tactons: Structured Tactile Messages for Non-visual Information Display. In *Proceedings of the Fifth Conference on Australasian User Interface (AUIC)*. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 15–23. http://dl.acm.org/citation.cfm?id=976310.976313

5. James McKeen Cattel. 1886. The Time Taken up by Cerebral Operations. *Mind* os-XI, 42 (1886), 220–242. DOI:http://dx.doi.org/10.1093/mind/os-XI.42.220

6. Qing Chen, Nicolas D. Georganas, and E.M. Petriu. 2007. Real-time Vision-based Hand Gesture Recognition Using Haar-like Features. In *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*. 1–6. DOI: http://dx.doi.org/10.1109/IMTC.2007.379068

7. Meredyth Daneman. 1988. How reading braille is both like and unlike reading print. *Memory & Cognition* 16, 6 (01 Nov 1988), 497–504. DOI: http://dx.doi.org/10.3758/BF03197051

8. N.H. Dardas and Nicolas D. Georganas. 2011. Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques. *Instrumentation and Measurement, IEEE Transactions on* 60, 11 (Nov 2011), 3592–3607. DOI: http://dx.doi.org/10.1109/TIM.2011.2161140

9. Robert O. Duncan and Geoffrey M. Boynton. 2007. Tactile Hyperacuity Thresholds Correlate with Finger Maps in Primary Somatosensory Cortex (S1). *Cerebral Cortex* 17, 12 (2007), 2878. DOI: http://dx.doi.org/10.1093/cercor/bhm015

10. T. Fan, C. Ma, Z. Gu, Q. Lv, J. Chen, D. Ye, J. Huangfu, Y. Sun, C. Li, and L. Ran. 2016. Wireless Hand Gesture Recognition Based on Continuous-Wave Doppler Radar Sensors. *IEEE Transactions on Microwave Theory and Techniques* 64, 11 (Nov 2016), 4012–4020. DOI: http://dx.doi.org/10.1109/TMTT.2016.2610427

11. Dennis F. Fisher. 1975. Reading and visual search. *Memory & Cognition* 3, 2 (01 Mar 1975), 188–196. DOI: http://dx.doi.org/10.3758/BF03212897

12. Pragati Garg, Naveen Aggarwal, and Sanjeev Sofat. 2009. Vision based hand gesture recognition. *World Academy of Science, Engineering and Technology* 49, 1 (2009), 972–977.

13. Robert H Gault. 1924. Progress in experiments on tactual interpretation of oral speech. *The Journal of Abnormal Psychology and Social Psychology* 19, 2 (1924), 155.

14. Frank A. Geldard. 1957. Adventures in tactile literacy. *American Psychologist* 12, 3 (1957), 115–124.

15. Przemys Glomb, Michal Romaszewski, Sebastian Opozda, and Arkadiusz Sochan. 2012. Choosing and Modeling the Hand Gesture Database for a Natural User Interface. In *Proceedings of the 9th International Conference on Gesture and Sign Language in Human-Computer Interaction and Embodied Communication (GW'11)*. Springer-Verlag, Berlin, Heidelberg, 24–35. DOI: http://dx.doi.org/10.1007/978-3-642-34182-3_3

16. Eric Gunther. 2001. *Skinscape: A tool for composition in the tactile modality*. Ph.D. Dissertation. Massachusetts Institute of Technology.

17. Jaehyun Han, Sunggeun Ahn, and Geehyuk Lee. 2015. Transture: Continuing a Touch Gesture on a Small Screen into the Air. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1295–1300. DOI: http://dx.doi.org/10.1145/2702613.2732849

18. Mokhtar M Hasan and Pramod K Mishra. 2012. Hand gesture modeling and recognition using geometric features: a review. *Canadian Journal on Image Processing and Computer Vision* 3, 1 (2012), 12–26.

19. Eve E. Hoggan, Sohail Anwar, and Stephen A. Brewster. 2007. *Mobile Multi-actuator Tactile Displays*. 22–33 pages. DOI: http://dx.doi.org/10.1007/978-3-540-76702-2_4

20. Barry Hughes, Arend W.A. Van Gemmert, and George E. Stelmach. 2011. Linguistic and perceptual-motor contributions to the kinematic properties of the braille reading finger. *Human Movement Science* 30, 4 (2011), 711 – 730. DOI:http://dx.doi.org/https://doi.org/10.1016/j.humov.2010.05.005 Special Issue: Progress in Graphonomics: A Perceptual Motor Skill Perspective.

21. Jacob H. Kirman. 1974. Tactile perception of computer-derived formant patterns from voiced speech. *The Journal of the Acoustical Society of America* 55, 1 (1974).

22. Kevin Larson. 2004. The Science of Word Recognition. *Advanced Reading Technology, Microsoft Corporation* (July 2004). https://www.microsoft.com/typography/ctfonts/WordRecognition.aspx

23. Hong Li, Wei Yang, Jianxin Wang, Yang Xu, and Liusheng Huang. 2016. WiFinger: Talk to Your Smart Devices with Finger-grained Gesture. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 250–261. DOI: http://dx.doi.org/10.1145/2971648.2971738

24. Yi-Chi Liao, Yi-Ling Chen, Jo-Yu Lo, Rong-Hao Liang, Liwei Chan, and Bing-Yu Chen. 2016. EdgeVib: Effective Alphanumeric Character Output Using a Wrist-Worn Tactile Display. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 595–601. DOI: http://dx.doi.org/10.1145/2984511.2984522

25. Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* 35, 4, Article 142 (July 2016), 19 pages. DOI: http://dx.doi.org/10.1145/2897824.2925953

26. G. Luzhnica, J. Simon, E. Lex, and V. Pammer. 2016. A sliding window approach to natural hand gesture recognition using a custom data glove. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. 81–90. DOI: http://dx.doi.org/10.1109/3DUI.2016.7460035

27. Granit Luzhnica, Sebastian Stein, Eduardo Veas, Viktoria Pammer, John Williamson, and Roderick Murray Smith. 2017. Personalising Vibrotactile Displays Through Perceptual Sensitivity Adjustment. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 66–73. DOI:http://dx.doi.org/10.1145/3123021.3123029

28. Granit Luzhnica and Eduardo Veas. 2017. Vibrotactile Patterns using Sensitivity Prioritisation. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 148–155. DOI: http://dx.doi.org/10.1145/3123021.3123032

29. Granit Luzhnica, Eduardo Veas, and Viktoria Pammer. 2016. Skin Reading: Encoding Text in a 6-channel Haptic Display. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM, New York, NY, USA, 148–155. DOI: http://dx.doi.org/10.1145/2971763.2971769

30. Susanna Millar. 2003. *Reading by touch*. Routledge.

31. Susanna Millar. 2004. *Reading by Touch in Blind Children and Adults*. Springer Netherlands, Dordrecht, 437–457. DOI: http://dx.doi.org/10.1007/978-94-017-1731-1_23

32. Kouichi Murakami and Hitomi Taguchi. 1991. Gesture Recognition Using Recurrent Neural Networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. ACM, New York, NY, USA, 237–242. DOI: http://dx.doi.org/10.1145/108844.108900

33. P. Neto, D. Pereira, J. Norberto Pires, and A.P. Moreira. 2013. Real-time and continuous hand gesture spotting: An approach based on artificial neural networks. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. 178–183. DOI: http://dx.doi.org/10.1109/ICRA.2013.6630573

34. Hugo Nicolau, João Guerreiro, Tiago Guerreiro, and Luís Carriço. 2013. UbiBraille: Designing and Evaluating a Vibrotactile Braille-reading Device. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. ACM, New York, NY, USA, Article 23, 8 pages. DOI: http://dx.doi.org/10.1145/2513383.2513437

35. Scott D. Novich and David M. Eagleman. 2015. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental Brain Research* 10 (2015). DOI: http://dx.doi.org/10.1007/s00221-015-4346-1

36. Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home Gesture Recognition Using Wireless Signals. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom '13)*. ACM, New York, NY, USA, 27–38. DOI: http://dx.doi.org/10.1145/2500423.2500436

37. Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124 (1998), 372–422. http://scholar.google.com/scholar.bib?q=info:kPKV2_0qLzcJ:scholar.google.com/&output=citation&hl=en&as_sdt=2000&as_ylo=1998&as_yhi=1998&ct=citation&cd=0

38. Keith Rayner, Barbara R. Foorman, Charles A. Perfetti, David Pesetsky, and Mark S. Seidenberg. 2001. How Psychological Science Informs the Teaching of Reading. *Psychological Science in the Public Interest* 2, 2 (2001), 31–74. DOI:http://dx.doi.org/10.1111/1529-1006.00004 PMID: 26151366.

39. Keith Rayner, Timothy J. Slattery, and Nathalie N. Bélanger. 2010. Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review* 17, 6 (01 Dec 2010), 834–839. DOI: http://dx.doi.org/10.3758/PBR.17.6.834

40. Gerald M. Reicher. 1969. Perceptual Recognition as a Function of Meaningfulness of Stimulus Material. *Journal of Experimental Psychology* 81, 2 (1969), 275.

41. Kailiang Ren, Sheng Liu, Minren Lin, Yong Wang, and Q.M. Zhang. 2008. A compact electroactive polymer actuator suitable for refreshable Braille display. *Sensors and Actuators A: Physical* 143, 2 (2008), 335 – 342. DOI: http://dx.doi.org/https://doi.org/10.1016/j.sna.2007.10.083

42. M. Romaszewski, P. Glomb, and P. Gawron. 2014. Natural hand gestures for human identification in a Human-Computer Interface. In *Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on*. 1–6. DOI: http://dx.doi.org/10.1109/IPTA.2014.7001997

43. A. Russomanno, S. OâĂŹModhrain, R. B. Gillespie, and M. W. M. Rodger. 2015. Refreshing Refreshable Braille Displays. *IEEE Transactions on Haptics* 8, 3 (July 2015), 287–297. DOI: http://dx.doi.org/10.1109/TOH.2015.2423492

44. Caitlyn E. Seim, David Quigley, and Thad E. Starner. 2014. Passive Haptic Learning of Typing Skills Facilitated by Wearable Computers. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2203–2208. DOI: http://dx.doi.org/10.1145/2559206.2581329

45. Frank Smith. 1969. Familiarity of configuration vs discriminability of features in the visual identification of words. *Psychonomic Science* 14, 6 (01 Jun 1969), 261–262. DOI:http://dx.doi.org/10.3758/BF03329112

46. Annette Sterr, Lisa Green, and Thomas Elbert. 2003. Blind Braille readers mislocate tactile stimuli. *Biological Psychology* 63, 2 (2003), 117 – 127. DOI: http://dx.doi.org/10.1016/S0301-0511(03)00051-6

47. Ian R Summers, Jon J Whybrow, Denise A Gratton, Peter Milnes, Brian H Brown, and John C Stevens. 2005. Tactile information transfer: A comparison of two stimulation sites. *The Journal of the Acoustical Society of America* 118, 4 (2005), 2527–2534.

48. Hong Z. Tan, Nathaniel I. Durlach, William M. Rabinowitz, Charlotte M. Reed, and Jonathan R. Santos. 1997. Reception of Morse code through motional, vibrotactile, and auditory stimulation. *Perception & Psychophysics* 59, 7 (01 Jan 1997), 1004–1017. DOI: http://dx.doi.org/10.3758/BF03205516

49. David Ternes and K E MacLean. 2008. Designing Large Sets of Haptic Icons with Rhythm. In *EuroHaptics*, M. Ferre (Ed.). Springer, Madrid, Spain, 199–208. http://www.cs.ubc.ca/labs/spin/publications/spin/ternes2008eh.pdf

50. Wouter Van Vlaenderen, Jens Brulmans, Jo Vermeulen, and Johannes Schöning. 2015. WatchMe: A Novel Input Method Combining a Smartwatch and Bimanual Interaction. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 2091–2095. DOI: http://dx.doi.org/10.1145/2702613.2732789

51. Francisco Vega-Bermudez and Kenneth O Johnson. 2001. Differences in spatial acuity between digits. *Neurology* 56, 10 (2001), 1389–1391.

52. Benjamin W. White, Frank A. Saunders, Lawrence Scadden, Paul Bach-Y-Rita, and Carter C. Collins. 1970. Seeing with the skin. *Perception & Psychophysics* 7, 1 (1970), 23–27. DOI: http://dx.doi.org/10.3758/BF03210126

53. RS Woodworth. 1937. Experimental Psychology. New York: Holt, 1938. *Department of Psychology Dartmouth College Hanover, New Hampshire* (1937).

54. Cheng Xu, Ali Israr, Ivan Poupyrev, Olivier Bau, and Chris Harrison. 2011. Tactile Display for the Visually Impaired Using TeslaTouch. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 317–322. DOI: http://dx.doi.org/10.1145/1979742.1979705

55. Chao Xu, Parth H. Pathak, and Prasant Mohapatra. 2015. Finger-writing with Smartwatch: A Case for Finger and Hand Gesture Recognition Using Smartwatch. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile '15)*. ACM, New York, NY, USA, 9–14. DOI: http://dx.doi.org/10.1145/2699343.2699350

56. Deyou Xu. 2006. A Neural Network Approach for Hand Gesture Recognition in Virtual Reality Driving Training System of SPG. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, Vol. 3. 519–522.

57. Xu Zhang, Xiang Chen, Wen-hui Wang, Ji-hai Yang, Vuokko Lantz, and Kong-qiao Wang. 2009. Hand Gesture Recognition and Virtual Game Control Based on 3D Accelerometer and EMG Sensors. In *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI '09)*. ACM, New York, NY, USA, 401–406. DOI: http://dx.doi.org/10.1145/1502650.1502708

58. Yixin Zhao, Parth H. Pathak, Chao Xu, and Prasant Mohapatra. 2015. Demo: Finger and Hand Gesture Recognition Using Smartwatch. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '15)*. ACM, New York, NY, USA, 471–471. DOI: http://dx.doi.org/10.1145/2742647.2745922