Gaze Interaction with Vibrotactile Feedback:

Review and Design Guidelines

Running Head: Supporting Gaze Interaction with Vibrotactile Feedback

Authors: Jussi Rantala, Päivä Majaranta, Jari Kangas, Poika Isokoski, Deepak Akkil, Oleg Spakov and Roope Raisamo

ABSTRACT

Vibrotactile feedback is widely used in mobile devices because it provides a discreet and private feedback channel. Gaze based interaction, on the other hand, is useful in various applications due to its unique capability to convey the focus of interest. Gaze input is naturally available as people typically look at things they operate, but feedback from eye movements is primarily visual. Gaze interaction and the use of vibrotactile feedback have been two parallel fields of human-computer interaction research with a limited connection. Our aim was to build this connection by studying the temporal and spatial mechanisms of supporting gaze input with vibrotactile feedback. The results of a series of experiments showed that the temporal distance between a gaze event and vibrotactile feedback should be less than 250 milliseconds to ensure that they are perceived as connected. The effectiveness of vibrotactile feedback was largely independent of the spatial body location of vibrotactile actuators. In comparison to other modalities, vibrotactile feedback performed equally to auditory and visual feedback. Vibrotactile feedback can be especially beneficial when other modalities are unavailable or difficult to perceive. Based on the findings, we present design guidelines for supporting gaze interaction with vibrotactile feedback.

CONTENTS

1. INTRODUCTION

Humans use the gaze to look at objects. This behavior can be used as a means to control interfaces in human-computer interaction by estimating the gaze point with the help of an eye tracker. Thanks to recent technological advancements and drop in price, eye tracking is no longer a niche technology only used in laboratories or by users with special needs. For example, with the price of an advanced game controller (~$100), players can enhance their gaming experience with eye tracking[1]. A gaze-aware game knows where the player's visual attention is at each moment and can offer optional input methods (Isokoski, Joos, Špakov, & Martin, 2009) and enhanced gaming experience (Vidal, Bismuth, Bulling, & Gellersen, 2015). At the same time, research on mobile eye tracking has been active. Simple eye-awareness is already included in some cell phone models (e.g. Samsung, 2013, p. 30), so that the phone "knows" when the user is looking at it. Research on pervasive and mobile gaze interaction has demonstrated how eye tracking can enhance the interaction with mobile phones (Dybdal, San Agustin, & Hansen, 2012; Miluzzo, Wang, & Campbell, 2010; Rozado, Moreno, San Agustin, Rodriquez, & Varona, 2015), tablets (Holland & Komogortsev, 2012), smartwatches (Akkil et al., 2015; Hansen et al., 2015; Hansen et al., 2016), smart glasses (Zhang et al., 2014), as well as smart environments and public displays (Zhang, Bulling, & Gellersen, 2013).

Since the eye is primarily a perceptual organ, using gaze as an intentional control method poses challenges for interaction design (Skovsgaard, Räihä, & Tall, 2012). Most importantly, viewing should not be misinterpreted as a voluntary command. In gaze interaction literature (Jacob, 1991), this problem is known as the "Midas touch" problem where viewed objects are unintentionally acted on. Feedback plays an essential role in informing the user how the system is interpreting the gaze. Gazing an object in real life provides naturally only visual feedback. Computers and smart devices can indicate if an object has been recognized as being pointed at, or being selected. Previous research has shown that visual and auditory feedback on gaze input significantly improve user performance and satisfaction (e.g. Majaranta, MacKenzie, Aula, & Räihä, 2006). However, the effects of haptic feedback in gaze interaction have remained largely unknown. We assume haptic feedback could provide a useful alternative to, at least the audio, as auditory and haptic perception are known to share similarities. For example, Jokiniemi, Raisamo, Lylykangas, and Surakka (2008) found that participants could perceive auditory and tactile rhythms more accurately than visual rhythms. Auditory and haptic feedback can be perceived independently from the gaze location. Unlike the distal senses of vision and hearing, touch is a proximal sense that provides information of things close to or in contact with us. How would the interplay of a distal (vision) and proximal (touch) sense work? For instance, instead of seeing a button change its appearance, the user could feel the click of a button after selecting it with gaze. Could this novel combination of modalities provide some benefits compared to visual and auditory feedback or is this unnatural combination of action and feedback perhaps incomprehensible? These were the questions that motivated us in the work reported in this paper.

Haptic feedback has become more common in consumer technology due to the emergence of mobile and wearable devices designed to be in contact with the skin.

---

[1] see e.g. http://www.tobii.com/xperience/apps/

The most popular form of haptic stimulation in mobile and wearable devices is vibrotactile feedback. For example, continuous vibration is an effective way to notify of incoming calls with mobile phones. Shorter vibration bursts are used on phones and tablets to replace the tactile feel of pressing a physical key when typing with a virtual keyboard. This has been shown to improve typing speeds (Hoggan, Brewster, & Johnston, 2008). Vibrotactile technology is also included in some smartwatches. In the Apple Watch[2], for instance, vibrotactile stimulation is used to mimic a heartbeat that can be sent to a close friend or family member. With multiple actuators, it is possible to create touch sensations that move on the wrist (J. Lee, Han, & G. Lee, 2015; Lee & Starner, 2010). To date, commercial smart glasses and other head-mounted devices have not utilized vibrotactile feedback. This is surprising since it is known that users can quite accurately localize which part of the head is stimulated with vibrotactile actuators (e.g. Myles & Kalb, 2013).

We were interested in studying how vibrotactile feedback could support gaze interaction. We conducted a series of experiments, where we focused on four main research questions: effectiveness of vibrotactile feedback (RQ1), temporal limits between gaze events and vibrotactile feedback (RQ2), effects of feedback location and spatial setup (RQ3), and vibrotactile feedback in comparison to other modalities (RQ4). Because our results are spread over more than 20 papers (14 of which are discussed here), this could make it difficult for future researchers to extract the main findings. The contribution of this paper is to summarize the research results in a compact form and serve as a collection of pointers to more detailed work in the original papers. The goal is to add to the understanding of how the two modalities of haptics and gaze can be utilized effectively in HCI.

The organization of the paper is as follows. We will first introduce gaze interaction and vibrotactile feedback. We then present results from the experiments before discussing lessons learnt from the studies. We end with general discussion and present design guidelines based on the accumulated knowledge and insights.

1.1. Gaze Interaction

Eye Movements and Gaze Tracking

There are different ways to implement gaze based human-computer interaction. We follow the classification of Møllenbach, Hansen, and Lillholm (2013) who used three categories based on how we use our eyes: fixation based, saccade based, and smooth pursuit based interaction.

When we look at things, we fixate our gaze on the object of interest. The duration of such a *fixation* can vary from tens of milliseconds to a few seconds (Holmqvist et al., 2011, p. 381), depending on the task. However, the typical average duration over several different tasks is about 200-350 ms (Rayner, 1998), with 100 ms as a typical minimum duration (Mulvey, 2012, p. 18). Due to the distribution of the sensory cells on the retina, only a small area of the scene can be seen accurately. The dense color-sensing cell population on a small area known as the fovea makes the visual acuity excellent, but only on a small 1.5-2 degree area at the center of the visual field. The need of humans to turn their eyes so that objects of interest are projected on the fovea

---

[2] http://www.apple.com/watch/

makes gaze tracking such a useful tool in tracking visual attention. Peripheral vision, on the other hand, is sensitive to changes and a movement in the periphery easily attracts our visual attention.

The eye movements between fixations are called *saccades*. These are rapid ballistic eye movements with high velocity (up to 500° per second) with average duration of about 30-50 ms, depending on the distance covered. For example, 12° movement takes about 50 ms (Gilchrist, 2011). Saccades can have quite wide amplitudes but most every day saccades remain below 15°. Our vision is suppressed during a saccade.

A third major type of eye movement is called *smooth pursuit*. Smooth pursuit allows us to continuously track a moving object and to track a stationary object while we ourselves are moving. When the head is moving, the eye movements are guided by the vestibule-ocular reflex, which utilizes information from balance organs in the head. When the head is not moving smooth pursuit is generated based on visual information only. Smooth pursuit movements are often accompanied with corrective saccades, to keep up with the motion.

There are also other eye movements, such as microsaccades, rotations, and vergence movements that are not relevant for this article. Further information can be found, e.g. in the handbook of eye movements by Liversedge, Gilchrist, and Everling (2011).

A number of methods have been used for tracking eye movements (Duchowski, 2007) and defining gaze position, or gaze vector (Hansen & Ji, 2010). The most common method is based on analyzing a video image of the eye, also known as video-oculography (VOG). For each captured frame, the tracking software detects a number of visual features, such as pupil size, pupil center, etc. VOG-based trackers typically require a calibration before the gaze point can be estimated. During the calibration, the user looks at dots (usually 5-9), one at a time. The calibration dots and the corresponding sets of visual features are used by the tracking software to calculate the *visual-features*-to-*gaze-point* transformation that is then used to estimate gaze point based on the eye images. An additional corneal reflection from a near-infrared light source (often used to provide stable eye illumination) can help in compensating for head movements, thus improving the tracking accuracy. The tracker can also be mounted on the head, e.g. by integrating it into eye glass frames. Tracking gaze in 3D poses additional challenges, such as how to map the gaze vector to the real life scene. Techniques include e.g. using a scene camera to recognize visual markers placed in the environment (see e.g. Pfeiffer & Renner, 2014).

In addition to VOG, eye movements can also be detected by electro-oculography (EOG), based on the cornea-retinal potential difference (Majaranta & Bulling, 2014). This method is most useful in detecting relative eye movements when the exact point-of-gaze is not needed. This is because the accuracy of the absolute gaze point position is not high. Earlier versions of EOG trackers were invasive, as they required sticky electrodes to be placed on the skin around the eyes. Most recent implementations hide the contact points, for example to the nose piece of eye glass frames (Ishimaru et al., 2014) or to ear pods (Manabe, Fukumoto, & Yagi, 2015), making the EOG a viable alternative to VOG. However, in our studies, we used the VOG based eye tracking.

Detecting saccades, fixations, and smooth pursuits from gaze data samples is not a trivial task, especially in real-time. As a general rule, it can be said that a fixation is detected when the eye remains within about one degree of vision for at least 100 ms, and a saccade when the movement exceeds two degrees (Mulvey, 2012, p. 18). Detecting smooth pursuit is challenging and several algorithms have been proposed (Larsson, Nyström, Andersson, & Stridh, 2015). Luckily, interactive applications can often be constructed without relying on explicit classification of samples. Instead, user interfaces can operate utilizing areas that are in focus without having to know whether a fixation or a smooth pursuit made the gaze stay on them. Also, smooth pursuit-based interaction does not really require smooth pursuit in a strict sense. Most algorithms can work regardless of whether smooth pursuit or small saccades cause the gaze trajectory correspond to the trajectory of a moving object. Consequently, in the work reviewed in this paper, names of the eye movements are used rather loosely. A fixation often means a relatively stable period in eye movements and smooth pursuit should be understood as a movement that could be a smooth pursuit, but could include something else as well.

There are multiple ways of using gaze in human-computer interaction. Gaze can be used as implicit input, where the system identifies the user's interest based on gaze pattern and adjusts system behavior accordingly. Alternatively, gaze can be used to provide explicit commands. For our studies, we chose three different approaches for explicitly using gaze as input; dwell-select, gaze gestures, and smooth pursuit based interactions. We will next describe these approaches in more detail.

Dwell-select

Dwell-select is based on fixations: the user fixates on a target in order to select it. However, the temporal and spatial thresholds for a dwell are different from a fixation. Since we use fixations for visual perception, dwell time should exceed the normal fixation duration to avoid the Midas touch problem. A suitable dwell time depends on the task. For example, a fairly short dwell time (e.g. <400 ms, Räihä & Ovaska, 2012) may be preferred in eye typing where dwell is used to select letters from an on-screen keyboard in a repetitive manner while more complex problem-solving tasks require longer dwell times (e.g. 1000 ms, Bednarik, Gowases, & Tukiainen, 2009).

Similarly to the dwell duration, the area in which the gaze should remain can be defined based on the task. Typically, for interaction purposes it is enough if the gaze remains within the borders of the interface control element, e.g. a button. The size of the fovea and the accuracy achieved by the VOG trackers limit the minimum size of a selectable object to about 1-2 degrees. Increasing the button size makes it easier to select (for methods to cope with the eye tracking inaccuracy, see e.g. Majaranta, Bates & Donegan, 2009).

To avoid false selections, some indication of the selection process is needed. First, an indication of gaze entering the target confirms that the tracker is correctly following the gaze and that the user's intention is correctly interpreted. When feedback on target entry is shown before the target is activated, it gives an opportunity to move the gaze away and thus prevent an unintended command. Finally, a separate feedback for the target activation (e.g. auditory click to confirm a button press) marks the end of a successful selection process.

Gaze Gestures

Gaze gestures exploit eye movements but they differ from natural saccades as they follow a specific pattern (Drewes & Schmidt, 2007; Istance, Hyrskykari, Immonen, Mansikkamaa, & Vickers, 2010). Huckauf and Urbina (2008) define gaze gestures as "sequences of fixation locations, which are not necessarily coupled to dwell times". In other words, even though there is a short fixation in between saccades, no dwelling is needed apart from a pause between consecutive gestures. The gestures can be simple, such as pointing downwards e.g. to scroll down, or complex sequences that are interpreted as abstract commands. For example, crossing predefined sequences of on-screen areas with the gaze can be recognized as commands by the system (Porta & Turina, 2008). One of the challenges in eye gesture design is the need to make gestures simple and fast to learn and execute (Huckauf & Urbina, 2008) but distinct enough to be easily distinguished from other eye movement activity (Drewes, De Luca & Schmidt, 2007). One of the proposed ways to achieve this are gestures that include off-screen targets (Isokoski, 2000) i.e. targets that are located outside the normal viewing area, such as computer or smartphone screen. Gestures do not need to be bound to locations. This means that they can be insensitive to spatial accuracy problems (Drewes & Schmidt, 2007). Eye trackers based on electro-oculography tend to have poor accuracy, but they can still be utilized with gesture systems, especially in mobile settings (Bulling, Roggen, & Tröster, 2009).

Visual feedback is useful for indicating the focus point on a computer screen. However, gaze gestures are often made in a fast sequence which may make it hard to perceive visual feedback. In addition, relative gaze gestures that are not bound on any specific location require different kind of feedback. Auditory feedback may be useful as it can be perceived independently of gaze behavior (MacKenzie & Ashtiani, 2011). The resulting action or command is of course a kind of feedback to the gaze-issued command. In addition, it may also be useful to provide feedback to the user while making the gesture to inform about its progress. Otherwise, it can become hard to interpret what kind of problem prevented the successful completion of the desired command by gaze: was it the tracker that was not receiving data, or one of the strokes that was not recognized, or did the user make a wrong gesture?

Smooth Pursuit Interaction

Smooth pursuit interaction is based on pursuit movement while tracking an object. One interesting feature of smooth pursuit is that it enables the user to draw by eyes a fairly smooth curve. Trying to draw a circle by the eyes without the pursuit results in rugged angular form, caused by numerous small saccades (Heikkilä & Räihä, 2009). Smooth pursuit eye movements enable spontaneous gaze interaction in situations where calibration is inconvenient or hard to accomplish. Instead of learning certain gaze gestures, the user simply follows one of the several moving objects to initiate a command. Smooth pursuit interaction can be exploited e.g. with large public screens (Vidal, Bulling, & Gellersen, 2013) or smartwatches (Esteves, Velloso, Bulling, & Gellersen, 2015).

Some visual feedback is inherent in the procedure of smooth pursuit interaction itself: the object to be tracked must be seen. Visual feedback is also useful in confirming that the system has noticed that the object is tracked by gaze, e.g. by changing the appearance of the control object (Esteves et al., 2015) or by playing

auditory feedback while the pursuit interaction is active (Kangas et al., 2016a; Špakov, Isokoski, Kangas, Akkil, & Majaranta, 2016).

In all of the above cases, proper feedback is essential for efficient and pleasant gaze interaction. Before going into combining gaze interaction with haptics, we will first introduce the main concepts of haptic and vibrotactile feedback.

1.2. Haptic and Vibrotactile Feedback

Haptics refers to "sensory and/or motor activity based in the skin, muscles, joints and tendons" (ISO, 2009). This activity is processed in the human somatosensory system which can be divided into proprioception, kinesthesis, and cutaneous senses (Goldstein, 1999). Proprioception and kinesthesis are related to the sense of position and movement of our limbs. They enable us to feel forces that can be created with haptic technology. However, many haptic force feedback devices are not suitable for mobile use because they are large and need to be attached to a surface. Cutaneous sensations, on the other hand, are mediated by the skin. In this article, we focus on tactile stimulation which refers to mechanical interaction with the skin (Kern, 2009). A typical tactile stimulation type is vibration.

Tactile stimulation is sensed via mechanoreceptors in the skin. The four main types of mechanoreceptors – Merkel receptors, Meissner corpuscles, Pacinian corpuscles, and Ruffini endings – each respond to different touch stimuli (Gardner, Martin, & Jessell, 2000) that can vary between pressure, taps, skin stretch, and vibration (Goldstein, 1999). Efficient use of these stimuli for communicating information requires knowledge of touch perception across different body parts.

Touch Perception

Pressure sensitivity of the skin has an influence, for example, on how easily we can sense a fly landing on our skin. This is easier in body parts with high sensitivity. The most sensitive areas to pressure are the forehead (face), trunk, and fingers, whereas the feet are the least sensitive areas (Weinstein, 1968). From tactile feedback viewpoint, information of pressure sensitivity is especially relevant when designing stimulation methods that use taps or other linear indentation of the skin. Another important sensitivity measure is vibration sensitivity. Using a vibrator placed against different parts of the skin, Wilska (1954) found hands and soles of the foot to be most sensitive. The least sensitive were the abdominal and gluteal regions of the body.

Furthermore, spatial sensitivity is related to our capability to identify and differentiate between touch locations. A classic measure to study this is the two-point threshold test which measures the smallest separation between two points on the skin that is perceived as two points rather than one (Goldstein, 1999). Weinstein (1968) reported that this spatial acuity is highest in the distal parts of the body (e.g. fingers and palm) and decreases when moving towards proximal parts (e.g. forearm and upper arm). The measured thresholds were lowest in fingers (2-3 mm), facial area (5-7 mm), and palm (10 mm). The least sensitive parts were the calf, thigh, upper arm, and back (all over 40 mm). In practice, the distance between two points should not be shorter than the thresholds to ensure that users can differentiate the stimuli.

Finally, there are also some temporal differences in perception of touch
stimulation across the human body. Because of neural processing, information of a
touch reaches the brain faster if the touched body part is close to the head. For
example, travel time for the toes is approximately 30 ms longer than for the nose
(Macefield, Gandevia, & Burke, 1989). Taken together, the most touch sensitive parts
of the human body are the fingers, palm, and facial area. This is also reflected in
earlier haptics research where especially fingers (e.g. Pietrzak, Crossan, Brewster,
Martin, & Pecci, 2009; Tan, Durlach, Reed, & Rabinowitz, 1999) and palm (e.g.
Rantala et al., 2011; Yatani & Truong, 2009) have been studied extensively. The
location of stimulation is only one parameter that can be varied when using tactile
stimulation in human-computer interaction. We will next extend our discussion to the
other parameters.

Stimulation Parameters

Vibration and other forms of tactile stimulation can be used to communicate
information to users by creating sensations that are perceptually different. This is
achieved by varying parameters of mechanical stimulation. In addition to spatial
location, the most important parameters are *frequency*, *amplitude*, *waveform*,
duration, and *rhythm*.

Frequency perception is mediated by mechanoreceptors that respond to
stimulation frequencies ranging from 0.3 Hz to over 500 Hz (Goldstein, 1999). The
human peak sensitivity of sensing vibration is approximately 250 Hz, and many
actuators are optimized for this frequency. In practice, it is recommended to use a
maximum of 3-5 different frequency levels because with more levels differentiating
them becomes difficult (Sherrick, 1985). Amplitude is related to the "loudness" of
stimulation, and it should be such that the stimulation can be detected by the user
while still being low enough not to be annoying or cause discomfort. Similarly to
frequency, the number of used amplitude levels should be kept relatively low (Brown
& Kaaresoja, 2006).

Waveform is the shape of the signal fed to a tactile actuator. A sine wave is the
recommended waveform for many actuators based on their electro-mechanical design,
and, therefore, not all actuators are capable of accurately reproducing other
waveforms (Brown, 2007). Duration of actuation defines the time the skin is
stimulated. The human skin is very sensitive to touch, and even stimulation lasting
only a few milliseconds can be perceived. In practical applications, it is beneficial to
use longer stimulation so that users are able to detect it even when mobile. On the
other hand, too long vibration can become annoying. Rather than use constant
vibration, it may be more efficient to create rhythmic stimulation consisting of several
temporally separated pulses. The design space of rhythm is vast, and it is possible to
create very large stimulus sets where individual stimuli are still perceptually different
(Enriquez & MacLean, 2008).

Lastly, as discussed above, spatial location is an effective parameter as long as the
spacing between actuators is sufficient for the stimulated body part. Sofia and Jones
(2013) reported that when nine actuators were placed on the palm, participants could
localize a single actuator with 85 % accuracy (inter-actuator spacing 22 mm). To
reach a comparable accuracy of 92 % on the abdomen, Cholewiak and Collins (2003)
placed eight actuators evenly with an inter-actuator spacing of 107 mm. For

stimulating the head, Myles and Kalb (2010) recommended to use no more than four vibrating actuators to ensure robust localization. While these findings can guide placing of tactile actuators, they should not be taken as absolute limits because many other factors such as the used actuator type can also affect the localization accuracy.

Vibrotactile Actuators

Vibration is the most commonly used tactile stimulus type in current consumer devices. This is partly because vibrotactile actuators are relatively simple to control and the technology is inexpensive. The actuators are typically the size of a coin or smaller and, therefore, highly suitable for mobile and wearable devices. Vibration is also a stimulation method that is currently better understood than other alternatives such as static pressure or skin stretch. For these reasons, we have used vibrotactile actuators exclusively in our work where we wanted to have the main focus on combined use of gaze and vibrotactile feedback.

The three main approaches to creating vibrotactile sensations are rotary electromagnetic actuators, linear electromagnetic actuators, and non-electromagnetic actuators (Choi & Kuchenbecker, 2013). Rotary electromagnetic actuators are typically based on an off-center mass which is attached to a shaft so that its rotation causes vibration when current is on (Figure 1a). These eccentric rotating mass (ERM) motors have been used in many mobile phones and game controllers. Limitations of ERM motors include latency in starting and stopping the mass. In addition, it is not possible to control both amplitude and frequency independently; changing one also affects the other.

Linear electromagnetic actuators are based on a magnet and movable coil that interact when current is applied to the coil. The same operating principle is used in audio speakers. Examples of linear electromagnetic actuators are shown in Figure 1. The C2 tactor (EAI, Engineering Acoustics, Inc.) has a contactor that moves perpendicularly to the skin (Figure 1b). Because the C2 has a cylindrical housing, it can be easier to embed in devices or garments than the HiWave HIHX14C02-8 (Figure 1c). The moving part can also be fully enclosed in a housing. Examples of this include the PMD C10-100 (Figure 1d) and the LVM8 (Figure 1e) actuators that have diameters of less than 10 mm. Linear electromagnetic actuators provide relatively precise control over stimulation parameters because amplitude and frequency can be defined independently.
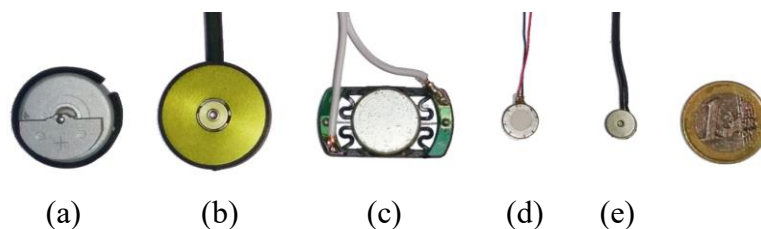

(a)   (b)   (c)   (d)   (e)

FIGURE 1. Examples of vibrotactile actuators: ERM motor (a), EAI C2 (b), HiWave HIHX14C02-8 (c), PMD C10-100 (d), and LVM8 (e). A one Euro coin is included for scale.

Many non-electromagnetic actuators such as piezoelectric actuators also offer very precise control over stimulation parameters. Piezoelectric actuators stimulate the

skin by utilizing thin layers that shrink or expand according to the driving signal polarity (Poupyrev & Maruyama, 2003). However, high activation voltage of around 100V is needed to create movement (Pasquero et al., 2007). Tactile stimulation can also be provided in mid-air using pneumatic air pressure systems (Sodhi, Poupyrev, Glisson, & Israr, 2013) or ultrasound transducers (Hoshi, Takahashi, Iwamoto, & Shinoda, 2010). These technologies are typically targeted to enhance interaction based on hand gestures (e.g. Sodhi, Poupyrev, Glisson, & Israr, 2013). Even though especially ultrasound-based solutions can already be miniaturized, we chose linear electromagnetic actuators because they are suitable also when stimulating body areas other than the hands.

In summary, our review focuses on vibrotactile stimulation of the human skin that is only a subset of the possible stimulation types in haptic interaction. Force feedback technology capable of stimulating also the muscles and tendons would offer a richer haptic experience, but enabling it in mobile and wearable settings is currently difficult. Furthermore, since haptics and gaze have not been studied together to date, starting to map the possible design space by focusing on one stimulation type was a practical choice.

2. METHODS

We conducted a series of experiments in controlled laboratory settings to start building an understanding of how to best combine gaze interaction and vibrotactile stimulation. We constructed prototype systems that utilized gaze trackers for estimating participant's gaze point and vibrotactile actuators to stimulate the participant's skin. The choice of eye tracker and vibrotactile actuators depended on the goal of each individual study (see Figure 2 for details).

Study	Gaze interaction technique	Device form factor	Main research question	Eye tracker	Vibro-tactile actuator	Location of stimulation
Zhi, 2014	Dwell	Desktop	RQ1: Effectiveness of vibrotactile feedback	Tobii T60	C2	Finger
Kangas et al., 2014d		Desktop	RQ2: Temporal limits	Tobii T60	LVM8	Finger
Rantala et al., 2014	Gaze gestures	Glasses	RQ3: Feedback location	Tobii T60	LVM8	Head
Kangas et al., 2014a	Gaze gestures	Mobile phone	RQ2: Temporal limits	Tobii T60	LRA of Nokia Lumia 900	Hand
Kangas et al., 2014b	Gaze gestures	Mobile phone	RQ1: Effectiveness of vibrotactile feedback	Tobii T60	LRA of Nokia Lumia 900	Hand
Kangas et al., 2014c	Gaze gestures	Glasses	RQ4: Feedback modality	Tobii T60	LVM8	Head

Käki et al., 2014	Dwell	Desktop	RQ1: Effectiveness of vibrotactile feedback	Tobii T60	C2	Finger
Akkil et al., 2015	Gaze gestures	Wrist watch	RQ4: Feedback modality	Ergoneer Dikablis	PMD C10-100	Wrist
Špakov et al., 2015	Directional cueing	Wearable	RQ3: Feedback location	Tobii T60	C2	Head, neck, and back
Rantala et al., 2015	Gaze gestures	Glasses	RQ1: Effectiveness of vibrotactile feedback	Tobii EyeX	LVM8	Head
Majaranta et al., 2016	Dwell	Desktop	RQ4: Feedback modality	Tobii T60	C2	Wrist and finger
Köpsel et al., 2016	Gaze gestures		RQ4: Feedback modality	SMI RED with iViewX	HiWave HIHX14 C02-8	Palm
Kangas et al., 2016a	Smooth pursuit	Desktop	RQ4: Feedback modality	Tobii EyeX	LVM8	Head
Kangas et al., 2016b	Gaze gestures	Glasses and tablet	RQ3: Feedback location	Tobii EyeX	LVM8	Head and fingers

FIGURE 2. Studies reviewed in this paper. For each study, we list the gaze interaction technique, device form factor, main research question, eye tracker, vibrotactile actuator, and location of vibrotactile feedback. See Figure 1 for images of the actuators.

The focus in each study was on studying the effects of vibrotactile feedback. However, all our user interface prototypes were multimodal because at least some visual feedback was always available. We made no special effort to block out sounds generated by the vibrotactile actuators because they would be present in real products as well. However, in some cases (Köpsel, Majaranta, Isokoski, & Huckauf, 2016; Majaranta et al., 2016) the actuators were placed on a pillow to prevent disturbing sound from the vibration. Participants were recruited mainly from the local University community, and all signed informed consent forms before proceeding to the experiments. If the used gaze interaction technique required tracker calibration, this was done in the beginning of the experiment. We also made sure that the stimulation amplitude level set in the piloting phase of each study was perceivable for participants. All experiments included within-subject comparisons, and we collected both quantitative and qualitative data. The ETU-Driver[3] was used for gaze data acquisition in some of the studies. Other studies utilized directly the SDKs provided by the tracker manufacturer, like Tobii EyeX SDK[4].

In testing for statistically significant differences, we often used a non-parametric permutation test (e.g. Dugard, 2014; Nichols, 2002). Permutation tests are not widely used in HCI. Therefore, we will briefly explain the rationale. The starting point is the

---

[3] http://www.sis.uta.fi/~csolsp/downloads.php
[4] http://developer.tobii.com/eyex-sdk/

observed value of the statistic in interest, the test statistic (e.g. the difference between means in two feedback conditions). This observed value of the test statistic is compared against a distribution of test statistics produced by re-sampling from the measurements assuming no difference between the test conditions (i.e. if the null hypothesis is true). The relevant *p*-value is the proportion of the distribution that is more extreme or equal to the observed value. In other words, given the observed data, how likely was the difference to occur by chance?

The argument is that if the null hypothesis is true, then all possible permutations of the data are equally likely, and the observed sample is just one of them and should appear as a typical value. If this does not seem to be the case (the observed test statistics value is rather extreme, there are only a few equally extreme or more extreme values) then the null hypothesis probably is not true. The permutation test principle is very general, and does not depend on assumptions on the normality of samples, random sampling, or independence of observations.

3. EFFECTS OF VIBROTACTILE FEEDBACK ON GAZE INTERACTION

This section presents results from our experiments organized according to the four main research questions. We start by introducing work that studied the effectiveness of added vibrotactile feedback (RQ1). We wanted to measure if vibrotactile feedback provides performance benefits in gaze interaction tasks. Next, we move on to studying the temporal limits in giving vibrotactile feedback for gaze events (RQ2). We expect that a too long delay after a gaze event makes it difficult to associate the event with subsequent feedback. We then investigate how body location of vibrotactile feedback affects the use of gaze input (RQ3). Since vibrotactile stimulation is typically absent in gaze interaction, there is no prior work on finding a suitable body location for feedback. We conclude the results section by focusing on comparisons between vibrotactile stimulation and other feedback modalities (RQ4).

3.1. Effectiveness of Vibrotactile Feedback

We were interested in learning whether added vibrotactile feedback has an effect on gaze interaction, and what this effect is. Gaze interaction with a mobile phone was chosen as the target application for the first experiment (N=12) originally reported in Kangas et al. (2014b). Vibrotactile feedback was given with the phone's built-in LRA actuator. Gaze gestures were used for input because they are more robust than dwell-select in mobile scenarios (Bulling et al., 2009; Drewes & Schmidt, 2007; Dybdal et al., 2012).

The task was to select a name from a contact list and make a simulated call (see Figure 3). The user could browse the list by making two-stroke gaze gestures. Each gesture started by looking at the phone screen. The user then made a gaze stroke outside of the screen, immediately (within a 500 ms threshold) followed by a returning stroke that took the gaze back to the screen. The list of names could be scrolled up and down by using gestures moving above or below the screen, respectively. The highlighted name in the middle could be selected by a gesture to the right. A gesture to the left cancelled the action and returned back to the list.

FIGURE 3. The mobile phone used in Kangas et al. (2014b) was held in front of a gaze tracker (left). The experimental interface consisted of a list of contact names (right). © 2014 Association for Computing Machinery, Inc. Reprinted by permission.

Four feedback conditions were tested:

- *No:* No vibrotactile feedback.
- *Out:* Vibrotactile feedback confirming a stroke out (from the device).
- *Full*: Vibrotactile feedback for the full gesture (to confirm the returning stroke).
- *Both*: Vibrotactile feedback for both strokes (out and return).

The results from a within subject study showed that vibrotactile feedback increased the efficiency of the interaction. Especially, task completion times in the *Out* condition were significantly shorter than in the *No* and *Full* conditions ($p \leq 0.05$ for both). Without vibrotactile feedback, participants also performed more gestures to complete the task than in other conditions. Participants appreciated the vibrotactile feedback; the condition which confirmed *both* strokes was most liked. None of the participants preferred the condition with no vibrotactile feedback.

Besides the convenience of having vibrotactile actuators embedded in the phone, vibrotactile feedback was well suited for this scenario also because visual feedback was not available outside of the device. Vibrotactile feedback indicated that the gesture segment ending outside the device was successfully recognized. In addition, feedback of a returning stroke could provide a confirmation of a full gesture even without visually inspecting the resulting changes on the screen. This is probably useful especially when several consecutive gestures are made at a fast pace. Visual feedback can also be hard to detect during rapid gesturing, where fixations for the purpose of observing feedback would just slow down the interaction (Istance et al., 2010).

We continued to study the effect of vibrotactile feedback on gaze gestures in another experiment (N=12) originally reported in Rantala et al. (2015) where the vibrotactile feedback was given via eye glass frames (Figure 4). The eye glass form factor of head-mounted gaze trackers provides multiple natural contact points to the skin, enabling to build an integrated system that tracks the gaze and also provides vibrotactile feedback to gaze events. Another advantage of the glass frames is that they can provide spatially congruent feedback: vibrotactile feedback can be given on the left or on the right, corresponding to the gesture direction.

FIGURE 4. Glasses with three vibrotactile actuators.

In this study, the participants' task was to make a gesture to the given direction on a computer screen. In the beginning, the participant focused on a marker in the center of the screen. The marker changed its appearance into an arrow symbol to indicate the gesture direction ('<' for the left and '>' for the right, see Figure 11). This time, there was a visible target box on the screen to mark the ending location of the "outgoing" gesture (the landing area used for tracking was bigger than the visible target box). The gesture started from the center box and ended when the gaze returned to the center box. The results showed that the task times were longer without vibrotactile feedback ($M = 510$ ms, $SD = 112$ ms) than with vibrotactile feedback ($M = 415$ ms, $SD = 104$ ms). Even though the difference was not statistically significant, the results indicated that vibrotactile feedback could potentially make gaze gestures faster also when stimulation is felt on the head.

In the studies described so far, vibrotactile feedback was used to confirm a gaze event (stroke or gesture). Alternatively, the feedback can be used to inform the user of an ongoing process or an upcoming event. Zhi (2014) studied the effect of vibrotactile feedback on dwell time progression in eye typing. The eye typing application had an on-screen keyboard that could be operated by dwell-select. Vibrotactile feedback was given on the participant's index finger. Three feedback conditions were selected to study whether added information of dwell time progression affects typing performance. In the *Ascending* feedback condition, the vibrotactile feedback faded in by increasing the amplitude of the feedback during the dwelling. In the *Warning* feedback condition, a short 50 ms notification was given to mark the start of the dwell duration. In the *No dwell* feedback condition no feedback of dwelling was given. In all conditions, the final selection was marked by a short (50 ms) vibrotactile feedback. The results of a user study (N=12) showed no benefits for the vibrotactile feedback on dwell progression in quantitative or qualitative measures. The only statistically significant result showed that the number of keystrokes per character increased in the *Ascending* condition compared to the *No dwell* condition, indicating that participants made more (erroneous) key activations when ascending vibrotactile feedback was given. One potential explanation for the increased error rate could be that the *Ascending* vibration made it hard to perceive the exact moment of selection. Earlier studies have shown that error rates may increase, if the feedback does not have a distinct point for selection (Majaranta et al., 2006). The main finding of this study was that vibrotactile feedback of dwelling provided no measurable benefits as evidenced

by the fact that participants performed equally well when given only the final
selection feedback in *No dwell* condition.

The possibility to use vibrotactile feedback as a "warning" was evaluated in a
study (N=12) where gaze was used to implement automatic scrolling of web pages
(Käki, Majaranta, Špakov, & Kangas, 2014). A short vibration was given to the user's
index finger to warn the reader that their gaze had entered an active scrolling area.
The results showed no statistically significant differences in task times between
conditions with and without vibrotactile feedback. However, post-experiment
interviews revealed an important design issue. Participants who understood the
purpose of the vibrotactile warning found it useful, while some had trouble
associating it with the scrolling. This was probably due to a too long delay of 700 ms
between the vibrotactile warning and the start of the scrolling. As described in the
next subsection, to be useful, the temporal distance (delay) between the gaze event
and the vibrotactile feedback should not be too long.

3.2. Temporal Limits between Gaze Events and Vibrotactile Feedback

Since the combination of gaze interaction and vibrotactile feedback had not been
studied before our research and some of the issues were not easily explored in the user
interface contexts described in the previous section, we also conducted experiments in
artificial laboratory tasks to find tolerable temporal limitations for giving vibrotactile
feedback for gaze events. The question of delay is of practical importance because
there are many potential sources of system delay in gaze-operated interactive systems.
These include, for example, the eye tracker sample rate, eye tracker video processing,
data transmission from the tracker to the main CPU of the device including network
delay, processing time, transmission delay of the vibrotactile pulse, as well as the rise
time of the vibrotactile actuator. As discussed in Section 1.2, the neural processing
time of touch stimulation also varies depending on the distance between the
stimulated body location and the brain. However, these perceptual delays are short
and beyond the control of the user interface designer. Therefore, it is important to
focus on the system delay. Because eye movements are fast and frequent, the question
of delay is especially relevant in gaze based interaction.

In order to find the time limits for vibrotactile feedback on gazing events, we
designed an experiment (N=12) where the user had to find a target object among non-
target distractors based on vibrotactile feedback (Kangas, Rantala, Majaranta,
Isokoski, & Raisamo, 2014). Five boxes were shown on a computer display (see
Figure 5). Our motivation was to simulate a scenario where the user would see objects
that can be activated by gaze. For example, with a head-mounted gaze tracker, it
could be possible to interact with nearby networked devices such as light switches or
volume controls by looking at them. In the interface, one of the boxes gave
vibrotactile feedback when the user pointed at it by moving her gaze to the box. The
task was to indicate this target by pressing the spacebar key on the keyboard as soon
as feedback was felt. In a real use scenario, such feedback could be useful in scanning
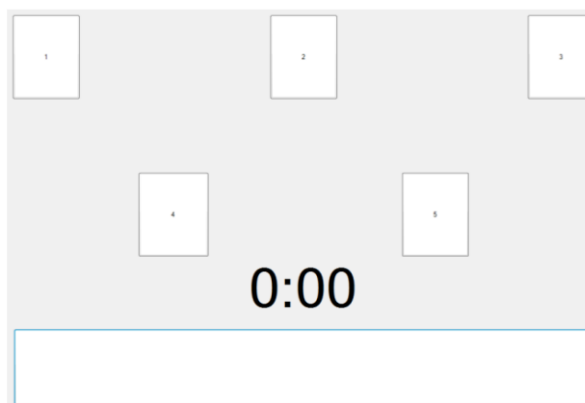the environment for objects that can be accessed by using gaze.

FIGURE 5. Experimental interface consisting of five target boxes (Kangas et al.,
2014d). One of the boxes was the target box that triggered vibrotactile feedback felt
via a handheld actuator. © 2014 Association for Computing Machinery, Inc.
Reprinted by permission.

The participant held a vibrotactile actuator between the thumb and the index
finger in the hand that did not operate the keyboard. We varied the delay between the
gaze event and the vibrotactile feedback to find out how much time we have for
giving the feedback before the (too long) delay starts to affect the user's behavior.
The results showed that there was a significant increase in the error rates (i.e. the gaze
+ button press did not match the target) with delays around 250 to 350 ms (see Figure
6). This falls within the typical range of average fixation durations reported in
literature (see e.g. Rayner, 1998; Holmqvist et al., 2011, pp. 381-382). Also, the error
rates started growing prominently around 250 ms from the start of the fixation. With a
longer delay, the gaze may have already moved away from one target to another,
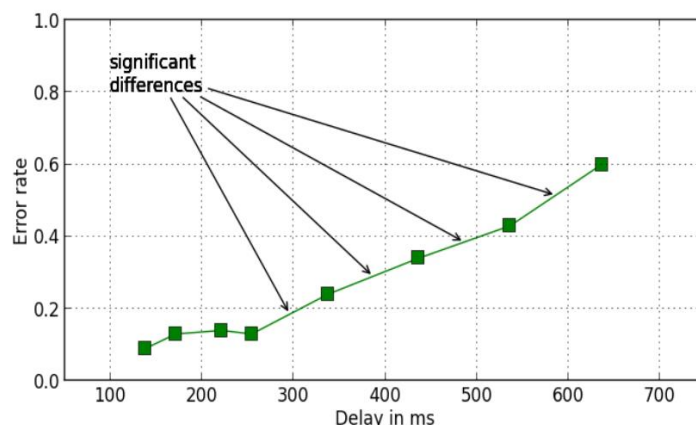causing incorrect associations and confusion.

FIGURE 6. Mean error rates for different delays in Kangas et al. (2014d). © 2014
Association for Computing Machinery, Inc. Reprinted by permission.

In a subsequent study (N=16), we investigated acceptable delays with gaze
gestures that exploit saccadic eye movements (Kangas et al., 2014a). Because the
movements can be quite fast, we could not rely on the results from the previous study
that was based on longer fixations at the target boxes. The setup and the task were
similar to the earlier study (Kangas et al., 2014b), where participants used gaze
gestures to browse a list of names and to make a simulated call (see Figure 3). We

varied the delay of the feedback from 100 to 450 ms with 50 ms steps. These delay
times affected both the vibrotactile feedback (felt via the phone, given for the
outgoing stroke) and the visual feedback (resulting action on the screen after the
whole gesture was made). It should be noted that participants did not have to wait for
the feedback but could continue to the next stroke or gesture immediately.

The results showed that the acceptable delay is shorter when gaze gestures are
used for interaction (see Figure 7). Task completion times were significantly lower
when the delay was 150 ms or less ($p < 0.05$). We also analyzed the average time of
completing a single gesture. The results indicated that the shortest delay of 100 ms
resulted in significantly faster gesture times compared to all other delays ($p < 0.05$).
Thus, in the case of gaze gestures, the shorter the delay the better. However, 200 ms
seemed to be the practical upper limit for smooth interaction. If the delay increased
above it, the use of gaze gestures became notably more difficult. This was also noted
by the participants in their subjective evaluations. Interestingly, Figure 7 shows that
the completion times decreased when the delay was longer than 300 ms. A possible
explanation for this finding is that with very long delays participants proceeded to the
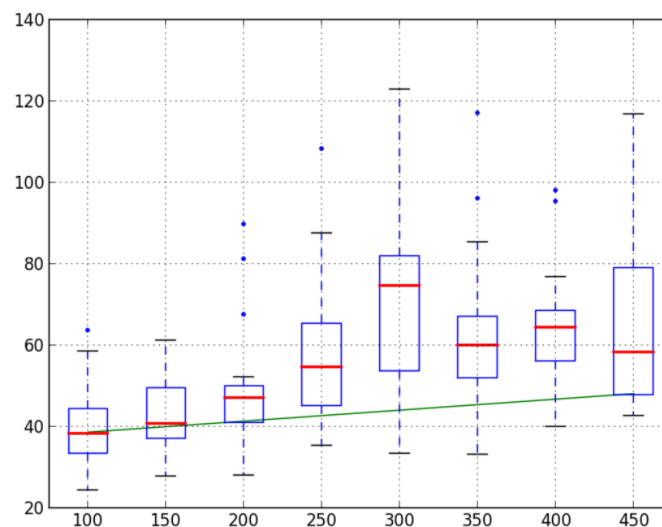next gaze gesture without waiting for the feedback.

FIGURE 7. Block completion times in seconds for different delays in Kangas et
al. (2014a). © Springer-Verlag Berlin Heidelberg 2014. Reprinted with permission of
Springer.

3.3. Effects of Feedback Location and Spatial Setup

The spatial location of vibrotactile feedback on gaze events requires consideration
because it is not clear where stimulation should be presented. In conventional haptic
input devices the action (e.g. pressing a button) and the feedback (haptic sensations
from the pressing action) are co-located. In gaze-operated interfaces the acting body
part is the eye, but it is unlikely that the haptic feedback should be directed at the eye.
In fact, considerable discomfort would be expected. Therefore, feedback that is not
co-located seems a more suitable design option in this context. To explore the
possible locations, we gave vibrotactile feedback to the palm of the *hand* (Kangas et
al., 2014a; Kangas et al., 2014b; Köpsel et al., 2016), *fingers* (Kangas et al., 2014d;
Käki et al., 2014; Majaranta et al., 2016; Zhi, 2014), *wrist* (Akkil et al., 2015;

Majaranta et al., 2016), *back* (Špakov, Rantala, & Isokoski, 2015), and *head* (Kangas
et al., 2014c; Rantala, Kangas, Akkil, Isokoski, & Raisamo, 2014; Rantala et al.,
2015; Špakov et al., 2015). We investigated how people experience and react to
vibrotactile feedback presented to different body locations.

We compared head and neck to back in a user study (N=12) where vibrotactile
stimulation was used as a cue for gaze direction (Špakov et al., 2015). Such cuing
could be used to indicate the direction of nearby gaze-interactive objects. This could
make it possible to locate and interact with objects when using a head-mounted gaze
tracker without a visual display. Also, providing cues for gaze direction could be used
for navigation assistance. Head and/or neck are potential areas because people often
wear eye glasses or necklaces, and head has been identified as a potential body
location for tactile navigation systems (Myles & Kalb, 2013). The back area could be
used, for example, with a car seat or in a wheelchair to support gaze interaction.

We ran pilot tests to find optimal locations where the stimulation would not be
unpleasant or too dominant, and at the same time would not get too weak because of
hair. The final locations are shown in Figure 8. By placing two actuators on the head
and two on the neck area, we could indicate directions using a similar four-actuator
configuration to that on the back. The design of the stimulation patterns was based on
the assumption that users' gaze direction is always nearly orthogonal to the plane of
the rectangle with actuators located in corners. This way, the stimulation should be
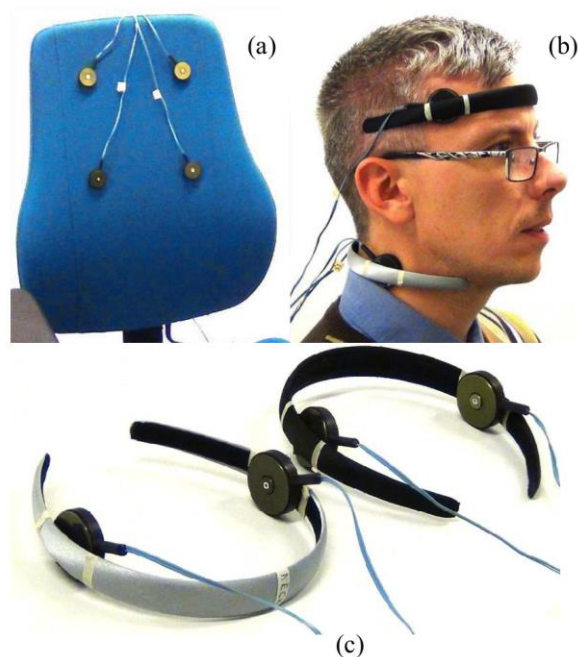perceived as occurring on the periphery of the visual field.

FIGURE 8. Locations of vibrotactile actuators on the back (a), neck and head (b)
in the gaze cueing study. Actuators were attached to two bands for the neck+head
condition (c). © 2015 IEEE. Reprinted, with permission, from Špakov et al. (2015).

We decided to use eight directions (see Figure 9) for cueing the gaze because this
was considered sufficient for our application scenarios of finding nearby gaze-
interactive objects and providing navigation assistance. To be able to do this with four
actuators, we developed encoding schemes for directions that are located in between

the physical actuators. Pilot tests led to two possible modes of vibrotactile stimulation: "parallel" where two actuators are stimulated at the same time and "sequential" where the actuators are stimulated sequentially, with a 50 ms interval. The participant's task was to move their gaze from the home position on the center to the direction of the vibrotactile cue. The results showed no statistically significant differences in selection error or reaction latency between the body locations. However, the participants made less errors when following cues given with sequential mode where only one point on the skin was stimulated at a time. The higher number of errors in parallel mode could be caused by the fact that when simultaneously stimulating multiple adjacent points on the skin, people typically sense a single focal sensation instead of multiple discrete points (Chen, Friedman, & Roe, 2003). It is possible that the mislocalized sensation did not communicate directional information as effectively as multiple discrete sensations in the sequential mode.
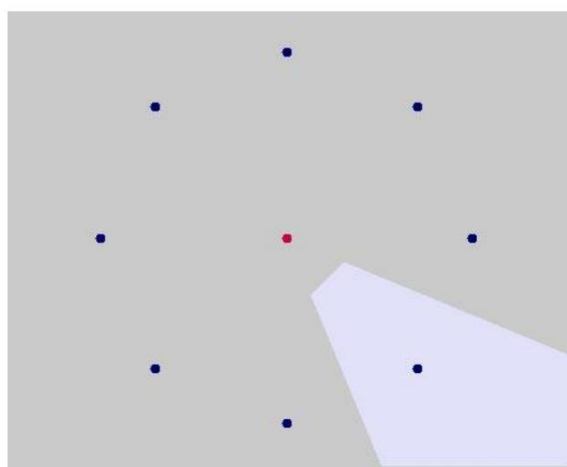
FIGURE 9. Experimental software showing the blue targets and red home-box in the gaze cueing study. © 2015 IEEE. Reprinted, with permission, from Špakov et al. (2015).

We compared the head and fingers as feedback locations in an experiment (N=10) where participants performed gaze gestures by looking at a tablet computer (Kangas et al., 2016b). With tablets, a convenient way to give vibrotactile feedback is to stimulate the back of the device that users naturally touch with their fingers (see dashed circles in Figure 10). If smart glasses (that perhaps include also eye tracking) are used, the glass frame provides possible locations for vibrotactile actuators (see solid circles in Figure 10). Because both the tablet and glasses can be equipped with multiple actuators, they are suitable form factors for providing spatially congruent feedback of simple left-right gaze gestures.
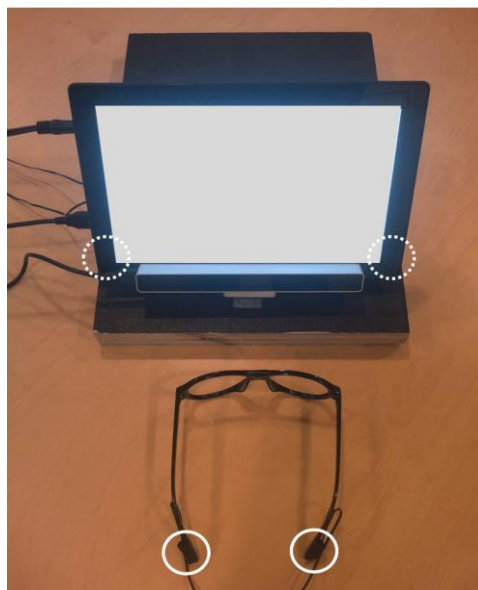
FIGURE 10. Locations of vibrotactile actuators on glasses (solid circles) and on tablet (dashed circles) in a study comparing vibrotactile feedback felt on fingers and head (Kangas et al., 2016b). © Springer International Publishing Switzerland 2016. Reprinted with permission of Springer.

We envision that gaze gestures could be used on a tablet, for example, to switch between active programs or scroll a web page by moving the gaze to the side of the display and back. For experimental purposes, we reduced this to a task where participants made gaze gestures from the center of the screen to the direction indicated by an arrow symbol ('<' or '>'). See Figure 11 for illustration of the required gesture. In addition to comparing vibrotactile feedback on the fingers and head, we were also interested in whether spatially congruent feedback affects user performance in gaze gesture tasks. Thus, we had five different feedback location conditions: *Spatial Head*, *Spatial Fingers*, *Non-Spatial Head*, *Non-Spatial Fingers*, and *No Feedback*. In the spatial conditions, feedback was given using one actuator so that it followed the gesture direction, and in non-spatial conditions using both actuators simultaneously regardless of the gesture direction. We also varied the number of successive gestures so that in a single trial the gesture was performed either once, twice, or three times. This was done to evaluate how the gesture complexity and vibrotactile feedback interact. For example, does feedback make the use faster if only one gesture is performed?
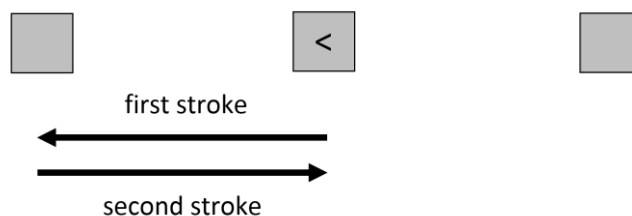

FIGURE 11. An illustration of the interface used in Kangas et al. (2016b). The middle box showed the gesture direction and number of repetitions with one to three arrows. Only the boxes were visible on the display. The texts and arrows illustrate the

gesture. © Springer International Publishing Switzerland 2016. Reprinted with
permission of Springer.

The results showed no statistically significant differences in task completion times
between the feedback locations when all gestures were included in analysis (see
Figure 12). However, when analyzing only gestures that were performed three times
in succession, task completion times in *No Feedback* condition were significantly
longer than in *Spatial Head* and *Non-Spatial Head* conditions. One possible
explanation for this could be that the efficiency benefit provided by feedback
accumulated over time when performing multiple gestures. Eight out of ten
participants preferred spatial over non-spatial feedback but there was no clear
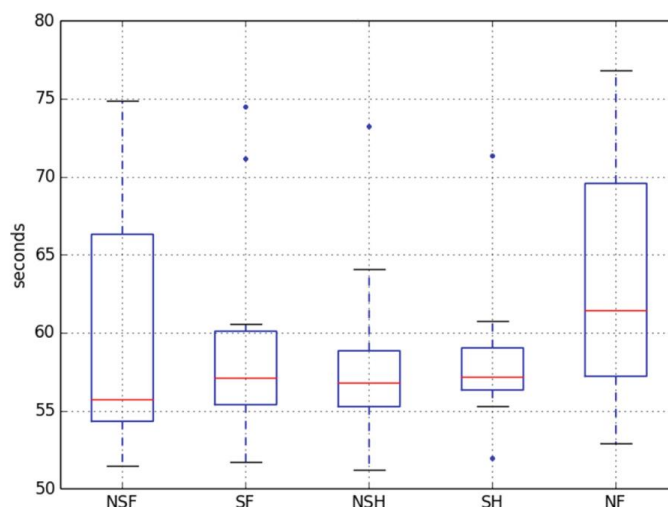preference between finger and head stimulation.

FIGURE 12. Task completion times of 12 trials by different feedback conditions
in Kangas et al. (2016b). The gesture was performed either once, twice, or three
times. © Springer International Publishing Switzerland 2016. Reprinted with
permission of Springer.

The results from our other studies without direct comparisons between multiple
body locations also support the finding that the body location for the vibrotactile
feedback is not as important as other design issues such as the properties of the
feedback signal itself. All the feedback locations that we tested in gaze interaction –
hand, fingers, wrist, back, and head – were generally feasible. Furthermore, we
compared legs and head in a study without gaze tracking. Nukarinen, Rantala, Farooq,
and Raisamo (2015) presented vibrotactile feedback using a seat and eye glass frames
in an experiment where directional cues were given to a driver to assist navigation.
Vibrotactile cuing was found useful and equally effective on both body sites in this
setup which required visual and auditory attention. Overall, we consider the finding
that many body sites are applicable in gaze interaction positive since this allows
designers to choose feedback location freely depending on the application and the
context (e.g. mobile/wearable).

In addition to the body location, also the spatial arrangement of the vibrotactile
actuators deserves some thought. As discussed above, some body locations such as
the head and neck may be more sensitive to vibrotactile stimulation (Špakov et al.,
2015). Furthermore, both the temporal as well as spatial distance of the vibrotactile

actuators may also affect how well the feedback is perceived (Köpsel et al., 2016; Špakov et al., 2015). Finally, as observed in the experiments above, not only can spatial vibrotactile stimulation be used to cue directions (Nukarinen et al., 2015; Špakov et al., 2015) but participants also prefer spatially congruent feedback in cases where the task itself includes spatial actions, such as gaze gestures (Kangas et al., 2016b).

3.4. Vibrotactile Feedback in Comparison to Other Modalities

In the sections above, we have shown how vibrotactile feedback can support gaze interaction. We were also interested in learning how vibrotactile feedback performs in comparison to other feedback modalities. As discussed earlier, visual and auditory feedback have been found to improve user satisfaction and performance in gaze interaction (e.g. Majaranta et al., 2006). In our studies, we wanted to measure how well vibrotactile feedback fared in comparison to these two more commonly utilized feedback modalities.

We started by comparing vibrotactile feedback alone to no feedback (Kangas et al., 2014b). When participants browsed a contact list and made simulated calls with gaze gestures, none of them preferred the condition with no vibrotactile feedback. If no vibrotactile feedback was given, the task was experienced as uncomfortable and more difficult. Similarly, in the study where participants controlled a bus timetable with gaze gestures (Kangas et al., 2014c), 9 out of 12 participants preferred the condition where vibrotactile feedback was given on the glass frames over the no vibrotactile feedback condition. Even though some participants disliked the vibrotactile feedback when it was not timed correctly or was presented when not intending to perform a gesture, they generally found that the feedback helped in confirming the action and made them feel more in control. These findings are in line with earlier work showing that vibrotactile feedback can provide benefits compared to no feedback. For example, vibrotactile feedback improved typing speeds on mobile phones when using a virtual keyboard (Hoggan et al., 2008).

Vibrotactile and visual feedback were compared in an experiment (N=12) with a gaze-aware smart wristwatch (Akkil et al., 2015). The watch knew when it was looked and could be operated by gaze gestures. Visual or vibrotactile feedback was used to confirm that the gaze was recognized and the watch was ready for interaction. Two experimental setups using applications with varying complexity were tested: a simple notification task where only a single gesture was required, and a menu navigation task that required multiple gestures to accomplish the task. Similarly to previous studies, also here we used off-screen gestures that started by looking at the screen. Visual feedback was given by highlighting the clock bar with a purple color for 500 ms as soon as the gaze was detected. Vibrotactile feedback was given as a 30 ms short tap using an actuator mounted on a layer of foam between the watch and the skin. The results showed no differences in feedback type preferences in the menu navigation task. However, in the notification task, 9 out of 12 participants preferred the vibrotactile feedback (2 preferred visual and 1 found them equal). Vibrotactile feedback was found clearer and more noticeable. Some considered visual feedback more appropriate because the user is already looking at the screen.

In another study, we compared vibrotactile feedback with auditory and visual feedback in an eye typing task where the feedback was used to confirm a selection (Majaranta et al., 2016). In the first, exploratory experiment (N=12), auditory feedback was given as a short click sound, vibrotactile feedback as a 100 ms vibration felt on the wrist, and visual feedback as a 100 ms "flash" by changing the background of the focused key. The results showed no statistically significant differences between the feedback modalities. However, some participants commented that they did not really need the confirmation on selection because the visual feedback on dwell time, shown as an animated circle that closed when the dwell time was over, was enough as a confirmation. Also, some participants did not like the vibrotactile feedback on their wrist but commented that a finger would be a more suitable location. Thus, we ran another experiment (N=12) with no hints for dwell time progression; we only highlighted the background of the key on fixation to indicate where the focus is. Vibrotactile feedback was given on the index finger instead of the wrist. We also modified the vibrotactile feedback so that it better resembled the auditory 'click' and felt more like a "tap" instead of a "buzz": the vibration started sharply with maximum amplitude and faded quickly. This time, vibrotactile feedback performed similarly to auditory feedback both in performance and subjective metrics. Visual feedback resulted in significantly slower typing speeds, higher error rates, and worse subjective ratings. Participants commented that auditory and vibrotactile feedback were easier to perceive; they could be heard or felt even if one blinked or had already started moving the eye to the next key.

The vibrotactile, auditory, and visual modalities were compared also in an experiment (N=16) studying calibration free smooth pursuit interaction (Kangas et al., 2016a). We developed widgets that could be useful for adjusting continuous values like sound volume or light level. We wanted to know how the feedback modality affects the interaction with those widgets. The experimental task was to adjust a level of gray color to approximately match the target color by looking at one of two moving targets. One target decreased the tone (lowest level 0, black) and another increased it (highest level 255, white). Feedback was given for every eighth adjustment step to confirm that the system is recognizing the gaze input and "things happen". Auditory feedback was a short 12 ms click and vibrotactile feedback a 20 ms 'tap' presented with glasses (see Figure 4). For visual feedback, we alternated symbols 'x' and '+' shown on the moving targets. We also had a mode with no other feedback than the final change of color for comparison. The results showed no statistically significant performance differences between the modalities but there was a clear user preference for vibrotactile and auditory feedback.

Finally, we ran a study (N=12) where auditory, vibrotactile, and visual feedback were compared in performing gestures either by gaze or by hand (Köpsel et al., 2016). The task was to enter numbers by using gestures composed of four stroke elements (up, down, left, and right). The feedback was spatially congruent with the gesture. For auditory feedback, we placed speakers in the laboratory so that in relation to the participant, they were above, below, on the left, or on the right side of the participant's head. Vibrotactile feedback was given on the palm of the hand using a

self-built pillow with four vibrotactile actuators (see Figure 13). Note that in the hand-gesture mode gestures were made using the hand that did not receive vibrotactile feedback. Visual feedback was shown on a computer screen that displayed four small flashing rectangles to resemble the directions. The rectangles were located approximately 6 cm away from the centre of the screen in the direction of the gesture. Since each gesture had four strokes, the feedback was given four times for each number.

FIGURE 13. Pillow with four actuators used in the study by Köpsel et al. (2016).

The results showed no statistically significant differences in performance between the feedback modalities. However, there were some differences between gestures made by hand or by gaze. Gaze was faster as expected but the hand input had lower error rate. It was also observed that using hand, participants could adjust the stroke based on the feedback; they got the feedback while moving the hand and could react and adjust the length of the stroke based on the feedback. Eye movements, on the other hand, are fast ballistic movements. Thus, in practice the feedback was received after the stroke had already been made. Furthermore, participants told they only noted there was some vibrotactile feedback when performing gestures but could not really differentiate the four directions on their palm. Interestingly, this was possible in the beginning of the experiment as we made sure that all participants could perceive the feedback from each actuator and understood its direction. The actuator placement should be revised to ensure that spatial differences in feedback could be felt also when attending to a primary task. In addition, considering future work, it would be interesting to study the effects of giving a more co-located feedback on the same hand that performs the gestures, e.g. by integrating the haptic actuators to a glove. We also ran a follow-up experiment (N=12) to compare the effects of feedback given after the whole gesture (with no feedback on strokes) but found no differences between feedback modes in it either (Köpsel et al., 2016).

4. DISCUSSION AND DESIGN GUIDELINES

We will present our main findings as design guidelines that are followed by a discussion relevant to each guideline. We conclude with general discussion.

1. *Consider vibrotactile feedback especially if the interaction involves no visual displays or only small displays that are difficult to visually observe when interacting by gaze.*

The research surveyed in this paper showed that the way vibrotactile feedback should be used in gaze based interaction depends on the gaze input type and use scenario. Because dwell-select and smooth pursuit interaction both require focusing the gaze on the target, this enables visual feedback on the object. In eye typing, if animated visual feedback on the key selection was available, participants relied on it and vibrotactile feedback on selection did not bring any additional benefits (Majaranta et al., 2016). Similarly, vibrotactile feedback did not provide performance benefits when using visual widgets based on smooth pursuit (Kangas et al., 2016a). However, in many scenarios visual feedback is difficult to provide or notice. When interacting with small displays (e.g. smartwatches) or with displays containing large amount of information, perceiving subtle visual feedback to gaze events may be hard. Other potential use cases for vibrotactile feedback are scenarios that do not involve a display (e.g. using gaze to interact with real-world objects) or scenarios that use gaze gestures or off-screen interactions similarly to Kangas et al. (2014b). We demonstrated faster use of off-screen gaze gestures with a mobile phone when vibrotactile feedback was available. Visual feedback may be problematic even with on-screen gestures, if users start to slow down gesturing in order to see the feedback on the strokes (Istance et al., 2010).

2. *Vibrotactile feedback is at least as good as auditory and visual feedback in task performance and user satisfaction.*

Vibrotactile feedback performed equally to auditory and visual feedback in supporting gaze interaction. In fact, vibrotactile and auditory feedback received more positive subjective ratings than visual feedback in two studies (Kangas et al., 2016a; Majaranta et al., 2016), and in one study vibrotactile and auditory feedback resulted in faster task performance (Majaranta et al., 2016). These findings are in line with earlier research indicating similarities between vibrotactile and auditory modalities (Jokiniemi et al., 2008). In practice, the choice of feedback modality in gaze based interaction depends also on the task, context, and user preferences. If visual feedback is available and implemented in a way that it does not hinder user performance, no additional feedback may be needed. If both visual and vibrotactile information are available, vision usually dominates (Ernst & Banks, 2002). On the other hand, because vibrotactile and auditory feedback do not require visual attention, they "free" the eyes for gaze input. Vibrotactile feedback also provides privacy that other modalities lack: the feedback is felt only by the user wearing or holding the device. Auditory feedback is private if played back via headphones or an earbud, but this also blocks environmental audio such as sounds of traffic that are important in avoiding accidents. Vibrotactile feedback, on the other hand, can be presented to different body locations without a negative effect on observing the environment. The actuated body location can also be changed to take into account environmental vibration. For example, when riding a bike on a bumpy road, perceiving vibrotactile feedback could be easier on the head than on the wrist. In general, it is best to provide users the

flexibility to choose the optimal feedback modality based on their personal and contextual preferences if multiple modalities are feasible.

> 3. *Delay between gaze event and the related vibrotactile feedback should not exceed 250 milliseconds.*

The temporal delay between a gaze event and vibrotactile feedback was shown to affect user performance. This is in agreement with the results of Viciana-Abad, Reyes-Lecuona, Poyade, and Escolano (2011) who demonstrated the importance of coherent sensory feedback in haptic force feedback systems. In their study, a feedback delay of 200 ms had a negative effect on pushing virtual buttons. Our results indicated that, as a general rule, the delay should not exceed 250 ms to ensure that users can associate the gaze event with its feedback (Kangas et al., 2014d). As the 250 ms limit falls within the typical range of average fixation durations (Rayner, 1998; Holmqvist et al., 2011), longer delays increase the risk of the user not perceiving feedback before the gaze point has already moved on to a different area of the visual field. However, it is important to note that the acceptable delay depends also on the task and eye movement type. Dwell-select based on fixations can tolerate longer delays, assuming that the used dwell time duration is at least 250 ms. On the contrary, with gaze gestures that require more rapid eye movements a negative effect on performance was observed already at 100-200 ms (Kangas et al., 2014a). While our experiments on the temporal aspects of vibrotactile feedback focused mainly on measuring user performance, it is also possible that too long delays add to user frustration if the link between an action and feedback does not appear natural. Due to inherent latency caused by the different components used in our experimental settings, we could not measure the effect of delays shorter than 100 ms. Pushing the delay close to zero could potentially lead to a situation where the user perceives the feedback already before perceiving visual information at gaze point. Even though such a situation remains largely theoretical, it does offer a possibility for future studies.

> 4. *Distance between the eyes and body location of vibrotactile feedback can be varied without affecting user's capability to associate a gaze event and the feedback.*

There were no performance differences in users' task completion times between gaze interaction tasks where vibrotactile feedback was given to fingers, head, and back. We hypothesized that the times required for neural processing of touch stimulation on different areas of the human body could introduce delays up to 30 ms to the interaction (Macefield et al., 1989). There are at least two possible reasons for why we did not find any differences when varying the spatial location of vibrotactile feedback. First, we did not study stimulation of the toes which supposedly should cause the longest delays. Second, as discussed above, the (to some extent random) delays in our experimental settings might have masked the possible effects caused by the feedback location. However, our findings did indicate that participants appreciated the congruence given by short spatial distance (achieved e.g. with glasses) as long as the stimulation presented relatively close to the eyes was perceived as pleasant. If the stimulation is too strong or startling, it could possibly lead to involuntary eye movements that affect user performance. Given that we did not observe such negative effects of head stimulation in our comparison studies (Kangas et al., 2016b; Špakov et al., 2015), the chosen amplitude and frequency levels seemed to be generally feasible. The fact that head stimulation received comparable subjective ratings to finger and

back stimulation is a valuable finding because the head is considered a highly touch
sensitive body part that has not been widely utilized in earlier research.

5. *Maintain spatial congruence by making the vibrotactile feedback follow the
direction of the gaze point.*

In addition, our findings showed that it is better to provide feedback that is
spatially congruent with gaze movement if vibrotactile feedback is presented to
multiple locations. For example, when a gaze gesture stroke is performed towards the
left hemifield, the feedback should ideally be given on the left side of the body.
Participants' comments in Kangas et al. (2016b) indicated that with spatially
congruent feedback it was easier to confirm that the stroke was performed
successfully even though analysis of task performance showed no statistically
significant differences between spatial and non-spatial conditions. The observed user
preference for spatially congruent feedback could be related to the role of visual and
tactile stimulation in spatial attention. Driver and Spence (1998) argue that tactile
stimulation of hand can draw visual attention to the location of the stimulation.
Following this notion, the location of vibrotactile feedback should be congruent with
the location of gaze point to avoid contradictory information between the senses. In
one of our studies (Köpsel et al., 2016), we presented vibrotactile feedback of gaze
gestures always on the palm of the left hand with four actuators while participants
performed four directional gaze gestures. As a result, participants had trouble
associating the gaze gestures with the stimulation they felt on their palm. It is possible
that the actuators were located too close to each other on the palm to be easily
distinguishable. Improved configurations could include increasing the space between
actuators on the palm, providing the feedback on the hand that executes the gesture
(e.g. via haptic glove), or even distributing the four actuators across the body (e.g.,
left and right hand) so that they would better match the directions of gaze movement.

6. *Pay attention to the amplitude, duration, and rhythm of vibrotactile
stimulation because they can influence the effectiveness and user acceptance
of interaction.*

The technical implementation of the feedback is a key factor to the success of
vibrotactile feedback. For example, in some tasks a discrete "tap" might be preferred
over a vibrating "buzz" (Majaranta et al., 2016). Stimulation amplitude should be set
based on the sensitivity of the body location and personal preferences to avoid
unpleasant feedback. For example, the neck is considered a sensitive area and fairly
low stimulation amplitudes should therefore be selected. If several vibrotactile
actuators are used together, sequential stimulation based on rhythmic variations is
easier to localize than stimulation of multiple locations at the same time (Chen et al.,
2003; Špakov et al., 2015). In terms of feedback duration, the optimal value depends
partly on the used feedback technology. We chose linear electromagnetic actuators for
our studies because they are easy to control, readily available and also simple to
attach to different wearable devices. With these actuators, we found 20-30 ms
stimulations to be suitable with gaze gestures that require short pulses if feedback
needs to be given and perceived already during the gesture. In dwell-based
interaction, longer feedback durations (e.g. 100 ms) may be tolerable (Majaranta et
al., 2016). Longer durations may also be needed in contexts where short taps may go
unnoticed.

7. *Aim for a balance where necessary information is given via vibrotactile stimulation without overburdening the feedback channel.*

In general, it would be important to use vibrotactile feedback in situations where it can provide objective or subjective benefits, but at the same time avoid overburdening the channel. For example, with two-stroke gaze gestures, it may be enough to present feedback of the first stroke to gain performance benefits. Furthermore, because too frequent feedback can be perceived as annoying, it could also take attention away from the main task (Witt, Lawo, & Drugge, 2008) or cause sensorial overload (Popescu, Burdea, & Trefftz, 2002).

8. *Out of multiple simultaneous vibrotactile stimuli the strongest dominates.*

In those experiments that involved multiple simultaneous vibrotactile stimuli (Rantala et al., 2014; Špakov et al., 2015), we noticed that the amplitude of the vibration was critical to the sensation. For example, when all three actuators in the eye glass frame were vibrating at the same time, it was very difficult to tell whether the front actuator was running or not. The actuators behind the ears dominated the sensation. If one of them was not in good contact with the skin it could feel as if only one actuator was on. This tendency to focus attention to the strongest stimulus makes feedback schemes that rely on multiple simultaneous signals very difficult to perceive in practice. It is better to run the actuators in a sequence (Špakov et al., 2015). A further benefit is that the sequences can encode additional information, for example directional information, conveniently.

Our aim in this paper has been to inform the design of interfaces that combine gaze interaction with vibrotactile feedback. Even with good knowledge of the characteristics of the feedback modalities, it is always advisable to run user trials. There may be unexpected effects, such as effects moderated by the task type and workload (Burke et al., 2006). For example, even though we found vibrotactile feedback to improve user performance and satisfaction when using gaze gestures to control a mobile phone (Kangas et al., 2014b), in another study (Köpsel et al., 2016), the added vibrotactile feedback did not bring any benefits. We believe the inconsistent results are explained by the difference in the task difficulty as well as by the way the vibrotactile feedback was implemented. In the former, the gaze gestures were simpler and the feedback only confirmed the strokes. In the latter, the gestures represented numbers and included more strokes, and the feedback also included information about the stroke direction. Thus, it seems that in gaze interaction vibrotactile feedback should ideally be used for communicating simple rather than abstract information.

There were a few limitations that need to be taken into account while generalizing the results and insights presented in this paper. First, all the studies used vibrotactile stimulation as the haptic feedback modality. Further research is required to understand how our results will apply to other types of tactile stimulation, such as skin stretch or pressure-based stimulations. Second, all the studies were of short durations and they did not provide insights from a long-term use perspective. It would be interesting to investigate if, for example, experienced gaze gesture users learn to maintain the performance improvement initially provided by vibrotactile feedback even when feedback is turned off. Third, our studies were conducted in controlled laboratory environments. In a real-world scenario, mobility and other vibrations in the environment could influence both gaze tracking accuracy and tactile perception.

Further research is required to understand user preference and acceptance of the combination of gaze and vibrotactile modalities in such situations.

We expect that the work presented in this review article will become increasingly relevant in the near future due to active interest in pervasive eye tracking and mobile eye-based interaction (Mardanbegi, Khamis, Jalaliniya, & Majaranta, 2016), eyewear computing (Bulling & Kunze, 2016), head-mounted displays (e.g., Microsoft HoloLens[5]), and gaze input in general (Hansen et al., 2015). At the same time, it is important to highlight that wider adoption of eyewear and smart glasses requires that social aspects, privacy, safety, reliability, and eye fatigue are properly addressed. Participants may not want to use gaze input or visual feedback in all use contexts, and therefore complementary modalities such as haptics could be used to convey information (Akkil et al., 2016). One potential application area for combining haptic and gaze interaction is navigation. Giannopoulos, Kiefer, and Raubal (2015) give an example of how pedestrian navigation could benefit from the combined use of eye tracking and vibrotactile feedback. In their application, a vibrating smartphone notified users when looking at the correct route in a virtual reality environment. This allowed users to verify where to navigate without having to look at their smart phone. Further, Deng, Kirkby, Chang, and Zhang (2014) foresee many opportunities for improved information presentation and better user engagement in serious games by utilizing gaze tracking and haptics. Combining haptics and gaze could also be beneficial in interpersonal communication. Qiu, Rauterberg, and Hu (2016) describe a prototype that detected gazing at a (blindfolded) discussion partner and translated this to vibrotactile stimulation felt by the partner. Their goal was to develop an interface that could enhance the level of engagement in face-to-face communication for visually impaired people. In conclusion, the number of recent studies on this topic suggests that haptic gaze interaction is an emerging field of HCI.

---

[5] https://www.microsoft.com/microsoft-hololens/en-us

ACKNOWLEDGMENTS

NOTES

REFERENCES

Akkil, D., Kangas, J., Rantala, J., Isokoski, P., Špakov, O., & Raisamo, R. (2015). Glance awareness and gaze interaction in smartwatches. *Proceedings of the CHI 2015 Conference on Human Factors in Computer Systems*. New York: ACM.

Akkil, D., Lucero, A., Kangas, J., Jokela, T., Salmimaa, M., & Raisamo, R. (2016). User expectations of everyday gaze interaction on smartglasses. *Proceedings of the NordiCHI 2015 Nordic Conference on Human-Computer Interaction* (accepted).

Bednarik, R., Gowases, T., & Tukiainen, M. (2009). Gaze interaction enhances problem solving: Effects of dwell-time based, gaze-augmented, and mouse interaction on problem-solving strategies and user experience. *Journal of Eye Movement Research, 3*, 1-10.

Brown, L. M. (2007). *Tactons: Structured vibrotactile messages for non-visual information display.* PhD Thesis, Department of Computing Science, University of Glasgow.

Brown, L. M., & Kaaresoja, T. (2006). Feel who's talking: Using tactons for mobile phone alerts. *Proceedings of the CHI 2006 Conference on Human Factors in Computer Systems.* New York: ACM.

Bulling, A., & Kunze, K. (2016). Eyewear computers for human-computer interaction. *interactions*, *23*, 70-73.

Bulling, A., Roggen, D., & Tröster, G. (2009). Wearable EOG goggles: eye-based interaction in everyday environments. *Proceedings of the CHI 2009 Conference on Human Factors in Computer Systems.* New York: ACM.

Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R., Coovert, M. D., Elliot, L. R, & Redden, E. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis. *Proceedings of the ICMI 2006 Conference on Multimodal Interaction.* New York: ACM.

Chen, L. M., Friedman, R. M., & Roe, A. W. (2003). Optical imaging of a tactile illusion in area 3b of the primary somatosensory cortex. *Science, 302*, 881-885.

Choi, S., & Kuchenbecker, K. J. (2013). Vibrotactile display: Perception, technology, and applications. *Proceedings of the IEEE, 101*, 2093-2104.

Cholewiak, R. W., & Collins, A. A. (2003). Vibrotactile localization on the arm: Effects of place, space, and age. *Perception & Psychophysics*, *65*, 1058-1077.

Deng, S., Kirkby, J. A., Chang, J., & Zhang, J. J. (2014). Multimodality with eye tracking and haptics: A new horizon for serious games? *International Journal of Serious Games, 1*, 17-34.

Drewes, H., De Luca, A., & Schmidt, A. (2007). Eye-gaze interaction for mobile phones. *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. New York: ACM.

Drewes, H., & Schmidt, A. (2007). Interacting with the computer using gaze gestures. *Proceedings of the INTERACT 2007 Conference on Human-Computer Interaction*. Berlin: Springer Heidelberg.

Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society B: Biological Sciences, 353*, 1319-1331.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice.* London: Springer-Verlag.

Dugard, P. (2014). Randomization tests: A new gold standard? *Journal of Contextual Behavioral Science, 3*, 65-68.

Dybdal, M. L., San Agustin, J., & Hansen, J. P. (2012). Gaze input for mobile devices by dwell and gestures. *Proceedings of the ETRA 2012 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Enriquez, M., & MacLean, K. (2008). The role of choice in longitudinal recall of meaningful tactile signals. *Proceedings of the HAPTICS 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. Washington, DC: IEEE Computer Society.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429-433.

Esteves, A., Velloso, E., Bulling, A., & Gellersen, H. (2015). Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. *Proceedings of the UIST 2015 Conference on User Interface Software and Technology.* New York: ACM.

Gardner, E. P., Martin, J. H., & Jessell, T. M. (2000). The bodily senses. In E. Kandel, J. Schwartz, & T. Jessell (Eds.), *Principles of neural science* (4th ed.). New York, NY: McGraw-Hill.

Giannopoulos, I., Kiefer, P., & Raubal, M. (2015). GazeNav: gaze-based pedestrian navigation. *Proceedings of the MobileHCI 2015 Conference on Human-Computer Interaction with Mobile Devices and Services*. New York: ACM.

Gilchrist, I. D. (2011). Saccades. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 85-94). Oxford University Press.

Gobbetti, E., & Scateni, R. (1998). Virtual reality: Past, present, and future. *Studies in Health Technology and Informatics*, *58*, 3-20.

Goldstein, E. B. (1999). *Sensation & perception* (5th ed.). Pacific Grove, CA: Brooks/Cole Publishing Company

Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 478-500.

Hansen, J. P., Biermann, F., Møllenbach, E., Lund, H., San Agustin, J., & Sztuk, S. (2015). A gazewatch prototype. *Proceedings of the MobileHCI 2015 Conference on Human-Computer Interaction with Mobile Devices and Services*. New York: ACM.

Hansen, J. P., Lund, H., Biermann, F., Møllenbach, E., Sztuk, S., & Agustin, J., S. (2016). Wrist-worn pervasive gaze interaction. *Proceedings of the ETRA 2016 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Heikkilä, H., & Räihä, K.-J. (2009). Speed and accuracy of gaze gestures. *Journal of Eye Movement Research, 3*, 1-14.

Hoggan, E., Brewster, S. A., & Johnston, J. (2008). Investigating the effectiveness of tactile feedback for mobile touchscreens. *Proceedings of the CHI 2008 Conference on Human Factors in Computer Systems.* New York: ACM.

Holland, C., & Komogortsev, O. (2012). Eye tracking on unmodified common tablets: Challenges and solutions. *Proceedings of the ETRA 2012 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* OUP Oxford.

Hoshi, T., Takahashi, M., Iwamoto, T., & Shinoda, H. (2010). Noncontact tactile display based on radiation pressure of airborne ultrasound. *IEEE Transactions on Haptics, 3,* 155-165.

Huckauf, A., & Urbina, M. H. (2008). On object selection in gaze controlled environments. *Journal of Eye Movement Research, 2*, 1-7.

Ishimaru, S., Kunze, K., Uema, Y., Kise, K., Inami, M., & Tanaka, K. (2014). Smarter eyewear: Using commercial EOG glasses for activity recognition. *Proceedings of the UbiComp 2014 Conference on Pervasive and Ubiquitous Computing.* New York: ACM.

ISO 9241-910 (2009). Ergonomics of Human-Computer Interaction – Part 910: Framework for Tactile and Haptic Interaction.

Isokoski, P. (2000). Text input methods for eye trackers using off-screen targets. *Proceedings of the ETRA 2000 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Isokoski, P., Joos, M., Špakov, O., & Martin, B. (2009). Gaze controlled games. *Universal Access in the Information Society, 8*, 323-337.

Istance, H., Hyrskykari, A., Immonen, L., Mansikkamaa, S., & Vickers, S. (2010). Designing gaze gestures for gaming: An investigation of performance. *Proceedings of the ETRA 2010 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Jacob, R. J. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems 9*, 152-169.

Jokiniemi, M., Raisamo, R., Lylykangas, J., & Surakka, V. (2008). Crossmodal rhythm perception. *Proceedings of the HAID 2008 Conference on Haptic and Audio Interaction Design.* Berlin: Springer-Verlag.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014a). Delayed haptic feedback to gaze gestures. *Proceedings of the EuroHaptics 2014 Conference.* Berlin: Springer-Verlag.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014b). Gaze gestures and haptic feedback in mobile devices. *Proceedings of the CHI 2014 Conference on Human Factors in Computer Systems.* New York: ACM.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014c). Using gaze gestures with haptic feedback on glasses. *Proceedings of the NordiCHI 2014 Nordic Conference on Human-Computer Interaction.* New York: ACM.

Kangas, J., Rantala, J., Majaranta, P., Isokoski, P., & Raisamo, R. (2014d). Haptic feedback to gaze events. *Proceedings of the ETRA 2014 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Kangas, J., Špakov, O., Isokoski, P., Akkil, D., Rantala, J., & Raisamo, R. (2016a). Feedback for smooth pursuit gaze tracking based control. *Proceedings of the AH 2016 Augmented Human International Conference.* New York: ACM.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2016b). Both fingers and head are acceptable in sensing tactile feedback of gaze

gestures. *Proceedings of the EuroHaptics 2016 Conference*. Berlin: Springer-Verlag.

Kern, T. A. (2009). Motivation and application of haptic systems. In T. A. Kern (Ed.), *Engineering Haptic Devices: A Beginner's Guide for Engineers* (pp. 5-18). Springer Berlin Heidelberg.

Käki, K., Majaranta, P., Špakov, O., & Kangas, J. (2014). Effects of haptic feedback on gaze based auto scrolling. *Proceedings of the NordiCHI 2014 Nordic Conference on Human-Computer Interaction.* New York: ACM.

Köpsel, A., Majaranta, P., Isokoski, P., & Huckauf, A. (2016). Effects of auditory, haptic and visual feedback on performing gestures by gaze or by hand. *Behavior & Information Technology, 35,* 1044-1062.

Larsson, L., Nyström, M., Andersson, R., & Stridh, M. (2015). Detection of fixations and smooth pursuit movements in high-speed eye-tracking data. *Biomedical Signal Processing and Control, 18*, 145-152.

Lee, J., Han, J., & Lee, G. (2015). Investigating the information transfer efficiency of a 3x3 watch-back tactile display. *Proceedings of the CHI 2015 Conference on Human Factors in Computer Systems.* New York: ACM.

Lee, S., & Starner, T. (2010). BuzzWear: Alert perception in wearable tactile displays on the wrist. *Proceedings of the CHI 2010 Conference on Human Factors in Computer Systems.* New York: ACM.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *The Oxford handbook of eye movements.* Oxford University Press.

Macefield, G., Gandevia, S. C., & Burke, D. (1989). Conduction velocities of muscle and cutaneous afferents in the upper and lower limbs of human subjects. *Brain, 112*, 1519-1532.

MacKenzie, I. S., & Ashtiani, B. (2011). BlinkWrite: Efficient text entry using eye blinks. *Universal Access in the Information Society, 10*, 69-80.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human-computer interaction. In S. Fairclough, & K. Gilleade, *Advances in Physiological Computing* (pp. 39-65). Springer.

Majaranta, P., Aoki, H., Donegan, M., Hansen, D. W., Hansen, J. P., Hyrskykari, A., & Räihä, K.-J. (2012). *Gaze interaction and applications of eye tracking: advances in assistive technologies.* IGI Global.

Majaranta, P., Bates, R., & Donegan, M. (2009). Eye tracking. In C. Stephanidis (Ed.), *The universal access handbook* (pp. 36:1-20), CRC Press.

Majaranta, P., Isokoski, P., Rantala, J., Špakov, O., Akkil, D., Kangas, J., & Raisamo, R. (2016). Haptic feedback in eye typing. *Journal of Eye Movement Research, 9*, 1-14.

Majaranta, P., MacKenzie, I. S., Aula, A., & Räihä, K. J. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society, 5*, 199-208.

Manabe, H., Fukumoto, M., & Yagi, T. (2015). Conductive rubber electrodes for earphone-based eye gesture input interface. *Personal and Ubiquitous Computing, 19*, 143-154.

Mardanbegi, D., Khamis, M., Jalaliniya, S., & Majaranta, P. (2016). The 6th international workshop on pervasive eye tracking and mobile eye-based interaction. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. New York: ACM.

Miluzzo, E., Wang, T., & Campbell, A. T. (2010). EyePhone: Activating mobile phones with your eyes. *Proceedings of the MobiHeld 2010 Workshop on*

Networking, Systems, and Applications on Mobile Handhelds. New York: ACM.

Mulvey, F. (2012). Eye anatomy, eye movements and vision. In P. Majaranta, A. Hirotaka, M. Donegan, D. Witzner, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 10-20). IGI Global.

Myles, K., & Kalb, J. T. (2010). Guidelines for head tactile communication. (Technical Report ARL-TR-5116). Aberdeen Proving Ground, MD: Army Research Laboratory.

Myles, K., & Kalb, J. T. (2013). Head tactile communication: Promising technology with the design of a head-mounted tactile display. *Ergonomics in Design: The Quarterly of Human Factors Applications, 21*, 4-8.

Møllenbach, E., Hansen, J. P., & Lillholm, M. (2013). Eye movements in gaze interaction. *Journal of Eye Movement Research, 6*, 1-15.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15*, 1-25.

Nukarinen, T., Rantala, J., Farooq, A., & Raisamo, R. (2015). Delivering directional haptic cues through eyeglasses and a seat. *Proceedings of the 2015 IEEE World Haptics Conference.* IEEE.

Pasquero, J., Luk, J., Levesque, V., Wang, Q., Hayward, V., & MacLean, K. (2007). Haptically enabled handheld information display with distributed tactile transducer. *IEEE Transactions on Multimedia*, *9*, 746–753.

Pfeiffer, T., & Renner, P. (2014). EyeSee3D: A low-cost approach for analyzing mobile 3D eye tracking data using computer vision and augmented reality technology. *Proceedings of the ETRA 2014 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Pietrzak, T., Crossan, A., Brewster, S. A., Martin, B., & Pecci, I. (2009). Creating usable pin array tactons for nonvisual information. *IEEE Transactions on Haptics, 2*, 61-72.

Popescu, G. V., Burdea, G. C., & Trefftz, H. (2002). Multimodal interaction modeling. In Kay M. Stanney (Ed). *Handbook of virtual environments: Design, implementation, and applications* (pp. 435-454). Mahwah, NJ: Lawrence Erlbaum Associates.

Porta, M., & Turina, M. (2008). Eye-S: A full-screen input modality for pure eye-based communication. *Proceedings of the ETRA 2008 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Poupyrev, I., & Maruyama, S. (2003). Tactile interfaces for small touch screens. *Proceedings of the UIST 2003 Symposium on User Interface Software and Technology*. New York: ACM.

Qiu, S., Rauterberg, M., & Hu, J. (2016). Tactile band: Accessing gaze signals from the sighted in face-to-face communication. *Proceedings of the TEI 2016 Conference on Tangible, Embedded, and Embodied Interaction.* New York: USA.

Rantala, J., Kangas, J., Akkil, D., Isokoski, P., & Raisamo, R. (2014). Glasses with haptic feedback of gaze gestures. *Proceedings of the CHI 2014 Conference on Human Factors in Computer Systems.* New York: ACM.

Rantala, J., Kangas, J., Isokoski, P., Akkil, D., Špakov, O., & Raisamo, R. (2015). Haptic feedback of gaze gestures with glasses: Location accuracy and

effectiveness. *Proceedings of the UbiComp 2015 Conference on Pervasive and Ubiquitous Computing.* New York: ACM.

Rantala, J., Myllymaa, K., Raisamo, R., Lylykangas, J., Surakka, V., Shull, P., & Cutkosky, M. (2011). Presenting spatial tactile messages with a hand-held device. *Proceedings of the 2011 IEEE World Haptics Conference.* IEEE.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rozado, D., Moreno, T., San Agustin, J., Rodriguez, F. B., & Varona, P. (2015). Controlling a smartphone using gaze gestures as the input mechanism. *Human–Computer Interaction, 30*, 34-63.

Räihä, K.-J., & Ovaska, S. (2012). An exploratory study of eye typing fundamentals: Dwell time, text entry rate, errors, and workload. *Proceedings of the CHI 2012 Conference on Human Factors in Computer Systems.* New York: ACM.

Samsung. (2013). *User Manual GT-I9295.* http://www.galaxys5s4manual.net/samsung-galaxy-s4-active-user-manual-gt-i9295/. Last accessed June 17, 2016.

Sherrick, C. E. (1985). A scale for rate of tactual vibration. *The Journal of the Acoustical Society of America, 78*, 78-83.

Skovsgaard, H., Räihä, K.-J., & Tall, M. (2012). Computer control by gaze. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä, *Gaze interaction and applications of eye tracking: advances in assistive technologies* (pp. 78-102). IGI Global.

Sodhi, R., Poupyrev, I., Glisson, M., & Israr, A. (2013). AIREAL: Interactive tactile experiences in free air. *ACM Transactions on Graphics*, *32* (Article 134).

Sofia, K. O., & Jones, L. A. (2013). Mechanical and psychophysical studies of surface wave propagation during vibrotactile stimulation. *IEEE Transactions on Haptics, 6*, 320-329.

Špakov, O., Isokoski, P., Kangas, J., Akkil, D., & Majaranta, P. (2016). PursuitAdjuster: An exploration into the design space of smooth pursuit–based widgets. *Proceedings of the ETRA 2016 Symposium on Eye Tracking Research and Applications.* New York: ACM.

Špakov, O., Rantala, J., & Isokoski, P. (2015). Sequential and simultaneous skin stimulation with multiple actuators on head, neck and back for gaze cuing. *Proceedings of the 2015 IEEE World Haptics Conference.* IEEE.

Tan, H. Z., Durlach, N. I. Reed, C. M., & Rabinowitz, W. M. (1999). Information transmission with a multifinger tactual display. *Perception & Psychophysics, 61*, 993-1008.

Viciana-Abad, R., Reyes-Lecuona, A., Poyade, M., & Escolano, J. (2011). The role of mismatches in the sensory feedback provided to indicate selection within a virtual environment. *Multimedia Tools and Applications, 55*, 353-378

Vidal, M., Bismuth, R., Bulling, A., & Gellersen, H. (2015). The royal corgi: Exploring social gaze interaction for immersive gameplay. *Proceedings of the CHI 2015 Conference on Human Factors in Computer Systems.* New York: ACM.

Vidal, M., Bulling, A., & Gellersen, H. (2013). Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. *Proceedings of the UbiComp 2013 Conference on Pervasive and Ubiquitous Computing.* New York: ACM.

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D. R. Kenshalo (Ed.), *The skin senses*, (pp. 195-222). Springfield, IL: Charles C. Thomas.

Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica, 31*, 285-289.

Witt, H., M. Lawo, & Drugge, M. (2008). Visual feedback and different frames of reference: The impact on gesture interaction techniques for wearable computing. *Proceedings of the MobileHCI 2008 Conference on Human Computer Interaction with Mobile Devices and Services.* New York: ACM.

Yatani, K., & Truong, K. N. (2009). SemFeel: A user interface with semantic tactile feedback for mobile touch-screen devices. *Proceedings of the UIST 2009 Symposium on User Interface Software and Technology.* New York: ACM.

Zhang, L., Li, X.-Y., Huang, W., Liu, K., Zong, S., Jian, X., Feng, P., Jung, T., & Liu, Y. (2014). It starts with iGaze: Visual attention driven networking with smart glasses. *Proceedings of the MobiCom 2014 Conference on Mobile Computing and Networking.* New York: ACM.

Zhang, Y., Bulling, A., & Gellersen, H. (2013). Sideways: A gaze interface for spontaneous interaction with situated displays. *Proceedings of the CHI 2013 Conference on Human Factors in Computer Systems.* New York: ACM.

Zhi, J. (2014). *Evaluation of tactile feedback on dwell time progression in eye typing.* M.Sc. Thesis, School of Information Sciences, University of Tampere, Finland. Available at: https://tampub.uta.fi/handle/10024/94973