Gaze Augmented Hand-Based Kinesthetic Interaction: What You See Is What You Feel

Zhenxing Li, Deepak Akkil, and Roope Raisamo

Abstract— Kinesthetic interaction between the user and the computer mainly utilizes the hand-based input with force-feedback devices. There are two major shortcomings in hand-based kinesthetic interaction: physical fatigue associated with continuous hand movements and the limited workspace of current force-feedback devices for accurately exploring a large environment. To address these shortcomings, we developed two interaction techniques that use eye gaze as an additional input modality: *HandGazeTouch* and *GazeTouch*. *HandGazeTouch* combines eye gaze and hand motion as the input for kinesthetic interaction, i.e. it uses eye gaze to point and hand motion to touch. *GazeTouch* replaces all hand motions in touch behavior with eye gaze, i.e. it uses eye gaze to point and gaze dwell time to trigger the touch. In both interaction techniques, the user feels the haptic feedback through the force-feedback device. The gaze-based techniques were evaluated in a softness discrimination experiment by comparing them to the traditional kinesthetic interface, *HandTouch*, which only uses the hand-based input. The results indicate that the *HandGazeTouch* technique is not only as accurate, natural, and pleasant as the traditional interface but also more efficient.

Index Terms— Kinesthetic interaction, gaze tracking, hand-eye coordination, force-feedback device, workspace, fatigue.

◆

1 INTRODUCTION

Our everyday interactions in the physical world are inherently multimodal [1]. Touch is one of our most important interaction senses, and it normally utilizes the visual sense in parallel [2]. Humans normally look at their target before a touch-based operation [3], [4] and typical touch manipulation involves close hand-eye coordination. Using the two senses simultaneously to inspect objects can help estimate many important properties. For example, visual cues can be used to know the shape and color of the object and, more importantly, to determine where to touch. Likewise, haptic cues can be employed to identify texture and hardness, for example.

In our typical everyday object manipulation tasks, we tend to look at the object, specifically at the landmark points on the object we will touch, as we initiate our hand movements for reaching. Further, we maintain our gaze on the location as our hands contact the target for touching. The visual feedback provided by our eyes helps guide our hands toward the target and monitor task progress [3], [4], [5]. In this process, the reaching and touching operations are necessary and interlinked because touch is a proximal sense that detects haptic cues from objects close to or in contact with us.

Providing kinesthetic cues in *Human-Computer Interaction* (HCI) follows a similar model to our everyday physical interactions. Virtual objects are modelled by the computer and displayed on the screen, and kinesthetic cues are produced by various force-feedback devices, such as SensAble Phantom [6] and Novint Falcon [7]. These hand-based devices have a mechanical arm that can be moved along three degrees of freedom within the physical workspace [8], and they allow natural hand-based reaching and touching behaviors. For example, a user can control the mechanical arm of the device along the x-y plane to indirectly control the onscreen *Haptic Interaction Point* (HIP) and then push the arm along the z-axis to touch the virtual object and feel the haptic response.

Two major issues limit the usability of current-generation force-feedback devices. First, force-feedback devices have a limited workspace, which practically limits the area of the onscreen objects that can be interacted with. Second, prolonged use of the device is associated with physical fatigue of the hand due to the frequent hand movements required to perform the touch interaction.

Unlike in real-world physical interactions, in HCI it is possible to decouple and replace the reaching and touching operations of hand-based interaction with other input modalities. We naturally look at the object that we are going to physically touch [2]. It is hence possible to use the gaze of the user as a pointing and triggering mechanism to augment traditional kinesthetic interaction.

We developed two interaction techniques that use the gaze of the user as an input. *HandGazeTouch* is an interaction technique that employs gaze input to substitute for the reaching operation. The user simply looks at the point of interest and moves the mechanical arm of the device along the z-axis for touching. *GazeTouch* uses the user's gaze to substitute for both reaching and touching operations. The user looks at a point of interest and stares at it to progressively touch the point. In both techniques, force feedback is felt using the mechanical arm that the user is holding (see Figure 1 for an indicative diagram of the user interaction).

There are multiple motivations for studying gaze in the context of kinesthetic interaction. First, human eyes can move very quickly in comparison to limbs [9]. Using eye movements to substitute for the reaching operation can potentially enable faster kinesthetic interaction, and more importantly it can reduce hand fatigue and the potential for injuries associated with the prolonged operation of force-feedback devices [10], [11]. Second, using the eyes as a complementary input modality to control the position of the HIP can help achieve an infinite workspace. Reaching and even touching the target object can be done using the gaze instead of hand movements with the
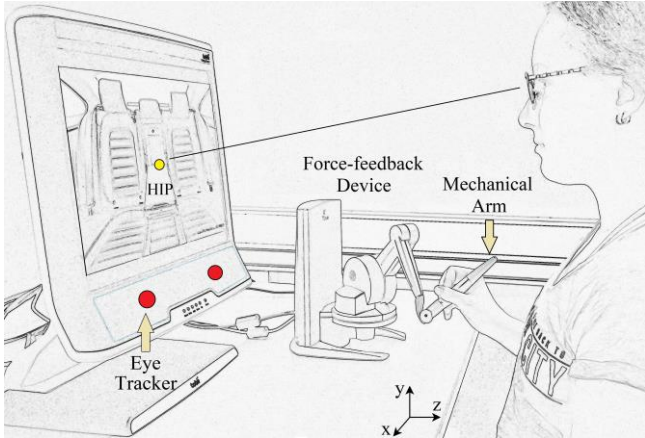
Figure 1: Indicative diagram for gaze-based kinesthetic interaction.

mechanical arm, which overcomes the limited workspace of current kinesthetic interface and has the potential to reduce the build complexity of traditional tabletop force-feedback devices. Third, thanks to technological advancements and a drop in price, eye tracking is no longer a niche technology used in the laboratory; it is widely available in the market [12]. The growing popularity and availability of eye tracking devices makes this a feasible input modality.

Gaze modality has been widely studied in non-haptic applications as a mean to improve user performance for target pointing and selection using a normal 2D display [9], [13], [14], [15] or a Virtual Reality (VR) head-mounted display [16], [17], [18]. In haptic applications involving tactile input devices such as touch screens, eye gaze has been used to improve object acquisition and manipulation, by completely replacing the hand motion [19], [20] or integrating with the hand input [21] to reach targets. Other studies in tactile interaction have investigated user performance while combining gaze input with tactile output [22], [23]. In kinesthetic interaction, gaze modality has been developed as an auxiliary function for solving technical and safety issues in tasks such as robotic surgery [24], [25], [26].

Replacing the reaching component of manual input with eye gaze to improve interaction has been studied for tactile interaction using the touch screen [19], [20]. However, no study extended this concept to kinesthetic interaction and used eye gaze to directly substitute for hand motions of kinesthetic input. Thus, the effects and performance of gaze in this research area have remained unexplored. Our study focused on the following research questions:

• RQ1: Is gaze a feasible input to substitute for hand-based reaching and touching operations in kinesthetic interaction?

• RQ2: How does the combination of gaze and hand modalities influence task performance in terms of efficiency?

• RQ3: Is there an effect on the accuracy of touch feeling, or are there other "side effects" to human perception?

We conducted an experimental study to evaluate two gaze-based interaction techniques, with the conventional hand-based technique (*HandTouch*) as the baseline. We compared the three interaction techniques in a softness discrimination test and focused on objective results, such as the efficiency of task completion and the accuracy of softness detection. In addition, we also recorded the participants' subjective responses, such as

naturalness, pleasantness, and physical and mental difficulties, through a questionnaire and interviews.

Our research has the following key novelties:

• The study explored the design space of combining eye gaze and hand motions as the input modality in kinesthetic interaction by developing two gaze-based interaction techniques (*HandGazeTouch* and *GazeTouch*).

• We experimentally compared *HandGazeTouch* and *GazeTouch* to the conventional kinesthetic interface that only uses hand-based input (*HandTouch*).

• Our study provides further theoretical understanding of human kinesthetic perception in identifying softness, and it extends a previous physical study [27] to a virtual environment using a force-feedback device.

The organization of the article is as follows: we first introduce the relevant previous studies in this research area, and then describe the two new interaction techniques (*GazeTouch* and *HandGazeTouch*) in greater detail. This is followed by the details of the experiment and the results. We finish with the discussion and practical implications of our results.

2 BACKGROUND

2.1 Hand-eye Coordination

In terms of hand-eye coordination, the eyes serve two distinct functions: locating the relevant task objects and guiding the appropriate motor actions [3]. The eyes and hands work in close synchrony in our everyday physical tasks. Foulsham [28] has noted that our eye movements are modulated by task characteristics, and the eyes fixate on the relevant objects at critical time points during the task. Bowman et al. [3] have observed that in tasks that require object manipulation, we look at the specific location of touch points before the contact happens. Land et al. [2] have found that human eyes fixate on the target object roughly half a second before its manipulation using the hand.

Similarly, previous studies have noted the systematic coordination between the hand and eyes when using indirect pointing devices such as a mouse [29]. However, few previous studies have leveraged natural hand-eye coordination to improve the kinesthetic interaction, and this is the focus of the current work.

2.2 The Limitations of Current Kinesthetic Interaction

As mentioned before, the mechanical arm of a force-feedback device not only allows the user to freely navigate a 3D environment, but also transfers force and torque to simulate the feeling of touch. However, the workspace of current force-feedback devices is limited by the length of the mechanical arm. Massie and Salisbury [8] have noted that a desktop force-feedback device such as Phantom only has a small wrist-centered workspace and the forearm is allowed only limited movement. Several studies [30], [31] have argued that for tasks requiring accurate positioning in a large environment, reaching the target is physically challenging using the current kinesthetic interface. To overcome the problems caused by the limited workspace of force-feedback devices, Conti and Khatib [30] have proposed the *Workspace Drift Controller*, which progressively centers the physical workspace of the device during the interaction.

Dominjon et al. [31] have proposed *Bubble*, which utilizes a spherical area around the HIP with a hybrid position and rate control for accurate and efficient haptic exploration in a large virtual environment.

Physical fatigue is another important issue in interactions that involve repetitive physical actions. Previous studies have shown that multiple factors cause fatigue in the use of force-feedback devices. For example, Ott et al. [10] have found that muscular fatigue could be caused by uncomfortable postures when using force-feedback devices. Hamam and Saddik [11] have demonstrated that repetitive kinesthetic tasks with a larger force or higher distance result in greater user fatigue.

Physical fatigue not only influences the user's experience but also negatively affects performance during interaction. Allen and Proske [32] have shown that muscle fatigue disturbs our sense of position and makes users commit more errors in estimating spatial positions. Cortes et al. [33] have demonstrated that the ability of users to perform a smooth and controlled physical action is influenced by physical fatigue. It is therefore important for designers of kinesthetic interfaces to strive to reduce user fatigue.

Another intuitive method to extend the workspace is to use gaze as an interaction modality. We normally look at the objects we are trying to touch [2]. We can practically achieve an infinite workspace by dynamically redefining the position of the HIP to the location of the gaze, and simultaneously hand fatigue may be reduced by replacing hand motions with eye gaze.

In addition, mental effort is not correlated with muscle movements, so most studies, such as [10], [11], had less concern for mental effort in haptic manipulation. Mental effort is considered in this study. We used both perceived physical and mental difficulty as measures to evaluate the development of the kinesthetic interaction techniques.

2.3 Tactile Interaction with Eye Gaze as An Input Modality

Tactile interaction focuses on applying cutaneous sensation to human-computer interactions [34]. The related studies with eye gaze can be categorized into two research areas: tactile output (e.g. vibration) and tactile input (e.g. touchscreen).

In the studies involving tactile output, some studies have addressed the value of eye gaze with vibrotactile feedback in a variety of devices, such as computers, mobile phones, and smartwatches [35]. Kangas et al. [22] have shown that gaze interaction with vibrotactile feedback increases the efficiency of interaction. Akkil et al. [23] have noted that vibrotactile feedback is a clearer and more noticeable modality for gaze events than visual feedback in small-screen devices such as smartwatches.

With tactile input devices, many have attempted to combine eye gaze with hand gestures onscreen as the input to improve the functions of object manipulation. Pfeuffer et al. [19] have developed *Gaze-touch*: an interaction technique that uses eye gaze for remote selection and touch gestures onscreen for object manipulation. This method decouples the reaching component from the hand-based input and replaces it with eye gaze, so users can remotely manipulate the objects without using their hands to reach for the target. In their following study [20], they developed *Gaze-shifting*: a generic mechanism for switching direct and indirect input modes in touch-based interaction based on the relative position of the gaze location and the touch operation. In addition, Stellmach and Dachselt [21] have developed *Look and Touch*, which integrates eye gaze with manual input instead of completely replacing it. Eye gaze is employed for coarse reaching, and the user selects the target by fine hand-based reaching, similar to the method used in Magic pointing [9].

Eye gaze as an input modality has been widely considered as a fast pointing and selection mechanism [13], [14], [15], and Pfeuffer et al.'s work [19] has demonstrated that using gaze to replace hand motions as the input is promising for tactile interaction. Our study extends this concept to kinesthetic interaction.

2.4 Kinesthetic Interaction with Eye Gaze as An Input Modality

Kinesthetic interaction is another important branch of haptic interaction, concentrating on motion sensations originating in the muscles, tendons, and joints [34]. Force-feedback devices (e.g. the Phantom device) are commonly used as haptic interfaces for kinesthetic interaction. Their mechanical arms often have three or six degrees of freedom [8], which is suitable for hand motions as the kinesthetic input, and they simultaneously transfer the haptic cues of virtual objects as kinesthetic output to the hand. Eye gaze has been considered an auxiliary input channel for kinesthetic interaction, e.g. to enhance safety in critical surgical tasks [24], [25], [26] or foster remote collaboration [36].

Previous studies [24], [25], [26] have employed a technique called *Gaze-Contingent Motor Channeling* (GCMC) for robotic surgery, which uses the gaze to set safety boundaries for the HIP in order to prevent the instrument from inadvertently penetrating the tissue during surgery. The safety boundaries were established by employing a spring force on the HIP toward the eye fixation point, with a magnitude based on the distance between the eye fixation point and the HIP.

Another study combining gaze and force-feedback devices was conducted by Leff et al. [36]. They developed a collaborative system that provides gaze awareness between remote partners during kinesthetic interaction. They used gaze as a channel to foster collaboration.

The focus of our study is to understand the utility of gaze in the reaching and touching operations of kinesthetic interaction. Can gaze be used to replace the reaching operation? Can gaze be used as a mechanism to initiate the touching operation? What are its effects on human kinesthetic perception? The objective of our study is to answer these questions.

3 INTERACTION TECHNIQUES AND RESEARCH HYPOTHESES

3.1 Interaction Techniques

The prototype system we developed supports three different interaction techniques: *HandTouch* (H), *HandGazeTouch* (HG), and *GazeTouch* (G). (The abbreviations are used in some of the figures and tables.) Figure 2 shows the three interaction techniques. In *HandTouch*, the user controls and manipulates the mechanical arm of the force-feedback device along all three degrees of freedom (x, y, z). In *HandGazeTouch*, the gaze of the user controls the position of the interaction point along the x-y plane. The user controls the touch behavior by manipulating the mechanical arm of the device along the z-axis (see also Figure 1). In *GazeTouch*, the gaze controls the interaction point along the x-y plane and the duration of gaze fixation controls its
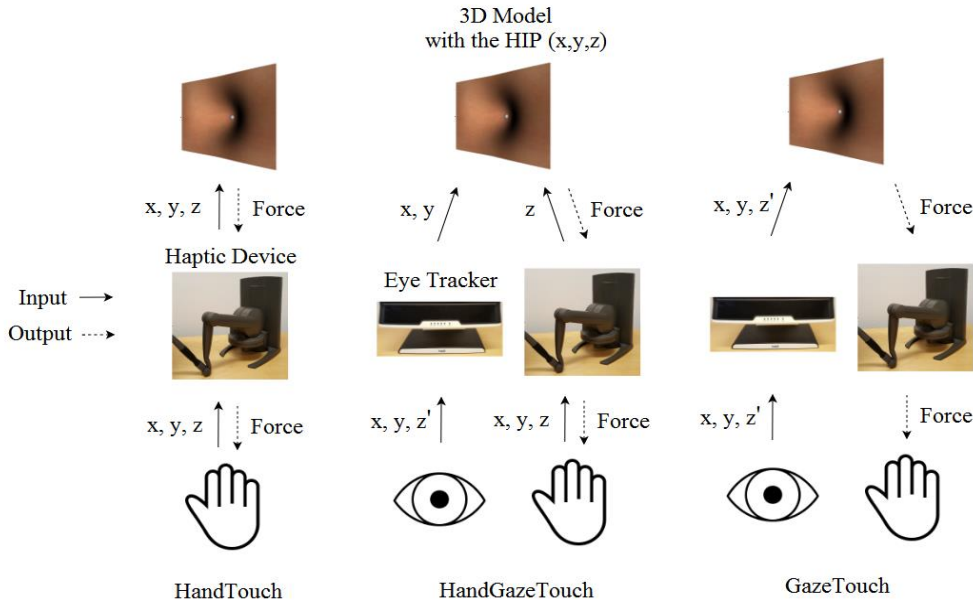
Figure 2: Concept map of the three interaction techniques. The hand movement coordinates respectively are x, y and z on the three axes. The eye gaze coordinates are x and y, and z' is the dwell time of the eye gaze.

TABLE 1.
SPECIFICATION OF INTERACTION TECHNIQUES.

Techniques	Interactive steps from the perspective of the users	Force-feedback device usage	Eye tracker usage	Input operating dimension
HandTouch	Uses the hand to reach and touch the target.	Yes	No	Hand: x, y, z Eyes: -
HandGazeTouch	Uses eye gaze to reach the target and the hand to touch.	Yes	Yes	Hand: z Eyes: x, y
GazeTouch	Uses eye gaze to reach and dwell time to trigger the touch.	Yes	Yes	Hand: - Eyes: x, y, z'

movement along the z-axis. The details of the three interaction techniques are shown in Table 1.

3.2 Hypotheses

Based on our research questions, we focused on the analysis of the three interaction techniques from five aspects: efficiency, accuracy, naturalness, fatigue, and pleasantness. We formulated the following hypotheses:

H1: *HandGazeTouch* and *GazeTouch* will be faster than *HandTouch*.
- o The eyes are faster at reaching the target object than the hand [2]. *GazeTouch* and *HandGazeTouch*, which use eye gaze, may hence be faster than *HandTouch*.

H2: *HandTouch* will be better than both *HandGazeTouch* and *GazeTouch* in the terms of the accuracy of kinesthetic perception.
- o Since people are used to employing the hands to touch objects, haptic cues caused by eye gaze may be unfamiliar to users. Users may be worse at discriminating differences in softness using *HandGazeTouch* and *GazeTouch*.

H3: *HandGazeTouch* and *GazeTouch* will cause less tiredness of the hand than *HandTouch*. However, *HandTouch* will be considered more natural and pleasant than the alternatives.

- o In both the gaze-based conditions, hand activities are replaced by eye gaze. This may lead to less fatigue of the hand for *HandGazeTouch* and *GazeTouch*. However, since the eyes are naturally used for perception and not for intentional control, *HandGazeTouch* and *GazeTouch* may be considered less natural, less pleasant, and more cognitively demanding by the users.

4 EXPERIMENT

4.1 Method

To test our hypothesis, we designed a controlled lab experiment that followed a within-subject design. Three interaction techniques were examined as the experimental conditions in a softness discrimination task. Softness discrimination was selected as the experimental test because softness is one of the important properties of physical objects and perceiving a change in softness could be an appropriate method to examine the effect of different interaction techniques in kinesthetic perception. In addition, common force-feedback devices such as Phantom only support single-point interaction, which is adequate for accurate softness discrimination [27]. Each experimental condition involved 36 repetitions of the softness discrimination task, and every task involved discriminating the softness between two onscreen square-shaped skin models presented as 2.5D models without considering the thickness. The participant had to touch

the two skin models and identify the harder of the two, then communicate the answer using the appropriate arrow key on the keyboard.

Since we are interested in both interaction efficiency and kinesthetic perception, a Fitts's Law [37] type of experiment only measuring pointing efficiency was not suitable for this study. Accurate kinesthetic perception may require participants to touch the same object repeatedly or compare the two models multiple times. Another experimental setup was motivated by the Fitt's Law study setup. The difficulty of touch behavior in softness discrimination relies on two factors: difficulty in reaching the target and difficulty in perceiving the difference in softness. Therefore, we designed the tasks to include two difficulty levels for reach (difficult and easy) and two difficulty levels for perception (difficult and easy).

The difficulty for reach was manipulated by controlling two variables of the skin models: the size of the skin models and the distance between two models displaced on the screen. The two skin models had the same size set at three possible levels: 2.8 cm, 4.8 cm, or 6.8 cm. The distance between the two models had three possible values: 8.1 cm, 16.2 cm, or 24.3 cm (measured from center to center). Each condition involved four occurrences of every combination of the size and distance levels (3 x 3 x 4 = 36 tasks). We used the Index of Difficulty (ID) of Fitts's Law as the tool to categorize the nine unique combinations of size and distance into the two distinct levels of difficulty. The grouping threshold was set at 1.75, with Fitts's ID calculated by the medium size (i.e. 4.8 cm) of the skin models and the small distance (i.e. 8.1 cm) between the models. The combinations with a higher ID belonged to the difficult level for reach, and the combinations with a lower or equal ID belonged to the easy level for reach.

The difficulty for perception was implemented by controlling the softness difference between the two skin models. The softness of each skin model is controlled by manipulating the stiffness coefficient (k) of the Spring Model used to implement the skin models, which is a variable with a range (i.e. 0.065–0.145) used in the software. The difference in values (Δk) of the stiffness coefficient for both the onscreen skin models (k1, k2) was manipulated to result in six different levels with varying difficulties for identifying the softness difference between the two skin models. The lowest value of the difference level was identified such that it is barely perceivable and requires close inspection to identify the harder model (i.e. k1 = 0.09, k2 = 0.12: Δk = 0.03). The highest value of difference level was chosen such that the difference is easy to perceive (i.e. k1 = 0.065, k2 = 0.145: Δk = 0.08). The six difference levels (i.e. 0.03–0.08) occurred six times in each condition (6 x 6 = 36 tasks). The six levels of difference were then categorized into two levels of perception difficulty (i.e. difficult: 0.03–0.05 and easy: 0.06–0.08) for ease of analysis.

For each task, the difficulties for reach and perception were independent of each other and randomly chosen from the list of possible combinations.

In the gaze-based conditions, the position of the HIP was dependent on the point of the user's gaze. Using raw gaze points caused the HIP to be jittery. We hence used a simple recursion-based filter to smoothen the gaze point, before displaying it onscreen, as also used in a previous study [38].

$$y(i) = W * x(i) + (1 - W) * y(i - 1). \quad (1)$$

where $y(i)$ is the ith smoothened gaze position and $x(i)$ is the ith actual gaze position produced by the eye tracker. The percentage weight for the actual gaze position W was selected as 0.1 in the study.

In the *GazeTouch* condition, the position of the HIP along the x- and y-axes was controlled by the gaze position onscreen, and the fixation duration was translated to the z-axis movement of the HIP. The user had to fixate for 1 second to initiate touch behavior and another second to reach the maximum touch force. During this process, the output force was linearly increased with time and continuously transferred to the mechanical arm. After 2 seconds of dwelling, the output force saturated and further dwelling did not lead to any changes. In addition, a large gaze movement (i.e. >1.4 cm) can reset the touch operation in *GazeTouch*, allowing the user to touch the same place again or a different place.

For the *HandTouch* and *HandGazeTouch* conditions, the position of the HIP along the x-y plane was respectively controlled by the hand and eye gaze. The distance to initiate touch in the z-axis for both conditions was set at a maximum of 6 cm, which is variable (normally < 6 cm) depending on how and where the user initially held the mechanical arm.

The study used both objective and subjective measures to understand the strengths and weaknesses of each interaction technique. The objective measures were the task completion time and the number of errors made in softness discrimination. The subjective measures were captured using a custom questionnaire. Table 2 shows the four questions with a 7-point Likert scale based on the subjective assessment questions used in NASA-TLX [39]. In addition, we used a post-test questionnaire in which the participants ranked the three interaction techniques based on the tiredness of their eyes and hand. Furthermore, the participants selected the preferred technique(s) from the three interaction techniques.

TABLE 2.
STATEMENTS IN THE QUESTIONNAIRE.

No.	Description
Q1	This interaction technique is mentally difficult.
Q2	This interaction technique is physically difficult.
Q3	With this interaction technique, it is natural to touch.
Q4	This interaction technique is pleasant.

4.2 Apparatus and Environment

The experiment was conducted on a Dell T3600 Windows 7 desktop computer with an Intel E5-1600 processor, NVIDIA Quadro 4000 graphics, and 8GB 1600MHz of memory. The experiment environment is shown in Figure 3. We used a Phantom Desktop [6] as the force-feedback device and a Tobii T60 [12] as both the display and the eye tracker. The software development kits were the open-source H3DAPI [40] for haptics and Tobii SDK for eye tracking. We also utilized TraQuMe [41], a tool to measure gaze data quality. A keyboard was used to select and record the answer for each task and move to the next task, and headphones were utilized to block out noise.

Figure 3: The experiment setup involved a Tobii T60 gaze tracker and a Phantom force-feedback device. The two skin models on the screen have the medium size of 4.8 cm (out of the three size levels) and the large distance of 24.3 cm between them (out of the three distance levels).

4.3 Pilot Study

A pilot study was conducted at first, employing six participants (three female and three male) aged between 24 and 42 years (Mean (M) = 31.8, Standard Deviation (SD) = 6.46) who had experience of using eye trackers and/or force-feedback devices. Based on the pilot tests, we further calibrated the system.

- In all conditions, six softness difference levels for the skin model (Δk: 0.03–0.08) were chosen, such that all were perceivable by the participants considering the sensitivity of the force-feedback device and human kinesthetic perception.

- In the *HandTouch* and *HandGazeTouch* conditions, the movement of the mechanical arm was transferred to the movement of the onscreen HIP without any scaling, e.g. in the *HandTouch* condition, a 1 cm lateral movement of the mechanical arm resulted in a 1 cm lateral movement of the HIP onscreen.

- In the *HandGazeTouch* and *GazeTouch* conditions, the weight of the current gaze position in the recursion filter used to smooth the gaze position was chosen to be 0.1 based on the good balance between jitteriness and the responsiveness of the gaze point.

- In the *GazeTouch* condition, the parameters for translating fixation duration to the z-axis movement of the HIP were selected. The values (1s initiation and 1s dwelling) were selected such that they overcome the *Midas Touch* problem [42] and at the same time do not require staring for too long for touching.

4.4 Procedure

The participants were first introduced to the study and the equipment used. The force-feedback device was placed in position based on the user's dominant hand. All participants signed an informed consent form and then filled in the background questionnaire. Each question in the questionnaire was explained to clarify its meaning, such as the difference between mental difficulty and physical difficulty, before proceeding to the experiment.

Before each of the gaze-based conditions, the eye tracker was calibrated using nine-point onscreen calibration. The quality of eye tracking was measured using a nine-point TraQuMe

evaluation [41]. We defined an objective criterion for recalibration. If any of the nine points showed more than 2 cm eye tracker offset, the participants were asked to recalibrate. The 2 cm threshold was defined considering the smallest size of the skin model used (i.e. 2.8 cm). If any of the tracking offset values were still beyond this threshold after multiple recalibration, the test was discontinued, and the data were not included in the analysis.

The participants were asked to finish each experimental task as accurately and efficiently as they could. Participants pressed the appropriate arrow key to record their answer for each softness comparison task, after which the system presented the next discrimination task. We did not present the participants with the feedback regarding the accuracy of their discrimination during the experiment.

There were three experimental conditions with 36 tasks in each condition. Before each condition, the participants had up to five minutes to familiarize themselves with the operation of each interaction technique. The order of the experimental conditions was counter-balanced. In the experiment, no hand-rest or elbow-rest equipment was used, and the height of chair/table was adjusted for the participants to make them face the screen and hold the mechanical arm of the force-feedback device horizontally. In addition, the participants were asked to wear headphones to block out the noise generated by the force-feedback device, because the noise level may indicate the magnitude of force and, thus, the softness of the onscreen skin models.

4.5 Participants

We recruited 24 participants from our university community (13 female and 11 male) aged between 19 and 42 years (M=26.5, SD=6.13). All participants had normal touch sensitivity. Seven participants had corrected vision and the remainder had normal vision. Only two participants used the left hand as their dominant hand; the remainder were right-dominant. Seven participants had used a similar eye tracker before ($\leq$ two times), and one participant had used the force-feedback device before (one time). Since their experience was limited, we included their data in this study.

The data for one of participants had to be replaced because of issues in gaze tracking. The participant could not pass the gaze tracking accuracy check using the TraQuMe tool, and thus could not complete the test. Another participant was invited to replace the original participant.

The mean gaze tracking accuracy for the 24 participants was 0.56 degrees (SD=0.17 degrees), which translated to 0.62 cm in screen distance (SD=0.19 cm).

5 RESULTS

5.1 Test Completion Time

For each condition with 36 tasks, we calculated the mean value of the task completion time for the different levels of reach difficulty and perception difficulty. The Shapiro–Wilk Normality test was conducted first. Since the data were not normally distributed (p < .001), we analyzed the data using a 3x2x2 Aligned Rank Transform (ART) repeated measures non-parametric ANOVA [43], and the post-hoc analysis was done using the Wilcoxon Signed-Rank test with Holm-modified

TABLE 3.
TESTS OF WITHIN-SUBJECTS EFFECTS ON COMPLETION TIME.

Sources	df	F	Sig.
Condition	2, 46	10.3	<0.001
Reach	1, 23	31.3	<0.001
Perception	1, 23	42.2	<0.001
Condition & Reach	2, 46	13.9	<0.001
Condition & Perception	2, 46	4.60	0.016
Reach & Perception	1, 23	0.41	0.52
Condition, Reach, & Perception	2, 46	0.60	0.54

Bonferroni correction [44] to control for family-wise type-1 error. All the p-values presented are after Holm-modified Bonferroni correction.

Table 3 shows the overall p-values for the completion times based on the conditions (*HandTouch*, *HandGazeTouch*, and *GazeTouch*), difficulty of reach (difficult and easy), and difficulty of perception (difficult and easy) through the repeated measure.

The ART ANOVA test results showed statistically significant main effects for all three factors: condition, reach, and perception. In addition, the ART ANOVA showed a statistically significant interaction effect for condition and reach as well as condition and perception. Other effects were not statistically significant.

Figure 4 shows the boxplot for overall completion time for the three interaction techniques. The mean value of task completion time, visualized as the y-axis in the boxplot, for *HandGazeTouch* (M = 5.92, SD = 2.21) was approximately 15% lower than for *HandTouch* (M = 6.82, SD = 2.92) and 29% lower than for *GazeTouch* (M = 7.61, SD = 1.86).

The Wilcoxon Signed-Rank test showed that *HandGazeTouch* was statistically significantly faster than both *HandTouch* (Z = -2.200, p = .048) and *GazeTouch* (Z = -3.743, p < .001). *HandTouch* and *GazeTouch* were not statistically significantly different from each other (Z = -1.914, p = .056), but the difference approached significance.

Unsurprisingly, reach and perception difficulty had a significant effect on the result. When the skin models were further apart or smaller in size, users took more time in movement (Z = -3.514, p < .001). Similarly, when the difference in softness between the two skin models was low, participants took more time in perception (Z = -4.200, p < .001). However,
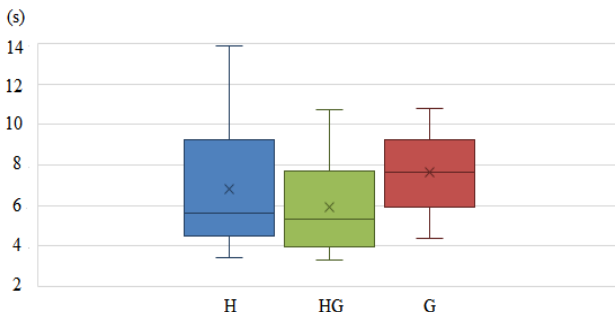

Figure 4: Completion time for the three interaction conditions. The line in the boxplot is the median value and the cross mark is the mean value (the following figures use the same marks).
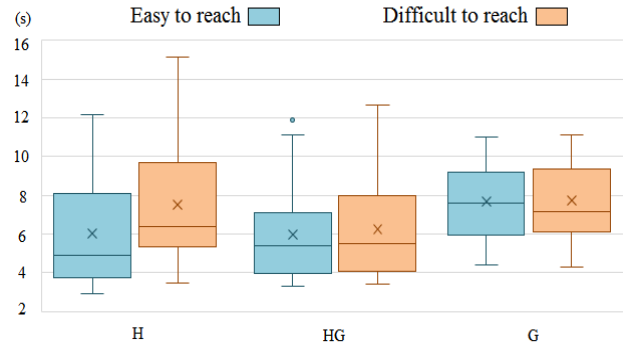

Figure 5: Task completion time of the three interaction techniques, based on reach difficulty levels.

TABLE 4.
MEAN TASK COMPLETION TIME USING THE THREE INTERACTION TECHNIQUES BASED ON REACH DIFFICULTY LEVELS.

Mean time (seconds)	Conditions		
	H	HG	G
Easy to reach	6.01 (SD=2.77)	5.95 (SD=2.45)	7.67 (SD=2.02)
Difficult to reach	7.51 (SD=3.16)	6.26 (SD=2.49)	7.72 (SD=2.01)

we are more interested in the interaction effect of these factors on the conditions to understand if the task completion times for the conditions were differentially affected by the perception and reach difficulty levels. Next, we analyze the interaction effects in detail.

Condition & Reach: Post-hoc analysis
Figure 5 demonstrates the completion time based on the condition and reach. Table 4 shows the mean completion time for the three interaction techniques based on the difficulty levels in the reach operations.

We analyzed the simple effect of reach difficulty for the three techniques using the Wilcoxon Signed-Rank test.

• *HandTouch* was significantly faster when the reach difficulty was low compared to when the reach difficulty was high (Z = -4.143, p < .001).

• The simple effect of reach difficulty was not statistically significant for the gaze-based conditions *HandGazeTouch* (Z = -1.057, p = .58) and *GazeTouch* (Z = - .029, p = .977).

Condition & perception: Post-hoc analysis
The completion time based on condition and perception is shown in Figure 6. Table 5 gives the mean completion time for each technique based on the perception difficulty levels.

We further analyzed the simple effect of perception difficulty levels for the three interaction techniques using the Wilcoxon Signed-Rank test.

• *HandTouch* was significantly faster when the perception difficulty was low than when the perception difficulty was high (Z = -2.371, p = .018): the mean value of task completion for *HandTouch* with a low perception difficulty was 6.63 seconds. This increased to 7.13 seconds, (a 7.5% increase) when the perception difficulty was high.

• The simple effect of perception difficulty was statistically significant for *HandGazeTouch* (Z = -3.600, p < .001): the mean
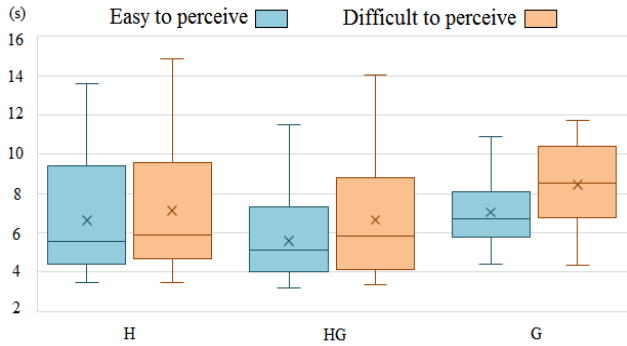
Figure 6: Task completion time of the three interaction techniques, based on perception difficulty levels.

TABLE 5.
MEAN COMPLETION TIME OF EACH TECHNIQUE BASED ON PERCEPTION DIFFICULTY LEVELS.

Mean time (seconds)	Conditions		
	H	HG	G
Easy to perceive	6.63 (SD=2.82)	5.58 (SD=2.20)	7.04 (SD=1.72)
Difficult to perceive	7.13 (SD=3.23)	6.65 (SD=2.85)	8.45 (SD=2.11)

value of the task completion time increased by 19%, from 5.58 seconds for low perception difficulty to 6.65 seconds for high perception difficulty.

• Similarly, the simple effect of perception difficulty was statistically significant for *GazeTouch* as well (Z = -3.943, p < .001): the mean value of the task completion time increased 20%, from 7.04 seconds for low perception difficulty to 8.45 seconds for high perception difficulty.

5.2 Error Analysis

Errors occurred when participants selected the wrong option after comparing the softness of the two skin models. Overall, there were a total of 109 errors, which is an average of 4.54 errors per participant, from a total of 108 tasks per participant (36 tasks per condition x 3 conditions). Of the total 109 errors, 85 errors occurred in tasks with the high perception difficulty. Since the number of errors were so few and most of them occurred in tasks with the high perception difficulty, we concentrated on the analysis of the three conditions with the perception difficulty.

The ART repeated measure 3x2 factorial ANOVA for error analysis showed a significant main effect of interaction techniques on error (F (2, 46) = 8.15, p = .001). The effect of perception difficulty was also statistically significant (F (1, 23) =
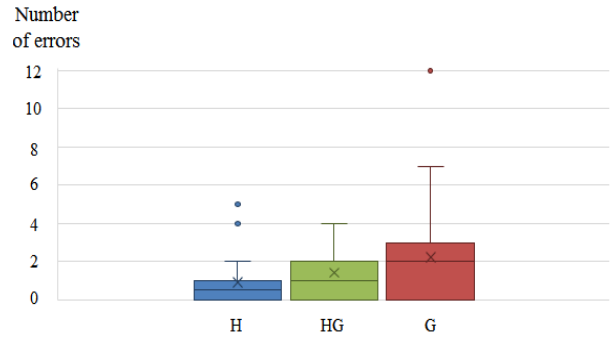

Figure 7: Error distribution of the three interaction techniques.

47.16, p < .001). There was no significant interaction effect between the perception difficulty and the conditions on the error rates (F (2, 46) = 2.26, p = .116).

Figure 7 provides the error distribution based on each technique. *HandTouch* had the least number of errors (median value below 1), followed by *HandGazeTouch* (median = 1), and *GazeTouch* (median = 2). The Wilcoxon Signed-Rank test showed that the difference between *HandTouch* and *GazeTouch* approached significance (Z = -2.251, p = .061). In addition, *HandGazeTouch* was not different from both *HandTouch* (Z = -1.452, p = .292) and *GazeTouch* (Z = -1.204, p = .229).

5.3 Results of Subjective Data

Subjective data from the questionnaire was evaluated to explore the results, including perceived mental and physical difficulties, naturalness, and pleasantness. Figure 8 shows the subjective results of the questionnaire; the data was analyzed with the Wilcoxon Signed-Rank test.

Mental difficulty: *HandTouch* was better in terms of mental difficulty than *HandGazeTouch* (Z = -3.241, p = .003) and *GazeTouch* (Z = -3.002, p = .006). There was no difference between *HandGazeTouch* and *GazeTouch* (Z = -.049, p = .961).

Physical difficulty: There were no differences in terms of physical difficulty among three interaction techniques (*HandTouch-HandGazeTouch*: Z = -1.677, p = .188; *HandTouch-GazeTouch*: Z = -.579, p = .563; *HandGazeTouch-GazeTouch*: Z = -1.064, p = .574).

Naturalness: Both *HandTouch* (Z = -2.485, p = .026) and *HandGazeTouch* (Z = -2.912, p = .012) were considered more natural than *GazeTouch*. There was no difference between *HandTouch* and *HandGazeTouch* (Z = -1.010, p = .313).

Pleasantness: *HandTouch* was considered more pleasant than *GazeTouch* (Z = -2.535, p = .033). However, *HandGazeTouch* was not statistically significantly different from the others (*HandGazeTouch-HandTouch*: Z = -1.311, p = .190;
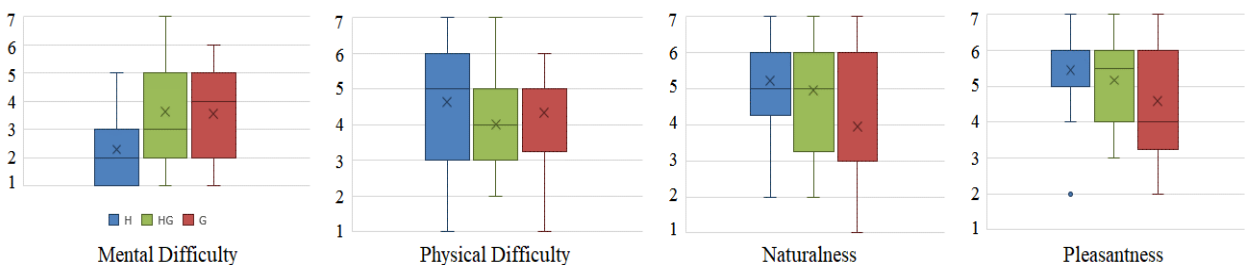

Figure 8: Subjective results of the study.

HandGazeTouch-GazeTouch: Z = -1.849, p = .130).

5.4 Overall Evaluation of the Interaction Techniques

At the end of the study, all 24 participants ranked the techniques based on the tiredness of the hand and eyes while using the techniques. Overall, *GazeTouch* was ranked to be the least tiring for the hand (21 votes), followed by *HandGazeTouch* (19 votes). *HandTouch* was considered the most tiring for the hand (21 votes).

In terms of eye tiredness, *HandTouch* was voted the least tiring for the eyes (22 votes), followed by *HandGazeTouch* (19 votes). *GazeTouch* was considered the most tiring for the eyes (18 votes).

For overall preference, the participants could select more than one choice if they liked more than one condition equally. Overall, 15 of the 24 participants preferred *HandTouch*, and an almost equal number of participants (14) preferred *HandGazeTouch*. *GazeTouch* was the least preferred interaction technique; it was preferred by only 4 participants.

The users' preferences and the reasons behind them were revealed in the free-form comments provided by the participants:

P1: "I can sense the differences in softness better with it [*HandTouch*]. I make a circular motion on the tissue to better understand the softness." (User preferred *HandTouch.)*

P11: "[I prefer *GazeTouch*] because I did not have to make much physical effort. The *HandTouch* method made my hands ache. The *HandGazeTouch* method felt too complicated because there was too much to do at the same time."

P18: "*HandGazeTouch* was the fastest and most pleasant to use, with a little practice." (User preferred *HandGazeTouch*.)

P20: "[I prefer *HandTouch*] because it is the closest to real life when we touch real objects."

P24: "The *HandGazeTouch* method felt reasonably natural and really fascinating." (User preferred *HandGazeTouch*.)

6 DISCUSSION

This study investigated the use of eye gaze in kinesthetic interaction. It demonstrates that eye gaze is a feasible and beneficial input modality in this context. We will now discuss the findings of the study in relation to our initial research hypotheses and state of the art.

6.1 H1: *HandGazeTouch* and *GazeTouch* will be faster than *HandTouch*

Our study partly supports this hypothesis. With the multimodal input shown in the results, *HandGazeTouch* was significantly faster than both *HandTouch* and *GazeTouch*, which only involved a single input modality.

HandGazeTouch leverages the natural hand-eye coordination, and it uses eye gaze to replace the hand-reaching component in kinesthetic interaction. Similarly, the *GazeTouch* technique uses the gaze to replace both the reaching and touching components. However, this condition had the largest task completion time.

Our task required the user to touch two soft skin models placed apart horizontally. Often, when the difference in softness was small, participants had to touch each model multiple times to evaluate the difference. In *HandGazeTouch*, reaching the soft model was fast and intuitive, as the participants simply had to

gaze at the target. The performance improvement in *HandGazeTouch* thus could be because of its improved efficiency in the reach operation. In *GazeTouch*, even though reaching the target was fast, the additional dwell time to cause haptic cues and overcome the *Midas Touch* problem likely slowed down the interaction.

The improved performance of *HandGazeTouch* in the reach operation is evident from our analysis of task completion times for reach difficulty. Our result (Figure 5) shows that the distance between the touch objects and their sizes influenced the task completion times in *HandTouch*. However, these variables did not have a noticeable effect in the gaze-based conditions (*HandGazeTouch* and *GazeTouch*). Saccadic eye movements, which are responsible for bringing an object of interest to our foveal vision, typically last 30–120 milliseconds, and the effect of the distance of the object on the time it takes for our eyes to focus on it (though almost linearly related) is minimal [45]. This, however, was not the case for the *HandTouch* technique. When the target was further away, using the hand took significantly more time to reach the targets.

Our analysis of task completion times for different perception difficulties shows an interesting result. Overall, participants took more time to complete the task in all conditions when the task perception difficulty was high. This suggests that when it was difficult to perceive the difference in softness, participants using all three techniques had to touch the same tissue multiple times or repeatedly alternate between the two models to clearly gauge the difference in softness. However, the increase in task completion times was less for *HandTouch* (only 7.5%) and substantially higher for the gaze-based conditions (19% and 20% for *HandGazeTouch* and *GazeTouch* respectively). A potential explanation is that even though gaze-based techniques have the advantage in the reach operation due to the saccadic eye movements, participants using *HandGazeTouch* and *GazeTouch* had to repeat the touch activity more times to accurately estimate the difference in softness between the two models. Previous research in touch perception has argued that purposeful hand operations modulate haptic perception [46]. Our results extend this finding to computer-mediated kinesthetic interaction. In tasks that require estimating subtle differences in softness, using gaze as a mechanism to substitute for the reaching or touching operations may lead to an increased amount of time in identifying softness compared to using *HandTouch*.

6.2 H2: *HandTouch* will be better than *HandGazeTouch* and *GazeTouch* in terms of the accuracy of kinesthetic perception.

Our study partially supports this hypothesis. The results give a preliminary indication that *HandTouch* may be better than *GazeTouch* in the accuracy of kinesthetic perception, but *HandGazeTouch* was not statistically significantly different from *HandTouch*.

The experimental task focused on softness discrimination between two skin models. Our participants committed a different number of errors in judging the softness difference based on the three interaction techniques (shown in Figure 7). The difference in the number of errors is interesting because it indicates that the kinesthetic interaction techniques may mediate our kinesthetic perception. Fewer errors in softness discrimination indicates that the generated kinesthetic cues were easier to interpret by the

somatosensory system. Similarly, a higher number of errors indicates that the kinesthetic cues were harder to interpret by the somatosensory system.

In *GazeTouch*, participants made noticeably more errors than in *HandTouch*. While the difference only approached statistical significance, we believe our results present preliminary evidence that active haptic interaction using hand motions may be more accurate than passively feeling the kinesthetic cues. Our results are consistent with our understanding of haptic perception in real-world interaction with physical objects. Lederman and Klatzky [47] have argued that to accurately perceive haptic cues provided by an object, both haptic cues generated by the object and the physical motion of the hand made to gain the haptic cues play a decisive role in our cognitive process. *GazeTouch*, which only uses gaze as the input without any hand motion, breaks the link between the human somatosensory system and hand movements, and it might result in lower accuracy in kinesthetic perception compared to *HandTouch*. This is also evident from the performances and comments of the participants. Most participants used specific strategies to sense softness when using *HandTouch* (e.g. moving the hand in a circular motion on the surface of the model or touching it multiple times). Such strategies were difficult to employ with the *GazeTouch* technique. Furthermore, the haptic cues generated from the models using a hand to touch are active haptic information (including both action and reaction forces), but the force feedback using gaze to touch is passive haptic information (including only reaction force). Lamotte [27] has demonstrated that softness discrimination is more accurate for active touch than passive touch.

HandGazeTouch, on the other hand, used gaze to reach but still required hand motion to touch, which could be considered a good compromise for both the fast and accurate exploration of haptic cues. It allowed the participants to easily touch the model multiple times and adjust the gaze position to, e.g. the left or right of the current contact point to simulate the sliding motion of the HIP, thus including both active and passive haptic information. For example, touching objects with a left-right or up-down movement was done by slightly moving the eye gaze, and thus the force feedback is passive; multiple touches on the same part of the model were done using hand movement, and thus the force feedback is active haptic information. This may explain why *HandGazeTouch* had a lower error rate than *GazeTouch* and had a higher error rate than *HandTouch*, but overall had a comparable performance to the other techniques in terms of the interaction accuracy of kinesthetic perception.

6.3 H3: *HandGazeTouch* and *GazeTouch* will cause less tiredness of the hand than *HandTouch*. However, *HandTouch* will be considered more natural and pleasant than the alternatives.

Our results partially support this hypothesis based on the subjective data and overall evaluation of the interaction techniques. Both gaze-based techniques lead to less hand fatigue. However, *HandGazeTouch* was still considered as natural and pleasant as *HandTouch*.

Based on the questionnaire data (Figure 8), there was no difference in overall physical difficulty among the three interaction techniques. However, gaze-based techniques caused less fatigue on the hand than the *HandTouch* technique, as is evident from the overall ranking of the conditions. Repetitive

hand-based force-feedback interaction can lead to fatigue of the hand [11]. Our results suggest that using the gaze to replace the hand motion can largely reduce the subjective perception of hand fatigue during the kinesthetic interaction. On the contrary, the use of the eyes as an input modality in gaze-based techniques led to increased fatigue of the eyes. *HandGazeTouch* led to less eye fatigue than *GazeTouch*. This may be because the eyes were used only as a "pointer" in *HandGazeTouch*, while the eyes were used both as a pointer and a mechanism for activation to cause haptic cues in *GazeTouch*. *HandTouch*, obviously, caused the least eye fatigue.

The same reason may also explain the mental difficulty of *GazeTouch*. In *GazeTouch*, the dual use of eye gaze to point and touch might also have induced a greater cognitive load. In *HandGazeTouch*, users needed to combine two senses (eye gaze and hand motion) as the kinesthetic input, which might incur additional cognitive load.

Furthermore, for the naturalness and pleasantness of gaze-based techniques, *GazeTouch* replaced both the reaching and touching operation with eye movements, which is very different from the way we interact with physical objects in the real world. On the other hand, humans are used to looking at an object prior to hand motions in physical tasks [2]. *HandGazeTouch* utilized our natural hand-eye coordination, and it is thus closer to real-world interaction than *GazeTouch*. This may explain our results in terms of the method's naturalness and pleasantness.

6.4 Limitations and Future Research

Our study has several limitations.

- We used a research-quality eye tracker and TraQuMe to ensure high-quality gaze data. We recalibrated the eye tracker when the offset in tracking in any screen area was above the threshold of 2 cm. While most of our participants did not have any problems in tracking accuracy (mean gaze offset value 0.62 cm), we noticed that despite of our best efforts, some of our participants still faced difficulties in interaction using the gaze-based techniques due to the accuracy of tracking. Overall, we think our eye tracking quality was good and may not be representative of the expected quality that could be anticipated in an everyday gaze-tracking scenario outside the lab (e.g. cheaper tracking hardware, frequent movement of the user, fewer user calibrations). Reduced gaze tracking accuracy will introduce additional complexities in the use of both the gaze-based techniques. For example, it would make gaze-based pointing more difficult, especially when the objects are small. Further research is required to better understand how the accuracy of tracking influences the use of the two gaze-based interaction techniques.

- Our experimental task only focused on softness discrimination. Other object properties, such as textural properties, were not part of this study. The performance of the three interaction techniques may turn out to be different in tasks that require discriminating the roughness or smoothness of object surfaces [48]. We propose to examine this in future research.

- Most of our participants had no experience in using gaze-tracking and force-feedback devices to interact with virtual objects. Our study involved a short-term evaluation, and therefore it provides little insight into the long-term use of the three interaction techniques. It is likely that our results on the mental and physical difficulty of the gaze-based interaction

techniques may be influenced by the participant selection. Previous work [49] on gaze-based human-computer interaction has shown that novice users suffer from eye fatigue associated with unnatural eye movements (such as staring). However, experienced users do not report any eye fatigue [50]. In addition, it is also likely that the effect of learning was differential for the three interaction techniques and related to the mental difficulty. A longitudinal study is needed to understand how users learn to use these techniques and how the user's opinion of the technique may change with extended use.

- The size of the display we used in the experiment was relatively small. We used a Tobii T60 as both the display and the eye tracker in the study. The width of the T60 screen is only 39.5 cm, which limited the maximum distance between the two skin models used in the experiment. A key question for future research is how the three interaction techniques will fare when the distance between the touch points are larger. Further work using large 2D displays or a VR head-mounted display is required to answer this question.

- The experiment involved predominantly young participants. Ruff and Parker [51] have shown that participants in different age groups have different performances in motor speed and hand-eye coordination. Older users are significantly slower than younger groups. Therefore, further research is required to understand the effect of participant selection in terms of the age-related aspects in our results.

- We studied two gaze-based interaction techniques that replaced different phases of the kinesthetic interaction with eye gaze input. This is by no means a complete exploration of the design space of combining gaze and hand-based input for kinesthetic interaction. For example, another way to augment kinesthetic interaction with gaze would be to use eye gaze for the large movement of the HIP, while allowing fine exploration using hand motions to provide better haptic cues from the object. Future studies should investigate other novel ways of augmenting kinesthetic interaction with gaze input.

We believe a key application area for gaze-augmented kinesthetic interaction would be in the VR environment. VR displays can provide an immersive and large 3D environment. Previous studies [16], [17], [18] have demonstrated the potential benefits of using gaze modality on target pointing and selection in such environments. For kinesthetic interaction, the 3D nature of the VR environment may introduce additional opportunities and challenges for gaze-based input, especially those associated with the depth of objects in space. In our study, the depth of the objects was fixed to one level. When there are multiple close objects of different depths, interacting with the objects becomes more complex, requiring the use of both conventional hand-based kinesthetic interaction and the gaze-augmented kinesthetic interaction that we presented in this paper. Previous research on 3D gaze estimation suggests that it is feasible to estimate the depth of visual focus based on the convergence of the individual's eyes [52]. Thus, gaze-augmented kinesthetic interaction for a 3D environment could potentially utilize the depth of focus as a method to control the position of the HIP along the z-axis. In *HandGazeTouch*, such an approach may enable a more robust and consistent interaction at different object depths. Future work could investigate this aspect.

7 CONCLUSION AND PRACTICAL IMPLICATIONS

This study explored the use of the eyes in kinesthetic interaction. We developed two kinesthetic interaction techniques based on the combination of gaze and hand-based input. Further, we conducted a comprehensive experimental study involving a softness discrimination task and analyzed the multi-faceted effect of the new interaction techniques on the efficiency of the interaction, the accuracy of kinesthetic perception, and the user's experience. Our results suggest that eye gaze as an input channel has both strengths and limitations in improving kinesthetic interaction. Below, we summarize our key findings and the practical implications of our results.

- Gaze-augmented kinesthetic interaction can help overcome two key limitations of conventional kinesthetic interaction: it reduces the fatigue of the hands and infinitely expands the workspace of the current haptic interface. Thus, gaze-augmented methods could be considered in kinesthetic interactions that involve a large interaction environment or involve sustained and repetitive actions.

- Utilizing gaze input as a mechanism to reach objects (*HandGazeTouch*) is better than using it for reaching and touching (*GazeTouch*) in gaze-augmented kinesthetic interaction. *HandGazeTouch* is not only faster than *GazeTouch*, but also more natural and noticeably more accurate in perceiving subtle differences in softness. Despite the limitations, *GazeTouch* is a feasible interaction technique, and may be suitable for specific disabled users.

- *HandGazeTouch* is comparable to conventional hand-only kinesthetic interaction (*HandTouch*) in terms of the accuracy, naturalness, and pleasantness of interaction. In addition, *HandGazeTouch* is more efficient than *HandTouch*, and it may thus be specifically suited to tasks that are time sensitive.

- *HandTouch* is a solid kinesthetic interaction technique that is considered natural, pleasant, and less cognitively demanding. *HandTouch* may be specifically suitable for interactions that are less frequent, non-repetitive, and involve a small haptic interaction space that is easy to navigate using hand motions.

- The suitability of the specific interaction techniques depends on the context of use. For example, for time-sensitive tasks, efficiency may be the key metric and, thus, *HandGazeTouch* may be the most suited of the three techniques. Similarly, for precision tasks, accuracy maybe more important, making *HandTouch* the best method in such contexts. On the other hand, for tasks that involve frequent and prolonged usage, the naturalness and pleasantness provided by both *HandTouch* and *HandGazeTouch* may be key. Our results suggest that it may be best to provide users with the flexibility to choose the interaction technique depending on the specific context of use.

REFERENCES

[1] M. Turk, "Multimodal interaction: A review," J. Pattern Recognition Letters, vol. 36, no. 1, pp.189-195. Jan. 2014. DOI: 10.1016/j.patrec.2013.07.003

[2] M. Land, N. Mennie and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *J. Perception*, vol. 28, no. 11, pp.1311-1328, Feb. 1999, doi:10.1068/p2935.

[3] M.C. Bowman, R.S. Johannson and J.R. Flanagan, "Eye-hand coordination in a sequential target contact task," *J. Exp. Brain Res.*, vol.

195, issue. 2, pp.273-283, May 2009, doi: 10.1007/s00221-009-1781-x.

[4] M.F. Land and M. Hayhoe, "In what ways do eye movements contribute to everyday activities?" *J. Vision Research*, vol. 41, issues. 25-26, pp.3559-3565, Nov.–Dec. 2001, doi: 10.1016/S0042-6989(01)00102-X.

[5] J.D. Crawford, W.P. Medendorp and J.J. Marotta, "Spatial Transformations for Eye–Hand Coordination," *J. Neurophysiol*, vol. 92, issue. 1, pp. 10–19, Jul. 2004, doi:10.1152/jn.00117.2004.

[6] SensAble Technologies, (n.d.). (Transferred to 3D systems). (2018, Oct 15). Available at https://www.3dsystems.com/

[7] NOVINT, (n.d.). (2018, Oct 15). Available at http://www.novint.com/.

[8] T.H. Massie and J.K. Salisbury, "The PHANToM Haptic Interface: A Device for Probing Virtual Objects," *Proc. of the ASME Dynamic Systems and Control Division*, vol. 1, DSC-Vol. 55-1, available at, https://alliance.seas.upenn.edu/~medesign/wiki/uploads/Courses/Massie94-DSC-Phantom.pdf, 1994

[9] S. Zhai, C. Morimoto and S. Ihde, "Manual and gaze input cascaded (MAGIC) pointing," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '99)*, pp. 246-253, May 1999, doi:10.1145/302979.303053.

[10] R. Ott, M. Gutierrez, D. Thalmann and F. Vexo, "Improving user comfort in haptic virtual environments through gravity compensation," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.401-409, Mar. 2005, doi:10.1109/WHC.2005.78.

[11] A. Hamam and A.E. Saddik, "User force profile of repetitive haptic tasks inducing fatigue," *2015 Seventh Int. Workshop on Quality of Multimedia Experience (QoMEX)*, May 2015, doi:10.1109/QoMEX.2015.7148127.

[12] TOBII, (n.d). (2018, Oct, 15). Available at https://www.tobii.com/

[13] L.E. Sibert and R.J.K. Jacob, "Evaluation of eye gaze interaction," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '00)*, pp.281-288, Apr. 2000, doi:10.1145/332040.332445.

[14] M. Kumar, A. Paepcke and T. Winograd, "EyePoint: practical pointing and selection using gaze and keyboard," *Conf. on Computer-Human Interaction (CHI) 2007*, pp.421-430, Apr/May 2007, doi:10.1145/1240624.1240692.

[15] P. Majaranta and KJ. Räihä, "Twenty years of eye typing: Systems and design issues," *Proc. of the 2002 Symp. on Eye Tracking Research & Applications (ETRA '02)*, pp.15-22, Mar. 2001, doi:10.1145/507072.507076.

[16] V. Tanriverdi and R.J.K. Jacob, "Interacting with eye movements in virtual environments," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '00)*, pp. 265-272, Apr. 2000, doi:10.1145/332040.332443.

[17] T. Piumsomboon, G. Lee, R.W. Lindeman and M. Billinghurst, "Exploring natural eye-gaze-based interaction for immersive virtual reality," *IEEE Symp. on 3D User Interfaces (3DUI)*, pp. 36-39, Mar. 2017, doi:10.1109/3DUI.2017.7893315.

[18] K. Pfeuffer, B. Mayer, D. Mardanbegi and H. Gellersen, "Gaze + pinch interaction in virtual reality," *Proc. of the 5th Symp. on Spatial User Interaction (SUI '17)*, pp. 99-108, Oct. 2017, doi:10.1145/3131277.3132180.

[19] K. Pfeuffer, J. Alexander, M.K. Chong and H. Gellersen, "Gaze-touch: Combining gaze with multi-touch for interaction on the same surface," *Proc. of the 27th Annual Symp. on User Interface Software and Tech. (UIST '14)*, pp. 509-518, Oct. 2014, doi:10.1145/2642918.2647397.

[20] K. Pfeuffer, J. Alexander, M.K. Chong and Y.X. Zhang, "Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze," *Proc. 28th Annual Symp. on User Interface Software and Tech. (UIST '15)*, pp. 373-383, Nov. 2015, doi:10.1145/2807442.2807460.

[21] S. Stellmach and R. Dachselt, "Look & Touch: Gaze-supported target acquisition," *Proc. of the SIGCHI conf. on Human Factors in Comput.*

[22] J. Kangas, D. Akkil, J. Rantala, P. Isokoski, P. Majaranta and R. Raisamo, "Gaze gestures and haptic feedback in mobile devices," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '14)*, pp.435-438, Apr/May 2014, doi:10.1145/2556288.2557040.

[23] D. Akkil, J. Kangas, J. Rantala, P. Isokoski, O. Spakov and R. Raisamo, "Glance Awareness and Gaze Interaction in Smartwatches," *Proc. of the 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Comput. Systems (CHI EA '15)*, pp.1271-1276, Apr. 2015, doi:10.1145/2702613.2732816.

[24] G.P. Mylonas, K.W. Kwok, A. Darzi and G.Z. Yang, "Gaze-contingent motor channelling and haptic constraints for minimally invasive robotic surgery," *Int. Conf. on Med. Image Comput. Comput. Assist. Interv. (MICCAI 2008)*, pp. 676-683, Sep. 2008, doi:10.1007/978-3-540-85990-1_81.

[25] D.R.C. James, F. Orihuela-Espina, D.R. Leff, G.P. Mylonas, K.W. Kwok, A.W. Darzi and G.Z. Yang, "Cognitive burden estimation for visuomotor learning with fNIRS," *Int. Conf. on Med. Image Comput. Comput. Assist. Interv. (MICCAI 2010)*, pp.319-326, Sep. 2010, doi:10.1007/978-3-642-15711-0_40.

[26] D.R.C. James, D.R. Leff, F. Orihuela-Espina, K.W. Kwok, G.P. Mylonas, T. Athanasiou, A.W. Darzi and G.Z. Yang, "Enhanced frontoparietal network architectures following "gaze-contingent" versus "free-hand" motor learning," *J. Neuroimage*, vol. 64, pp. 267-276. Jan. 2013, doi:10.1016/j.neuroimage.2012.08.056.

[27] R.H. LaMotte, "Softness discrimination with a tool," *J. Neurophysiol.*, vol. 83, no. 4, pp.1777-1786, Apr. 2000, doi: 10.1152/jn.2000.83.4.1777.

[28] T. Foulsham, "Eye movements and their functions in everyday tasks," *Cambridge Ophthalmological Symp. 2014*, Eye 29, pp.196-199, Nov. 2014, doi:10.1038/eye.2014.275.

[29] D.J. Liebling and S.T. Dumais, "Gaze and mouse coordination in everyday work," *Proc. of the 2014 ACM Int. Joint Conf. on Pervasive and Ubiquitous Comput.: Adjunct Publication (UbiComp '14 Adjunct)*, pp.1141-1150, Sep. 2014, doi:10.1145/2638728.2641692.

[30] F. Conti. And O. Khatib, "Spanning large workspaces using small haptic devices," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.183-188, Mar. 2005, doi:10.1109/WHC.2005.118.

[31] L. Dominjon, A. Lécuyer, J.M. Burkhardt, G. Andrade-Barroso, S. Richir, "The "bubble" technique: Interacting with large virtual environments using haptic devices with limited workspace," *First Joint Eurohaptics Conf. and Symp. on Haptic Interfaces for Virtual Envir. and Teleoperator Systems. World Haptics Conf. 2005*, pp.639–640, Mar. 2005, doi:10.1109/WHC.2005.126.

[32] T.J. Allen and U. Proske, "Effect of muscle fatigue on the sense of limb position and movement," *J. Exp Brain Res*, vol. 170, no. 1, pp.30-38, Mar. 2006, doi:10.1007/s00221-005-0174-z.

[33] N. Cortes, J. Onate and S. Morrison, "Differential effects of fatigue on movement variability," *J. Gait Posture*, vol. 39, no. 3, pp.888-893, Dec. 2013, doi:10.1016/j.gaitpost.2013.11.020.

[34] A.E. Saddik, M. Orozco, M. Eid, and J. Cha, "Haptics Technologies: Bringing Touch to Multimedia," *Springer Series on Touch and Haptic Systems*, M. Ferre, M.O. Ernst and A. Wing eds., Berlin: Springer-Verlag, 2011.

[35] J. Rantala, P. Majaranta, J. Kangas, P. Isokoski, D. Akkil, O. Špakov and R. Raisamo, "Gaze interaction with vibrotactile feedback: review and design guidelines," *J. HCI.*, Mar. 2017, doi:10.1080/07370024.2017.1306444.

[36] D.R. Leff, D.R. James, F. Orihuela-Espina, K.W. Kwok, L.W. Sun, G. Mylonas, T. Athanasiou, A.W. Darzi and G.Z. Yang, "The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills," *J. Front. Hum. Neurosci.*, vol. 9, issue. 526, Oct. 2015, doi:10.3389/fnhum.2015.00526.

[37] P.M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp.381–391, Jun. 1954, doi:10.1037/h0055392.

[38] D. Akkil, J.M. James, P. Isokoski and J. Kangas, "GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks," *Proc. of the 2016 CHI Conf. Extended Abstracts on Human Factors in Comput. Systems (CHI EA '16),* pp.1151-1158, May 2016, doi: 10.1145/2851581.2892459.

[39] S.G. Hart and L.E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *J. Advances in Psychology*, vol. 52, Human mental workload, pp.139-183, 1988, doi:10.1016/S0166-4115(08)62386-9.

[40] H3DAPI, (n.d.). (2018, Oct 15). Available at http://www.h3dapi.org/ .

[41] D. Akkil, P. Isokoski, J. Kangas, J. Rantala and R. Raisamo, "TraQuMe: A Tool for Measuring the Gaze Tracking Quality," *Proc. of the 2014 Symp. on Eye Tracking Research & Applications (ETRA '14)*, pp.327-330, Mar. 2014, doi:10.1145/2578153.2578192.

[42] R.J.K. Jacob, "The use of eye movements in human-computer interaction techniques: what you look at is what you get," *J. ACM Transactions on Information Systems (TOIS) - Special issue on CHI*, vol. 9, issue. 2, pp.152-169, Apr. 1991, doi:10.1145/123078.128728.

[43] J.O. Wobbrock, L. Findlater, D. Gergle and J.J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '11)*, pp.143-146, May 2011, doi:10.1145/1978942.1978963.

[44] S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scand. J. Statist.*, vol. 6, no. 2, pp. 65–70, Jan. 1979, doi: 10.2307/4615733.

[45] R.J.K. Jacob, "Eye Tracking in Advanced Interface Design," *Virtual environments and advanced interface design*, W. Barfield and T. A. Furness, eds., New York: Oxford Univ. Press, pp.258-288, 1995.

[46] G. Juravle, F. McGlone and C. Spence, "Context-dependent changes in tactile perception during movement execution," *J. Front. Psychol.,* 4: 913, Dec. 2013, doi:10.3389/fpsyg.2013.00913.

[47] S.J. Lederman and R.L. Klatzky, "Hand Movements: A window into haptic object recognition," *J. Cognitive Psychol.*, vol. 19, issue. 3, pp.342-368, Jul. 1987, doi:10.1016/0010-0285(87)90008-9.

[48] W.M.B. Tiest, "Tactual perception of material properties", *J. Vision Research*, vol. 50, issue 24, pp.2775-2782, Dec. 2010, doi:10.1016/j.visres.2010.10.005.

[49] P. Majaranta, U.K. Ahola and O. Špakov, "Fast Gaze Typing with an Adjustable Dwell Time," *Proc. of the SIGCHI conf. on Human Factors in Comput. Systems (CHI '09)*, pp.357-360, Apr. 2009, doi:10.1145/1518701.1518758.

[50] N. Chitty, "User Fatigue and Eye Controlled Technology," PhD dissertation, OCAD Univ. Toronto, Ontario, Canada, 2013.

[51] R.M. Ruff and S.B. Parker, "Gender- and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the finger tapping and grooved pegboard tests," *J. Percept. Mot. Skills*, 76 (3 Pt 2), pp.1219-1230. Jun. 1993, doi:10.2466/pms.1993.76.3c.1219.

[52] A.T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, Springer International Publishing, pp.77-81, 2017.
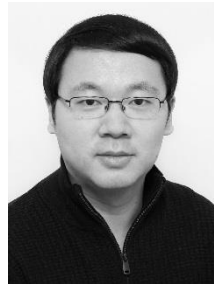
Zhenxing Li received his MSc degree in Radio Frequency Electronics from Tampere Univerisy of Technology and MSc degree in User Interface Software Development from Univeristy of Tampere, Finland, in 2009 and 2013 respectively. He currently is a researcher in Tampere Unit for Computer-Human Interaction (TAUCHI) research center at Tampere University and working toward PhD degree in Interaction Technology. His research interests include haptic interaction and rendering, multimodal interaction, gaze tracking, mathematical modelling and algorithms.

Deepak Akkil received his MSc degree in Human-Technology Interaction from University of Tampere in 2014. He is currently pursuing PhD degree in Interaction Technology at Tampere Unit for Computer-Human Interaction (TAUCHI) research center in Tampere University. His research interests include gaze tracking, haptic interaction, computer-supported cooperative work (CSCW) and user experience design.

Roope Raisamo is a professor of computer science in Tampere University, Faculty of Information Technology and Communication Sciences. He received his PhD degree in computer science from the University of Tampere in 1999. He is the head of TAUCHI Research Center (Tampere Unit for Computer-Human Interaction), leading Multimodal Interaction Research Group. His research is focused on multimodal interaction, XR, haptics, gaze, gestures, interaction techniques, and software architectures.