Spatial-Temporal Characteristics

of Multisensory Integration

by

Sung Hun Sim

A Dissertation Presented in Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

Approved September 2017 by the

Graduate Supervisory Committee:

Bing Wu, Chair

Nancy Cooke

Robert Gray

Russell Branaghan

ARIZONA STATE UNIVERSITY

December 2017

ASTRACT

We experience spatial separation and temporal asynchrony between visual and haptic information in many virtual-reality, augmented-reality, or teleoperation systems. Three studies were conducted to examine the spatial and temporal characteristic of multisensory integration. Participants interacted with virtual springs using both visual and haptic senses, and their perception of stiffness and ability to differentiate stiffness were measured. The results revealed that a constant visual delay increased the perceived stiffness, while a variable visual delay made participants depend more on the haptic sensations in stiffness perception. We also found that participants judged stiffness stiffer when they interact with virtual springs at faster speeds, and interaction speed was positively correlated with stiffness overestimation. In addition, it has been found that participants could learn an association between visual and haptic inputs despite the fact that they were spatially separated, resulting in the improvement of typing performance. These results show the limitations of Maximum-Likelihood Estimation model, suggesting that a Bayesian inference model should be used.

Keywords: visual-haptic, stiffness, temporal delay, spatial separation, interaction speed

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Our brain continuously receives information from the environment through the five senses and combines these inputs into a unified experience. Multiple sensory inputs can come at the same time. For instance, when we push a pillow with our hand, we see the deformation of the pillow, and feel the softness of the pillow by touch. How information is integrated across sensory channels is the focus of this research. Particularly, two senses will be covered, vision and touch. They might be the two most important senses for us: It has been estimated that we gain approximately 80% of information by seeing (Woodson & Conover, 1964), whereas touch plays a critical role in control of our action. More specifically, the topic will focus on the perception of stiffness by integrating touch and vision.

Stiffness is a mechanical property that describes how strongly an object resists being deformed (Fung, 1981). It is a higher-order percept and no sensory organ or receptor can directly perceive stiffness, whereas location and size can be directly perceived. Stiffness can be perceived through active interaction, and it is related to feedback for motor planning. Accurate perception of stiffness is required in many tasks. In medicine, for example, breast cancer tumors can be discovered by palpation based on changes in the felt stiffness. In surgery, surgeons use the perceived stiffness to plan and perform delicate procedures with high precision. A better understanding of how visual and haptic (i.e., active touch, see Chapter 2.3.1) cues interact to produce a perception of stiffness will ultimately feed into clinical practice to facilitate training of surgeons and improve their performance.

It can also be applied to industrial and engineering applications, for example, the development of multimodal interfaces.

The Maximum-Likelihood Estimation (MLE) model is the commonly used model to describe the processing of inputs during multisensory perception. In the MLE model, the inputs' reliability determines the degree of weight given to the inputs. It does not consider other factors such as 'what the source is' and 'where it comes from'. Important factors that influence the integration of visual and haptic information are temporal synchrony and spatial coincidence (Helbig & Ernst, 2007). In the real world, we see and touch an object at the same time, and there is no spatial separation between the seen and the felt object. However, temporal asynchrony between visual and haptic information is created by visual delays in virtual-reality (VR), augmented-reality (AR), and teleoperation applications. Also, spatial incongruence between vision and touch is inevitable in minimally-invasive surgery (MIS) in which surgeons rely on a distant monitor for visual information. So far, only a few studies investigated how these factors influence our stiffness perception, therefore, further exploration is required.

This paper is structured as follows. Chapter 2 provides background for this paper. Chapter 3 presents how visual latency affects visual-haptic perception of stiffness. Chapter 4 investigates the effects of visual delay and interaction speed on stiffness perception. Chapter 5 examines if participants could learn an association between visual and haptic inputs with spatial separation, and if this learning could facilitate participants' typing performance. Chapter 6 summarizes this paper, and moreover, provides applications and future directions.

CHAPTER 2

BACKGROUND

This chapter reviews some recent experimental findings with respect to stiffness perception and then presents some models that describe the integration of visual and haptic information. Chapter 2.1 briefly introduces the definition of stiffness. Chapter 2.2 presents some psychophysical methods commonly used to measure the perception of stiffness. Chapter 2.3 summarizes the experimental findings about human abilities to estimate and discriminate stiffness by touch, vision or both senses. The modeling of multisensory stiffness perception is described in Chapter 2.4. Chapter 2.5 concludes this chapter.

2.1 Physical aspects of stiffness

Stiffness is a property of all solid objects. It is the extent that it resists deformation in response to an applied force (Fung, 1981). Although the definition sounds very technical and abstract, the perception of stiffness is not. For example, when you firmly press the surface of a mattress with your hand and see how much you can push down, you are feeling the stiffness of the mattress: The deeper you can push down, the less stiff it is. By definition, stiffness is measured as the ratio of the applied force with respect to the corresponding deformation it produces. Take a spring as an example, its stiffness is given by Hooke's law as

$$k = f / d, \tag{1}$$

where k is the stiffness, f is the force applied to the spring, and d is the resulting deformation or displacement produced by the applied force f. The unit of stiffness is N/m in the

International System of Units (SI). The inverse of stiffness is called "compliance", "flexibility", or "softness" with a unit of m/N.

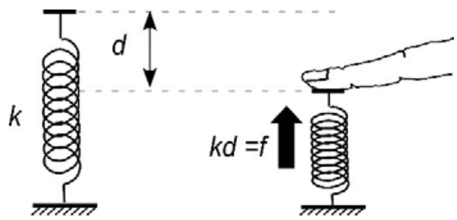


Figure 1. Stiffness (*k*) refers to the spring-like behavior of an object, which can be estimated from perceived force (*f*) and deformation (*d*).

Although stiffness is sometime used as interchangeable with the term "elasticity", strictly speaking, the two terms refer to different concepts. Elasticity is the property of a given material. It is characterized as the ratio of stress (i.e., force per unit area) to strain (i.e., deformation per unit length) and known as Young's modulus. The elasticity of a material is often measured by tensile tests. Its SI unit is Pa (i.e., N/m^2). On the other hand, stiffness refers to the spring-like behavior of an object. It is determined by the material(s) of which the object is comprised as well as by the object's geometry and internal structure. As to the stiffness of an object, we perceive it by touch and also vision.

2.2 Psychophysical measures of stiffness perception

Human perception of stiffness has been extensively studied using both neurophysiological and psychophysical approaches. In this chapter, I will briefly review some common psychophysical methods used to measure how well people can differentiate and estimate stiffness (for a review, see Kingdom & Prins, 2010).

Consider first the Just-Noticeable-Difference (JND, or differential threshold, or difference limen). JND indicates the minimum difference that people can barely detect between two similar stimuli. According to the Weber's law (Weber, 1834/1996), the JND would be approximately proportional to the magnitude of the reference. Therefore, JNDs are often expressed as a percentage or proportion of the reference (i.e., Weber fractions) in literature. However, research has shown that Weber's law does not quite hold for stiffness (see Jones (2000) for a review; see also Chapter 2.3). The reported JND varies significantly depending on the range of the reference stiffness, how the participant interacts with the stimuli, and the availability of sensory cues. These experimental findings of JND will be reviewed in Chapter 2.3.

The following three methods are commonly used to measure JNDs: the method of adjustment, the method of limits, and the method of constant stimuli (Marks & Gescheider, 2002; Kingdom & Prins, 2010). When using the method of adjustment, the participant is presented with two stimuli, a reference and a comparison that is obviously different from the reference. Then the participant is asked to adjust the comparison to make it feel equal to the reference. The results tend to be variable because of the influence of biases. In order to reduce the variance, multiple descending/ascending adjustment sessions are usually conducted, and the average is adopted as the JND.

The method of limits is the same as the method of adjustment, except that no manual adjustment of the comparison is performed by the participant, eliminating some sources of response bias. In this method, the participant decides whether the comparison is stronger or softer than the reference, and then the comparison is accordingly adjusted with small and predetermined steps. For example, in the staircase limits method (Dixon & Mood, 1948; Kaernbach, 2001), the difference between comparison and reference is reduced every time a correct judgment is made, or increased for an erroneous judgment, forcing the comparison to eventually converge to a threshold.

Among the three methods, the method of constant stimuli is most accurate. In this method, a set of comparison stimuli is employed that are equally spaced and range from obviously-softer to obviously-stiffer than the reference. The comparison and the reference are presented in pair to the participant multiple times, and the order of presentations is randomized. On each trial, the participant judges if the comparison is stronger (or softer) than the reference. The proportion of positive responses (i.e., the "yes" response) is calculated for each comparison stimulus and plotted against the stiffness values. Such data is then fitted by a psychometric curve using the cumulative Gaussian distribution. The JND is calculated as the stimulus difference between the 50%-chance performance and the 75%-level performance (Wickens, 2001).

To study how the perceived stiffness grows with the physical stiffness, the participant's perception is often assessed using magnitude estimation methods (Marks & Gescheider, 2002). In such experiments, the stimuli are springs, material samples, or virtual stimulations that have a wide range of stiffness. Participants are required to feel the stimuli, one at a time, and then give numerical estimates of the felt stiffness. The task may be done with or without the presence of a standard stimulus. If a standard stimulus is present, ratio magnitudes with respect to the standard are obtained. In this case, Hellman and Zwislocki (1961) suggested that the standard stimulus be chosen by the participant, instead of being designated by the experimenter, so as to avoid subjective biases. If the standard is omitted, absolute magnitudes are obtained. Participants are asked not to compare the present stimulus to previous ones, and a subsequent normalization process is needed to remove the variation due to participants' use of different scales (Stevens, 1975).

2.3 Perception of stiffness by touch, vision, or both

2.3.1 Haptic perception of stiffness

The term "*haptics*" comes from the Greek, meaning "*to touch*". According to Loomis and Lederman (1986), the sense of touch involves two neurologically and functionally distinct systems: one is the cutaneous system that senses stimulations from the external world using the receptors in the skin, and the other is the kinesthetic system that senses the position, motion and tension of the body and body parts using the receptors in muscles, tendons, and joints. Haptic perception involves both systems. It refers especially to the experience obtained via *active touch* (c.f., being passively touched) where the perceiver can manipulate or interact with an object or environment under his or her own control (Lederman & Klatzky, 2009; Klatzky & Lederman, 2002). Thus, haptic experience is based on the integration of the afferent inputs from the cutaneous and kinesthetic receptors, as well as an efferent input from the motor cortex that produces the exploratory actions.

Stiffness is primarily perceived through haptic interaction by actively applying a force to the target object and detecting how it responds. Klatzky, Lederman, and Reed

(1989) suggested that "the haptic perceptual system makes use of stereotyped motor patterns," termed *exploratory procedures*, to sense different object properties. For stiffness, the optimal exploratory procedure is tapping or squeezing: we usually squeeze an object with hands or poke it with a finger or a rigid tool to feel how soft or rigid it is. It is believed that such an exploratory procedure can maximize the sensory inputs corresponding to the object property to be perceived and increase ease of encoding.

Our ability to perceive stiffness by touch, however, is quite poor. For example, Cholewiak, Tan, and Ebert (2008) reported that only 2-3 levels of stiffness, that is, soft, strong, and possibly a middle level, could be reliably identified for stiffness between 0.2 and 3.0 N/mm, equivalent to a capacity of transmitting only 1.46 bits of information. The difference threshold, as measured by the JND, could be as high as 99%, and the exact JND value was significantly influenced by factors like the range of stimuli, the methods for measuring judgments, how the participant interacted with the stimuli, and the richness of haptic cues (Tan et al., 1992, 1993, 1995; Jones, 2000). Roland and Ladegaard-Pedersen (1977) reported a JND of ~17% for participants to finger-grasp and compare the stiffness of springs that were enclosed in cylinders with rigid ends. Jones and Hunter (1990) used a contralateral limb matching procedure and found an average JND of 23% for comparing the stiffness of simulated springs.

In clinical procedures, limits in stiffness perception are indicated by the low sensitivity in palpation screening. Palpation is the most common method for self-examining and diagnosing breast cancer: By pressing on the breast, physicians try to distinguish tumors from normal tissue via the subtle differences in displacement and resisting force. Although cancer can be up to 7 times as stiff as normal tissue (Sarvazyan et al., 1995), the tumor detection rate is only about 39-59% by palpation (Shen & Zelen, 2001), depending on the proximity of the tumor to the body surface, the density of normal breast tissue, and also the physician's experience. Tumors which are as small as 2-3 mm in diameter can be detected after physicians get sufficient training (Adams, Hall, & Pennypacker, 1976; Bloom et al., 1982). However, such an increase in detection rate is often accompanied with an increase in false positives (Campbell et al., 1990). That is, the doctor is prone to diagnose more breast anomalies as tumors.

The low sensitivity in stiffness perception might be attributable to the fact that it is a higher-order percept derived from the perceived force and deformation (Klatzky & Wu, 2014). Research has shown that the change in force that a person can reliably discriminate is 7-10% over a range of 0.5-200 N (Jones, 2000). The differential thresholds measured for limb movement, position and speed are between 5-8% (Jones, Hunter, & Irwin, 1992; Tan et al., 1995, McKee, 1981, Orban, de Wolf, & Maes, 1984). When stiffness is judged by relating the perceived deformation to the perceived force or vice versa, such process causes a significant reduction in perceptual resolution and an increase in JND, as illustrated in a study reported by Tan et al. (1995). In this study, the participants were asked to hold and squeeze a virtual spring (two rigid plates driven by a computer-controlled linear motor) using their thumb and index finger. The stiffness JND was found to be ~22% when the displacement varied across trials and thus both force and displacement had to be judged in order to estimate the stiffness. In contrast, the JND reduced to ~8% in a fixed-displacement condition when the task could be reduced to a force-discrimination task.

With respect to subjective stiffness, the perceived magnitude of stiffness is found to be best related to physical stiffness by a power function. Harper and Stevens (1964) asked their participants to rate the stiffness of rubber samples and reported a compressive growth curve of estimation with an exponent coefficient of 0.8. In contrast, Varadharajan, Klatzky, Unger, Swendsen, and Hollis (2008) used simulated 3D virtual springs (12.0 N/mm to 48.0 N/mm) as the stimuli and found a linear relationship between the estimated and rendered stiffness.

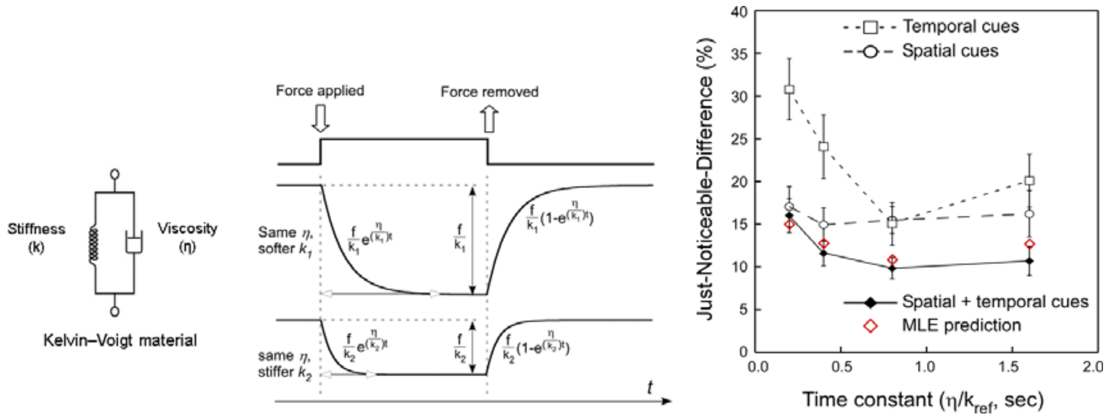
2.3.2 Visual perception of stiffness

One may think it is impractical to visually perceive stiffness because no visual cues seem informative with regard to force. But contrary to such intuitions, force can be visually judged. Michaels and De Vries (1998) asked their participants to watch videos, in which an actor pulled a handle with different forces without moving his feet, and then the participants guess the forces. High correlations were found between the participants' visual judgments and the actual forces exerted by the actor. Such feat might be accomplished through learning (White, 2012). Because our actions and experiences of force normally occur in the context of visual experiences, the haptic and visual experiences have become associated in long-term memory. Thus kinematic features in a visual percept can be matched to the stored haptic experiences and used to infer the interaction force. In addition, some visual cues might be directly associated with force. For example, a pressure at the finger pad affects blood flow and causes a change of color in the fingernail area. By analyzing such color responses, Sun, Hollerbach, and Mascaro (2008) report that the force applied to the fingertip can be estimated up to 10 N with an accuracy of 5-10%. Although it remains undecided how effective these visual cues are, we can safely assume that they provide at least heuristic information for estimating forces when perceiving stiffness.

Not surprisingly, stiffness can be visually estimated. Drewing, Ramisch, and Bayer (2009a) asked their participants to watch how another person pressed a soft rubber specimen with his or her index finger and then estimate the specimen's stiffness. The estimated stiffness was well correlated with the physical stiffness ($r^2 = 0.67$). But as compared to the touch-based estimates, the specimens were judged to be ~15% softer by vision.

The pattern of deformation may be particularly informative for the perception of stiffness. Deformation can be perceived by both touch and vision, and vision usually plays a more important role because of rich visual depth cues. With some knowledge about the applied force from perception, cognition, or both, it is possible to visually judge the stiffness from the amount of deformation perceived. Wu, Klatzky, Hollis and Stetten (2012) presented to their participants the simulated ultrasound videos that depicted the compression of virtual materials under a constant load. They found that the judgments of stiffness were significantly influenced by the visual quality of the simulations, with the JNDs ranging from 12% to 17% for different levels of noise generated by the imaging process and the regularity of the material's structure.

In addition, the perception of stiffness can be improved by using both spatial and temporal cues in the deformation pattern. Essentially, all solid materials have elastic and also viscous (i.e., resistance to change over time) properties, particular for biological materials (Fung, 1981). When such an object is compressed (or released), the compression (or rebound) gradually builds up, and the time-varying change is determined by both elasticity and viscosity, as shown in Figure 2(a). Wu, Sim, Hibbard, and Klatzky (2014) used Kelvin-Voigt models to simulate the behavior of virtual materials that had same viscosity but slightly different stiffness. In the temporal-cues condition, the simulations were manipulated so that the amount of deformation was identical but the duration of deformation varied with stiffness. The observed JNDs were 15-31%, and higher than those obtained in the spatial-cues condition, suggesting that the temporal cues are less effective than the spatial ones. When both types of cues were presented, the observed JND of stiffness reduced to ~10%. More importantly, such a reduction in JND could be explained very well by a linear combination of the temporal- and spatial- JNDs using a Maximum Likelihood Estimation (MLE) model. The integration process and the model will be further discussed in Chapter 2.4.



(a) (b)

Figure 2 (a). The linear viscoelastic behavior of Kelvin-Voigt materials. Wu, Sim, Hibbard, and Klatzky (2014) analyzed the time-varying changes and suggested that both spatial (e.g., the amount of deformation) and temporal (e.g., the duration of deformation) could be effective cues to stiffness. (b) JNDs obtained in the temporal-cues, spatial-cues, and spatiotemporal-cues conditions. Temporal and spatial cues were found to contribute to perception of stiffness collectively and could be linearly integrated in a statistically optimal fashion as described by a Maximum-Likelihood Estimation (MLE) model.

To briefly summarize, although the sense of touch is the predominant basis for perceiving stiffness, vision also offers important cues to the perception of force, deformation, and stiffness. The question is then, how well we can perceive stiffness by using both haptic and visual cues, and how these cues are integrated within and across sensory modalities.

2.3.3 Haptic-visual perception of stiffness

In theory, the perception of stiffness should be facilitated by the addition of visual information to haptic exploration. However, mixed results have been reported due to differences in the experimental tasks, the richness of visual and haptic cues, and the coherence and synchrony of sensory inputs. Varadharajan et al. (2008) investigated the haptic and visual contributions to stiffness perception. Using a haptic force-feedback device along with a visual display, they created a high-fidelity simulation of 3D virtual springs and used it to measure how well people could estimate and discriminate stiffness in the haptic-only and haptic-visual conditions. Their results revealed a significant impact of visual inputs on stiffness discrimination: The average JND was 14.2% in the haptic-visual condition, and it increased to 17.2% in the haptic-only condition where vision was eliminated. However, no significant contribution of vision to the judgments of stiffness magnitude was found in their magnitude-estimation experiment. The estimated stiffness was found to be linearly correlated with the rendered stiffness, and almost identical estimates were obtained in the haptic-only and haptic-visual conditions. Couroussé, et al. (2006) also reported similar results. They found that additional visual information did not assist people in estimating stiffness, but it induced changes in the pattern of haptic exploration. The amplitude of hand movement, the mean manipulation speed, and the amount of exerted force were significantly reduced when visual feedback was available. In contrast, Drewing, Ramisch, and Bayer (2009b) found significant impact of visual information on the perceived stiffness but no effects on stiffness discrimination. When estimating the stiffness of soft rubber specimens under a visual-haptic condition, participants' judgments were shifted half-way from haptic estimates towards visual estimates: The slopes of the linear regressions performed on the estimated and physical stiffness were 0.84, 1.02, and 0.90, respectively, for visual, haptic, and bimodal estimates. Instead, no significant improvement was observed in stiffness JND after the addition of visual information. Although the above-cited studies reported inconsistent results, we may still conclude that the brain uses visual information in different ways when performing different tasks.

Sometimes, vision is even found to dominate touch. Srinivasan et al. (1996) asked their participants to compare two springs that were rendered in a virtual visual-haptic environment. The relationship between visual and haptic stiffness was experimentally manipulated. The visual deformation may be consistent with the haptic stiffness felt by the participants (i.e., a small deformation was depicted in conjunction with a haptically hard spring, and a large visual deformation for a soft spring). Or the visual experience was contradictory to the haptic experience (i.e., a large deformation was rendered for a haptically hard spring and a small deformation for a haptically soft spring). They found that participants' judgments primarily corresponded to what was seen rather than what was felt by touch.

2.4 Modeling the multisensory perception of stiffness

Several models have been proposed to provide a framework for understanding how visual and haptic information is integrated in the perception of stiffness. I will start with the Maximum-Likelihood Estimation (MLE) model, which integrate the inputs from different sensory modalities by averaging them with weights that are assigned on the basis of the inputs' reliability. Next I will present two processing models that explain specifically how visual and haptic information is integrated in stiffness perception. Lastly, I will briefly review how stiffness perception is influenced by the spatiotemporal incongruence among the visual and haptic inputs, and present a work that explained such effects using the MLE model.

2.4.1 The Maximum-Likelihood Estimation (MLE) model

In the literature, the linear combination model is commonly used to describe the processing of information during multisensory perception. Specifically, the weighted averaging of visual and haptic information can be expressed as

$$S_{vh} = w_v S_v + w_h S_h \,, \quad \text{where } w_v + w_h = 1. \tag{2}$$

Here S_v/S_h denote respectively the visual and haptic inputs, and w_v/w_h are the weights assigned to them. S_{vh} is the integrated multisensory estimate, which is a perceptual continuum ranging from "vision-based" at one end, to "bimodal" in the middle, and to "touch-based" at the other end, depending on how the weights are assigned to the two senses. Then the question is, how does our brain decide the weights?

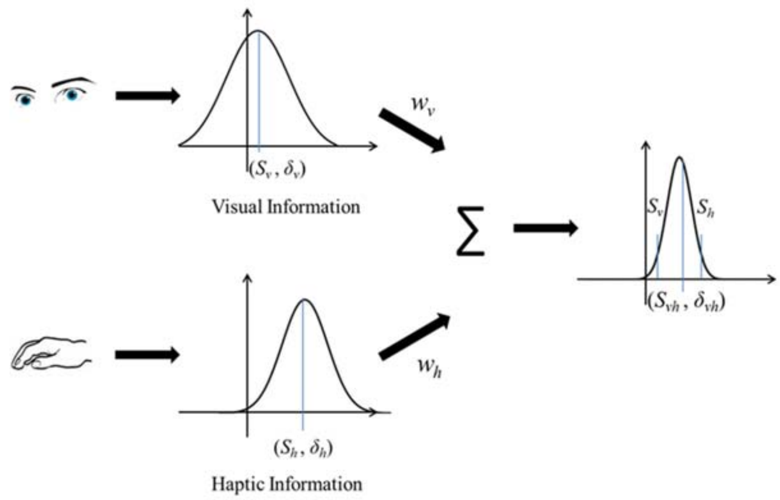


Figure 3. Maximum-Likelihood Estimation (MLE) model of visual-haptic integration. See text for details.

In the MLE model, the inputs are weighted by their reliability. That is, a more reliable source of information will be given a greater weight. For simplicity, let us assume that the visual and haptic inputs are two statistically independent variables with Gaussian distributions. The reliability of each input is defined as the reciprocal of its variance. The weights in the MLE are

$$w_v = \frac{\frac{1}{\delta_v^2}}{(\frac{1}{\delta_h^2}+\frac{1}{\delta_v^2})} = \frac{\delta_h^2}{(\delta_h^2+\delta_v^2)}, \quad w_h = \frac{\frac{1}{\delta_h^2}}{(\frac{1}{\delta_h^2}+\frac{1}{\delta_v^2})} = \frac{\delta_v^2}{(\delta_h^2+\delta_v^2)}. \qquad (3)$$

It has been proved that the combined estimate using such weights has the *minimum* variance (Cochran, 1937), which equals to

$$\delta_{vh}^2 = \frac{\delta_h^2\delta_v^2}{(\delta_h^2+\delta_v^2)}\ , \text{ or } \ \frac{1}{\delta_{vh}^2} = \frac{1}{\delta_v^2} + \frac{1}{\delta_h^2}\ . \qquad (4)$$

Minimum variance means the best perceptual ability to discriminate between stimuli. That is why the MLE model is considered as the optimal way to integrate the sensory inputs. [NOTE1]

[NOTE 1] It is safe to assume that statistical independence holds for the signals from different sensory modalities. For the within-modality cues, they may be encoded as de-correlated inputs in the sensory systems to maximize the efficiency of information transmission (Barlow & Földiák, 1989; Barlow, 2001). Even if correlations exist, the MLE model still holds but the weights need to be calculated in a slightly different way by taking into account the covariance (Oruç, Maloney, & Landy, 2003), and the outcome is still an optimal integration with the minimum variance. Another concern may be how the reliability is assessed by the nervous system. It has been suggested that the reliability may be implicitly encoded using neural population codes (Ma, Beck, Latham, & Pouget, 2006; Fetsch, DeAngelis, & Angelaki, 2013).

Two testable predictions are derived from the MLE model. First, since JNDs are defined to be the standard deviation of the noise, Equation (4) can be rewritten as $\frac{1}{JND_{vh}^2} = \frac{1}{JND_v^2} + \frac{1}{JND_h^2}$. That is, the JND of the multimodal perception can be estimated from the unimodal JNDs, and it should be no greater than any unimodal JND. Second, the multimodal percept will be estimated from the unimodal percepts using Equation (2) and the weights calculated using Equation (3).

These predictions have been experimentally validated in several studies. Ernst and Banks (2002) demonstrated that the integration of visual and haptic information in judging the size of an object follows such an MLE model. In their experiments, participants grasped a virtual object with their thumb and index finger and felt its size in the touch-only condition. The haptic JND was found to be ~8.5%. In the visual-only condition, the virtual object was rendered using stereoscopic images, and different levels (0, 67, 133, or 200%) of noise were added to vary the reliability of visual signals. Accordingly, the visual JND was found to increase from ~4% to ~20% with increased levels of noise. In the visual-haptic condition, the JND was significantly reduced by the additional visual information even if it was noisy. Moreover, the observed visual-haptic JNDs matched well with the predictions made from the unimodal visual and haptic JNDs using Equation (4). The perceived object size, as measured by the Point of Subjective Equality, also gradually shifted from the visual estimate to the haptic estimate as the visual noise increased, consistent with the predictions of the model. In other relevant studies, it has been demonstrated that the MLE model holds for processing other types of sensory information, for example, the integration of visual and vestibular information (Gharahmani, Wolpert, &

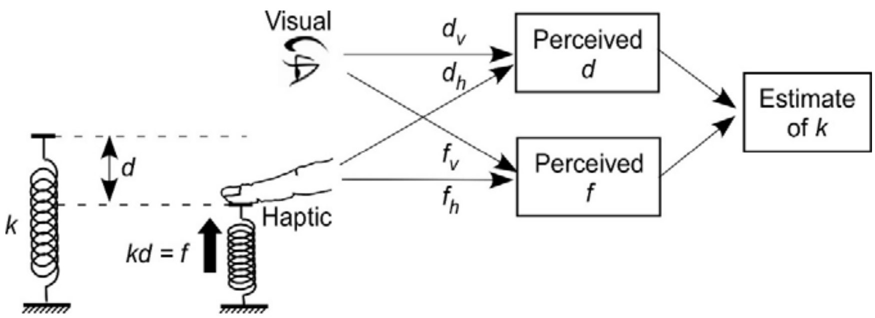
Jordan, 1997; Butler, Smith, Campos, & Bülthoff, 2010), the integration of visual and proprioceptive information (van Beers, Sittig, & Denier van der Gon, 1998, 1999), and the integration of visual and auditory information (Ghahramani, Wolpert, & Michale, 1997; Knill & Saunders, 2003; Alais & Burr, 2004; Hillis, Watt, Landy, & Banks, 2004).

The MLE model can also be used to describe the integration of information within a sensory modality, for example, the combination of visual depth cues like texture, motion, and stereo disparity (Young, Landy, & Maloney, 1993; Johnston, Cumming, & Landy, 1994; Landy & Kojima, 2001; Knill & Saunders, 2002). It also holds for the visual judgment of stiffness using the temporal and spatial cues. As illustrated in Figure 2 (b), the observed JNDs in the both-cues conditions (black filled diamonds) were well predicted by the integration of temporal- and spatial- JNDs using a MLE model (red open diamonds).

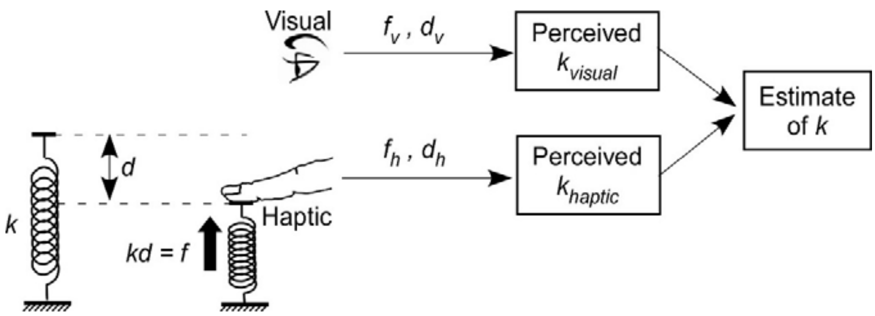
2.4.2 The processing models of visual-haptic integration in stiffness perception

It is important to note that all above-cited work investigated the information processing underpinning the direct perception of object or environment properties from sensory inputs. In contrast, stiffness is a higher-order percept that is related to two component properties, displacement and force. The natural questions to follow are: Can the MLE model be applied to the multisensory perception of stiffness? If so, how are haptic and visual cues processed and combined into the final percept of stiffness?

Figure 4 illustrates two models of the visual-haptic perception of stiffness. The *component-based processing* model suggests that the multisensory integration occurs in the perception of each component, which is deformation and force in this case. As illustrated in Figure 4(a), the visual and haptic cues are integrated, presumably, by using the MLE rule so that displacement and force can be better perceived. Stiffness is then estimated from the perceived force and displacement (Srinivasan, et al., 1996). Alternatively, the *modality-based processing* model suggests that sensory information constituting the component variables is processed in a modality-specific manner (Kuschel, Di Luca, Buss, & Klatzky, 2010), and then multi-sensory processing follows, as depicted in Figure 4(b). That is, stiffness may be directly perceived by touch and vision as an invariant in changing stimulation (Gibson, 1979; Walker et al. 1980; Kuschel, Buss, Freyberger, Farber, & Klatzky, 2008), and the two estimates are then integrated into a single percept.



(a)



(b)

Figure 4. Two processing models of visual-haptic perception of stiffness. (a). The *component-based processing* model suggests that force and deformation are estimated multimodally and then used in the estimation of stiffness. (b). The *modality-based processing* model suggests that stiffness is estimated in a unimodal way by vision or touch, followed by a process of cross-modal integration.

Kuschel et al. (2010) tested these two models using a visual-haptic virtual reality system. In their experiment, the stimuli were virtual cubes with different stiffness, which could be seen and pinched by participants with their thumb and index finger. In the control condition, both visual and haptic cues were provided to participants for estimating force and deformation: By vision, deformation (d_v) could be judged from the distance between two spheres representing the squeezing digits, and force (f_v) from the cube's surface curvature. By touch, displacement (d_h) could be felt through finger motion, and force (f_h) through cutaneous and kinesthetic receptors. In the cue-reduced experimental condition, the participant passively felt the forces exerted by two haptic devices while keeping the fingers fixed in their original place. Visually, the cube was seen to be squeezed by an amount calculated using a spring model, but no surface deformation was rendered. By this way, some cues to force and deformation were eliminated: For the judgment of deformation, the haptic cues (d_h) were removed by restricting the motion of fingers so that deformation could only be visually perceived (d_v). For the judgment of force, no visual cues (f_v) were available so that force could only be haptically perceived (f_h). Importantly, the eliminated cues were non-dominant (i.e., visual cues of force and haptic cues of deformation), which should produce weak effects on the judgments of force and deformation.

As to participants' performance in the cue-reduced condition, the two models gave very different predictions. According to the *component-based processing* model, the elimination of non-dominant cues should have little impact on the perception of stiffness because force and deformation could still be judged, respectively, by remaining haptic and visual cues. Given that touch is the predominant basis for estimating force and vision for deformation, one would expect that stiffness perception would not change too much as compared to the control condition. In contrast, according to the *modality-based processing* model, the cue-reduced condition should essentially preclude stiffness from being perceived. This is because both deformation and force are needed to compute the within-modality stiffness estimate. In the cue-reduced condition, visual cues are available for deformation but not for force, while haptic feedback provides only information of force but no finger displacement. Therefore, one would expect a significant increase in stiffness JND in the cue-reduced condition.

The results of the experiment agreed with the prediction of the *modality-based processing* model: The stiffness JND was 29% in the control condition but it almost tripled (83%) in the cue-reduced condition. These findings suggest that two estimates of stiffness are first obtained separately by touch and vision, and then combined into a multi-modal percept by an integration process.

The next question is, does the integration process follow the maximum-likelihood rule? To answer this question, Kuschel et al. (2010) conducted another experiment, in which the visual rendering of displacement was purposely distorted and was set to be double, equal, or half the haptic displacement. The visual cues to force (i.e., the surface curvature) remained unchanged. Thus, the visual estimate of stiffness would be doubled (or halved) for the halved (or doubled) displacement. Then when the visual and haptic estimates of stiffness were linearly integrated using Equation (2), the final percept would shift in the same direction as the visual estimate. These patterns were observed. But the weights were found not to be determined solely by the variances as in the MLE model. According to Equation (4), the bimodal JND should be no greater than any unimodal JND. This was actually violated in the magnified-visual-displacement condition: The stiffness JND was 26%, much higher than 17% in the haptic-only condition. Further analysis has showed that the weights used to integrate the visual and haptic inputs were optimal only in the no-visual-distortion condition.

Drewing, Ramisch, and Bayer (2009) reported similar findings. When estimating the compliance of soft rubber specimens under a visual-haptic condition, participants' judgments were shifted half-way from haptic-only estimates towards vision-only estimates. However, the reliability of judgments, as measured by the standard deviations of individual's estimates, was not improved by the addition of visual information. To sum up, these studies suggest that visual and haptic cues are collectively used in stiffness perception, but when the two inputs are not in congruence, they may not be integrated in an optimal fashion as described by the MLE model.

2.4.3 Impact of visual-haptic incongruence on stiffness perception

From a different perspective, neurophysiological research has provided us with important insights into how the brain combines information from different senses. Based on their studies of the visual-auditory neurons in the superior colliculi, Stein and Meredith (1993) suggested several rules for multisensory integration. One is the *spatiotemporal* rule which states “the spatial register among the receptive fields of multisensory neurons and their temporal response properties provide a neural substrate for enhancing responses to stimuli that co-vary in space and time and for degrading responses that are not spatially and temporally related” (p. 172). These rules provide a conceptual framework for studying how multisensory perception is influenced by the incongruence among sensory inputs. Previous research has revealed the existence of spatial and temporal gradients in multisensory response. For example, a strong Rubber Hand Illusion (i.e., the misperception of a rubber hand as part of one’s own body. See Botvinick & Cohen (1998), Tsakiris & Haggard (2005) for reviews) can be achieved if the latency between the vision and tactile stimulations is less than 300 ms. The illusion fades away as the latency increases beyond 500 ms (Shimada, Fukuda, & Hiraki, 2009). It also reduces significantly if the distance between the real and fake hands is larger than 30 cm (Lloyd, 2007). Such spatial and temporal effects may be explained by the responses of the bimodal visual-tactile neurons in the premotor and posterior parietal regions, which are believed to be involved in the Rubber Hand Illusion (Gentile, Petkova, & Ehrsson, 2011).

Note that stiffness is a higher-order percept and it is doubtful that “stiffness” neurons exist. Then can such spatiotemporal rule be applied to the perception of stiffness, and will spatial and temporal gradients be observed in the visual-haptic perception of stiffness? Gepshtein, Burge, Ernst, and Banks (2005) examined the impact of spatial separation on the visual-haptic perception of object size. In their experiments, the bimodal JND was measured in four conditions where the visual display was displaced by 0, 3, 6, or 9 cm relative to the haptic device. The bimodal JND was found to be lowest when there was no separation, and its value matched with the prediction by the MLE model. As the separation increased, the bimodal JND increased as well. When the separation was larger than ~3 cm, the integration seemed no longer to be achieved: in this case, the bimodal JND was no lower than the visual or haptic JND. To the author's knowledge, no study has been published studying the influence of spatial incongruence on stiffness perception. We may expect to see a gradual breaking-down of the visual-haptic integration with increasing incongruence, but further research is needed.

Much effort has been devoted to studying the influence of temporal asynchrony on the visual-haptic integration, possibly because of the practical importance of this problem. In multimodal virtual-reality, augmented-reality, and tele-operation systems, a temporal asynchrony between the visual and haptic feedback often exists due to different delays in sampling, processing, transmitting, and rendering the two types of signals. Studies have shown that such asynchrony can cause inaccuracy in stiffness perception: stiffness will be overestimated (or underestimated) if visual feedback lags behind (or leads ahead of) haptic feedback (Pressman, Welty, Karniel, & Mussa-Ivaldi, 2007; Maher & Adams, 1996; Di Luca, Knoerlein, Ernst, & Harders, 2011; Sim, Wu, & Klatzky; 2014). For example, Di Luca et al. (2011) found that as a visual delay increased from 0 to 198 ms, the perceived stiffness increased accordingly in a linear fashion. Sim, Wu, and Klatzky (2014) also reported similar results.

Di Luca et al. (2011) applied the MLE model to explain such effects. In their experiments, the stimuli were virtual springs that could be felt by the participants using loading (i.e., to actively exert a force on a spring to compress it), unloading (i.e., to release a compressed spring and passively feel the force), or both types of actions. A visual delay was found to produce opposite effects on stiffness felt with the loading and unloading actions: The stiffness was overestimated when judged using the loading action, but underestimated using the unloading action. Moreover, the JND was found to be much lower in the loading-action condition than in the unloading-action condition. Therefore, when the participant could freely explore the spring using both types of actions, the two estimates would be combined. According to the MLE model, more weights could be assigned to the loading-action estimate, causing an overestimation of stiffness. The results of Experiments 3 and 4 in Di Luca et al. (2011) have provided evidence in support of such an explanation.

2.5 Conclusion

To conclude, I have briefly reviewed the studies aimed at understanding how the visual and haptic information is integrated into the percept of stiffness. The converging findings in the Chapter are that visual and haptic cues can be collectively used in stiffness perception, and the bimodal perception is generally more accurate and less variable than any unimodal perception.

CHAPTER 3

IMPACT OF VISUAL LATENCY ON VISUAL-HAPTIC EXPERIENCE OF STIFFNESS

A conference proceeding titled "Effects of visual latency on visual-haptic experience of stiffness" was used in this chapter with only very minor edits. As the second listed co-author, I participated in this work from beginning to end. I was involved in coming up with specific research questions, designing experiments, recruiting participants and running the experiments, analyzing the data of the experiments, discussing the results with co-authors, and other things. The results of this work were consistent with my anticipation.

3.1 Introduction

Virtual-reality (VR), augmented-reality (AR), and teleoperation systems with haptic feedback have come into widespread use in many fields, including robotic surgery, medical training, industrial control, and military applications. While the addition of haptic feedback dramatically increases the amount of information that can be conveyed to users, it brings about an important issue to be considered: A temporal asynchrony often exists between visual and haptic feedback due to differences in sampling, transmitting, processing, and rendering the two types of signals. In this study, we investigated how such haptic-visual asynchrony would influence our perceptual experience with virtual objects. We focused on the perception of stiffness because accurate perception of stiffness has immense importance particularly in medicine. For example, tumors like breast cancer can be discovered by palpation based on felt changes in stiffness. In surgery, surgeons use the perceived stiffness to plan and perform procedures with high precision. A better understanding of how haptic and visual information is processed in stiffness perception not only has importance for physicians and surgeons, but also can help engineers improve multimodal interfaces for robotic surgical systems and medical simulators.

Stiffness is primarily perceived through haptic interaction. However, our ability to perceive stiffness by touch is limited (for review, see Klatzky & Wu, 2014). Cholewiak et al. (2008) reported that only 2-3 levels of stiffness could be reliably identified for stiffness between 0.2 and 3.0 N/mm, equivalent to a capacity of transmitting only 1.46 bits of information. In clinical practice, limits in stiffness perception are indicated by the low sensitivity in palpation screening. Although cancer can be 7 times stiffer than normal tissue (Sarvazyan et al., 1995), the tumor detection rate is only about 39-59% by palpation (Shen & Zelen, 2001), depending on the proximity of the tumor to the body surface, the density of normal tissue, and the physician's experience. Tumors that are as small as 2-3 mm in diameter can be detected after physicians get sufficient training (Bloom, Criswell, Pennypacker, Catania, & Adams, 1982). But such an increase in detection rate is often accompanied with an increase in false positives (Campbell, Fletcher, Pilgrim, Morgan, & Lin, 1990)

The perception of stiffness can be facilitated by adding visual information to haptic exploration. Although it seems counterintuitive, we actually can judge stiffness by vision with reasonable accuracy (Drewing, Ramisch, & Bayer, 2009; Wu, Klatzky, Hollis, & Stette, 2012). Our actions and experiences of force normally occur in the context of certain visual experiences, and the haptic and visual experiences eventually become associated in long-term memory. Thus visual features in a deformation pattern can be matched to the stored haptic experiences and used to infer the force (Michaels & de Vries, 1998) and stiffness (Drewing et al., 2009; Wu et al., 2012). When both visual and haptic information is available, stiffness perception is usually improved. Varadharajan, Klatzky, Unger, Swendsen, and Hollis (2008) investigated the haptic and visual contributions to stiffness perception. Their results revealed a significant impact of visual inputs on stiffness discrimination: The average differentiation threshold was 17.2% in the haptic-only condition and it reduced to 14.2% in the haptic-visual condition. It has been suggested that the two types of information are combined in a statistically optimal fashion that produces a maximum-likelihood estimate of stiffness (Di Luca, Knörlein, Ernst, & Harders, 2011).

An important factor that affects the perceptual integration of haptic and visual information is temporal synchrony. Naturally, sensations of seeing and touching an object occur at the same time. But such simultaneity is often disrupted in multimodal VR/AR and teleoperation applications. While the updating rate of haptic feedback is at least 1000 Hz, the refresh rate of visual information is usually less than 60 Hz. Significant delays are often created when high-definition images and videos are transmitted and processed. The impact of such visual delay on task performance has been extensively studied. For example, Kim, Zimmerman, Wade, and Weiss (2005) examined the influence of delayed visual feedback on telerobotic surgery and found that the task completion time increased with delay. When the delay was longer than 400 ms, the operators usually switched to a move-and-wait strategy. In contrast, relatively few studies have been devoted to investigating the effects of visual delay on perception. Di Luca et al. (2011) used an AR system to examine the effects of visual delays on stiffness perception. Their results showed that the perceived stiffness increased in a linear fashion as the visual delay increased from 0 to 198 ms. One limitation in Di Luca et al. (2011) is that all visual delays were constant whereas in real world situations, delays may vary significantly with computational load or network traffic. In addition, there was an intrinsic end-to-end visual latency of 66 ms in their AR system, which could also produce some effects on stiffness perception.

In this Chapter, three psychophysical experiments were conducted using a state of the art haptic-visual VR system to assess the impact of visual delay on the user's subjective judgments of stiffness. Experiment 1 measured how well people could perceive visual delays, and Experiments 2 & 3 respectively examined how the perception of stiffness was influenced by constant and variable delays.

3.2 Exp 1. Perception of visual delays in visual-haptic simulations

Mixed results have been reported regarding our ability to detect visual-haptic asynchronies (Vogels, 2004; Shi, Hirche, Schneider, & Muller, 2008). Vogels (2004) conducted three experiments to measure the detection threshold for delays between the visual and haptic feedback. Participants were asked to hold a force-feedback joystick, use it to move a virtual object (a black square shown on a computer screen), and collide it with a virtual wall (a horizontal line on the screen). The haptic rendering of collision was an impulse force of 5.5 N applied to the joystick handle. When a temporal offset existed between the visual and haptic feedback, the average threshold for detecting visual delays was ~45 ms. Active motor control was found to impede the perception of asynchrony: The lowest threshold was obtained when the collision was passively viewed and felt by participants with no operation of the joystick. Shi et al. (2008) conducted a similar experiment using simulated collisions with a feedback force of 2.0 N. They reported an average detection threshold of ~40 ms. But in contrast to Vogels (2004), they showed that active motor control improved such temporal discrimination.

Clearly, the detection of temporal delays between visual and haptic inputs can be influenced by many factors such as the system's hardware limitations, the type of experimental stimuli, the quality of visual and haptic feedback, and the execution of visuo-motor actions. In this study, all experimental stimuli were the simulations of virtual springs that could be compressed or released, producing a gradually changing resistance force. The visualization was a simulated ultrasound video, in which the visual change was also continuous and gradual over time. Therefore, we expected that the detection of visual delays in such simulations would be more challenging as compared to the above-cited work. This experiment was run to measure the detection threshold.

3.2.1 *Method*

Participants: Twenty graduate and undergraduate students (twelve males and eight females, aged 18-44 years) participated with informed consent. All were naïve to the purposes of this study.

Experimental setup, procedure & design: The experimental stimuli were virtual elastic springs that were simulated using a multimodal virtual-reality system shown in Figure 5. The system consisted of a magnetic levitation haptic interface (Model# Maglev-200, Butterfly Haptics LLC. Pittsburgh, PA, http://www.butterflyhaptics.com) for rendering haptic feedback, a 27-inch LCD (Model# VG278H, ASUSTeK Computer Inc., http://www.asus.com, resolution: 1920x1080 @ 120 Hz) for providing visual feedback, a computer for controlling stimuli and acquiring data, and a keypad for the participant to enter responses.

A multi-threaded software was implemented in C++ and ran on the control computer to generate the visual and haptic feedback. The haptic effects were rendered using the Maglev-200 interface, which uses Lorentz forces arising from current-carrying coils to float and move a "flotor" within a strong magnetic field. The device uses no motors, gears, bearings, or linkages and is thus free of static friction. It can generate forces with a resolution of 0.02 N and move the flotor with a resolution of 2 μm. A handle was firmly attached to the flotor, which was held by the user to interact with the virtual springs and feel the force feedback. In this study, the user's interaction with the springs was restricted to vertical hand movements. When a virtual spring was compressed, the displacement of the handle was calculated relative to a predetermined resting position and used to proportionally produce a resistance force according to the Hooke's Law. The updating rate of the haptic rendering was 1000 Hz.

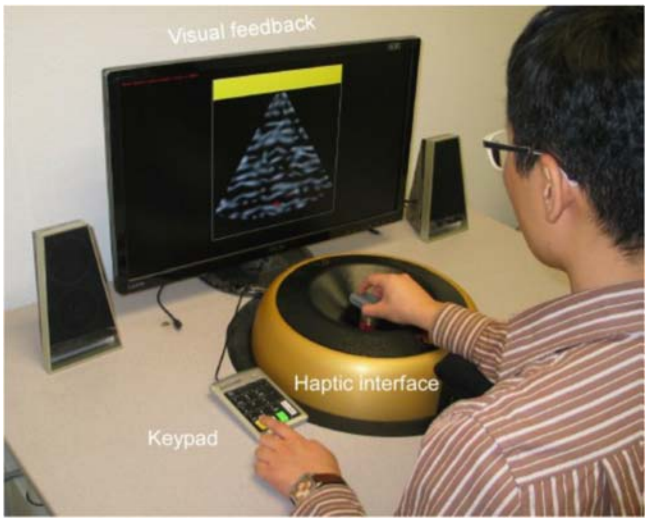


Figure 5. The experimental set-up. The haptic feedback was presented using a magnetic levitation haptic interface. The visual feedback was presented on a 27-inch LCD. A keypad was used to record the participant's responses. A Windows PC was (not shown in the picture) was used for controlling stimuli and acquiring data.

The deformation of virtual springs was visualized using simulated B-mode ultrasound videos to mimic the image-guided surgical operations. In addition to the consideration of potential clinical applications, one practical reason for using simulated ultrasound was that such 2D simulation could be run and updated at a rate higher than 120 Hz by using the algorithm described in Perreault and Auclair-Fortier (2007) and an NVIDIA Quadro K4000 graphics card. The presentation of the visual feedback then could be synchronized with the refresh of the screen, allowing for accurate timing of the experimental stimuli. Moreover, the use of such simulation enabled us to experimentally manipulate variables like the amount of speckle noise and the geometrical regularity of visual structure. Previous work has shown that visual cues in ultrasound images can be effectively used to judge stiffness (Wu et al., 2012).

In this study, the experimental manipulation was the visual delay relative to the haptic feedback. The position of the handle of the haptic device was updated at a rate of 1000 Hz by the haptic rendering thread and provided to the visualization thread to calculate the amount of deformation and accordingly update the ultrasound image. An additional delay could be introduced into the visual feedback by purposely using an old position data. In this experiment, the amount of delay ranged from 1 to 15 frames in a step of 2 frames (8.3 ms to 125 ms at a refresh rate of 120 Hz). In addition, there was an intrinsic delay of 1 frame (8.3 ms) because of double buffering.

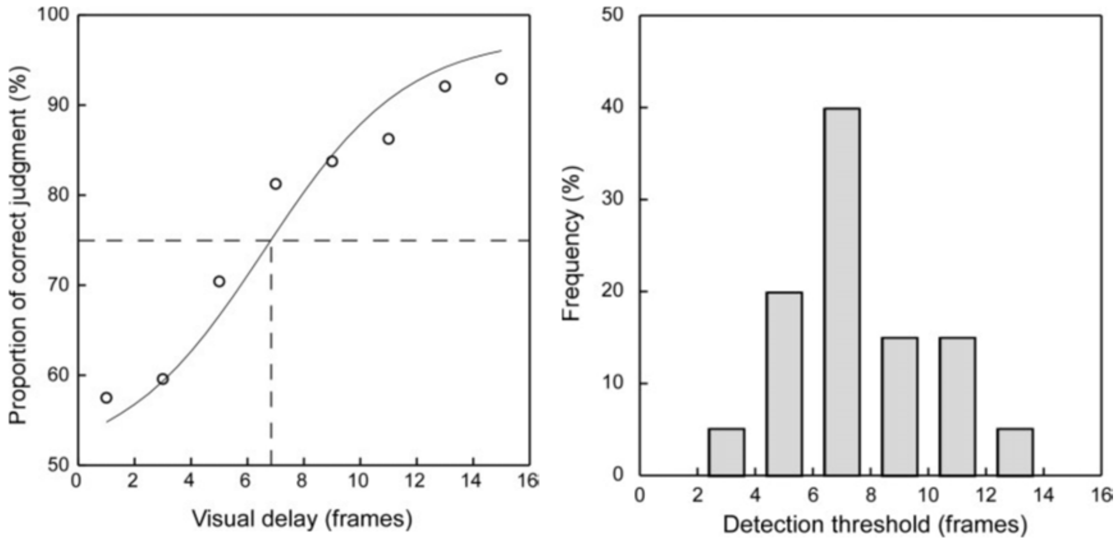
Participants' perception of visual delay was measured using a Two-Alternative-Forced-Choice (2AFC) procedure. On each trial, a pair of simulations was presented, a reference with no delay added and a comparison with an additional delay of 1-15 frames. The simulated stiffness was 125 N/m in both reference and comparison simulations. The two simulations were presented sequentially in a random order. The participant was asked to interact with them, pay attention to the visual feedback, judge which simulation had longer visual latency, and then report the judgment by a key press. He or she could switch between the two simulations as many times as desired. A transition phase with a random duration between 0.4 and 1.0 sec was inserted between the switches, during which the screen was masked with a checkerboard pattern. No feedback was provided to the participants about the accuracy of their judgments.

Each of the eight delays was tested ten times, yielding a total of 80 experimental trials. The trials were randomized and blocked into 3 sessions. Typically, a participant finished one trial in less than 20 seconds and a session in less than 15 minutes. There was a break of about 5 minutes for rest between the sessions. In addition, six practice trials were run before the experimental trials to familiarize the participant with the task. The whole experiment took about one hour.

3.2.2 *Results & Discussion*

The participants were tested individually, and their data were analyzed separately. For each participant, a psychometric curve was constructed using the PSIGNIFIT toolbox (Wichmann & Hill, 2001) by fitting a cumulative Gaussian function to the proportions of correct judgments. The delay that corresponded to the 75% correct point on the psychometric curve was taken as the detection threshold (halfway between the guessing rate of 50% and the perfect performance).

Figure 6(a) plots the average proportions of correct detections against the visual delay in the comparison simulation. The detection rate started from a value close to the chance rate of 50% when the delay was as short as 1 frame, and gradually increased as the delay became larger and more noticeable. On the psychometric curve, the 75% correct point corresponded to a delay of ~6.9 frames (57.4 ms @ 120 Hz). The individual differences among participants were also apparent, as shown in Figure 6(b). The threshold varied across the participants from 24.2 ms (2.9 frames) to 100.1 ms (12.1 frames), but for most participants, their thresholds were 6-8 frames.



(a) (b)

Figure 6. (a) The mean proportions of correct detections as a function of visual delay in frames. The curve is the average of the cumulative-Gaussian psychometric functions obtained from the participants. (b) Histogram of participants' thresholds.

The threshold we obtained here was larger than the values reported in Vogels (2004) and Shi et al. (2008). This could be accounted for by the difference in visual and haptic feedback. In Vogels (2004) and Shi et al. (2008), virtual collisions were visually indicated by the contact between the object and wall, and haptically by the onset of an impulse force. In contrast, the changes in the visual and haptic feedback were gradual over time in this experiment. Therefore, the task was harder and the threshold was higher.

3.3 Exp 2. Effects of constant visual delays on the perception of stiffness

The purpose of this experiment was to assess the impact of a constant visual delay on the perception of stiffness. Theoretically, when a visual delay was long enough, the visual and haptic feedback would be perceived as separate events. This then could deteriorate perception if the two sources of information were processed separately rather than being perceptually combined. To examine such effects, we purposely set the amount of visual delay to a range of values from sub-threshold to supra-threshold based on the results of Experiment 1, and measured the impact of visual delays on stiffness perception.

3.3.1 *Method*

Participants: Another group of twelve right-handed participants were tested (seven males and five females, aged 20-38 years) with informed consent. All were naïve to the purposes of this study.

Experimental setup, procedure & design: The experimental setup and procedure were the same as in the previous experiment. The difference was the task. On each trial, the participant was asked to interact with a pair of virtual springs, feel their stiffness, and judge which spring felt stronger. The pairs of virtual springs consisted of a reference spring of 125 N/m that was rendered with an additional visual delay of 0, 41, 83, or 166 ms (0, 5, 10, or 20 frames @ 120 Hz), and a comparison spring that had eight possible levels of stiffness (113.75, 118.25, 122.75, 127.25, 131.75, 136.25, 140.75, 145.25 N/m) and was rendered with no visual delay. Each combination of the reference and comparison springs was tested ten times, yielding a total of 320 trials (4 x 8 x 10). These trials were randomized and blocked into 6 sessions containing 53 or 54 trials each session. The participants usually finished one trial in less than 30 seconds and a session in less than 25 minutes. To avoid muscle fatigue, they could take a break at any time by withholding the response.

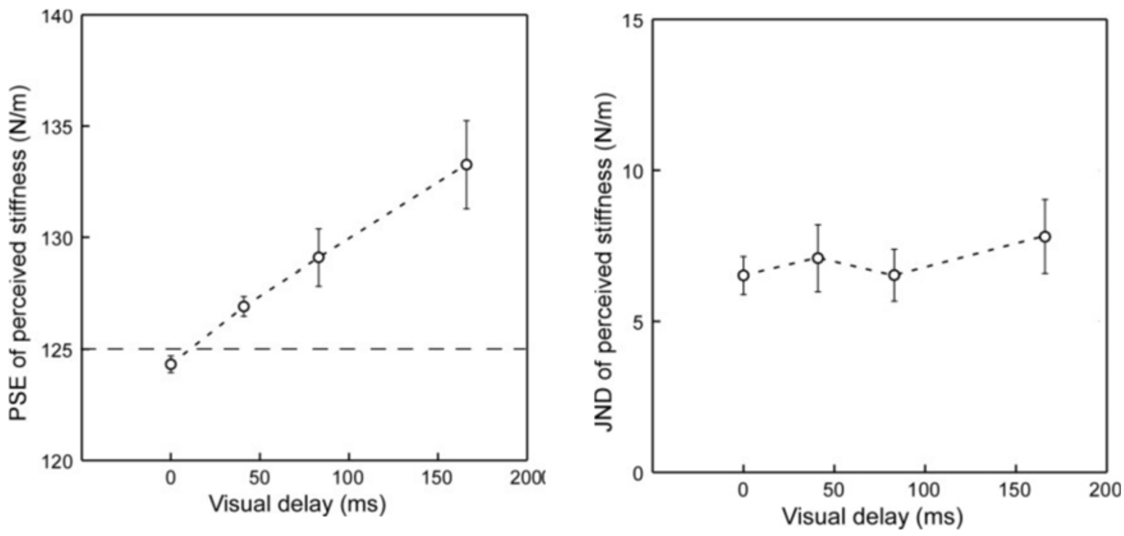
Additionally, there was a break of at least 5 minutes for rest between the sessions. The experiment was run for two days, and it took about one and half hour each day.

3.3.2 *Results & Discussion*

The participants' data were analyzed in the similar way as in the previous experiment. For each participant, a psychometric curve was fit to the proportions of the eight comparisons judged to be stronger than the reference. The perceived stiffness of the reference material was measured by the PSE that corresponded to the 50% point on the psychometric curve. The participant's ability to differentiate stiffness, quantified as the JND, was measured as the difference between the 50% and 84% points on the psychometric curve.

The measured PSEs, averaged across participants, are shown in Figure 7(a). The impact of visual delay on the perceived stiffness was evident. As the delay gradually increased from 0 to 166 ms, the perceived stiffness of the reference spring linearly increased: On average, the perceived stiffness increased by 1% for every delay of 25 ms. One-way repeated-measures ANOVA confirmed that such effect was significant ($F(3,33)=15.632$, $p<0.001$). Further comparisons with Bonferroni corrections revealed that such perceptual change was significant even in the shortest delay condition ($t(11)=3.448$, $p<0.04$) where the delay (41 ms) was below threshold (57.4 ms) and largely unperceivable.

As shown in Figure 7(b), the participants' ability to discriminate stiffness was little influenced by the visual delay. The observed JNDs were similar across all conditions. One-way repeated-measures ANOVA found no significant effect of visual delay ($F(3,33)=1.536$, $p=0.22$).



(a) (b)

Figure 7. The mean PSEs (a) and JNDs (b) as functions of the visual delay. The error bars stand for ±1 standard error.

Di Luca et al. (2011) proposed an explanation for these effects. Note that one consequence of visual delay was to change the force-displacement relation to be hysteretic and bi-phasic. When a force was applied during the loading phase (i.e., to actively exert a force on a spring to compress it), the visually perceived displacement lagged behind the force. This produced an “effective stiffness” larger than the physical stiffness. The effect was opposite during the unloading phase (i.e., to release a compressed spring and passively feel the force): the reduction of the force was faster than the displacement perceived from vision, producing an "effective stiffness" less than the physical stiffness. When the loading and unloading estimates were averaged and if a larger weight was assigned to the loading estimate, an overestimation of stiffness was then produced. The experiments in Di Luca et al. (2011) provided evidence in support of such an explanation. More importantly, they found that the visual-haptic integration was optimal even with visual delays. This might explain why the JNDs were similar across all conditions.

In summary, the findings from this experiment showed that when there was a constant delay between the visual and haptic feedback, such delay yielded an increased perception of stiffness, but had little impact on the user's ability to discriminate stiffness. Then the next question was, how would the perception be changed by variable visual delays?

3.4 Exp 3. Effects of variable visual delays on the perception of stiffness

In this experiment, the amount of visual delay varied according to a normal distribution. This introduced another source of noise to the perceptual system. Then an intuitive guess was that the visual-haptic integration would be hindered as the noise increased.

3.4.1 *Method*

Participants: Eight right-handed participants were tested (four males and four females, aged 19-38 years) with informed consent. All were naïve to the purposes of this study.

Experimental setup, procedure & design: The experimental setup and procedure were identical to Experiment 2. The eight comparison springs were also the same. The reference spring had the same stiffness of 125 N/m, which was rendered in four conditions: the no-delay condition, the constant-delay (83 ms) condition, and two variable-delay conditions in which the visual delay varied with a normal distribution around a mean of 83 ms and with a standard deviation of 14 or 20 ms. As in Experiment 2, each combination of the reference and comparison springs was tested ten times. A total of 320 trials were randomized and blocked into 6 sessions, which were tested in two days.

3.4.2 ***Results & Discussion***

The participants' data were analyzed in the same way as in Experiment 2 to estimate the PSE and JND for each participant. Two participants also participated in Experiment 2. Their results did not significantly differ from other participants' results, and so all data were pooled in the statistical analysis.

Figure 8(a) shows the mean PSEs, which are significantly different across four conditions (One-way repeated-measures ANOVA, $F(3,21)=4.919$, $p=0.01$). Compared to the no-delay condition, the stiffness was significantly overestimated in the constant-delay condition ($t(7)=5.369$, $p<0.01$). This replicated the results of Experiment 2. As compared to the constant-delay condition, the effects of visual delay gradually decreased as the temporal variance in the delay increased. The difference between the constant-delay and the largest-variance-delay condition was significant ($t(7)=2.471$, $p=0.04$). Figure 8(b) shows the mean JNDs. No significant difference was found across the four conditions ($F(3,21)= 1.514$, $p=0.24$).

These results suggested that the visual-haptic integration gradually broke down and the participants changed to rely more on the haptic feedback in judgments of stiffness as the variance in the visual delay increased. Again, this might be explained by the maximum-likelihood estimation model (Di Luca et al., 2011), which suggested that the visual and

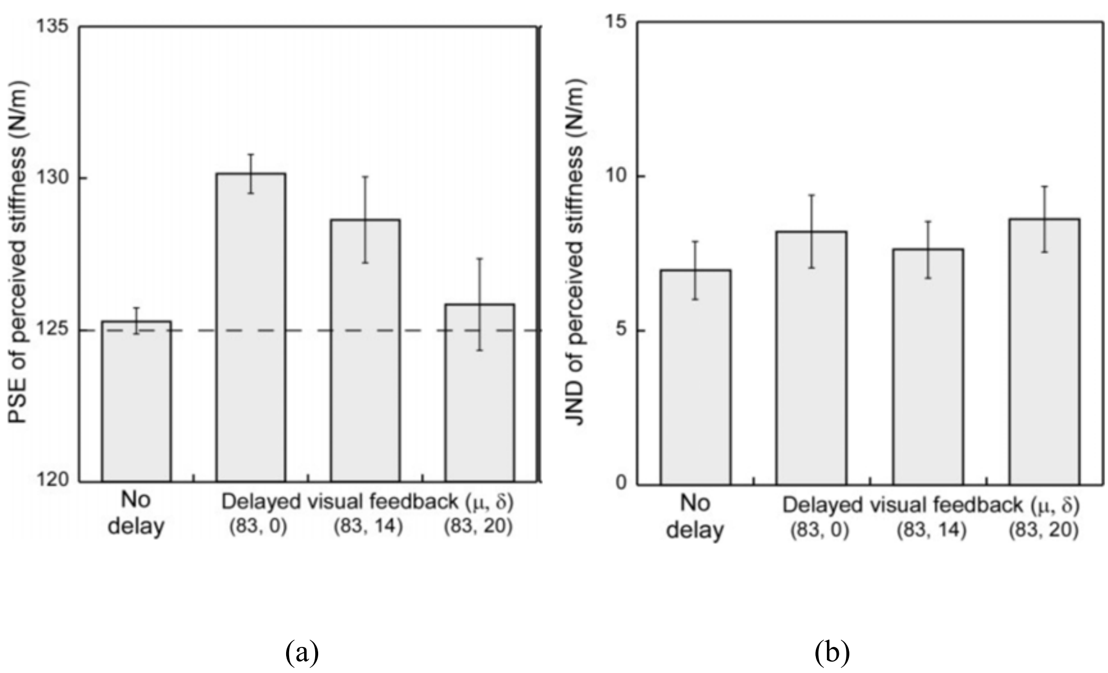


Figure 8. The mean PSEs (a) and JNDs (b) as functions of four visual-delay conditions. The error bars stand for ±1 standard error. The amount of visual delay varied according to a normal distribution in the variable-delay conditions.

haptic inputs, when being combined, were weighted by their reliability. Thus as the temporal variance of the visual input increased, the perceived reliability of visual cues was reduced. Accordingly, less (more) weight would be assigned to the visual (haptic) input, causing the perception to be biased towards haptic sensations. This was also confirmed by the participants' subjective experience. In anecdotal post-experiment interviews, almost all participants agreed that it was hard to extract useful information from the ultrasound visualization if the video was "jittering" or "unstable".

3.5 General Discussion

Whereas considerable effort has been devoted aiming to the reduction of visual delays in VR/AR or teleoperation systems, delays often still exist. In this study, we show that even a subthreshold, unperceivable delay could significantly influence our perceptual experience and cause an overestimation of stiffness. Such perceptual effects then should be considered in the design and implementation of visual-haptic interfaces and applications. For example, if the amount of visual delay can be determined, its perceptual effects might be compensated in the modeling of object stiffness. Alternatively, visual cues can be augmented to reduce such effects of delay. For example, visual magnification could lead to underestimation of stiffness, which may be used to counteract the overestimation caused by visual delay.

When the visual delay is not constant, our experiments show that the temporal variance could break down the process of visual-haptic integration. That is, the perception becomes more haptically dominant. Although this could be beneficial in some applications where the haptic feedback is more critical and accurate, we still argue that such variance in visual delay should be avoided because it is usually beneficial to use the information from all sensory inputs. In real-world applications, if visual delays are unavoidable, our results suggest that a constant delay would be preferable to a variable one. A variable delay caused by the fluctuation of internet speed is common in real-world. In addition, although the current study focused on the perception of stiffness, the experimental findings should also be applicable to other types of multimodal experience such as the perception of size, shape, and weight of virtual objects that involves the similar mechanism of multisensory integration.

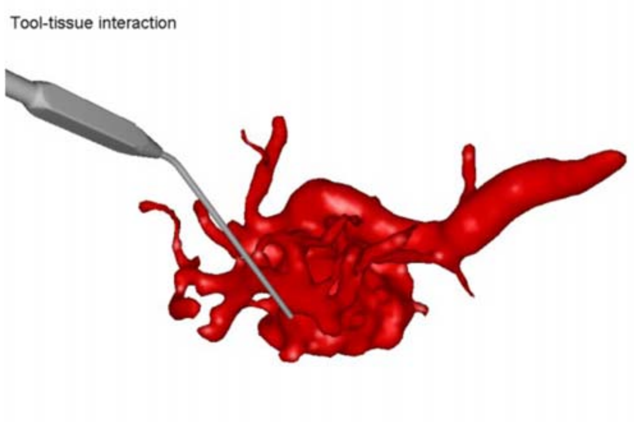


Figure 9. The neural surgical simulator.

In this study, we have shown that the threshold for detecting a visual delay is about 60 ms and stiffness perception can be significantly influenced by a sub-threshold delay. In many industrial or medical applications, tasks can be well performed even if there is a visual delay much longer than 60 ms. For example, the first trans-Atlantic operation was performed successfully with a mean time lag of 155 ms between the surgeon's movements and the video image (Wichmann & Hill, 2001). Mackenzie and Ware (1993) found that performance was affected when delays exceeded 75 ms. High tolerance of visual delays might be contributed to some cognitive mechanisms that may be used to predict and compensate some perceptual effects of visual delays. To further study the role of such cognitive mechanisms, we are conducting more studies using a surgical simulator shown in Figure 9.

CHAPTER 4

EFFECTS OF VISUAL DELAY AND INTERACTION SPEED ON STIFFNESS PERCEPTION

4.1 Introduction

Whereas the haptic senses are the predominant basis for judging stiffness, vision also affects the perception (see Klatzky & Wu, 2014 for a review). Though the neural mechanisms of the visual-haptic integration still remain unclear, psychophysical studies have shown that people can judge stiffness with better precision and accuracy using both visual and haptic cues. For example, Varadharajan, Klatzky, Unger, Swendsen, and Hollis (2008) reported that the average JND (Just-Noticeable-Difference) for stiffness discrimination decreased from 17.2% in the haptic-only condition to 14.2% in the haptic-visual condition. Drewing, Ramisch, and Bayer (2009) found that the slope of the visual-haptic stiffness estimates was close to the average of the two unimodal slopes. On the other hand, misperception may arise from inter-sensory incongruences such as asynchrony. If visual feedback lags are behind (or leads ahead of) haptic senses, stiffness is overestimated (or underestimated; Pressman, Welty, Karniel, & Mussa-Ivaldi, 2007; Maher & Adams, 1996; Di Luca, Knoerlein, Ernst, & Harders, 2011; Wu, Sim, Enquobahrie, & Ortiz, 2015). In this study, we examined the effects of delays in visual feedback that was common in many teleoperation or multimodal virtual-reality systems. A delay in force feedback often impairs the system stability. It also disrupts the causal relation between action and perception. Our previous work found that an unperceivable delay shorter than the threshold of visual-haptic simultaneity could lead to a significant overestimation of stiffness, and the amount of overestimation increased linearly for visual delays of 0-166 ms (Wu et al. 2015).

Di Luca et al. (2011) examined the perceptual effects of visual delay on subjective stiffness and explained their findings using a maximum likelihood estimation (MLE, Ernst & Banks, 2002) model. In their experiments, the stimuli were virtual springs that were felt by the subjects using loading (i.e., to compress a virtual spring), unloading (i.e., to release a compressed spring), or both types of actions. As illustrated in Figure 10, the force-displacement relation, if perceived through asynchronous visual and haptic feedback, was no longer linear. A visual delay produced opposite effects during the loading vs. unloading phases: Stiffness was overestimated when judged with the loading action, but underestimated with the unloading action. In Experiment 3, Di Luca et al. measured the stiffness JND and found that the loading JND was much lower than the unloading JND (16.8% vs. 34.8%). According to the MLE model, when the force and displacement information was integrated over the whole interaction trajectory, a larger weight would be assigned to the loading estimate than to the unloading estimate (0.78 vs. 0.22), causing overestimation of stiffness.

As illustrated in Di Luca et al.'s work and other research (for a review, see Lederman & Klatzky, 1996, 2009), action and perception are so tightly coupled in haptics that the hand acts as not only a manipulator exploring objects but also a sensor collecting information. Thus, it is important to understand how the action parameters affect the perceptual outcome. Di Luca et al. (2011) contrasted the role of loading vs. unloading actions in stiffness perception. Here we looked at another important parameter: the interaction speed. Previous research has reported some effects of interaction speed on subjective stiffness. Blair and Coppen (1942) asked people to squeeze rubber specimens slowly or fast with different durations of 0.5-4.0 seconds. Their results revealed a trend such that the longer the duration of squeezing and hence the slower the interaction speed, the softer the judged stiffness. It has also been suggested that the rate-hardness, the ratio of the change rate of force to the velocity of displacement, may be an effective cue to stiffness (Lawrence, Pao, Dougherty, Salada, & Pavlou, 2000 ; Han & Choi, 2010).

To date, no research has systematically examined how people's perception of stiffness would be affected by interaction speed under the practical condition of asynchronous visual-haptic feedback. It can be seen in Figure 10, for a given amount of visual delay, the force-displacement relation became more hysteretic as the interaction speed increased. Mathematically, such hysteresis can be modeled using the Maxwell model. For simplicity, we characterize the flexion/extension of hand at wrist joint by using a sinusoidal function, $\sin(\omega t)$, where ω denotes the loading/unloading frequency of hand-spring interaction. When a unit sine-wave force (f) is applied to a linear spring (k), the physical spring deformation is

$$f = \sin(\omega t)\,,\ d = \frac{f}{k} = \frac{1}{k} \cdot \sin(\omega t). \tag{1}$$

If the deformation is perceived with a constant delay of Δt, the perceived deformation then will be

$$d' = \frac{\sin(\omega(t+\Delta t))}{k} = \frac{\cos(\omega \Delta t)}{k} \cdot \sin(\omega t) + \frac{\sin(\omega \Delta t)}{k} \cdot \cos(\omega t). \tag{2}$$

d' consists of two components: an in-phase component corresponding to a purely elastic spring of effective stiffness of $\frac{k}{\cos(\omega\Delta t)}$ and an out-of-phase component corresponding to a purely viscous damper of $\frac{k}{\sin(\omega\Delta t)}$. This is the behavior of Maxwell materials that are modelled by a spring and a damper connected in series. That is, the visual delay desynchronizes the perceived force and deformation, changing an elastic spring into a Maxwell material.

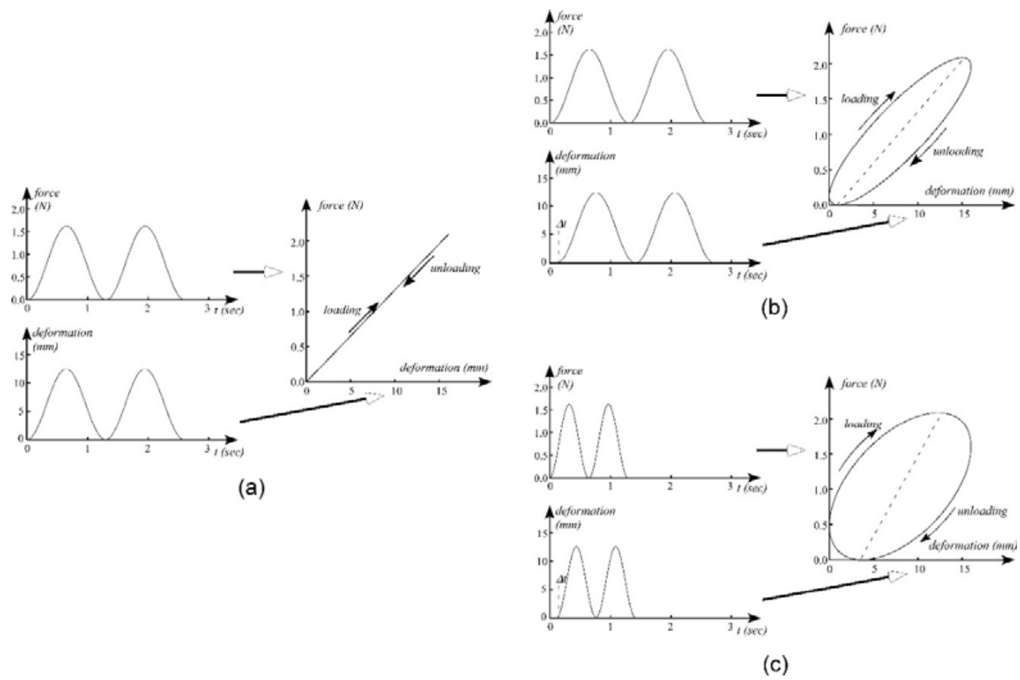


Figure 10. Effects of visual delay and interaction speed on the force-deformation profile. (a). The force-deformation profile of a purely elastic spring. No visual delay is assumed. (b). The hysteretic force-deformation profile for the same spring when there is a visual delay of Δt in the perceived deformation. (c). The force-deformation profile becomes more hysteretic for the same spring if the interaction speed is double while a visual delay (Δt) remains unchanged.

The above analysis can explain some observed effects of the delayed visual feedback on stiffness perception. For example, the elastic component in Eq.(2) produces an effective stiffness of $\frac{k}{\cos(\omega\Delta t)}$. As the delay (Δt) increases, $\cos(\omega\Delta t)$ decreases and hence the effective stiffness increases. And the effective stiffness increases in a nearly linearly fashion with increasing Δt for short-enough delays. This may explain the experimental findings that subjective stiffness increased in proportion to the visual delay for up to 198 ms (Di Luca et al., 2011; Wu et al., 2015) assuming that stiffness perception reflects the effective stiffness. Moreover, it suggests the effects of interaction speed. For a given visual delay (Δt), its effects are modulated by the interaction speed (ω): The faster a person interacts with a spring, the larger the ω, thus the smaller $\cos(\omega\Delta t)$ and the larger the effective stiffness of $\frac{k}{\cos(\omega\Delta t)}$. That is, when there exists a delay in visual feedback, one may expect that stiffness should be overestimated more with increasing interaction speed.

What makes the above analysis more complicated is the fact that a delay in visual feedback alters motor behavior. Numerous studies utilizing tracking or teleoperation paradigms have shown that the delayed visual feedback causes a decrease in interaction speed (Fujisaki, 2012; Miall, Weir, & Stein, 1985; Langenberg, Hefter, Kessler, & Cooke, 1998; Sheridan & Ferrel, 1963). Brady, Wu, Sim, Enquobahrie, et al. (2016) tested participants in a 2D mouse-pointing task with delays up to 133 ms in visual feedback. They found that the mean movement speed decreased with delay, and the movement time increased proportionally to the amount of delay across all targets of different difficulty.

Kim, Zimmerman, Wade, and Weiss (2005) assessed the influence of delayed visual feedback on teleoperation, and found that the task completion time eventually increased with delay, and for visual delays longer than 400 ms, the operators switched to a "move-and-wait" strategy. Then for stiffness that is felt via interaction, changes in motor control may alter perception and hence the effects of visual delay.

In this study, we tested the above analysis and conducted three experiments to measure the effects of visual delay and interaction speed on stiffness perception. We used a visual-haptic VR system (see Figure 11 for illustration) to generate virtual springs as the stimuli, in which the deformation of stimulus springs was visualized using simulated ultrasound with or without a delay of 166ms relative to the haptic feedback. In Experiments 1 & 2, the subjects were instructed by using a visual guidance to interact with the stimulus springs with a controlled speed. As shown in Eq.(2), delayed visual feedback should produce overestimation in the perceived stiffness, and the overestimation should increase with the interaction speed. Experiment 1 was designed to test these predictions. In Experiment 2, the effects of delayed visual feedback were measured over a range of stiffness. We expected that the perceptual change caused by a given visual delay should be comparable across stiffness levels. In Experiment 3, the subjects freely explored the stimuli. The force-displacement trajectories were recorded, analyzed, and used to further explore the relation between the perceived stiffness and the interaction speed.

4.2 Exp 1. Perception of visually-delayed stiffness under controlled speed

In this experiment, two independent variables were manipulated, namely, the interaction speed and the delay of visual feedback relative to haptic feedback. The visual feedback was rendered with a latency of 166 ms or in synchrony with the haptic feedback. Additionally, visual guidance was provided to the subjects concerning how fast they should use their hand to press or release the stimuli. The perceived stiffness was measured using a Two-Alternative-Forced-Choice (2AFC) procedure and quantified by the PSE (point of subjective equality). We expected that the visually-delayed stiffness would be overestimated and the amount of overestimation should increase with interaction speed.

4.2.1 *Method*

Participants: Twelve graduate and undergraduate students (8 males and 4 females, aged 18-38 years) participated with informed consent. The sample size was calculated using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) based on the effect sizes reported in our previous work (Wu, et al., 2015) and Di Luca et al. (2011) to achieve a power greater than 0.80 at the 0.05 level of significance. All subjects had normal or corrected-to-normal vision. All were right-handed by self-report and performed the experimental task using the right hand. They were naïve to the purposes of this study.

Apparatus & Stimuli: The experimental stimuli were virtual springs that were simulated using a haptic-visual simulator shown in Figure 11. The haptic effects were rendered using a Maglev-200 haptic interface (Model# Maglev-200, Butterfly Haptics LLC.). The device uses Lorentz forces to float and move a handle within a strong magnetic field. It can exert a force up to 40 N in the vertical dimension with a resolution of 0.02 N and track the handle movement with a spatial resolution of 2 μm. In this study, all virtual springs were linearly elastic, that is, their behavior was determined by Hook's law. Subjects' interactions with the springs were restricted to vertical hand movements. When a virtual spring was compressed, the displacement of the handle was calculated relative to a predetermined resting position and used to proportionally produce a resistance force. The updating rate of the haptic rendering was 1000 Hz.

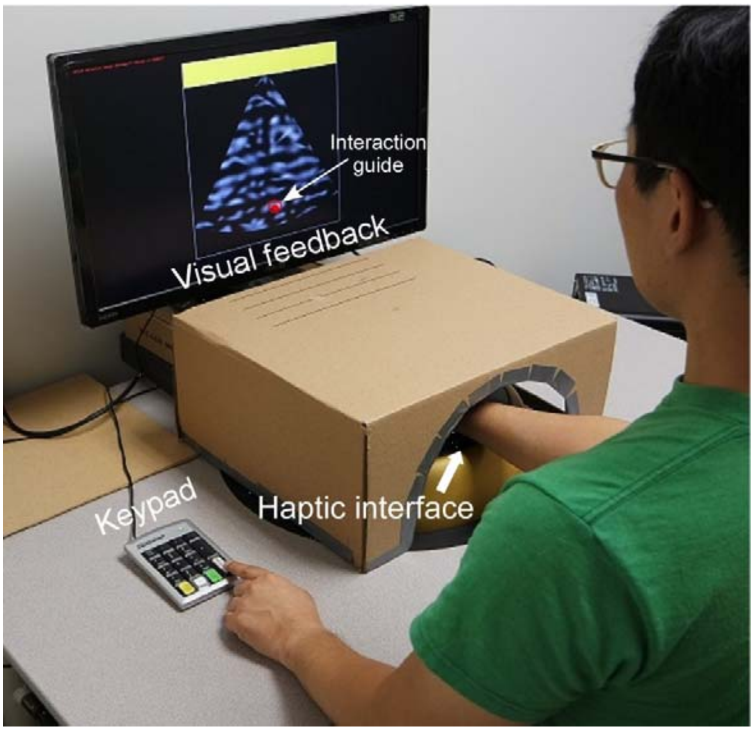


Figure 11. Illustration of the experimental setup. The interaction guide was used in Experiments 1&2 to control the interaction speed. It was removed in Experiment 3 to allow the participants to freely interact with the stimuli.

The participant's hand was covered with a cardboard box. The hand motion and hence the compression of the virtual spring was visualized using an ultrasound-like video. Our previous work has shown that visual cues in such images can be effectively used to judge stiffness (Wu, Klatzky, Hollis, and Stetten, 2012). In addition to the consideration of potential clinical applications, another practical reason for using simulated ultrasound was that such simulation was two-dimensional and could be updated in real time and synchronized with the refresh of the LCD screen at a rate of 120 Hz by using the algorithm described in Perreault & Auclair-Fortier (2007) and a powerful NVIDIA Quadro K4000 graphics card.

In this study, we experimentally delayed the visual feedback relative to the haptic feedback. This was implemented in the simulation software. When a visual delay was needed, an old reading of handle position was used to calculate the spring compression and generate the ultrasound image (c.f., the force feedback was always calculated using the current value of handle displacement without delay). Double buffering technology was used to perform flicker-free screen updates: an ultrasound image was written to a back buffer and shown in the next frame. Thus, in addition to the experimental visual delays, there existed an additional system delay of up to 8 ms (1 frame @ 120 Hz) in visualization.

Another variable was the speed at which subjects interacted with virtual springs. The desired interaction speed was shown by presenting a circle in the ultrasound video that vertically moved up and down in a periodic sinusoidal pattern. In addition, a solid dot was drawn to indicate the current position of the haptic device's handle. The subjects were instructed to operate the haptic device, move its handle as the circle indicated, and maintain the dot inside the circle. Two motion profiles were used, which guided the subjects to compress virtual springs by 20 mm with a motion period of 4.8 or 1.6 s/cycle, respectively, producing an average interaction speed of 16.6 or 49.7 mm/s.

Experimental procedure & design: The participants were tested individually. Their perception of stiffness was measured using a 2AFC procedure. On each trial, a pair of virtual springs were presented, a reference and a comparison. The reference spring had a stiffness of 125 N/m, and it was rendered with a visual delay of 0 or 166 ms. The comparison spring was rendered with no visual delay, and its stiffness could be one of eight values ranging from 107.5 N/m to 156.5 N/m in a step of 7.0 N/m. The two springs were presented sequentially in a random order, one at a time, along with a randomly assigned color label (yellow or green). The subject was asked to feel the springs, judge which seemed stronger, and then report the judgment by a keypress. He or she could switch between the two springs as many times as desired by first removing the hand from the handle of the haptic device and then pressing a key labeled “SWITCH”. A transition phase with a random duration between 0.4 and 1.0 second was inserted between the switches, during which the screen was masked with a checkerboard pattern, the handle of the haptic device was reset to the resting position, and the simulated stiffness was gradually changed to the new stimulus value.

A 2x2 within-subjects design was implemented with two independent variables: Visual-delay and Interaction-speed. The amount of visual delay was 0 or 166 ms (i.e., 0 or 20 frames @ 120 Hz). The predetermined hand-motion profiles were sinusoidal with an amplitude of 20 mm and a period of 4.8 or 1.6 s/cycle, corresponding to an average interaction speed of 16.6 (slow interaction) or 49.7 (fast interaction) mm/s, respectively.

The reference (125 N/m) was paired with and compared to 8 comparisons (107.5 - 156.5 N/m). Each pair was tested 10 times for each combination of Visual delay and Interaction speed, constituting a total of 320 trials. The trials were blocked by Interaction-speed and grouped into 6 sessions containing 53 or 54 trials that were presented in a randomized order within each session. Typically, the subjects finished one trial in less than 30 seconds and a session in less than 25 minutes. To avoid muscle fatigue, they could take breaks at any time by withholding the response. There were breaks of at least 5 minutes for rest among the sessions. In addition, the experiment was run on two consecutive days with one speed tested each day and the test order of two speeds was counterbalanced across subjects. The test took about 1.5 hours each day.

Six practice trials were run before each day's experiment to familiarize the subject with the task. Practice trials followed the same procedure as the experimental trials, except that different settings of stiffness were used and all virtual springs were rendered without visual delay. Throughout the experiment, no feedback was provided to the subjects about the accuracy of their judgments.

4.2.2 *Results & Discussion*

Figure 12 plots the mean proportions of the "comparison-stronger-than-reference" judgments, averaged across all subjects, for four combinations of Visual-delay and Interaction-speed. The psychometric curves are cumulative Gaussian functions fitted to the data using the PSIGNIFIT toolbox (Wichmann & Hill, 2001). The impact of visual delay on stiffness judgments was evident: When the reference was rendered with a visual delay of 166 ms, it was more likely to be judged as "stronger" than the comparisons, causing the psychometric curve to shift rightward. Moreover, the effects of visual delay were much larger in the fast-interaction conditions: The psychometric curve shifted much more in the right panel than in the left panel.

For each subject, psychometric curves were fitted to his or her data. The perceived stiffness of the reference spring (125 N/m) was measured by the PSE that corresponded to the 50% point on the psychometric curve. The mean PSEs are shown in Figure 13. When a delay of 166 ms was introduced to the visual feedback, the perceived stiffness increased from a value very close to 125 N/m in the 0ms-delay conditions to 127.7 N/m and 135.6 N/m, respectively, at the slow and fast interaction speeds. A two-way repeated-measures ANOVA found significant main effects for both Visual-delay ($F(1,11) = 16.87$, $p = 0.002$, partial $\eta^2=0.61$) and Interaction-speed ($F(1,11) = 6.71$, $p = 0.03$, partial $\eta^2=0.38$). The interaction between the two variables was also significant ($F(1,11) = 20.11$, $p = 0.001$, partial $\eta^2=0.65$). Planned contrasts with Bonferroni corrections further revealed that the perceived stiffness was significantly larger in the "166ms-delay/fast-interaction" condition as compared to other conditions ($t(11) > 3.44$, $p<0.01$).

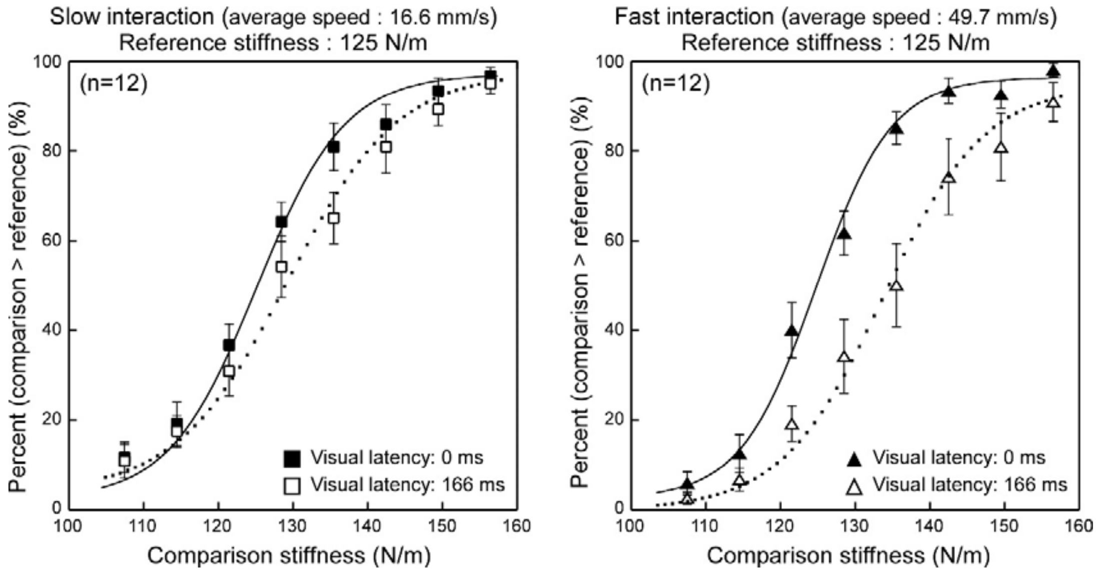


Figure 12. The mean proportions of "comparison-stronger-than-reference" judgments as functions of the comparison stiffness. The curves plot the cumulative Gaussian functions fitted to the data.

Briefly, the results so far supported the analysis outlined in Introduction: Consistent with previous studies (Di Luca et al., 2011; Wu et al., 2015), we found overestimation of stiffness caused by visual delay; Also, we showed that the effects of visual delay were effectively modulated by the interaction speed.

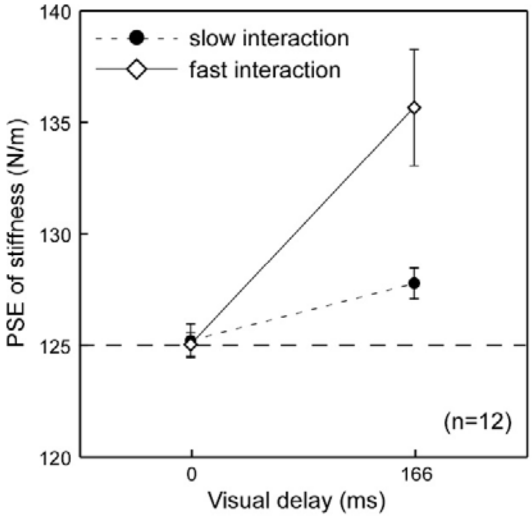


Figure 13. Mean perceived stiffness as a function of visual delay and interaction speed. The perceived stiffness was quantified by the Point of Subjective Equality (PSE). Error bars indicate between-subject standard error of the means.

4.3 Exp 2. Perception of visually-delayed stiffness over stiffness range

This experiment aimed to further examine the relationship between subjective stiffness and the effective stiffness defined in Eq.(2). In our analysis, we implicitly assumed that the effective stiffness formed the basis, at least in part, for judging the visually-delayed stiffness. A closer examination of Eq.(2) suggests that as compared to the physical stiffness, the effective stiffness is "magnified" by $\frac{1}{\cos(\omega\Delta t)}$. Then if our perception was based on the effective stiffness, the amount of overestimation should be similar across stiffness levels for a given visual delay and interaction speed. Here we tested this using two new levels of reference stiffness (75 N/m and 200 N/m) and the interaction speed was set to the "fast" speed used in the previous experiment. We would compare the observed stiffness overestimation to that observed in the previous experiment.

4.3.1 *Method*

Participants: Another ten right-handed students (6 males and 4 females, aged 20-38 years) participated in this experiment. All had normal or corrected-to-normal vision. All were naïve to the purposes of the study.

Experimental setup & Procedure: The experimental setup was identical to the previous experiment. Participants' perception of stiffness was measured using the same 2AFC paradigm and the stimuli were generated and presented in the same way as in the previous experiment. The only difference was the stimuli: The reference stiffness was set to 75 N/m or 200 N/m and paired with two different sets of comparisons, as detailed in the next section. In addition, the interaction speed was controlled and set to be 49.7 mm/s (i.e., the fast speed in the previous experiment).

Design: A 2x2 within-subjects design was used, with Visual delay (0 or 166 ms) and Reference stiffness (75 N/m or 200 N/m) as the two independent variables. Like the previous experiment, each reference stiffness was compared to a set of 8 comparisons. The comparison stiffness ranged from 65 N/m to 93 N/m in a step of 4 N/m for the reference of 75 N/m. They varied from 180 N/m to 236 N/m in a step of 8 N/m for the reference of 200 N/m. Each pair of the reference and comparison was tested 10 times, constituting a total of 320 trials. The trials were blocked by Reference stiffness and grouped into 6 sessions containing 53 or 54 trials. Within each session, the presentation order of the trails were randomized. As in Experiment 1, this experiment was run for two days with one Reference stiffness tested each day in a counterbalanced order across subjects. The test took about 1.5 hours each day.

Again, the interaction speed was controlled by showing a circle that moved vertically on the screen in a periodic sinusoidal pattern. The period of motion and the average interaction speed were set to 1.6 s/cycle and 49.7 mm/s (i.e., the fast interaction speed tested in the previous experiment). Throughout the experiment, no feedback was provided to the subjects about the accuracy of their judgments.

4.3.2 *Results & Discussion*

The participants' data were analyzed in the same way as in Experiment 1. As shown in Figure 14, significant effects of visual delay were observed on the subjects' perception of stiffness: The psychometric curves shifted rightward for both reference stiffness when a visual delay of 166 ms was introduced.

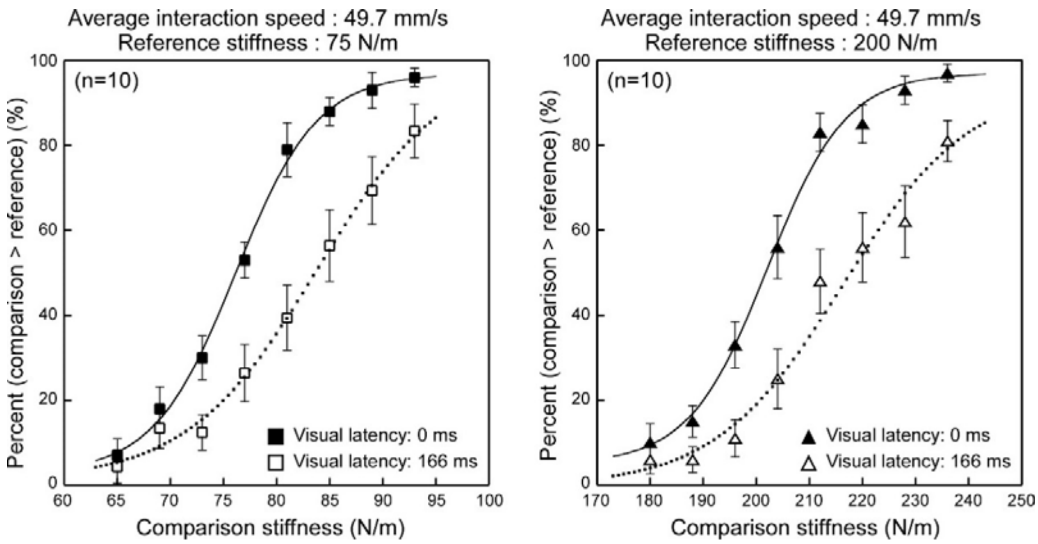


Figure 14. The mean proportions of "comparison-stronger-than-reference" judgments as functions of the comparison stiffness for the two references tested. The curves plot the cumulative Gaussian functions fitted to the data.

The PSE was calculated from each participant's data using the same method as in the previous experiment. To compare the effects of visual delay across stiffness, the PSEs were converted to the percentages of the reference stiffness. As shown in Figure 15, the mean PSE increased to 108% and 110%, respectively, for the reference stiffness of 75 N/m and 200 N/m with the presence of a visual delay of 166 ms. A two-way repeated-measures ANOVA found a significant main effect of Visual delay ($F(1,9) = 26.97$, $p = 0.001$, partial $\eta^2=0.996$). But neither the main effect of Reference stiffness ($F(1,9) = 3.34$, $p = 0.10$, partial $\eta^2=0.37$) nor the interaction of (Visual delay x Reference stiffness) ($F(1,9) =1.78$, $p = 0.21$, partial $\eta^2=0.22$) reached significance.

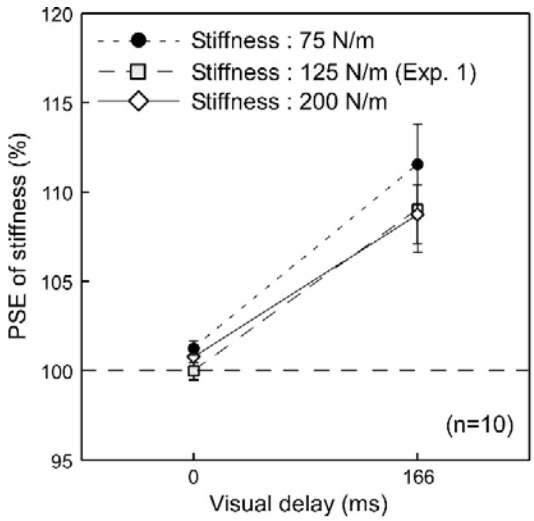


Figure 15. The mean percentage PSEs as a function of visual delay for the reference stiffness tested. The gray squares re-plot the “fast-interaction” data in Figure 15 in percentage values. The error bars stand for ±1 standard error.

Clearly, this experiment showed comparable effects of visual delay on stiffness perception over a range of physical stiffness at a predetermined interaction speed, consistent with the predictions based on the effective stiffness in Eq.(2). Next, we moved on to remove the control on the interaction speed and measure the subjects' perception of stiffness with visual delay under free haptic exploration.

4.4 Exp 3. Perception of visually-delayed stiffness with free exploration

In contrast to Experiments 1&2, the participants were told to freely interact with the virtual springs in this experiment, and no restriction was imposed on the interaction speed. The participants' hand movements were recorded and the force-displacement data were analyzed off-line. A regression analysis was then performed to evaluate the relation between the interaction speed and the perceived stiffness.

4.4.1 *Method*

Participants: A total of 24 new right-handed students (15 males and 9 females, aged 19-32 years) were tested. All had normal or corrected-to-normal vision. All were naïve to the purposes of the study.

Experimental setup, procedure & design: The experimental setup and procedure was identical to Experiment 1. The stimuli were also generated and presented in the same way, except that the interaction guide (i.e., the dot and circle shown in Figure 11) was removed and the subjects were told to freely explore the stimulus with comfortable movements. The exploratory hand-motion was sampled at a rate of 120 Hz and recorded for offline analysis.

The reference stiffness (125 N/m) was tested with or without a visual delay of 166 ms. It was paired with and compared to 8 comparisons (113.75 - 145.25 N/m). Each pair was tested 8 times for each level of visual delay, constituting a total of 128 trials. The trials were randomly intermixed and grouped into 4 sessions, 32 trials each. Typically, the subjects completed the whole experiment in about one hour. No feedback was provided to them about the accuracy of their judgments.

4.4.2 *Results & Discussion*

The PSE was calculated for all but one subject using the same method as in the previous experiments. That subject was excluded because for unknown reasons, he judged the reference as being stiffer in almost all 166ms-delay trials regardless of the comparison stiffness, making it impossible to fit a psychometric curve and estimate the PSE from his data. Thus, the total number of subjects was 23 in statistical analyses. As shown in Figure 16, consistent with the previous two experiments, a significant effect of Visual delay ($F(1,22) = 30.93$, $p<0.001$, partial $\eta^2=0.58$) was observed: the mean PSE increased from 125.3 N/m to 131.4 N/m after a visual delay of 166 ms was introduced.

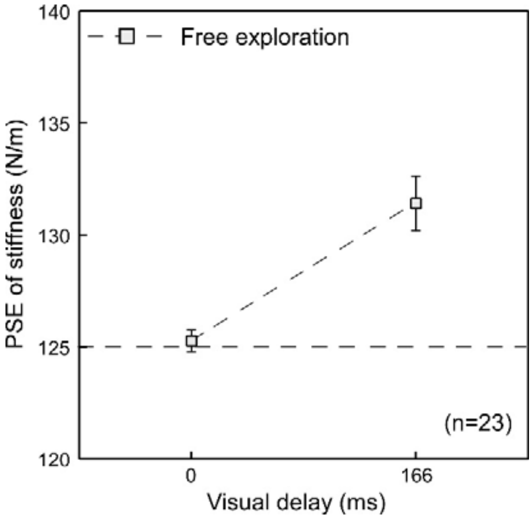


Figure 16. The mean PSE as a function of visual delay. The error bars stand for ±1 standard error.

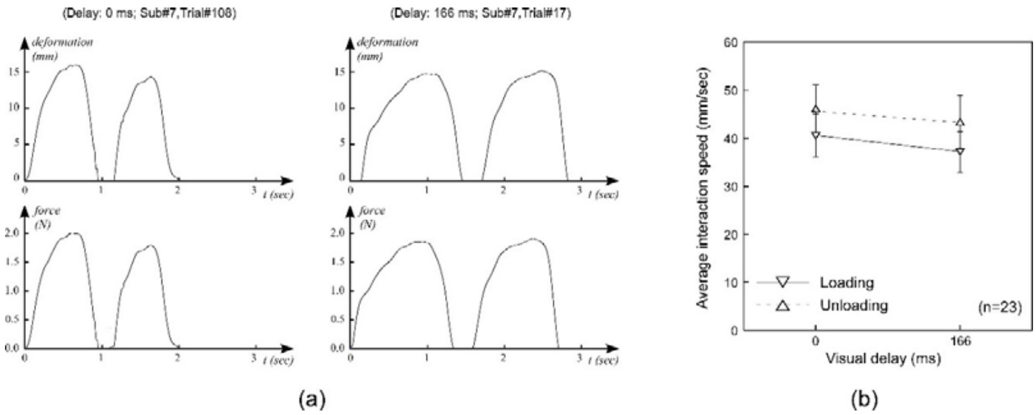


Figure 17. (a) Part of two force-deformation trajectories showing the subject's loading-unloading interaction with the reference spring. The trajectories were recorded from a no-delay trial (left) and a 166ms-delay trial (right), respectively. (b) The average interaction speed as a function of visual delay and type of interaction (loading vs. unloading).

The force-displacement trajectories were recorded and analyzed. Here our analysis examined how fast the subjects compressed (i.e., the loading action) and then released (i.e., the unloading action) the reference spring. Figure 17(a) shows two example trajectories, one recorded from a no-delay trial (left) and another from a 166ms-delay trial (right). It can be seen in the trajectories that (1) the delayed visual feedback effectively reduced the subject's speed of interaction with the virtual spring: As compared to the left trajectory, the loading-unloading cycle shown in the right trajectory was much longer as a delay of 166ms was presented in the visual rendering of deformation. (2) the loading (i.e., to compress) speed was much slower than the unloading (i.e., to release) speed. These two patterns held for all subjects (Figure 17(b)), though there were huge individual differences with an average interaction speed ranged from 17.7 to 72.5 mm/s. A two-way repeated-measures ANOVA found significant main effects for both Visual-delay ($F(1,22) = 15.14$, $p = 0.001$, partial $\eta^2=0.41$) and Interaction-type ($F(1,22) = 26.87$, $p < 0.001$, partial $\eta^2=0.55$). The interaction between the two variables was insignificant ($F(1,22) = 0.50$, $p = 0.49$, partial $\eta^2=0.02$). Thus, we used the overall interaction speed in the subsequent regression analysis, which was computed using both loading and unloading interactions.

As shown in Eq.(2), the influence of visual delay on stiffness perception was modulated by the interaction speed: The faster a person interacts with a spring, the more the stiffness would be overestimated. And the modulation effect would increase almost linearly if the delay was small (Di Luca et al., 2011; Wu et al., 2015). To test this, we conducted a regression analysis predicting the amount of stiffness overestimation experienced by a subject from the speed of interaction he or she acted. Stiffness overestimation was quantified as the difference between the PSEs obtained from the 166ms-delay and no-delay conditions. As shown in Figure 18, we found a significant positive correlation between the interaction speed and stiffness overestimation ($r2 = 0.31$, $p = 0.005$). Interestingly, the data obtained from Experiment 1 (the filled circle and open diamond) were also close to the regression line.

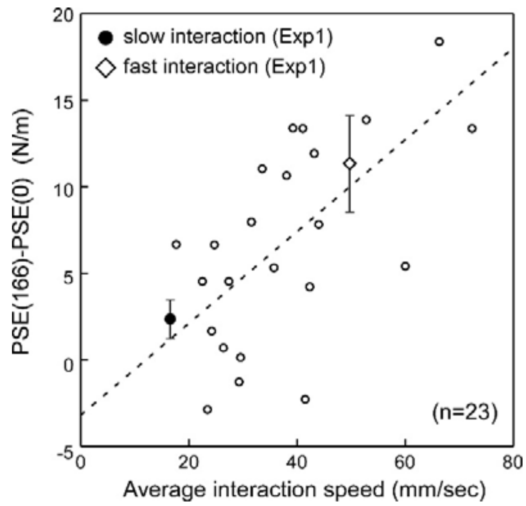


Figure 18. Correlation between the interaction speed and stiffness overestimation. For comparison, data from Experiment 1 were also plotted where the mean interaction speed was set to 16.6 or 49.7 mm/s. Error bars represent the standard error of the mean.

4.5 General Discussion

We investigated the effects of interaction speed and delayed visual feedback on stiffness perception. We found in all three experiments that the visually-delayed stiffness was overestimated and the amount of overestimation was comparable across three levels of stiffness tested (Experiments 1& 2). Our results also revealed a modulatory effect of interaction speed on stiffness overestimation: Stiffness was judged stiffer at faster speeds through free exploration (Experiment 3) or under controlled speed (Experiment 1), and a positive correlation was revealed between interaction speed and stiffness overestimation (Experiment 3).

Physically, stiffness is determined by Hooke's law as the force-to-deformation ratio. While the ratio has also been suggested as the invariant underlying stiffness perception (Gibson, 1966), we are still unclear of the processes through which stiffness is judged using all available sensory information. For example, the integration of visual and haptic information can occur at different stages of processing. The visual-haptic integration may occur in the perception of force and deformation, and stiffness is estimated from these two variables (Srinivasan, Beauregard, & Brock, 1996). Alternatively, the integration process may occur after the modality-specific process. For stiffness, it can be directly perceived by touch or vision. For example, Drewing et al. (2009) showed that people could judge rubber specimen's stiffness by just watching another person pressing the specimens (the regression slope of the perceived vs. physical stiffness: 0.84, $r^2 = 0.67$). Then the visual-haptic perception of stiffness is an amodal percept derived from the visual and haptic estimates, presumably using the optimal integration strategy as described in the MLE model. In Drewing et al.'s (2009) work, they found a regression slope of 0.90 in the bi-modality condition, close to an average of the visual-only (0.84) and haptic-only (1.02) slopes.

It is challenging to differentiate the above two integration mechanisms because the force and displacement information is perfectly correlated within and across sensory modalities. One way to break such coupling is to temporally desynchronize sensory inputs, as we did in this study. Then given asynchronous sensory inputs, the integration-of-modality-specific-estimates model predicts null effects: The visual or haptic stiffness is judged using only information available from seeing or touching. The estimates and their variances should not be affected by the intersensory asynchrony and hence the consequent outcome of amodal integration. Clearly, this prediction is not supported by our results. In contrast, we calculated an effective stiffness in Eq.(2) using the simultaneously perceived force and deformation, which predicted stiffness overestimation caused by visual delay and also modulatory effects of interaction speed. Our results were consistent with these predictions, suggesting that the perception may be based on the effective stiffness that was derived from the multisensorily perceived force and deformation.

While the MLE model has been successfully applied to account for many multisensory phenomena, it should be noted that the model has its own limits. The model takes a bottom-up approach, in which the integration process and the outcome are determined only by the statistical characteristics of individual sensory inputs. Note that stiffness is perceived through action. It remains unclear that how well active perception under the guidance of cognitive processes can be modelled by the MLE model without considering any motor parameters. Endo (2016) measured the exerted force while participants actively pressed soft materials to judge stiffness. The exerted force was found to be adjusted according to target stiffness: When comparing two stiff samples, people tended to compress them with similar forces and the stiffer, the more similar the pressing forces. When the visual feedback was delayed, we showed that the interaction speed was reduced (see Figure 17(b)). Thus to account for such action-perception interaction, a more general Bayesian model may be more appropriate. Ernst and Bülthoff (2004) proposed an inference model, in which the Bayes' rule is applied to generate perception by combining sensory inputs with prior knowledge or other cognitive processes and also to guide action for seeking information from the environment. Further research is needed to test this model in active perception of stiffness.

Our research aims at the application of visual-haptic teleoperation, VR, and AR in medical settings. For such applications, it is crucial to render deformable objects such as soft tissues with high fidelity. Given the current technology, delays are often unavoidable particularly for rendering graphics in high resolutions. Although such delay alters a user's perception of virtual objects and also motor performance, our work suggested an approach to design-in such perceptual-motor effects. For example, we may compensate stiffness overestimation by measuring the visual delay and user's interaction and then adjusting the simulated stiffness.

CHAPTER 5

LEARNING AN ASSOCIATION BY COMBINING INFORMATION ACROSS SPACE

5.1 Introduction

Minimally invasive surgery (MIS) is a surgery minimizing surgical incisions to lessen damage to the body. In MIS, an endoscope is used to visually guide the surgery, and the surgeon performing MIS views inside the patient's body on a monitor. Tiny surgical instruments are inserted through small incisions. As the monitor is usually located away from the patient's body, the source of visual information is spatially separated from that of haptic information. Thus, surgeons perform their task in a "doing-here and looking-there" fashion. Although two types of sensory inputs are provided in a spatially separated state, surgeons see the consequence of their actions, and such causality promotes the visual-haptic integration.

In some cases, spatial displacement is not overcome, resulting in incomplete integration of sensory information. For instance, the Rubber Hand Illusion (Botvinick & Cohen, 1998), a perceptual phenomenon of misperceiving a rubber hand as part of one's own body, has a spatial limit. Lloyd (2007) tested the illusion with six levels of distance from 17.5 cm to 67.5 cm, and found that the illusion decayed significantly if the distance between the real and fake hands was above 37.5 cm. Gepshtein, Burge, Ernst, and Banks (2005) revealed that a large-enough spatial separation could cause multisensory integration to break down. However, when prior knowledge was provided to participants that they were seeing and touching the same object, Helbig and Ernst (2007) showed that people could still integrate visual and haptic information in a near-optimal manner even if the inputs came from different locations.

In this study, we investigated if participants could learn an association between two sensory inputs spatially asynchronous, and if this learning could affect participants' typing performance. Specific questions were: Is an association between them automatically formed or needs to be learned? What is learned? Can the learning be carried over to subsequent trials? A key on a flat keyboard was presented on a LCD as visual input, and tactile key-identity information was provided on a participant's thumb as haptic input. The level of learning an association between visual and haptic inputs was evaluated by recognition tests. Four experiments were conducted to examine if key-identity information could be effectively conveyed via tactile stimulation and if the tactile information could be integrated with visual information to facilitate typing.

When multiple sensory inputs are present, multisensory reaction time (RT) is often found to be shorter than unisensory RTs. For example, Todd (1912) reported that the RT to a visual stimulus was shortened by 20-80 ms when it was paired with a simultaneous or slightly-delayed auditory stimulus. Several models have proposed to explain such cross-sensory facilitation. Raab's model of statistical facilitation (1962) suggested that the effect might be based on the selective processing of one input which happened to arrive first and provide the earliest available information on a particular trial. Nickerson (1973) suggested two possible mechanisms at different levels of multisensory processing. The energy-summation hypothesis suggests that the inputs from different senses fuse in early stages, producing a stronger signal for faster detection and more rapid processing. The preparation-enhancement hypothesis argues that the pathway for processing the imperative input may be facilitated by the accessory input and become more ready for producing a response. Although more research is needed to understand the mechanisms underpinning multisensory perception and action, it is safe to conclude that the more converging information from different senses, the faster the response time and the more accurate the response.

When tactile feedback is presented along with other types of sensory feedback, mixed results have been reported as to its effectiveness. Calhoun et al. (2003) measured the operators' RTs in an unmanned aerial vehicle control station simulation and found that tactile warning signals could lead to faster response than visual alerts, but when presented in concert with visual and auditory alerts, tactile feedback yielded no additional benefits. Ma et al. (2015) tested three types of tactile key-click feedback using a Microsoft Touch Cover and found that the typing speed was improved and the error rate was reduced, regardless of whether the key-click feedback was delivered to the typing finger, to the five fingers of the typing hand, or to all ten fingers. Moreover, when an auditory key-click was added, the addition of auditory feedback yielded no further improvement in typing performance. That is, the tactile information dominated performance in their study.

The discrepancy may be attributed to the type of information conveyed through tactile feedback. In Ma et al. (2015), the simulated key-click mainly provided users information on whether and when a key had been pressed. In contrast, the tactile warning signals in Calhoun et al. (2003) were associated with more information content. Kim and

Tan (2015) defined the following three types of information that can be conveyed to users in a typing task:

1) Key-entry information: It gives the user a confirmation whether the action of key-pressing is successful. The confirmation can be provided by vision (e.g., asterisks or dots when entering a password), audition (e.g., clicking sound while typing on smartphones), or touch (e.g., short vibrations with each keystroke).

2) Key-correctness information: It gives the user a confirmation whether the correct key has been pressed. Such correct-or-wrong feedback information can also be provided by means of visual, auditory or tactile signals (e.g., two different colors, sounds, or vibration patterns).

3) Key-identity information: It tells the user which key has been pressed. Such information can be provided to the user by displaying the entered letter on the screen or playing the sound of the letter. When using real keyboards, such information is also available from haptic sensations: On conventional QWERTY keyboards, two "home keys" (F & J) are marked by a raised hyphen. Skilled typists can use the two keys as the anchors and find other keys using the position information perceived from the separation and extension of the fingers.

Most of previous research has focused on how to improve typing performance on flat real or virtual keyboards by providing users the key-entry or key-correctness feedback information by different sensory means (Ma et al., 2015; Hoggan, Brewster, & Johnston, 2008; Lee & Zhai, 2009; Bender, 1999; Kim & Tan, 2014). Little work has examined the worth of the key-identity feedback information and whether such information can be effectively conveyed through tactile feedback. Actually, our tactile perceptual system has amazing capabilities of picking up and processing complex information. For example, after training, deaf-blind individuals can use the Tadoma method to understand speech by using their hands to feel the movement of the speaker's lips and vibrations of the vocal cords, and the information transmission rate can be as high as 12 bits/sec (Reed, Durlach, & Delhorne, 1992). Here we would like to capitalize on such capacity of the tactile perceptual system to investigate if the key-identity information could be delivered by playing the sounds of the letters at users' fingertips and help them to improve typing performance. We hypothesized that the tactile presentation of sounds might be a natural way for users to identify the letters thanks to the familiarity of sound patterns.

In this study, two sets of experiments were carried out using a flat-surface, pressure-sensitive keyboard. In a single-letter-typing task, participants saw a target letter and were asked to press that letter on the keyboard as quickly and accurately as possible. The key-identity information was conveyed by displaying the pressed key on a LCD screen and also by playing the letter sound at the participant's right thumb using a voice coil. That is, visual and haptic information were presented at different spatial locations as shown in Figure 19(a). The tactile stimulation was presented as part of the stimulus in Experiments 1a & 1b (Figure 20) or as response-feedback in Experiments 2a & 2b (Figure 22). In Experiments 1a & 2a, participants went through four training sessions to form an association between the letters and their vibration patterns. The target set consisted of 6 letters, which should be small enough to fit the capacity of human working memory and allow fast learning. In the post-training tests, the participants' typing performance was measured with a meaningless 250-Hz buzz (1-beat 200ms sine-wave buzz as in Tacton applications such as Gunther, Davenport, & O'Modhrain, 2002; Brewster & Brown, 2004; Hoggan, Brewster, & Johnston, 2008). If the key-identity information could be effectively delivered through tactile feedback and used to guide the keypress response, we expected to see reduced performance in the post-training tests after such information had been removed. In contrast, participants received no key-identity information during the training in Experiments 1b & 2b, and the letter sounds were played at their fingertips only during the post-training tests. In these two experiments, we may see improved performance in the post-training tests when additional information was provided through tactile feedback.

5.2 Exp set 1. Effectiveness of the key-identity information as part of the stimulus

In this set of two experiments, the tactile stimulation was presented as part of the stimulus. That is, the voice coil would start to vibrate at the same time as the target letter ('A' in Fig 20) appeared on the screen. In Experiment 1a, participants were trained to learn an association between the letters and their tactile sounds: On each training trial, participants would see a letter and at the same time feel the vibration of its sound at his or her fingertip. The vibration was changed to a meaningless 250-Hz buzz in the post-training test. The main difference between the buzz and letter sounds was the information content: Whereas both signaled the presentation of a stimulus, only the sounds conveyed the identity of a target letter. Assuming that the two types of tactile stimulations were used only for stimulus detection, we would expect to see some further improvement, or at least no worsening in performance in the post-training test. In contrast, if the key-identity information could be perceived from the tactile stimulation and combined with the visual stimulus to guide the keypress response, we expected that participants' performance would become worse after the tactile key-identity information had been eliminated in the post-training test.

In Experiment 1b, participants received no key-identity information during the training since the tactile stimulation would be a 250-Hz buzz for all letters. The letter sounds were used only in the post-training test. If some key-identity information could be perceived from the tactile stimulation in the post-training test, we expected to see an improvement in participants' performance.

5.2.1 *Method*

Participants: A total of 24 graduate and undergraduate students (15 males and 9 females with an average age of 20.9 years) participated with informed consent. They were randomly assigned to Experiment 1a or 1b. To eliminate the possible effects of handedness, all participants were right-handed by self-report and performed the experiment using their right hand. They were naïve to the purposes of this study.

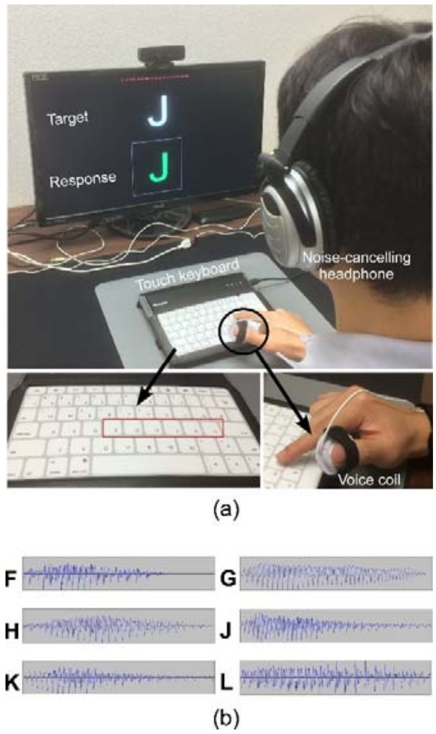


Figure 19. (a) The experimental setup. The bottom images show close up views of the touch keyboard and the voice coil. (b) The letter sounds used as the tactile stimulations.

Apparatus & Stimuli: As shown in Fig 19(a), the experimental setup consisted of a 27-inch LCD (resolution: 1920x1080 @ 120 Hz, Model# VG278HE, ASUSTeK Computer Inc., http://www.asus.com) for showing the target letter (the white "J") and providing a visual feedback of the participant's response (the green "J" inside the box), a Kensington Keylite Touch keyboard (Model# K39598US, Kensington Computer Products Group, San Mateo, CA, http://www.kensington.com) for the participant to enter responses, a Xenics Vibeholic voice coil (Model# Vibeholic-001, Xenics Corpoartion, Seoul, Korea, http://www.xenics.co.kr) for delivering the tactile stimulation, and a Dell Precision T5500

Workstation (Model# T5500, Dell Inc., Round Rock, TX, http://www.dell.com) for controlling the experimental stimuli and recording the accuracy and response time of keypresses. The Kensington keyboard has a flat, pressure-sensitive surface and thus provides no haptic feedback to users. It measures 110 x 225 x 5 mm in overall size, and its keys measure 16 x 16 mm with a key-spacing of 2 mm. The keyboard was connected to the Dell workstation via Bluetooth. The Vibeholic voice coil was connected to and driven by a Sound Blaster sound card (Model#: Sound Blaster Live!, Creative Technology Ltd., Jurong East, Singapore, http://www.creative.com). The sound level was set to 15% of the full volume, which was selected based on three participants' subjective opinions from a pilot test: Three participants including two authors tried different levels of volume, felt the vibration of the voice coil plate at the fingertip, and agreed that the tactile feedback could be clearly felt at that volume level but could not be heard owing to the headphones' sound isolation. In this and all subsequent experiments, the voice coil was attached to the participant's right thumb while they typed using the index finger.

The experimental stimuli consisted of 6 upper-case letters (F, G, H, J, K, & L). These letters are adjacent keys so that they can be hit with comparable motor effort. Importantly, their sounds are dissimilar when being felt by touch (Figure 19b). The sounds were generated using the AT&T Lab's Natural Voices Text-To-Speech Demo (http://www.corp.att.com/attlabs/technology/demos.html, Mike's voice in American English accent with a sampling rate of 16k Hz), and then filtered and down-sampled to 2400 Hz using Audacity (http://www.audacityteam.org), an open-source audio recording and editing program. The sound duration ranged from 227 to 362 ms. They were aligned with a zero voice-onset-time for accurate control of timing. As a result, these sounds were quite different from the natural sounds of the letters. But they were still quite discernible and could be differentiated by the buildup or decay of the vibrations as well as the slight difference in duration. In addition, active noise-cancelling headphones were worn by the participants to further prevent them from hearing rather than tactually feeling the sounds.

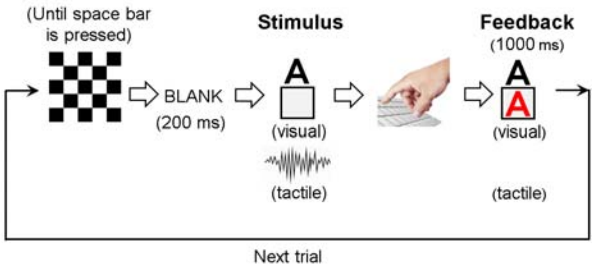


Figure 20. The sequence of events during the experimental paradigm for each trial in Experiments 1a & 1b. In these two experiments, the tactile stimulation was part of the stimulus.

Procedure & Design: Figure 20 shows the sequence of events during an experimental trial in Experiments 1a & 1b. Participants were tested individually in a single-letter-typing task. Each trial began with a checkerboard mask that remained on the screen until the space bar was pressed by the participant. After 200 ms (24 frames @ 120 Hz), a target letter was presented on the screen in white color along with a tactile stimulus. The presentation of the visual and tactile stimuli was synchronized. The participant was instructed to press the target letter on the keyboard as quickly and accurately as possible using his or her right index finger. After a keypress had been detected, the pressed key would be shown in green color for 1 second inside the box underneath the target. Next, the checkerboard mask appeared again, and the next trial began.

A participant first went through 24 practice trials to familiarize himself or herself with the keyboard and the task. Practice trials followed the same procedure as the experimental trials, except that different letters and sounds were used. The participant then went through a total of six experimental sessions as described below. Each session consisted of 50 trials, including 8 repetitions of six target letters in a random order and two warm-up trials. The experimental sessions were:

Pre-training session: The voice coil was worn by the participant but turned off in this session. The participant's typing performance was measured without tactile feedback. This provides a baseline for comparison with the conditions with the tactile buzz or letter-sounds.

Training sessions: On each trial, the participant would see a target letter and at the same time feel a vibration at his or her fingertip. The tactile stimulation was the letter's sound pattern (See Figure 19(b)) in Experiments 1a, or a meaningless 250Hz, 200-ms buzz in Experiments 1b. Four training sessions were run.

Post-training session: The tactile stimulation was switched to a 250-Hz, 200-ms buzz in Experiments 1a. In Experiments 1b, the target letter's sound was played.

In addition to these single-letter-typing-task sessions, two sound-recognition tests were carried out immediately before and after the training sessions. On each trial, a sound pattern was played and felt by the participant at his or her fingertip after the space bar was pressed. No target was visually displayed, and the participant was asked to guess what the sound pattern was on the basis of their tactile feelings. The participants were clearly instructed that only 6 letters and their sounds were used as the stimuli.

The RT and accuracy were measured for typing responses and the recognition rate was measured for the sound-recognition tests. Typically, the participants finished one trial in less than 3 seconds and a whole experimental session in less than 3 minutes. There was a break of 3 minutes for rest between the experimental sessions. The entire experiment took approximately one hour.

5.2.2 *Results & Discussion*

In both experiments, all participants showed high accuracy rates (>96%) in the single-letter-typing task: only two of 24 subjects made more than 2 errors within an experimental session. Therefore the error rate was not statistically tested. Instead, our analyses focused on the RT as well as the recognition rate of tactile sounds.

Consider first Experiment 1a. As shown in Figure 21(a), the presentation of tactile sounds improved the participants' typing performance gradually over the course of training. Compared to the pre-training test, no immediate improvement was shown in RT in the first training session (964.9 vs. 1028.4 ms, paired t-test: $t(11) = 1.16$, $p = 0.27$, Cohen's d = 0.70). Over the four training sessions, the mean RT gradually reduced from 1028.4 ms to 902.1 ms (one-way repeated measures ANOVA: $F(3,33) = 11.53$, $p < 0.001$, partial η^2= 0.51). The recognition rate of tactile sounds was also found to significantly increase from a chance level of 16.7 ± 1.4(SE)% to 34.7 ± 3.6(SE)% (paired t-test: $t(11) = 4.98$, $p < 0.001$, Cohen's $d = 3.0$), indicating that an association had been formed between the letter identity and the tactile sounds. Thus not surprisingly, the mean RT was found to increase to 971.2 ms in the post-training test when the tactile stimulation was changed to a meaningless 250-Hz buzz (paired t-test: $t(11) = 2.36$, $p = 0.038$, Cohen's $d = 1.42$).

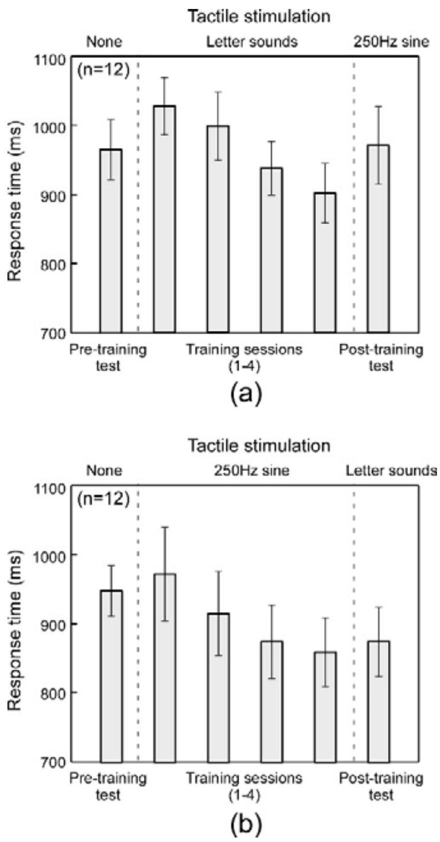


Figure 21. Mean RTs of typing response measured in Experiments 1a (a, top panel) & 1b (b, bottom panel). Error bars represent ±1 standard error. In these experiments, the tactile stimulation was presented as part of the stimulus.

Significant reductions of RT were also observed in Experiment 1b across the training sessions: While the mean RT was comparable between the pre-training and first-training sessions (946.2 vs. 972.5 ms, paired t-test: $t(11) = 0.35$, $p = 0.73$, Cohen's $d = 0.21$), it gradually reduced to 859.3 ms at the end of training (Figure 21(b), one-way repeated measures ANOVA: $F(3,33) = 6.50$, $p = 0.001$, partial $\eta^2 = 0.37$). Clearly, the multimodal stimuli led to faster typing reactions. In contrast, no significant change was found in the recognition rate of tactile sounds (19.3 ± 2.4(SE)% vs. 20.8 ± 3.0(SE)%, paired t-test: $t(11) = 0.73$, $p = 0.48$, Cohen's $d = 0.44$), and the post-training recognition rate (20.8 ± 3.0 (SE)%) was not better than chance. This was expected because here a 250-Hz buzz was played for all letters during the training and the buzz carried no key identity information. The post-training RT was similar to that obtained in the last training session (874.5 ms vs. 859.3 ms, paired t-test: $t(11) = 0.63$, $p = 0.54$, Cohen's $d = 0.38$). Clearly, although the letter sounds were presented in the post-training trials, no key-identify information had been automatically extracted from the tactile stimulation and used to guide the keypress response.

5.3 Exp set 2. Effectiveness of the key-identity information as response feedback

This set of two experiments used the same design and method as in Experiments 1a & 1b. The only difference was that the tactile stimulation was presented as response feedback after keypresses (Figure 22). Again, participants were trained to form an association between the letters and their vibration patterns in Experiment 2a by simultaneously presenting the response letter by both vision and touch, while in Experiment 2b no such training was provided. We expected to observe similar patterns in the results as in Experiments 1a & 1b.

5.3.1 *Method*

Another group of 26 graduate and undergraduate students (17 males and 9 females with an average age of 21.5 years) were tested, 13 each in the two experiments. None of them participated in Experiments 1a & 1b. All participants were right-handed by self-report and performed the experiment using their right hand. They were naïve to the purposes of this study.

The apparatus, stimuli, and procedure were the same as those in the previous experiments. The design of Experiment 2a/2b was identical to Experiment 1a/1b. The only change was the time at which the tactile stimulation was played. As shown in Fig 22, a letter sound or a 250-Hz, 200-ms buzz was played immediately after a keypress was detected.

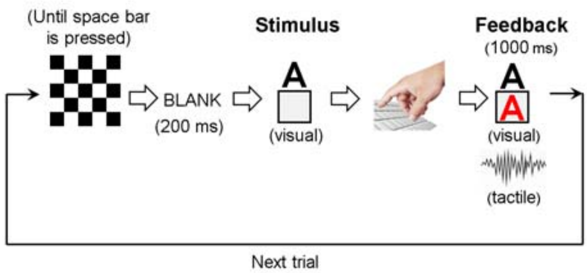
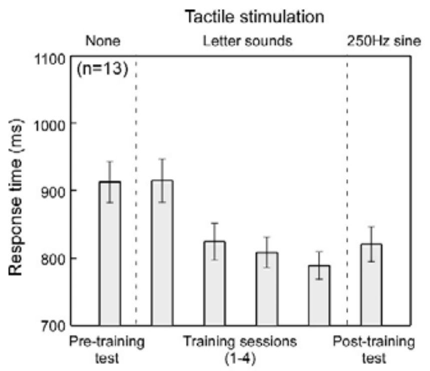


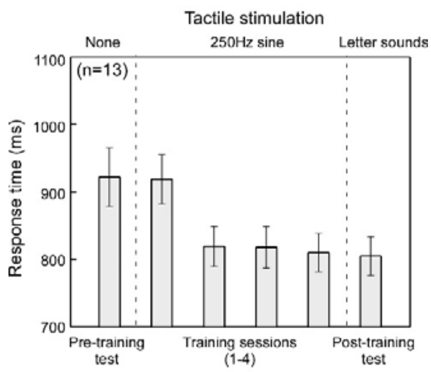
Figure 22. The sequence of events for each trial in Experiments 2a & 2b. In these two experiments, the tactile stimulation served as a response feedback and was played after the keypress.

5.3.2 *Results*

The results of Experiment 2a were quite similar to those observed in Experiment 1a. As shown in Figure 23(a), there was no significant RT difference between the pre-training and first-training sessions (912.6 vs. 914.5 ms, paired t-test: $t(12) = 0.06$, $p = 0.95$, Cohen's $d = 0.03$). The effects of training were evident: the mean RT reduced from 914.5 ms to 789.0 ms (one-way repeated measures ANOVA: $F(3, 36) = 15.40$, $p < 0.001$, partial $\eta^2 = 0.56$) and the recognition rate of tactile sounds also increased from 18.6 ± 2.8(SE) % to 31.2 ± 3.5 (SE) % (paired t-test: $t(12) = 2.91$, $p = 0.013$, Cohen's $d = 1.68$). In the post-training test, the RT increased to 820.3 ms since the tactile sound was replaced with a meaningless 250-Hz buzz (paired t-test: $t(12) = 2.57$, $p = 0.025$, Cohen's $d = 1.48$).



(a)



(b)

Figure 23. Mean RTs of typing response measured in Experiments 2a (a, top panel) and 2b (b, bottom panel). Error bars represent ±1 standard error. In these experiments, the tactile stimulation was presented as a response-feedback.

The results observed in Experiment 2b were also quite similar to those in Experiment 1b. The mean RT was 921.7 and 918.3 ms, respectively, in the pre-training and first-training sessions (paired t-test: $t(12) = 0.07$, $p = 0.95$, Cohen's $d = 0.04$). It reduced from 918.3 ms to 809.2 ms over the training (Figure 23(b), one-way repeated measures ANOVA: $F(3,36) = 8.53$, $p = 0.013$, partial $\eta^2 = 0.42$). The post-training RT was very similar to that measured in the last training session (809.2 ms vs. 804.1 ms, paired t-test: $t(12) = 0.31$, $p = 0.76$, Cohen's $d = 0.18$). As in Experiment 1b, no improvement was observed in the recognition rate of tactile sounds (16.7 ± 1.8(SE) % vs. 18.4 ± 2.2 (SE) %, paired t-test: $t(12) = 0.90$, $p = 0.39$, Cohen's $d = 0.52$).

5.4 General Discussion

The present study investigated if participants could learn an association between visual and haptic information presented at spatially separated locations, and if participants' typing performance would be improved by this learning. Visual information was a key shown on a LCD, and as haptic information, the tactile stimulation identifying a specific key was provided on a participant's thumb. Recognition rates indicated the level of learning an association between visual and haptic information. We examined if the information about key identity could be effectively delivered via tactile stimulation and if such information could be used to facilitate typing on a flat keyboard. Our experiments found that the key identity information could be conveyed by using vibration patterns, but such information would hardly be perceived unless sufficient training was provided. Once an association was learned between the key identity and the tactile stimulation, the key-identity information was effectively used to facilitate typing performance in terms of shorter response times. These suggest that participants could overcome spatial distance between visual and haptic inputs in integrating them.

One finding from this study is that similar facilitative effects were observed when the tactile stimulation was presented at different time for different purposes. In Experiments 1a & 1b, the tactile stimulation was part of stimulus and presented when the target letter was shown. In Experiments 2a & 2b, the tactile stimulation served as a response feedback and was played after a keypress had occurred. Although such feedback was too late to be useful to correct the keypress action in the current trial, the mean RT was found to be shortened by an amount similar to that observed in Experiments 1a & 1b. This may suggest that some memory mechanisms are involved in multisensory processing and so the information extracted from the tactile feedback can be carried over to the subsequent trials to produce a facilitative outcome. Such cognitive factors are not included in the MLE model, however, our results suggest that those factors should also be taken into account in modeling multisensory integration. Another possible explanation may be the preparation-enhancement hypothesis. The key-identity information from the response feedback may alter the readiness of the visual-motor system and make it more ready for producing a response in the next trial.

There is no doubt that our tactile perceptual system has amazing capabilities of processing complex information and performing challenging tasks such as the Tadoma speechreading. However, such capabilities have to be eventually developed through extensive training, rather than being readily available. For example, deaf-blind individuals usually take years of training to learn Tadoma (Reed et al., 1992; Dinsmore, 1959). In this study, although the small target set consisted of only six letters and their sounds, the participants' tactile recognition of the sounds was still quite poor after four training sessions. Extended training will improve the recognition rate and make the tactile key-identity information more efficient to use. On the other hand, it should be noted that the sound stimuli used in this study were not optimally designed. For example, high frequencies were filtered from the stimuli (Figure 19b). But given that the tactile system has limited temporal resolution, we believe that the loss of high-frequency sounds should not be the major cause of difficulty in the tactile sound recognition. Instead, the difficulty may be partially attributed to other factors, for example, the removal of the voice-onset-time that is known to play important roles in categorizing speech sounds (Ganong, 1980). Previous research has identified several ways to enhance information transmission. For example, though our ability to process unidimensional stimuli may be limited, the range can be extended by employing multidimensional stimuli. For tactile communications, correlated changes in the frequency and amplitude enhance perception as compared to varying amplitude or frequency alone (Murray, Klatzky, & Khosla, 2003).

CHAPTER 6

SUMMARY, APPLICATIONS AND FUTURE DIRECTIONS

This investigation considered the spatial and temporal characteristic of multisensory integration. We found that participants could integrate visual and haptic information despite spatial or temporal asynchrony. Chapter 5 showed that although visual and haptic inputs were spatially separated, an association between them could be learned by participants, leading to the improvement of typing performance. In chapter 3 and 4, it has been shown that participants could integrate information from two sensory modalities with visual delay, however, this delay affected participants' perception of stiffness. Chapter 3 revealed that a constant visual delay increased the perceived stiffness, while a variable visual delay made participants depend more on the haptic sensations in stiffness perception. This could be explained by the MLE model. Chapter 4 showed that participants judged stiffness stiffer at faster speeds, and interaction speed was positively correlated with stiffness overestimation. The MLE was not enough to explain the results of chapter 4 in which action-perception interaction was included, therefore, a more general Bayesian model was required.

According to a Bayesian inference model proposed by Ernst and Bülthoff (2004) (Figure 24), a perception is generated from sensory inputs through a two-stage processing. The sensory data are first integrated within or across modalities using a MLE estimator. The outcome of the MLE estimator is further combined with prior knowledge or other cognitive inputs, yielding a final percept. The Bayes' rule is applied to maximize the a-posteriori probability of the final percept.

This model can be applied to explain the results in chapter 4. In this model, perception is considered as an active process: The perception is used to guide action, and the action brings new information to be processed by perceptual system, which in turn provides an updated guidance for the action. Considering that stiffness is perceived through action, such action-perception interaction can be explained more appropriately by this model.

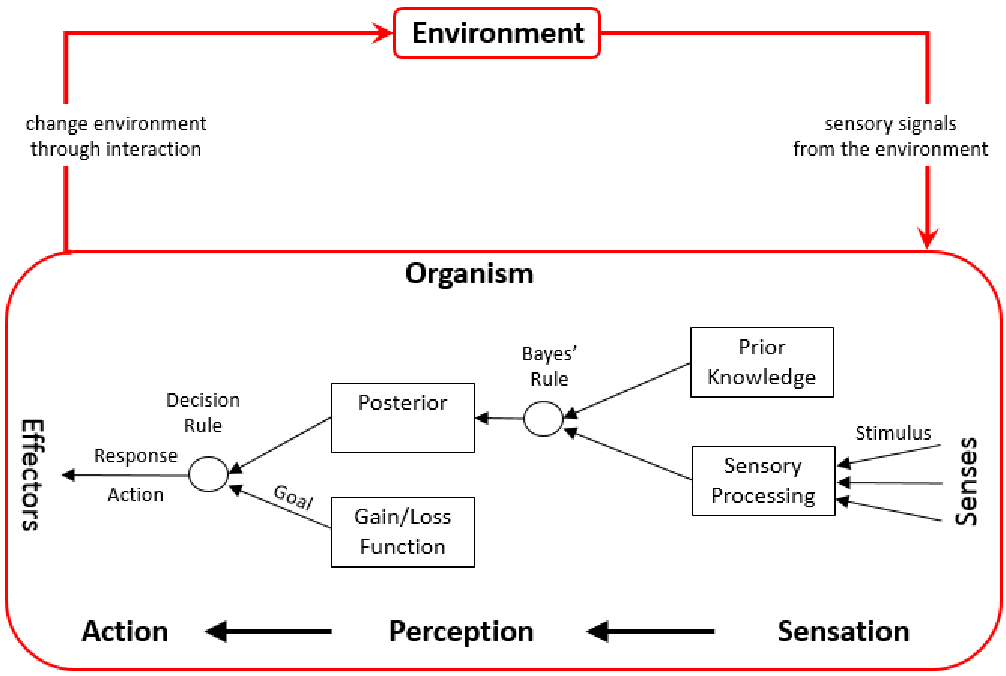


Figure 24. The Bayesian inference model (adapted from figure 1 in Ernst & Bülthoff, 2004). According to this model, the sensory inputs are first integrated using a MLE estimator and then combined with prior knowledge to produce a final perception (i.e., the posterior in the model). The perception is used to guide action, and the action then generates new sensory data to update the perception, forming a perception-action loop. The Bayes' rule is applied to all stages of information processing.

This work suggests applications in the medical field, such as VR, AR, or teleoperation systems. Given the existing temporal asynchrony and spatial disparity in such settings, accurate perception of stiffness would be hindered. This work shows a possibility of compensating stiffness overestimation. For instance, the magnification of visual cues or adjusting the simulated stiffness by measuring the visual delay and a user's interaction can be considered.

In chapter 5, multisensory integration with spatial asynchrony was studied, however, it was not in the context of stiffness perception. The impact of spatial separation and temporal delay on participants' perception of stiffness needs to be examined. Investigating the influence of top-down factors in the multisensory integration will also be the future directions. Prior knowledge, experience, and attention affect the perception of stiffness. For example, if a surgeon believes that a tumor is present from other test results, it would affect that surgeon's perception of stiffness. Studying these factors would provide solutions for many theoretical and practical problems.

REFERENCES

Adams, C. K., Hall, D. C., Pennypacker, H. S., Goldstein, M. K., Hench, L. L., Madden, M. C., Stein, G. H., & Catania, A. C. (1976). Lump detection in simulated human breasts. *Perception & Psychophysics*, *20*(3), 163-167.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, *14*(3), 257-262.

Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, *12,* 241-253.

Barlow, H. B., & Földiák, P. (1989). Adaptation and decorrelation in the cortex. In R. Durbin, C. Miall, & G.Mitchison (Ed.). *The Computing Neuron* (pp. 54-72). Wokingham, England: Addison-Wesley.

Bender, G. T. (1999). *Touch screen performance as a function of the duration of auditory feedback and target size*, Doctoral dissertation, Wichita State University, College of Liberal Arts and Sciences.

Blair, G. S., & Coppen, F. M. V. (1942). The subjective conception of the firmness of soft materials. *The American Journal of Psychology*, 215-229.

Bloom, H. S., Criswell, E. L., Pennypacker, H. S., Catania, A. C., & Adams, C. K. (1982). Major stimulus dimensions determining detection of simulated breast lesions. *Perception & psychophysics*, *32*(3), 251-260.

Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature*, 391, 756.

Brady, K., Wu, B., Sim, S., Enquobahrie, A., Ortiz, R., & Arikatla, S. (2016). Modeling Reduced User Experience caused by Visual Latency. In *7th International Conference on Applied Human Factors and Ergonomics*, Orlando, Florida.

Brewster, S., & Brown, L. M. (2004, January). Tactons: structured tactile messages for non-visual information display. *Australasian User Interface Conference*, 18-22.

Butler, J.S., Smith, S.T., Campos, J.L., & Bülthoff, H.H. (2010). Bayesian integration of visual and vestibular signals for heading. *Journal of Vision*, *10* (11), 1-13.

Calhoun, G., Draper, M., Ruff, H., Fontejon, J., & Guilfoos, B. (2003, October). Evaluation of tactile alerts for control station operation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 47(20), 2118-2122, Los Angeles, CA.

Campbell, H. S., Fletcher, S. W., Pilgrim, C. A., Morgan, T. M., & Lin, S. (1990). Improving physicians' and nurses' clinical breast examination: a randomized controlled trial. *American journal of preventive medicine*, *7*(1), 1-8.

Cholewiak, S. A., Tan, H. Z., & Ebert, D. S. (2008, March). Haptic identification of stiffness and force magnitude. In *2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 87-91). IEEE.

Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society,* 102-118.

Di Luca, M., Knörlein, B., Ernst, M. O., & Harders, M. (2011). Effects of visual-haptic asynchronies and loading-unloading movements on compliance perception. *Brain Research Bulletin*, *85*(5), 245-259.

Dinsmore, A. B. (1959). *Methods of communication with deaf-blind people*. New York: American Foundation for the Blind.

Dixon, W.J., & Mood, A.M. (1948). A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association*, *43*, 109–126.

Drewing, K., Ramisch, A., & Bayer, F. (2009, March). Haptic, visual and visuo-haptic softness judgments for objects with deformable surfaces. In *Proceedings of Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 640-645). IEEE.

Drewing, K., Ramisch, A., & Bayer, F. (2009). Multisensory softness perception of deformable objects, *Perception*, *38s,* 144.

Endo, H. (2016). Pressing movements and perceived force and displacement are influenced by object stiffness, *Physiology & Behavior*, *163*, 203-210.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162–169.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*(6), 429-442.

Fujisaki, W. (2012). Effects of delayed visual feedback on grooved pegboard test performance. *Frontiers in psychology*, *3*, 61.

Fung, Y.C. (1981). *Biomechanics: Mechanical Properties of Living Tissue,* 1st Edition. New York: Springer.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110 – 125.

Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). Integration of visual and tactile signals from the hand in the human brain: an FMRI study. *Journal of neurophysiology*, *105*(2), 910-922.

Gepshtein, S., Burge, J., Ernst, M. O., & Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity. *Journal of Vision*, *5*(11), 1013-1023.

Ghahramani, Z., Wolpert, D. M., & Michale I, J. (1997). Computational models of sensorimotor integration. *Advances in Psychology*, *119*, 117-147.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception,* Boston: Houghton Mifflin.

Gunther, E., Davenport, G., & O'Modhrain, S. (2002). Cutaneous grooves: composing for the sense of touch. In *Proceedings of Conference on New Instruments for Musical Expression*, Dublin, IR, 1-6.

Han, G., & Choi, S. (2010). Extended rate-hardness: A measure for perceived hardness. *Haptics: generating and perceiving tangible sensations*, 117-124.

Harper, R., & Stevens, S. S. (1964). Subjective hardness of compliant materials. *Quarterly Journal of Experimental Psychology*, *16*(3), 204-215.

Helbig, H. B., & Ernst, M. O. (2007). Knowledge about a common source can promote visual- haptic integration. *Perception-London*, *36*(10), 1523-1534.

Hellman, R. P., & Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *The Journal of the Acoustical Society of America*, *33*(5), 687-694.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, *4*(12), 1.

Hoggan, E., Brewster, S. A., & Johnston, J. (2008, April). Investigating the effectiveness of tactile feedback for mobile touchscreens. In *Proceedings of the SIGCHI conference on Human factors in computing systems*,1573-1582.

Horeman, T., Rodrigues, S. P., van den Dobbelsteen, J. J., Jansen, F. W., & Dankelman, J. (2012). Visual force feedback in laparoscopic training. *Surgical endoscopy*, *26*(1), 242-248.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision research*, *34*(17), 2259-2275.

Jones, L. A., & Hunter, I. W. (1990). A perceptual analysis of stiffness. *Experimental Brain Research*, *79*(1), 150-156.

Jones, L. A. (2000). Kinesthetic sensing. In Cutkosky, M., Howe, R., Salisbury, K., & Srinivasan, M. (Eds.), *Human and Machine Haptics*, MIT Press.

Jones, L. A., Hunter, I. W., & Irwin, R. J. (1992). Differential thresholds for limb movement measured using adaptive techniques. *Perception & psychophysics*, *52*(5), 529-535.

Kaernbach. C. (2001). Adaptive threshold estimation with unforced-choice tasks. *Perception and Psychophysics*, *63*, 1377–1388.

Kim, J. R., & Tan, H. Z. (2014, February). A study of touch typing performance with keyclick feedback. *Haptics Symposium*, 227-233.

Kim, J. R., & Tan, H. Z. (2015, June). Effect of information content in sensory feedback on typing performance using a flat keyboard. In *World Haptics Conference*, 228-234.

Kim, T., Zimmerman, P. M., Wade, M. J., & Weiss, C. A. (2005). The effect of delayed visual feedback on telerobotic surgery. *Surgical Endoscopy and Other Interventional Techniques*, *19*(5), 683-686.

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A Practical Introduction*. Academic Press: London.

Klatzky, R. L., & Lederman, S. J. (2002). Perceiving texture through a probe. In M. L. McLaughin, J. Hespanha, & G. Sukhatme (Eds.), *Touch in virtual environments* (pp. 180-193). Upper Saddle River, NY: Prentice Hall PTR.

Klatzky, R. L., Lederman, S. J., & Reed, C. (1989). Haptic integration of object properties: texture, hardness, and planar contour. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 45-57.

Klatzky, R. L., & Wu, B. (2014). Visual-haptic compliance perception. In M. Di Luca (Ed.), *Multisensory Softness: Perceived Compliance from Multiple Sources of Information*. (pp. 17-30). London, England: Springer.

Knill, D. C., & Saunders, J. (2002). Humans optimally weight stereo and texture cues to estimate surface slant. *Journal of Vision*, *2*(7), 400.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision research*, *43*(24), 2539-2558.

Kuschel, M., Buss, M., Freyberger, F., Farber, B., & Klatzky, R. L. (2008, March). Visual-haptic perception of compliance: fusion of visual and haptic information. In *Haptic interfaces for virtual environment and teleoperator systems,* (pp. 79-86). IEEE.

Kuschel, M., Di Luca, M., Buss, M., & Klatzky, R. L. (2010). Combination and integration in the perception of visual-haptic compliance information. *Haptics, IEEE Transactions on*, *3*(4), 234-244.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of Optical Society of America*, *18*(9), 2307-2320.

Langenberg, U., Hefter, H., Kessler, K. R., & Cooke, J. D. (1998). Sinusoidal forearm tracking with delayed visual feedback I. Dependence of the tracking error on the relative delay. *Experimental brain research*, *118*(2), 161-170.

Lawrence, D. A., Pao, L. Y., Dougherty, A. M., Salada, M. A., & Pavlou, Y. (2000). Rate-hardness: A new performance metric for haptic interfaces. *IEEE Transactions on Robotics and Automation*, *16*(4), 357-371.

Lederman, S. J., & Klatzky, R. L. (1996). *Action for perception: Manual exploratory movements for haptically processing objects and their features*. San Diego, CA: Academic Press.

Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, *71*(7), 1439-1459.

Lee, S., & Zhai, S. (2009, April). The performance of touch screen soft buttons. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 309-318.

Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and cognition*, *64*(1), 104-109.

Loomis, J. M., & Lederman, S. J. (1986). Tactual perception. *Handbook of perception and human performances*, *2*, 2.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature neuroscience*, *9*(11), 1432-1438.

Ma, Z., Edge, D., Findlater, L., & Tan, H. Z. (2015, June). Haptic keyclick feedback improves typing speed and reduces typing errors on a flat keyboard. In *World Haptics Conference*, 220-227.

MacKenzie, I. S., & Ware, C. (1993, May). Lag as a determinant of human performance in interactive systems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 488-493). ACM.

Maher, C.G., & Adams, R.D. (1996). Stiffness judgments are affected by visual occlusion. *Journal of Manipulative Physiological Therapeutics*, *19*(4), 250–256.

Marescaux, J., Leroy, J., Rubino, F., Smith, M., Vix, M., Simone, M., & Mutter, D. (2002). Transcontinental robot-assisted remote telesurgery: feasibility and potential applications. *Annals of surgery*, *235*(4), 487-492.

Marks, L. E., & Gescheider, G. A. (2002). Psychophysical scaling. In H. Pashler (Ed. In Chief) & J. Wixted (Vol. Ed.), *Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (pp. 91–138). New York: John Wiley & Sons.

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision research*, *21*(4), 491-500.

Miall, R. C., Weir, D. J., & Stein, J. F. (1985). Visuomotor tracking with delayed visual feedback. *Neuroscience*, *16*(3), 511-520.

Michaels, C. F., & de Vries, M. M. (1998). Higher order and lower order variables in the visual perception of relative pulling force. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 526.

Murray, A. M., Klatzky, R. L., & Khosla, P. K. (2003). Psychophysical characterization and testbed validation of a wearable vibrotactile glove for telemanipulation. *Presence: Teleoperators and Virtual Environments*, *12*(2), 156-182.

Nickerson, R. S. (1973). Intersensory facilitation of reaction time: energy summation or preparation enhancement?. *Psychological review*, 80, 189-509.

Orban, G. A., de Wolf, J., & Maes, H. (1984). Factors influencing velocity coding in the human visual system. *Vision research*, *24*(1), 33-39.

Oruç, I., Maloney, T. M., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, *43,* 2451–2468.

Perreault, C., & Auclair-Fortier, M. F. (2007, May). Speckle simulation based on B-mode echographic image acquisition model. In *Fourth Canadian Conference on Computer and Robot Vision* (pp. 379-386). IEEE.

Pressman, A., Welty, L. J., Karniel, A., & Mussa-Ivaldi, F. A. (2007). Perception of delayed stiffness. *The International Journal of Robotics Research*, *26*(11-12), 1191-1203.

Raaja, O. H. (1962). Statistical facilitation of simple reaction time. *Transctions of the New York Academy of Sciences*, 24, 574-590.

Reed, C. M., Durlach, N. I., & Delhorne, L. A. (1992). The tactual reception of speech, fingerspelling, and sign language by the deaf-blind. In *Digest of Technical Papers of the Society for Information Display International Symposium*, 23,102-105.

Roland, P. E., & Ladegaard-Pedersen, H. (1977). A quantitative analysis of sensations of tension and of kinaesthesia in man: Evidence for a peripherally originating muscular sense and for a sense of effort. *Brain*, *100*(4), 671-692.

Sarvazyan, A. P., Skovoroda, A. R., Emelianov, S. Y., Fowlkes, J. B., Pipe, J. G., Adler, R. S., Buxton, R. B., & Carson, P. L. (1995). Biophysical Bases of Elasticity Imaging. In J. P. Jones (Ed)., *Acoustical imaging*, *21*. (pp. 223-240). New York, NY: Springer.

Shen, Y., & Zelen, M. (2001). Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. *Journal of Clinical Oncology*, *19*(15), 3490-3499.

Sheridan, T. B., & Ferrell, W. R. (1963). Remote manipulative control with transmission delay. *IEEE Transactions on Human Factors in Electronics*, *1*, 25-29.

Shi, Z., Hirche, S., Schneider, W. X., & Muller, H. (2008, March). Influence of visuomotor action on visual-haptic simultaneous perception: A psychophysical study. In *2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 65-70). IEEE.

Shimada, S., Fukuda, K., & Hiraki, K. (2009). Rubber hand illusion under delayed visual feedback. *PLoS ONE*, *4*(7), e6185.

Sim, S. H., Wu, B., & Klatzky, R. (2014, November). Influence of haptic-visual asynchrony on stiffness perception. In *55th Annual Meeting of the Psychonomic Society*, Long Beach, CA.

Srinivasan, M. A., Beauregard, G. L., & Brock, D. L. (1996, November). The impact of visual information on the haptic perception of stiffness in virtual environments. In *ASME Winter Annual Meeting*, 58, 555-559.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects.* New York: Wiley.

Sun, Y., Hollerbach, J. M., & Mascaro, S. A. (2008). Predicting fingertip forces by imaging coloration changes in the fingernail and surrounding skin. *Biomedical Engineering, IEEE Transactions on*, *55*(10), 2363-2371.

Todd, J. W. (1912). Reaction to multiple stimuli. *Archives of Psychology*, 3, 145.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 80.

van Beers, R. J., Sittig, A. C., & van der Gon, J. J. D. (1998). The precision of proprioceptive position sense. *Experimental Brain Research*, *122*(4), 367-377.

van Beers, R. J., Sittig, A. C., & van der Gon, J. J. D. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology*, *81*(3), 1355-1364.

Varadharajan, V., Klatzky, R., Unger, B., Swendsen, R., & Hollis, R. (2008, March). Haptic rendering and psychophysical evaluation of a virtual three-dimensional helical spring. In *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 57-64). IEEE.

Vogels, I. M. (2004). Detection of temporal delays in visual-haptic interfaces. *Human factors: The journal of the Human Factors and Ergonomics society*, *46*(1), 118-134.

Weber, E. H. (1834/1996). De pulsu, resorptione, auditu et tactu. *Annotationes anatomicae et physiologicae*. Leipzig: Koehler.

White, P. A. (2012). The experience of force: the role of haptic experience of forces in visual perception of object motion and interactions, mental simulation, and motion-related judgments. *Psychological bulletin*, *138*(4), 589-615.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & psychophysics*, *63*(8), 1293-1313.

Wickens, T.D. (2001). *Elementary Signal Detection Theory*. New York, NY: Oxford University Press.

Woodson, W. E., & Conover, D. W. (1964). Human engineering guide for equipment designers. University of California Press.

Wu, B., Klatzky, R. L., Hollis, R., & Stetten, G. (2012, November). Visual perception of viscoelasticity in virtual materials. In *53rd annual meeting of the psychonomic society*, Minneapolis, MN.

Wu, B., Sim, S. H., Enquobahrie, A., & Ortiz, R. (2015, May). Effects of visual latency on visual-haptic experience of stiffness. In *Seventh International Workshop on Quality of Multimedia Experience* (pp. 1-6). IEEE.

Wu, B., Sim, S. H., Hibbard, J., & Klatzky, R. L. (2014, November). Optimal integration of spatiotemporal cues in visual perception of stiffness. In *55th Annual Meeting of the Psychonomic Society*, Long Beach, CA.

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*(18), 2685–2696.

APPENDIX A

PERMISSION TO USE PUBLISHED WORKS

I, Sung Hun Sim, am the second listed co-author of a conference proceeding titled "Effects of visual latency on visual-haptic experience of stiffness," and this work was used in Chapter 3 of this dissertation with only very minor edits. All co-authors have granted their permissions to use this work in this dissertation.

I, Sung Hun Sim, am the first listed co-author of a meeting presentation titled "Influence of haptic-visual asynchrony on stiffness perception," and this work was included as a part of chapter 4 of this dissertation with only very minor edits. All co-authors have granted their permissions to use this work in this dissertation.

I, Sung Hun Sim, am the first listed co-author of a conference proceeding titled "Improved typing on a flat keyboard via tactile key-identity feedback," and this work was used in Chapter 5 of this dissertation with only very minor edits. All co-authors have granted their permissions to use this work in this dissertation.

APPENDIX B

IRB APPROVAL DOCUMENT

Consent Form: Social Behavioral

Title of research study:

Haptic/visual perception of virtual viscoelastic tissue

Investigator:

Dr. Bing Wu, Assistant Professor in Dept. of Human Systems Engineering, Ira A. Fulton Schools of Engineering, Arizona State University.

Why am I being invited to take part in a research study?

We invite you to take part in a research study because you (1) age between 18 and 55 years, (2) are NOT pregnant (if female), (3) have normal or corrected-to-normal vision and stereoacuity of at least 40 arc-sec, and (4) have no physical/mental disorders.

Why is this research being done?

In many clinical practices, it is necessary for a physician to feel tissue mechanical properties, such as stiffness, in order to diagnose pathologies like lumps or tumors and to perform fine surgical procedures with minimal disturbance to the patient. But still, we have little understanding of human perception of mechanical properties of soft tissues. In this study, we use simulation techniques and psychophysical methods to investigate how people judge such properties, more specifically, elasticity (i.e. stiffness) and viscosity (i.e. thickness) of soft materials from the forces and torques arising from interactions, along with visual cues to deformation. The knowledge obtained from this study will be applied to the development of technologies for robot-assisted telesurgical systems and surgical training systems to facilitate medical training and surgical performance.

How long will the research last?

We expect that individuals will spend about 1-2 hour participating in the proposed activities.

How many people will be studied?

We expect about 360 people will participate in this research study.

What happens if I say yes, I want to be in this research?

You are free to decide whether you wish to participate in this study. If you decide not to participate, there will be NO penalty to you, and you will NOT lose any benefits or rights to which you are entitled. If you agree to participate, you will be asked to sign this consent form.

The study will be conducted in Dr. Wu's research laboratory in Simulator Building. In this experiment, the stimuli will be simulations of soft objects, which are implemented as computer software models and presented using a visual-haptic simulator. The properties of virtual objects such as viscosity, elasticity, size, shape, etc. will be manipulated to examine their influence on human perception. The experiment will consist of multiple trials. On each trial, you will grasp the handle of the visual-haptic simulator, and use it to feel a virtual object through force feedback provided by the simulator. You will actively control how much force is applied. Meanwhile, you will see a simulated ultrasound video that provides the visual cues to deformation. You will be asked to undergo some or all experimental procedures listed below:

(1) Target-detection tasks: You are required to judge if there exists a target or not in the stimulus and then press the “Y” or “N” keys to indicate your judgment.

(2) Magnitude-estimation task: You are required to freely assign a number to the stimulus, with the rule that higher numbers mean stiffer or more viscous (stiffness and viscosity will be explained by daily life examples).
(3) Identification/Localization task. You are required to explore a virtual material freely and find the boundaries of a hidden target which is defined by a difference in viscosity or elasticity. The task is to verbally report the shape of the target or draw a figure to show your judgment.

There will be no cost to you if you participate in this study. You will receive either $__________ ($10/hour, cash) or __________ psychology experiment credit (1 credit/hour) for your time and effort of participation. You will be paid by the duration of participation, even that you do not complete the study.

What happens if I say yes, but I change my mind later?

Participation in this study is completely voluntary. It is ok for you to say no. Even if you say yes now, you are free to say no later, and withdraw from the study at any time. Refusal to participate or withdrawal of your consent or discontinued participation in the study will NOT result in any penalty or loss of benefits or rights to which you might otherwise be entitled.

Is there any way being in this study could be bad for me?

There are no known risks from taking part in this study.

Will being in this study help me in any way?

We cannot promise any direct benefits to you or others from your taking part in this research. Your participation will help us to better understand how physicians perceive and interact with soft tissue using their sense of touch, together with visual information from medical imaging. The knowledge can foster the development of technologies for surgical training and assist physicians toward effective interaction.

What happens to the information collected for the research?

All information obtained in this study is strictly confidential. The results of this research study may be used in reports, presentations, and publications, but the researchers will not identify you. In order to maintain confidentiality of your records, your data and consent form will be kept separate. Your consent form will be stored in a locked cabinet in Dr. Bing Wu's office (Santa Catalina Hall 150E) and will not be disclosed to third parties. Computerized data files will be encrypted. Paper data files will be kept in locked locations accessible only to authorized researchers. Your name, address, contact information and other direct personal identifiers in your consent form will NOT be mentioned in any publication or dissemination of the research data and/or results. In this study, you will be assigned a case number and your identity on all research records will be indicated only by that number. We will NOT collect or save any information that may associate that number with your identity.

Efforts will be made to limit the use and disclosure of your personal information, including research study records, to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the University board that reviews research.

What else do I need to know?

This research is being funded by the National Institutes of Health.

Who can I talk to?

If you have questions, concerns, or complaints, please talk to

Bing Wu, Ph.D.
Dept. of Human Systems Engineering
Ira A. Fulton Schools of Engineering
Arizona State University
Santa Catalina Hall, Room 150E
7271 E. Sonoran Arroyo Mall
Mesa, AZ 85212
(412) 256-8168
(Bing.Wu@asu.edu)

This research has been reviewed and approved by the Social Behavioral IRB. You may talk to them at (480) 965-6788 or by email at research.integrity@asu.edu if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research participant.
- You want to get information or provide input about this research.

Your signature documents your permission to take part in this research.

Signature of participant	Date
Printed name of participant	
Signature of person obtaining consent	Date
Printed name of person obtaining consent	