

Temporal perception of visual-haptic events in multimodal telepresence system

Zhuanghua Shi, Heng Zou and Hermann J. Müller
*Allgemeine und Experimentelle Psychologie
Ludwig-Maximilians-Universität München
80802, München, Germany*

1. Introduction

Multidomal telepresence systems have been used in remote or hazardous environments, such as telesurgery, orbit or underwater teleoperation, etc. (Ballantyne, 2002; Draper et al., 1998; Hirzinger et al., 1993). In a typical teleoperation task, local action commands are transmitted to a remote teleoperator, which then executes the commands and sends back multimodal sensory information, such as visual, auditory and haptic signals. Due to the communication, data encoding, and control scheme, time delays of this information feedback are inevitable in multimodal telepresence systems. These delays can vary from dozens of milliseconds to seconds. For example, the feedback latency for an intercontinental teleoperation via the Internet is around 300 ms (Peer et al., 2008), while the latency can be up to 5–10 seconds for teleoperation tasks in space (Hirzinger et al., 1993).

The effects of time delay on performance have been investigated in many studies (Ferrell, 1966; Held, 1993; Jay et al., 2007; Kim et al., 2005; Sheridan & Ferrell, 1963). For example, examining the effect of visual-feedback delay on user's task completion time, Mackenzie and Ware (1993) found that performance was affected by delays exceeding 75 ms, with completion time thereafter increasing linearly with time delay (> 75 ms) and task difficulty. A recent study examining the effects of delayed visual feedback on telerobotic surgery (Kim et al., 2005) revealed a similar picture: completion time increased with time delay and dexterity was affected when the delay was longer than 250 ms; in addition, starting at a delay around 400 ms, the operators came to gradually adopt a move-and-wait strategy. Delays in haptic feedback have been found to have a similar influence on performance. For example, in a positioning task, delayed haptic feedback resulted in an increase in completion times and error rates (Ferrell, 1966). A recent study of the effects of continuous (as opposed to pulse) force feedback with delays revealed task performance to be very sensitive to the haptic-feedback delay: the error rate started to increase at delays of 25 ms and then rose more steeply than with visual-feedback delays (Jay et al., 2007).

Since time delay can vary for different types of information feedback, remote multimodal synchronous events, such as a visual-haptic collision, may be turned into local asynchronous incidents. Furthermore, crossmodal asynchrony may also arise from observer-internal

perceptual processes. Crossmodal synchrony and asynchrony perception has been shown to depend on many factors ranging from low-level neural encoding and to higher-level cognitive processes. For example, auditory (speech) stimuli are typically processed faster by the brain than visual stimuli, so that observers can detect audiovisual asynchrony more easily when the sound is physically presented before the image (Dixon & Spitz, 1980). Similarly, experiments requiring observers to make temporal-order judgments (TOJ) of visual-haptic events have revealed that the visual-haptic simultaneity window is located asymmetrically around the point of physical synchronicity (Spence et al., 2001). Although there is a large literature on multisensory simultaneity perception, most of these studies have used a purely psychophysical approach in which observers are passively presented with crossmodal events and then have to make temporal-order or asynchrony judgments. There has been considerably less research on the perception of multisensory simultaneity in interactive telepresence systems. However, multisensory perception in such systems may be rather different from the situation in pure psychophysical studies: since the human operator continuously interacts with the remote operator through the human-system interface, perception and action are closely linked. Previous studies have demonstrated that visuo-motor action is an influential factor for perception (Noë, 2005; Witney et al., 1999). Consequently, perception and action must be taken into account together for understanding multisensory temporal perception in telepresence systems.

In this chapter, we present an overview of multisensory temporal perception, with particular focus on crossmodal simultaneity, in telepresence systems, based on relevant work as well as our recent studies of visual-haptic simultaneity in the closed perception-action loop and scenarios with communication delays involving packet loss. Since feedback from the events of touching and manipulating objects is of vital importance for humans performing teleoperation tasks, our studies were mainly concerned with the temporal perception of visual-haptic 'collision' events. Based on experimental findings and relevant models of human sensorimotor control, we then go on to propose a process model for multisensory temporal perception. This model permits some general guidelines to be derived for the optimal (human-factors) design of multimodal telepresence systems.

2. Perception of simultaneity and related psychophysical methods

Crossmodal simultaneity is of vital importance in our daily life. It provides the most salient temporal cue to form coherent percepts of objects and events, and it helps to segregate targets from background (Moutoussis & Zeki, 1997). There are two important parameters characterizing the perception of simultaneity: the point of subjective simultaneity (PSS) and the temporal just noticeable difference (JND). The PSS is the time interval between the onsets of two sensory stimuli at which the two stimuli are reported to be most synchronous. The JND, on the other hand, indicates the resolution of the temporal discrimination. These parameters are typically determined using two different types of task (Spence et al., 2001; Vogels, 2004): temporal-order judgment (TOJ) and synchrony/asynchrony judgment (SAS) task, respectively. In TOJ tasks, the observer is asked to determine the temporal order of two stimuli. The PSS is then defined by the 50%-threshold and the JND is calculated as half the difference between the lower (25%) and upper (75%) bound of the threshold. In contrast, in SAS tasks, the PSS is defined by the maximum of the distribution of 'synchronous'

responses (plotted as a function of the stimulus onset asynchrony, SOA, between the two events) and the JND is estimated from the standard deviation of the distribution (see below).

To estimate the PSS and JND, psychometric functions, such as the logistic function, are often applied to model the binomial response data (Collett, 2002) in the TOJ task:

$$P(t) = \frac{1}{1 + \exp\left(\frac{\alpha - t}{\beta}\right)}, \quad (1)$$

where $P(t)$ is the probability function of the TOJ response (event A before event B) over the crossmodal SOA t ; α is a parameter indicating the location of the threshold at 50%; and β is a parameter indicating the slope of the psychometric function. The PSS and JND can be then estimated as follows:

$$P\hat{S}S = \hat{\alpha}, \quad (2)$$

$$J\hat{N}D = \ln 3 \cdot \hat{\beta}. \quad (3)$$

In the SAS task, on the other hand, a Gaussian function is often fitted using maximum likelihood estimation (Stone et al., 2001):

$$P(t) = a \cdot \exp\left(\frac{-(t - \mu)^2}{2\sigma^2}\right), \quad (4)$$

where a is the maximum proportion of 'synchronous' responses, μ is the location of the maximum 'synchronous' responses, and σ is the standard deviation of the Gaussian function. The PSS and JND are then estimated as follows:

$$P\hat{S}S = \hat{\mu}, \quad (5)$$

$$J\hat{N}D = \sqrt{\ln 4} \cdot \hat{\sigma}. \quad (6)$$

3. Crossmodal simultaneity in multimodal telepresence system

3.1 Influences of visuo-motor action on crossmodal simultaneity

Although crossmodal simultaneity has been intensively investigated (Dixon & Spitz, 1980; Fujisaki et al., 2004; Levitin et al., 2000; Spence et al., 2001; Stone et al., 2001; van Erp & Werkhoven, 2004; Vatakis & Spence, 2006), how crossmodal simultaneity is influenced by time delays and self-generated actions in multimodal human-system interfaces has thus far received only little attention. Recently, in an important study, Vogels (2004) examined the effects of time delay on visual-haptic asynchrony judgments, using a task that involved a visual object touching a virtual wall. Participants were asked to move a joystick such that a visual object approached and then collided with this wall. The SOA between the visual collision and the haptic collision (the latter was rendered by a constant 5.5 N force feedback)

were varied across trials. Moreover, a further time delay (of 0 to 80 ms) was randomly added between the action and the feedback to simulate the ubiquitous communication delays in telepresence system. In addition, no visual feedback of the initial (joystick-generated) object movement was provided for a random interval of 50-150 ms. The results showed that, regardless of action-feedback delay, participants could easily detect the visual-haptic asynchrony when the SOA exceeded 50 ms. In another experiment, participants received touch and visual stimulus events passively while holding the joystick steady. The sensitivity for visual-haptic asynchrony (indicated by the standard deviation of the response distribution) was found to vary between 15 and 38 ms, which was smaller compared to that in the active-movement condition. Vogels' study represents a pioneering study of the effects of time delay and visuo-motor movement on visual-haptic simultaneity perception. However, her finding of an advantage, in terms of sensitivity, of passive touch over active movement was likely confounded by the experimental setup (as will be elaborated below). Indeed, there is a large body of evidence showing that active perception yields more information and enhances performance compared to passive perception (Heller & Myers, 1983; Wexler & Klam, 2001). Critically however, in Vogels' study, the space of the hand movement was separated from that of the visual representation - which may have impeded the crossmodal movement prediction and required more attentional resources. As a result, it is likely that, in her setup, the uncertainty of the asynchrony judgments was actually higher in the active-movement condition compared to the passive condition.

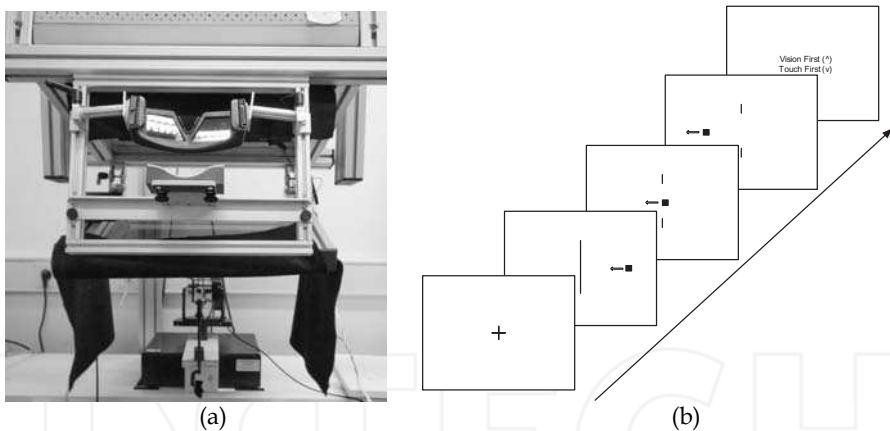


Fig. 1. (a) Collocated visual-haptic virtual environment used in Shi et al. (2008). Visual stimuli were presented on a CRT monitor and mirrored to the haptic space (below the mirror), where haptic stimuli were generated by means of a PHANToM device. (b) Schematic illustration of a trial sequence in Shi et al. (2008). A small box was moved from the left to the right, or vice versa, either by the participant's active finger movement or automatically by a program-generated (passive) movement. When the moving object approached the middle line (the virtual 'wall'), the line broke into two pieces, signalling a visual collision. Meanwhile, with a varying time delay relative to (i.e., preceding or following) the visual collision, there was a short haptic pulse force feedback in the direction opposite to the movement, signalling the haptic collision. The participant's task was to discriminate which event: the visual or the haptic collision, had happened first.

This potential confounding in Vogels' study by separated manual and visual spaces was addressed in a recent study by the current authors (Shi et al., 2008), who implemented a collocated visual-haptic virtual environment to (re-)examine the influence of the visuomotor closed loop on visual-haptic simultaneity (see Figure 1). The visuomotor closed loop has been demonstrated to be important for controlling fine movements (Keele, 1986). To disassociate the influence of visual-motion feedback and of sensorimotor (hand movement) processing on the crossmodal simultaneity judgments, Shi et al. manipulated these two factors in a full factorial design (manual movement: active vs. passive; visual-motion feedback: with vs. without feedback). One of the resulting conditions simulated a common telepresence scenario: active hand movement with visual-motion feedback; another condition, passive hand movement without visual-motion feedback (except for the visual breaching of the line and the haptic-collision feedback), is comparable to the situation in classical psychophysical studies.

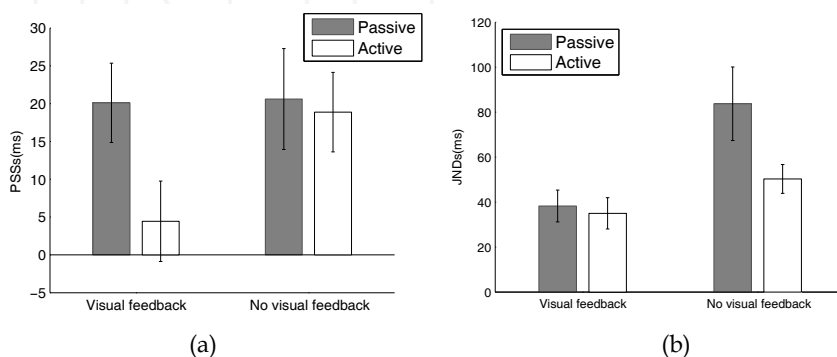


Fig. 2. (a) mean PSSs, and (b) mean JNDs for the four conditions in Shi et al. (2008).

Consistent with the results from classical visual-haptic studies, the PSS was shifted towards event sequences in which the visual stimulus physically occurred first, indicating that the haptic (collision) event was perceived earlier than a synchronous visual (collision) event. On average, the visual event had to occur 20 ms before the haptic event to be perceived as 'synchronous' - a value that is smaller than the typical PSS obtained in classical psychophysical studies (e.g., 50 ms in Spence et al., 2001). Interestingly, in the visuomotor closed-loop condition, the PSS was even smaller: only 4 ms (see Figure 2a). The small PSSs obtained in the interactive environment suggest that additional information, such as sensorimotor and visual-motion information, can help the central nervous system (CNS) to compensate for the difference in crossmodal latencies and improve the temporal-order judgment. More interestingly perhaps, the results of Shi et al. (2008), in contrast to those reported by Vogels (2004), showed that active movement does improve the sensitivity of asynchrony detection: the mean JND was reduced by 18 ms from passive to active visual-haptic collisions. In addition, the visual-motion feedback also greatly improved the sensitivity of crossmodal asynchrony judgments (Figure 2). Thus, the perceptual system can use the movement trajectory and active motor control to predict the forthcoming collision, thus improving the temporal discrimination. The performance improvement conferred by

the visuomotor closed loop is consistent with previous studies on spatial positioning, in which the motor command, in conjunction with internal models of both hand and visual feedback, has been demonstrated to be useful for anticipating the resulting load force and the position of the object (van Beers et al., 1999; Wolpert & Ghahramani, 2000; Wolpert et al., 1995). The discrepancy between the studies of Vogels and Shi et al. may come from the different spatial setups. In the latter study, the visual and haptic spaces were collocated in a single space and multisensory events were generated in a natural way, permitting sensorimotor and visual feedback to provide additional sources of information for discerning temporal order.

In summary, these results indicate that the temporal perception of visual-haptic events can be influenced by additional information such as sensorimotor and visual feedback. A similar influence of the perception-action closed loop has also been found in haptic-audio asynchrony detection, and action-to-visual-feedback-delay detection (Adelstein, Begault et al., 2003; Adelstein, Lee et al., 2003). Thus, for the design of telepresence systems, this body of work strongly suggests that the perception-action loop should be taken into account when making considerations as to human operator’s capacity for multimodal simultaneity perception.

3.2 Influences of packet loss on visual-haptic simultaneity

In multimodal telepresence system, crossmodal temporal perception is not only influenced by the perception-action loop, but also by inevitable communication delays and disturbances. Telepresence systems operating over large geographical distances are subject to packet loss and network communication delays, so that physically ‘synchronous’ events may be turned into ‘asynchronous’ incidents. Packet loss is a common issue in communication network using the DHCP service. Phenomenally, packet loss in video streams reduces image quality and interrupts video continuity. However, how packet loss influences the perception of visual-haptic simultaneity is, as yet, largely unknown. With regard to visual-packet loss, the current authors (Shi et al., 2009) recently examined this issue in a series of experiments. The task in these experiments was similar to the temporal-discrimination task used by Shi et al. (2008, see Figure 1), while adding frame-based packet loss to the visual feedback. The packet loss in the experiments was generated by a 2-state Gilbert-Elliot model (Elliot, 1963; Gilbert, 1960). This model can be wholly described by two transition probabilities between packet loss state (L) and packet no-loss state (N): $P_{n,l}$ and $P_{l,n}$ (See Figure 3). With two probabilities, two important features of the packet loss process, namely: the mean loss rate r_p and the mean burst length l_t , can be easily calculated (Eq. 7 and 8).

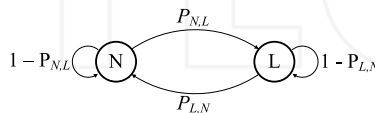


Fig. 3. Illustration of the 2-state Gilbert-Elliot model. ‘N’ and ‘L’ denote the states of ‘No packet loss’ and ‘Packet loss’, respectively.

$$r_p = \frac{P_{n,l}}{P_{n,l} + P_{l,n}}, \tag{7}$$

$$l_i = \frac{1}{P_{l,n}}. \quad (8)$$

In Experiment 1 of Shi et al. (2009), four different mean packet loss rates ($r_p = 0, 0.1, 0.2$ and 0.3), with a constant mean burst length of 33 ms, were examined. The 33-ms burst length was chosen as it is slightly above the critical flicker fusion (CFF) rate, thereby ensuring that, on average, the packet loss was perceivable to the observers. The results demonstrated that visual-haptic simultaneity was influenced by the visual-packet loss rate: with increasing loss rate, the PSS was linearly shifted towards visual delay and away from haptic delay, indicating that observers tended to judge a video stream with packet loss as a delayed video stream. On average, the visual delay increased by 25 ms for each 10%-increment in visual-packet loss. Furthermore, the JND was found to increase when the packet loss rate increased, indicating that the simultaneity judgments became more difficult with higher packet loss rates. In part, these shifts in visual-haptic temporal perception were due to the packet loss disturbing the perception of the visual collision (i.e., the visual collision was 'blacked-out'). More interestingly, however, both trends, in PSSs and JNDs, remained the same even when these parameters were re-estimated based on only those trials on which visual-haptic collision events remained intact (i.e., on which the packet loss did not occur at the visual collision; see Figure 4). Shi and colleagues concluded from these results that visual-haptic simultaneity is influenced by prior exposure to packet loss, more precisely: when the perceptual system adapts to visual feedback degraded by packet loss, the internal estimation of forthcoming crossmodal simultaneity is biased towards visual delay. A similar adaptation effect has also been found in a study concerned with the recalibration of audiovisual asynchrony (Fujisaki et al., 2004). In this study, after exposure to asynchronous audiovisual events for several minutes, observers displayed a shift in their subjective-simultaneity responses toward the particular asynchrony to which they adapted. Our study showed that such recalibration processes can take place even more rapidly: packet loss just prior to the collision already influenced the visual-haptic simultaneity judgment within that trial.

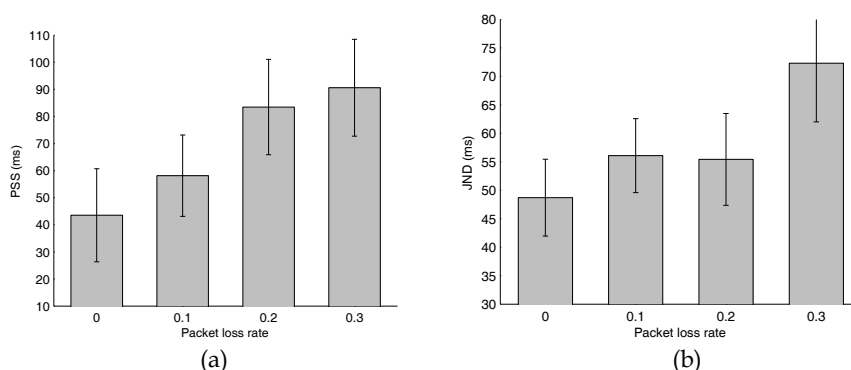


Fig. 4. (a) PSSs and (b) JNDs as a function of the visual-packet loss rate in Experiment 1 of Shi et al. (2009). The mean values were estimated based on only those trials on which the packet loss did not 'mask' the visual collision.

3.3 Influences of prior information on visual-haptic simultaneity

The study of Shi et al. (2009) suggests that the perceptual system may use past information, such as from visual feedback, to predict the forthcoming events. However, how rapidly past information can be used for this prediction is still an open question. From the perspective of system design, the update rate of the internal temporal percept is an important factor, since it describes the temporal resolution of the dynamic adjustment of crossmodal simultaneity. Thus, to further examine the update rate of prior information on crossmodal temporal perception, we conducted a new experiment on visual-haptic temporal discrimination with packet loss in the visual feedback. In this experiment, we kept the packet loss rate constant at 0.2 for the initial movement prior to the collision event. The experimental design and task were similar to Shi et al. (2009). On a typical trial, the observer moved his/her finger from the left to the right (or vice versa) and made a collision with the 'wall'. When the visual moving object (represented by a small dot, which was controlled by the observer's index finger) approached the wall, visual-packet loss was 'switched off' at certain distances before reaching the wall (i.e., from the respective distance onwards, there was no longer a chance of a packet loss occurring). Four different switch-off distances (i.e., distance from the position of the moving object to the wall at the moment packet loss was switched off) were examined in the experiment: 5, 30, 60 mm, and the whole movement trajectory (in the latter condition, there was no packet loss at any distance; see Figure 5).

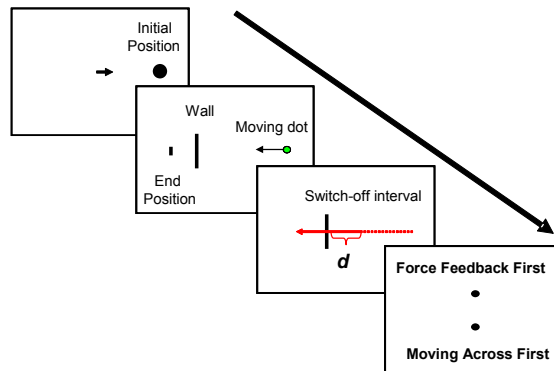


Fig. 5. Schematic illustration of a trial sequence. The movement trajectory is denoted by the long red arrow. The dashed line of the trajectory denotes visual feedback with packet loss, and the solid line visual feedback without packet loss. The packet loss switch-off distance is denoted by d .

The mean PSSs were 106.8, 87.3, and 80.1 ms for switch-off distance of 5, 30, and 60 mm, respectively; the mean PSS for the no-packet-loss condition was 79.5 ms. A repeated-measures ANOVA revealed the effect of switch-off distance to be significant, $F(3,30) = 4.68$, $p < 0.01$. A further contrast tested showed the PSS to decrease linearly with increasing switch-off distance, $F(1,10) = 5.82$, $p < 0.05$. The fact that, with increasing switch-off distance, the PSS approached the level achieved in the no-packet-loss condition suggests that 'no-packet-loss' information between the switch-off and the collision led to a gradual updating of the internal prediction of the forthcoming visual event. To estimate the internal update rate, we converted the switch-off distances into switch-off time intervals using observers' movement

speeds; these intervals were, on average, 14 ms, 85 ms, and 172 ms for 5-mm, 30-mm, and 60-mm distances, respectively. The relationship between PSS and switch-off time interval is shown in Figure 6. The 95% confidence intervals revealed that the PSS was significantly larger, relative to the (no-packet-loss) baseline, at a switch-off interval of 87 ms (30-mm distance), while the PSS at a switch-off interval of 172 ms (60-mm distance) was no different from the baseline. This means that a complete update with prior visual feedback took between 85 and 172 ms. In other words, the internal update rate was in-between 6 to 12 Hz.

In summary, the above results demonstrate that prior information does not immediately impact on the internal representation. The internal processing requires some time to update and adapt to changes of the external world. The time required by the internal processing is in the range of a hundred or so milliseconds, which may relate to the short-duration working memory involved in crossmodal temporal processing. In the design of telepresence systems, it would be advisable to use this update rate for the implementation of assistive functions.

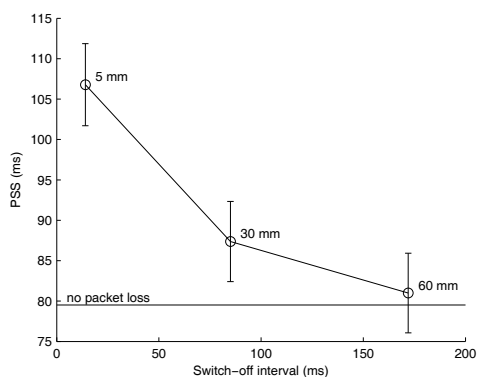


Fig. 6. PSS as a function of the switch-off time interval. The switch-off time intervals were estimated from the movement velocity. Error bars indicate 95% confidence intervals, which were estimated from 1000-sample bootstrapping.

4. Process model of crossmodal temporal perception

The studies discussed above showed that crossmodal simultaneity in an explorative environment is not only influenced by crossmodal temporal inconsistency, but also by many other sources of information, such as the visuomotor movement, the quality of the feedback signal, prior adaptation, etc. A recent study by Adelstein and colleagues (Adelstein, Lee et al., 2003) on head tracking latency also suggested that in virtual environments (with a head-mounted display), observers might use 'image slip' rather than the explicit time delay between input head motion and its displayed consequences to detect the asynchrony. Similarly, it has been found previously in audio-visual simultaneity judgments that, in relatively large environments, the brain may take sound velocity and distance information into account in the simultaneity perception of audio-visual events (Sugita & Suzuki, 2003). All available evidence converges on the view that the CNS may use additional information

to predict, or infer, the external forthcoming events. Predicting the next states has been shown to be useful for compensating for the slow speed of updating in the visuomotor control system (Wolpert, 1997; Wolpert et al., 1995). This capacity for prediction has been attributed to an internal model that is assumed to underlie the nervous system's remarkable ability to adapt to unknown or underdetermined changes in the environment (Tin & Poon, 2005).

Inspired by this idea of an internal model for the sensorimotor system, we suggest that dynamic multisensory temporal perception can be described in an analogous way. Figure 6 illustrates such an internal model of multisensory temporal perception. When there are only individual (unrelated) multisensory inputs, the CNS may use the resolution in the individual sensory channels to estimate the onset (or offset) time of events and from this determine crossmodal simultaneity. However, such passive forward estimation may suffer from differences in the neural latencies among different modalities. For example, an auditory event is usually perceived as 'earlier' than a synchronous visual event (Dixon & Spitz, 1980). When additional information is available, such as sensorimotor information, visual-motion trajectories, or visuo-proprioceptive discrepancies, the CNS may use this information to make a fine prediction and provide for crossmodal compensation in anticipating the forthcoming events. Using this model, one can easily explain the small PSS found in the visuo-motor closed-loop condition in Shi et al. (2008). The visuo-motor closed-loop helps the CNS to make a fine prediction of the forthcoming visual events, thus partially compensating for the delay inherent in the visual processing. The prediction mechanism can also be applied to account for the results of the packet loss experiments (Shi et al., 2009). The visual-feedback signal was disturbed by the packet loss, which made the video stream appear stagnant from time to time. Such prior 'delay' information is used by the CNS for predicting the timing of the forthcoming visual-haptic events. As a result, the PSS was shifted towards visual delay. Note, however, that the use of prior information by the CNS to adjust the crossmodal temporal representation is not immediate: the experiment outlined above (in section 3.3) suggests that the update rate of using prior information is only of the order of 6-12 Hz.

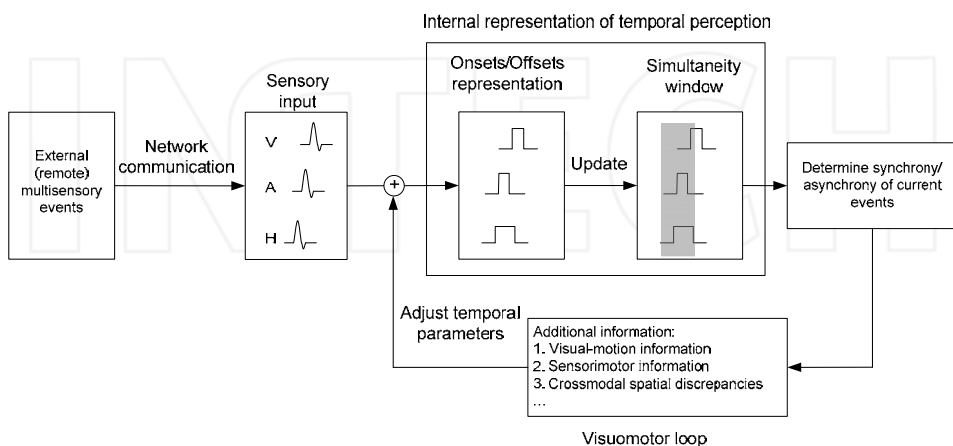


Fig. 6. Internal model of multisensory temporal perception.

5. Conclusion

In summary, we have provided an overview of studies concerned with visual-haptic simultaneity perception in multimodal telepresence system. It is clear that the perception of visual-haptic simultaneity is dynamic. In general, visual events are perceived as 'later' than physically synchronous haptic events. The visual-haptic simultaneity window (indicated by the PSS and JND parameters) may vary from dozens to hundreds of milliseconds. In interactive virtual environments such as telepresence systems, the crossmodal simultaneity window is influenced by other sources of information, such as sensorimotor feedback, packet loss in the feedback signal, and prior adaptation. Packet loss in visual feedback can bias visual-haptic judgments towards visual delay and such biases may influence even the perception of intact (visual-collision) events. In addition, prior information may also influence crossmodal simultaneity, however, this information is effectively taken into account only after one hundred milliseconds or so. Finally, based on the range of empirical evidence reviewed, we proposed that multisensory temporal perception involves an internal process model. The results, and the proposed framework model, can be used to derive guidelines for the design of the multimodal telepresence systems, concerning the crossmodal temporal perception of the human operator.

6. References

- Adelstein, B. D., Begault, D. R., Anderson, M. R., & Wenzel, E. M. (2003) Sensitivity to haptic-audio asynchrony, *Proceedings of the 5th international conference on Multimodal interfaces* (pp. 73-76). Vancouver, British Columbia, Canada.
- Adelstein, B. D., Lee, T. G., & Ellis, S. R. (2003) Head Tracking Latency in Virtual Environments: Psychophysics and a Model, *Human Factors and Ergonomics Society Annual Meeting Proceedings* (pp. 2083-2087): Human Factors and Ergonomics Society.
- Ballantyne, G. H. (2002) Robotic surgery, telerobotic surgery, telepresence, and telementoring. Review of early clinical results. *Surgical Endoscopy* 16(10), 1389-1402.
- Collett, D. (2002) *Modelling Binary Data* (2 ed.): Chapman & Hall/CRC.
- Dixon, N. F., & Spitz, L. (1980) The detection of auditory visual desynchrony, *Perception* (Vol. 9, pp. 719-721).
- Draper, J. V., Kaber, D. B., & Usher, J. M. (1998) Telepresence. *Human Factors* 40(3), 354-375.
- Elliot, E. O. (1963) A Model of the Switched Telephone Network for Data Communications, *Bell System Technical Journal* (Vol. 44, pp. 89-109).
- Ferrell, W. R. (1966) Delayed force feedback, *Human Factors* (Vol. 8, pp. 449-455).
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004) Recalibration of audiovisual simultaneity. *Nature Neuroscience* 7(7), 773-778.
- Gilbert, E. N. (1960) Capacity of a burst-noise channel, *Bell System Technical Journal* (Vol. 39, pp. 1253-1265).
- Held, R. (1993) Telepresence, time delay and adaptation. In S. R. Ellis, M. K. Kaiser & A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 232-246): Taylor and Francis.
- Heller, M. A., & Myers, D. S. (1983) Active and passive tactual recognition of form. *Journal of General Psychology* 108(2d Half), 225-229.

- Hirzinger, G., Brunner, B., Dietrich, J., & Heindl, J. (1993) Sensor-based space robotics-ROTEX and its telerobotic features, *IEEE Transactions on Robotics and Automation* (Vol. 9, pp. 649-663).
- Jay, C., Glencross, M., & Hubbard, R. (2007) Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment, *ACM Transactions on Computer-Human Interaction* (Vol. 14, Article 8, 1-31)
- Keele, S. W. (1986) Motor control. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance, Cognitive processes and performance* (Vol. II, pp. 30-60): Wiley.
- Kim, T., Zimmerman, P. M., Wade, M. J., & Weiss, C. A. (2005) The effect of delayed visual feedback on telerobotic surgery. *Surgical Endoscopy* 19(5), 683-686.
- Levitin, D. J., Maclean, K., Mathews, M., & Chu, L. (2000) The perception of cross-modal simultaneity, *International Journal of Computing and Anticipatory Systems* (pp. 323-329).
- Mackenzie, S. I., & Ware, C. (1993) Lag as a determinant of human performance in interactive systems, *CHI '93: Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems* (pp. 488-493). New York, NY, USA: ACM.
- Moutoussis, K., & Zeki, S. (1997) Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society B: Biological Sciences* 264(1387), 1407-1414.
- Noë, A. (2005) *Action in Perception*. Cambridge: MIT Press.
- Peer, A., Hirche, S., Weber, C., Krause, I., Buss, M., Miossec, S., et al. (2008) Intercontinental cooperative telemanipulation between German and Japan, *Proceedings of the IEEE/RSJ International Conferences on Intelligent Robots and Systems* (pp. 2715-2716).
- Sheridan, T. B., & Ferrell, W. R. (1963) Remote Manipulative Control with Transmission Delay, *IEEE Transactions on Human Factors in Electronics* (Vol. 4, pp. 25-29).
- Shi, Z., Hirche, S., Schneider, W., & Muller, H. J. (2008) Influence of visuomotor action on visual-haptic simultaneous perception: A psychophysical study, *2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (pp. 65-70).
- Shi, Z., Zou, H., Rank, M., Chen, L., Hirche, S., & Müller, H. J. (2009) Effects of packet loss and latency on temporal discrimination of visual-haptic events. *IEEE Transactions on Haptics*, in press.
- Spence, C., Shore, D. I., & Klein, R. M. (2001) Multisensory prior entry, *Journal of Experimental Psychology: General* (Vol. 130, pp. 799-832).
- Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001) When is now? Perception of simultaneity, *Proceedings of the Royal Society B: Biological Sciences* (Vol. 268, pp. 31-38).
- Sugita, Y., & Suzuki, Y. (2003) Audiovisual perception: Implicit estimation of sound-arrival time. *Nature* 421(6926), 911.
- Tin, C., & Poon, C. S. (2005) Internal models in sensorimotor integration: perspectives from adaptive control theory. *Journal of Neural Engineering* 2(3), S147-163.
- van Beers, R. J., Sittig, A. C., & Gon, J. J. (1999) Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology* 81(3), 1355-1364.
- van Erp, J. B. F., & Werkhoven, P. J. (2004) Vibro-tactile and visual asynchronies: Sensitivity and consistency, *Perception* (Vol. 33, pp. 103-111).

- Vatakis, A., & Spence, C. (2006) Evaluating the influence of frame rate on the temporal aspects of audiovisual speech perception, *Neuroscience Letters* (Vol. 405, pp. 132-136).
- Vogels, I. M. (2004) Detection of temporal delays in visual-haptic interfaces. *Human Factors* 46(1), 118-134.
- Wexler, M., & Klam, F. (2001) Movement prediction and movement production. *Journal of Experimental Psychology: Human Perception and Performance* 27(1), 48-64.
- Witney, A. G., Goodbody, S. J., & Wolpert, D. M. (1999) Predictive motor learning of temporal delays. *Journal of Neurophysiology* 82(5), 2039-2048.
- Wolpert, D. M. (1997) Computational approaches to motor control. *Trends in Cognitive Sciences* 1(6), 209-216.
- Wolpert, D. M., & Ghahramani, Z. (2000) Computational principles of movement neuroscience. *Nature Neuroscience* 3 Suppl, 1212-1217.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995) An internal model for sensorimotor integration. *Science* 269(5232), 1880-1882.

INTECH

INTECH



Advances in Haptics

Edited by Mehrdad Hosseini Zadeh

ISBN 978-953-307-093-3

Hard cover, 722 pages

Publisher InTech

Published online 01, April, 2010

Published in print edition April, 2010

Haptic interfaces are divided into two main categories: force feedback and tactile. Force feedback interfaces are used to explore and modify remote/virtual objects in three physical dimensions in applications including computer-aided design, computer-assisted surgery, and computer-aided assembly. Tactile interfaces deal with surface properties such as roughness, smoothness, and temperature. Haptic research is intrinsically multi-disciplinary, incorporating computer science/engineering, control, robotics, psychophysics, and human motor control. By extending the scope of research in haptics, advances can be achieved in existing applications such as computer-aided design (CAD), tele-surgery, rehabilitation, scientific visualization, robot-assisted surgery, authentication, and graphical user interfaces (GUI), to name a few. *Advances in Haptics* presents a number of recent contributions to the field of haptics. Authors from around the world present the results of their research on various issues in the field of haptics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Zhuanghua Shi, Heng Zou and Hermann J. Muller (2010). Temporal Perception of Visual-Haptic Events in Multimodal Telepresence System, *Advances in Haptics*, Mehrdad Hosseini Zadeh (Ed.), ISBN: 978-953-307-093-3, InTech, Available from: <http://www.intechopen.com/books/advances-in-haptics/temporal-perception-of-visual-haptic-events-in-multimodal-telepresence-system>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821