The Effect of Voluntary Movement on Audio-Haptic Temporal Order Judgment

A fundamental study for timing control in human-machine interaction

A. Nishi , M. Yokoyama , T. Ogata , T. Nozawa , Y. Miyake

Abstract— **In this study, we investigated the effect of active/passive motion on audio-haptic temporal order judgment. We used a psychophysical method to measure the differences in Point of Subjective Simultaneity (PSS) and Just Noticeable Difference (JND). PSS and JND were measured under "Voluntary" condition, "Involuntary" condition, and "No-movement" condition. Except Voluntary condition, when the haptic stimulus was presented before the auditory stimulus, the two stimuli were more likely to be perceived as simultaneous. PSS and JND of Voluntary condition were relatively smaller than those of Involuntary and No-movement conditions. Each PSS and JND under Involuntary condition was not significantly different from under No-movement condition. These results suggest that the proprioceptive information and the efference copy in voluntary motion accelerates haptic perception and improves the resolution of temporal order judgment for audio-haptic stimulus, while the proprioceptive information alone does not achieve significant effect on the judgment in involuntary motion.**

I. INTRODUCTION

People integrate cross-modal sensory information from the environment and act to the environment in real-time despite many delays including sensory integration, sensory motor coordination, and interaction with the environment. It is important to understand the characteristics of the interaction and the internal mechanism for it not only in cognitive psychology but also in the field of human-machine interface and virtual reality where the interaction between human and machine need to be dealed with [1], [2]. It is not clear yet how people integrate tempolarily cross-modal information during real-time interaction with the environment.

Several studies have reported that the integration of different types of information, originating from an event, requires the temporal simultaneity of the sensory inputs [3], [4]. Therefore, we focused the simultaneity of an event as one of the simplest sensory integration. Simultaneity Judgment (SJ) task [5] - [10] and Temporal Order Judgment (TOJ) task [11] - [19] are often used to investigate cross-modal simultaneity perception. These studies have shown that participants perceived pairs of visual and auditory stimuli and pairs of visual and haptic stimuli as simultaneous when the visual stimuli come earlier; additionally, participants perceived pairs of auditory and haptic stimuli as simultaneous when haptic stimuli preceded auditory stimuli [14], [15]. Such asymmetry in Point of Subjective Simultaneity (PSS) was reported to be affected by stimulus intensity and selective attention [8], [17] in addition to spatial location [18], [19]. Furthermore, in a study on audiovisual synchrony perception, participants exposed to a fixed audiovisual time lag for several minutes exhibited PSS shifts toward the lag and Just Noticeable Difference (JND) increase [6].

Previous studies, however, have primarily focused on simultaneity perception in the situation where the participants receive the stimuli passively. And simultaneity perception in the situation where the participants obtain the stimuli voluntarily such as "active touch" [20] or "active head movement" [21] were marginally investigated. In the field of virtual reality research [22], a TOJ task using visual and haptic stimuli revealed that the PSS decreased, and the JND was narrowed under conditions with active motor control. This finding suggests that perceived simultaneity is influenced by active motor control, although the nature and extent of this influence has yet to be resolved.

Motor-related factors that may affect synchrony perception include the efference copy of a motor command and the

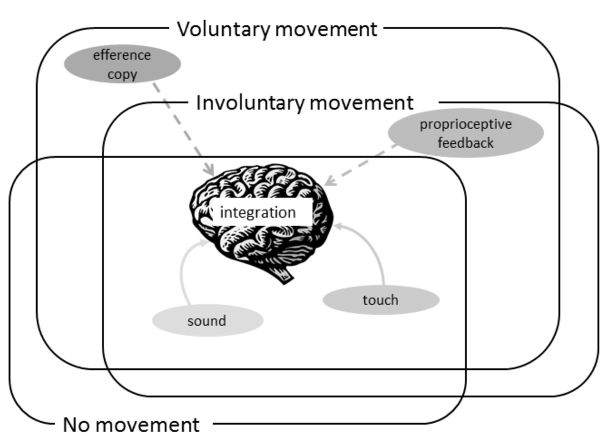

Fig. 1. Possible factors that affect synchrony perception.

A. Nishi, M. Yokoyama, Y. Miyake are with Department of Computational Inteligence and Systems Science, Tokyo Institute of Technology, 4259 Nagatsuda, Midori, Yokohama 226-8502, Japan (e-mail: nishia@myk.dis.titech.ac.jp, yokoyama121@gmail.com, miyake@dis.titech.ac.jp)

T. Ogata is with the University of Tokyo, Intelligent Modeling Laboratory, Yayoi 2-11-16, Bunkyo, Tokyo 113-0032, Japan and Department of Computational Inteligence and Systems Science, Tokyo Institute of Technology, 4259 Nagatsuda, Midori, Yokohama 226-8502, Japan (e-mail: ogata@iml.u-tokyo.ac.jp)

T. Nozawa is with Tohoku University, Smart Ageing International Research Center, Seiryo 4-1, Aoba, Sendai 980-8575, Japan (e-mail: taka.nozawa@gmail.com)

Fig. 2. Illustration of the haptic device. The participant's right index finger was strapped into the haptic device.

proprioceptive sensation which represents body orientation and movement. The efference copy signal is thought to influence activity in the sensory areas indirectly [23]. Libet and colleagues [9] have suggested that the efference copy signal for an active motor control occurs around 250 ms before a movement. The efference copy may, therefore, be used to predict the consequences of the movement [24]. Winter et al. claimed, however, that the efference copy does not affect simultaneity perception based on the result of simultaneity judgment of active/passive touch [25]. Information derived from proprioceptive sensation, on the other hand, has been suggested to be used to judge whether sensory stimulus provides feedback information of the body movement [10].

The effects of these motor-related factors on the perception of simultaneity should be analyzed in a differentiating manner (Fig. 1). This study examined whether the audio-haptic TOJ is influenced by voluntary finger movement or not. Specifically, we investigated the effect of efference copy and proprioceptive information on the audio-haptic TOJ. We made choice of audio-haptic TOJ, because little is known about the effects of active motor control on audio-haptic synchrony detection. We ever investigated JND on TOJ task with active movement [26]. The present study expanded our focus to JND and PSS.

II. METHOD

A. Participants

The one author and five paid participants (males; mean age of 23.3) attended the experiment. They were all right handedness, had an appropriate auditory threshold and normal touch, and exhibited no problems in moving their right index finger. Four participants had TOJ experiment. This experiment was approved by the ethics committees from Tokyo Institute of Technology.

B. Stimuli

The participants were presented with sinusoidal wave (2000 Hz, 50 dB, 15 ms) in both ears through earphones (MHP-EP5, JESTAX, Japan). The timing of the presentation was controlled to an error margin of 1 ms. The PHANToM® Desktop haptic device (SensAble Technologies, USA) was used to provide haptic stimuli (3N, 15 ms, rectangular pulse). The movement of the haptic device was also controlled within an error margin of 1 ms. These sensory stimulation systems were operated by computer programs installed on a PC workstation (HP xw4600/CT, Hewlett-Packard, USA), which were developed using the OpenHaptics software development toolkit (SensAble Technologies, USA) on the Microsoft® Visual C++ 2008 platform (Microsoft, USA). Tests were conducted in a sound-attenuated room free from noises that could possibly interfere with the auditory stimulation. The participants wore sound-insulating earmuffs over the earphones during the experiments. In addition, right index finger was held in a brace, to control the arm movement.

C. Procedure

The audio-haptic TOJ tasks were performed under three conditions: Voluntary condition, Involuntary condition, and No-movement condition.

Voluntary condition (Fig. 3(a)):

The participants were seated in front of the stimulation systems with the palmar side of participant's right index finger touching the haptic test device (Fig. 2). For each run of trials, a single tone was generated to announce that the recording was ready. The participants started to flex their right index finger voluntarily at their own timing. On a preliminary experiment, the temporal gap between the presentation of the single tone and the start of the arm movement was 1300 to 2800 ms. The start time of motion was defined as the time when the finger moves 10mm from the initial position. A haptic stimulus was presented at 500 ms from the start of the arm movement. Additionally, the high-pitched tone stimulus was presented in sync with the haptic stimulus. The participant was then given a two-alternative forced choice test to provide the temporal discrimination of the auditory and haptic stimulus pairs by answering which stimulus was perceived earlier. The preceding time of the auditory stimulus onset relative to that of the paired haptic stimulus was selected from the following stimulus onset asynchrony (SOA) values: -200, -90, -60, -30, 0, +30, +60, +90, and +200 ms (where the negative values indicate that the haptic stimulus preceded the auditory stimulus).

Involuntary conditions (Fig. 3(b)):

Similar to the Voluntary conditions, a single tone was generated to indicate the start of the recording. The haptic test device started to move the participant's right index finger 1300 to 2800 ms after the tone. This temporal gap between the presentation of the single tone and the start of the device finger movement was determined to reproduce the variance in the onset timing of voluntary movement in a preliminary experiment. A haptic stimulus was presented at 500 ms from the start of the arm movement. The speed

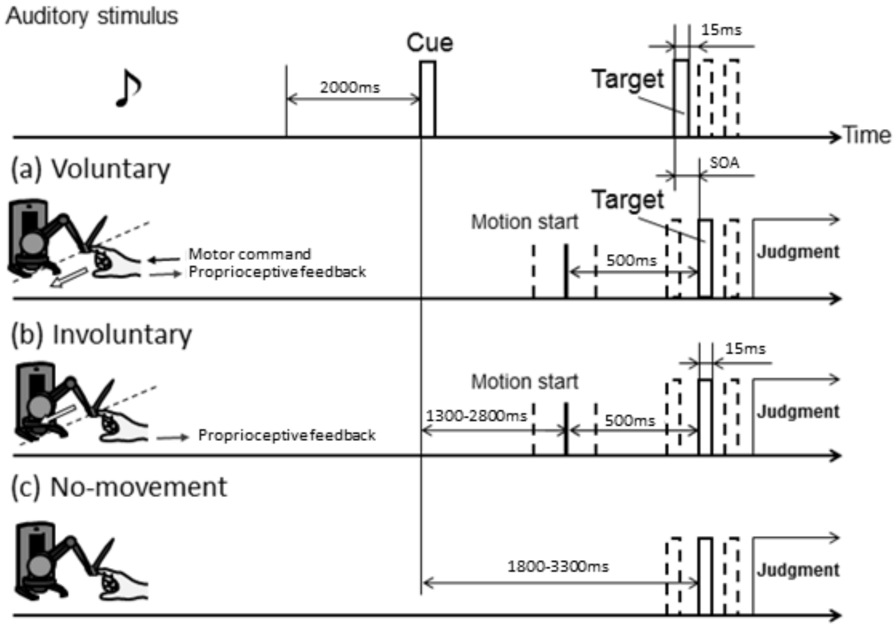

Fig. 3. Schematic illustration of the three conditions: Voluntary, Involuntary, and No-movement. Under Voluntary condition (a), the participants started to move their right arm voluntarily at their own timing after single tone. Under Involuntary condition (b), the haptic test device started to move the participant's right index finger 1300 to 2800 ms after the tone. Audio-haptic stimulus was presented at 500 ms from the start of the finger movement under Voluntary/Involuntary conditions. Under No-movement condition (c), the audio-haptic stimulus was presented after a 1800 to 3300 ms delay from the presentation of the tone. The SOA ranged from -200 ms to 200 ms.

of the finger movement was chosen for each experimental run from 76, 88, 100, 112, and 124 mm/s, whose occurrence rates were calculated from the distribution of data collected under the voluntary conditions in preliminary experiments. The procedure for evaluating the temporal discrimination, and the SOA values were the same as those used for the Voluntary condition.

No-movement condition (Fig. 3(c)):

A single tone indicating the start of the recording was generated, and a haptic stimulus was presented after a 1800 to 3300 ms (1300+500 to 2800+500 ms) delay from the presentation of the tone signal.

The experimental design was developed to make the following comparisons: 1) results of the No-movement condition and the Voluntary condition to reveal the effect of voluntary movement on the audio-haptic TOJ; 2) results of the Voluntary and Involuntary condition to clarify the effect of the efference copy; 3) results of the No-movement condition and the Involuntary condition to examine the effect of the proprioceptive sensation.

In this experiment, the participants completed five blocks (each block consisting of 45 trials, that is 5 trials for each SOA). The sequential order of the blocks was chosen randomly. There was a 2000 ms interval between trials. In order to learn to move participant's finger at a speed as close to 100 mm/s as possible, the participant underwent one block of practice sessions for the Voluntary condition before embarking on the formal test trials. In addition, they conducted practice runs of 5 trials just before each block under the Voluntary condition. During the practice sessions, only the haptic stimulus was presented, and no auditory stimulus was delivered for temporal judgment. In order to allow the participant's experience with the TOJ task, they were also given practice sessions consisting of one block each for all test conditions before starting the formal data collection trials. It took approximately five minutes for them to complete one block of trials. They were given several minutes of rest between blocks. They completed a total of 880 runs (including practice runs), and the entire procedure took roughly three hours. In order to eliminate confounding effects by visual stimuli, they were instructed to close their eyes during the experiments. Additionally, we asked them to pay constant attention to the haptic stimuli during the trials in order to control for the 'prior entry' effect [8], [17] on the test results under different testing conditions, which relatively facilitates the processing of an attended stimulus compared with an unattended stimulus.

D. Data analysis

The ratio of the answers indicating the earlier presentation of the auditory stimulus was calculated for each SOA. We conducted logistic regressions using a generalized linear model with the ratio data of each experiment [27]. The following equation was applied to the regression analysis:

$$y = \frac{1}{1 + e^{\frac{(\alpha - x)}{\beta}}} \tag{1}$$

where α represents the estimated PSS, x denotes SOA, and

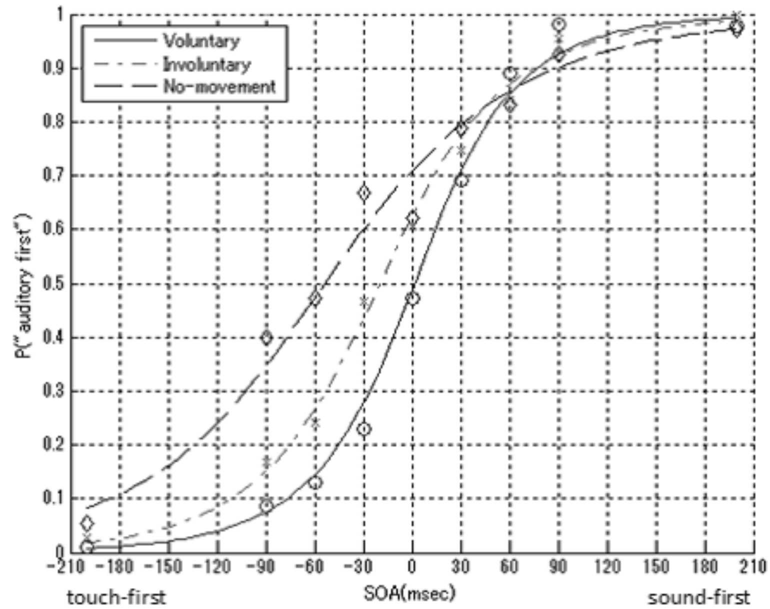

Fig. 4. Average psychometric functions across all experiments under the three conditions.

β is related to JND as shown in the following:

$$JND = \frac{x_{75} - x_{25}}{2} = \beta \log 3 \quad (2)$$

where x_p represents the SOA with p percent of 'auditory first' responses. MATLAB Statistics Toolbox® (MathWorks, USA) was used for the statistical regression calculation and graphic presentation of the results. Under the Voluntary condition, the data with 60 mm/s to 140 mm/s finger velocity were used for the analysis in accordance with the previous study [22].

III. RESURTS

As illustrated in Figure 4, psychometric curves were fitted to the distribution of the mean TOJ data for the Voluntary, Involuntary, and No-movement condition. We determined the JND and PSS values for each participant using the regression analysis (Eq.(1) and (2)), and further processed the data statistically to obtain the mean and standard error values for each condition. As shown in Figure 5, The JND under the Voluntary condition was smaller than other two conditions. In addition, Figure 6 shows the mean PSS on the Voluntary condition was almost close to zero, though other conditions resulted in negative values.

The result of JND and PSS were examined by Friedman test, and difference of JND among three conditions was significant ($p < 0.05$). In the PSS, significant difference was also observed ($p < 0.01$). The between-group differences each JND and PSS values were evaluated by Scheffe's paired comparison analysis. The results (Fig. 5) indicated that the JND value under Voluntary condition was smaller than the No-movement conditions ($p < 0.05$). Additionally, the results (Fig.6) also shows that the Voluntary condition produced smaller the absolute value of PSS compared with the No-movement conditions ($p < 0.05$). The presence/absence of voluntary movement probably marks the essential difference between the Voluntary and the No-movement condition.

IV. DISCUSSION

Our results on the JND corroborate the results of a preceding study by Shi et al., [22] between the Voluntary and No-movement conditions despite the difference of modality combination. Shi et al., examined the influence of visuomotor interaction on visual-haptic simultaneous perception, participants could make predictions by the combination of visual information, proprioceptive information, and motor efference copy. In our study, on the other hand, participants predicted the motion based only on proprioceptive information and efference copy. And as illustrated in Fig. 5, the JNDs under Involuntary and No-movement conditions were not significantly different. This suggests that the proprioceptive sensation alone does not affect prediction of the time of arrival of the stimuli, and therefore, not improve a resolution of the JND. Thus, in the present study, the predictive performance improved with both proprioceptive information and efference copy.

Our results on the PSS under No-movement condition was correspondence with the most studies which investigated audio-haptic TOJ without hand movement. Previous studies [14], [15] showed that in audio-haptic TOJ, the haptic stimuli had to be presented prior to auditory stimuli to reach PSS. On the other hand, as illustrated in Fig. 6, the PSS under the Voluntary condition shifted to the point which both stimuli were almost presented simultaneously. This result corresponded to the privious TOJ study with movement [22]. Shi et al., conducted TOJ under four conditions: active motor control with additional visual feedback, active motor control and no additional visual feedback, no movement with additional visual feedback, and no movement and no additional visual feedback. In the situation of no visual feedback, PSS under active motor control condition was not different from under no movement condition. On visual feedback conditions, however, PSS decreased under active motor control condition. From this comparison, it was suggested that the feedback information was not enough to change PSS in this experiment. The present study, however, resulted that feedback information from active motor control changed PSS. It indicated that the combination of modality affect PSS. And we also showed that action that decreases PSS sufficiently, was not involuntary movement with only proprioceptive information, but voluntary movement with efference copy and proprioceptive information.

The present study suggests that the efference copy is one essential factor which affects PSS and JND. However, based on the result of simultaneity judgment of active/passive touch, Winter et al. concluded that the efference copy did not affect simultaneity perception [25]. The disagreement between our view and Winter et al.'s view possibly stems from the difference of the experimental tasks. In their experiment, the SJ task was direct comparison of active and passive touch, always involving active motor control. Therefore, they mainly focused on PSS, and did not study the difference of JND between active and passive motor conditions. Besides, they reported that the difference between PSS and physical zero point was not significant, suggesting that the information processing speed of active touch differs little from that of passive touch. This result confirms our result on PSS.

In conclusion, our findings demonstrate that the predictive functions of proprioceptive information and motor command copies in voluntary motion accelerates haptic perception and improves the resolution of TOJ for audio-haptic stimulus, while the proprioceptive information alone does not have significant effect on the judgment in involuntary motion.

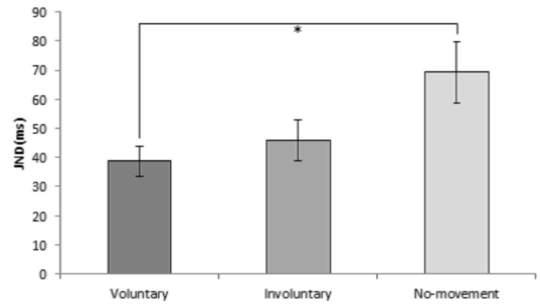

Fig. 5. The mean JNDs under the three conditions. Error bars represent the standard errors of means. The asterick indicates that the Voluntary condition produced smaller the value of JND compared with the No-movement condition ($p < 0.05$).

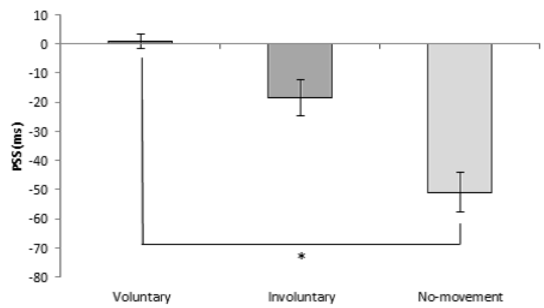

Fig. 6. The mean PSSs under the three conditions. Error bars represent the standard errors of means. The asterick indicates that the Voluntary condition produced smaller the absolute value of PSS compared with the No-movement condition ($p < 0.05$).

REFERENCES

[1] M. Claypool and J. Tanner, "The Effects of Jitter on the Peceptual Quality of Video," *Proceedings of the seventh ACM international conference on Multimedia (Part 2)* pp. 115-118, 1999.

[2] D. Wijesekera, J. Srivastava, A. Nerode, and M. Foresti, "Experimental Evaluation of Loss Perception in Continuous Media," *Multimedia Systems* vol. 7, no. 6, pp. 486-499, 1999.

[3] D.A. Slutsky, G.H. Recanzone, "Temporal and spatial dependency of the ventriloquism effect," *Neuroleport* vol. 12, pp. 7-10, 2001.

[4] M.A. Meredith, "On the neuronal basis for multisensory convergence: a brief overview," *Cognitive brain research* vol. 14, pp. 31-40, 2002.

[5] M. Zampini, S. Guest, D.I. Shore, C. Spence. "Audio-visual simultaneity judgments," *Perception & Psychophysics* vol. 67, pp. 531-544, 2005.

[6] W. Fujisaki, S. Shimojo, M. Kashino, S. Nishida, "Recalibration of audiovisual simultaneity," *Nature neuroscience* vol. 7, pp. 773-778, 2004.

[7] J.V. Stone, N.M. Hunkin, J. Porrill, R. Wood, V. Keeler, M. Beanland, M. Port, N.R. Porter, "When is now? Perception of simultaneity," *Proceedings of the Royal Society B: Biological Sciences* vol. 268, pp. 31-38, 2001.

[8] M. Zampini, D.I. Shore, C. Spence, "Audiovisual prior entry," *Neuroscience Letters* vol. 381, pp. 217-222, 2005.

[9] B. Libet, C.A. Gleasoon, E.W. Wright, D.K. Pearl, "Time of conscious intention to act in reaction to onset of cerebral activity(readinesspotential). the unconscious initiation of a freely voluntary act," *Brain* vol. 106, pp. 623-642, 1983.

[10] P. Guillaume, "Imitation in Children", University of Chicago Press, 1971.

[11] C. Parise, C. Spence, "Synesthetic congruency modulates the temporal ventriloquism effect," *Neuroscience Letters* vol. 442, pp. 257-261, 2008.

[12] E. Poliakoff, D.I. Shore, C. Lowe, C. Spence, "Visuotactile temporal order judgments in ageing," *Neuroscience Letters* vol. 396, pp.207-

211, 2006.
[13] M. Kanabus1, E. Szelag, E. Rojek, E. Pöppel, "Temporal order judgement for auditory and visual stimuli," *Acta Neurobiologiae Experimentalis* vol. 62, pp. 263-270, 2002.
[14] M. Zampini, T. Brown, D.I. Shore, A. Maravita, B. Röder, C. Spence, "Audiotactile temporal order judgments," *Acta Psychologica* vol. 118, pp. 277-291 , 2005.
[15] J.V. Hanson, J. Heron, D. Whitaker, "Recalibration of perceived time across sensory modalities," *Experimental Brain Research* vol. 185, pp. 347-352, 2008.
[16] S. Yamamoto, S. Kitazawa, "Reversal of subjective temporal order due to arm crossing," *Nature neuroscience* vol. 4, pp. 759-765, 2001.
[17] C. Spence, D.I. Shore, R.M. Klein, "Multisensory prior entry," *Journal of Experimental Psychology* vol. 130, pp. 799-832, 2001.
[18] Y. Sugita, Y. Suzuki, "Audiovisual perception: Implicit estimation of sound-arrival time," *Nature* vol. 421, pp. 911, 2003.
[19] N. Kitagawa, M. Zampin, C. Spence, "Audiotactile interactions in near and far space," *Experimental Brain Research* vol. 166, pp. 528-537, 2005.
[20] J.J. Gibson, "Observations on active touch," *Psychological Review* vol. 69, pp. 477-491, 1962.
[21] M.B. Cowan, L.R. Harris, "Temporal processing of active and passive head movement," *Experimental Brain Research* vol. 214, pp. 27-35, 2011.
[22] Z. Shi, S. Hirche, W.X. Schneider, H. Müller, "Influence of visuomotor action on visual-haptic simultaneous perception:A psychophysical study," *Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems* pp. 65-70, 2008.
[23] S.J. Blakmore, D.M. Wolpert, C.D. Frith, "Central cancellation of self-produced tickle sensation," *Nature Neuroscience* vol. 1, pp. 635-640, 1998.
[24] M.C. Morrone, J. Ross, D. Burr, "Saccadic eye movements cause compression of time as well as space," *Nature Neuroscience* vol. 8, pp. 950-954, 2005.
[25] R. Winter, V. Harrar, M. Gozdzik, LR. Harris, "The relative timing of active and passive touch," *Brain Research* vol. 1242, pp. 54-58, 2008.
[26] M. Yokoyama, S. Yoshida, H. Ora, Y. Miyake, "Effect of active movement in audio-haptic simultaneity perception," *Proceeding of Annual Conference of SICE System Integration(SI2009)* pp. 1923-1926, 2009(in japanese).
[27] D.J. Finney, "Probit analysis : a statistical treatment of the sigmoid response curve," Cambridge Univ. Press, 1952.