Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, Kobe International Conference Center, Kobe, Japan, December 15-17,

Visual motion information influences audiovisual simultaneity

Jinhwan Kwon, Kenichiro Ogawa, Taiki Ogata, JongHwan Kim and Yoshihiro Miyake

Abstract— **Motion information plays an important role in our daily life and temporal information is influential in dynamic environment. We therefore investigated how visual apparent motion information influences temporal perception. Participants performed audiovisual temporal order judgment (TOJ) task under apparent motion condition and non-apparent motion condition. We showed that visual apparent motion has resulted in faster visual processing and higher temporal resolution compared with non-apparent motion. Our findings suggest motion information accelerates temporal perception and improves temporal resolution.**

I. Introduction

We take a disciplined action by temporally integrating sensory information from environment. The environment is a dynamic environment in general and we perceive motion as information of dynamic environment. Therefore, the relation between motion information and temporal perception is important in dynamic environment. Also, the relation has a significant effect on man-machine interface. However, it is not clear whether motion information influences temporal perception. In this study, we focused on how motion information affects temporal perception.

In terms of temporal perception, temporal order judgment (TOJ) task is known as a psychophysical method for measuring the characteristics of multimodal integration on temporal dimension [1]. In particular, point of subjective simultaneity (PSS) and just noticeable difference (JND) are used as the quantities for measurement [2], [3]. The PSS represents the point of subjective simultaneity between the two senses and it means which sensory information is sooner perceived. The JND is used as an indicator that discriminates the temporal resolution in cross-modality.

There are little evidences considering motion information in TOJ task. Most studies have used single or static stimulation so far [1], [4]. Also, although some researchers have reported outcomes which are related to the intervals they do not take into account the presence or absence of motion [6], [7]. Even if the temporal characteristics of single stimulus are important, the temporal perception of motion is essential for cross-modal simultaneity when considering dynamic environment in daily life. Therefore, we need to investigate quantitatively temporal perception of motion.

The purpose of this study is to examine how motion information affects temporal perception. To clarify the influences of motion, we used visual apparent motion as motion stimulation. Motion could be separated into temporal and spatial elements and apparent motion is a phenomenon that the spatiotemporal characteristics are well presented. Apparent motion is an optical illusion that motion appears by the appropriate spatiotemporal intervals even despite two discrete stimuli [8, 9]. Also, TOJ task was used as measurement for temporal perception. So, participants performed TOJ task between apparent motion condition and non-apparent motion condition.

J. Kwon, K. Ogawa, Y. Miyake are with Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, 4259 Nagatsuda, Midori, Yokohama 226-8502, Japan (e-mail: kwon@myk.dis.titech.ac.jp, ogawa@dis.titech.ac.jp, miyake@dis.titech.ac.jp)

T. Ogata is with Research into Artifacts, Center for Engineering(RACE), The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba, 277-8568, Japan (e-mail: ogata@race.u-tokyo.ac.jp)

J. Kim is with Graduate School of Information Security, Korea Advanced Institute of Science and Technology, 291 Daehak-ro(373-1 Guseong-dong), Yuseong-gu, Daejwon 305-701, Korea (e-mail: zzoru@kaist.ac.kr)

II. Methods

A. Participants

Twelve participants (one author, 10 males, with a mean age of 24.3) took part in experiment. All participants had normal hearing and normal or corrected-to-normal visual acuity. Informed consent was obtained, and this experiment was approved by the ethics committee of the Tokyo Institute of Technology.

B. Apparatus and stimuli

All experiments of TOJ task were conducted in a dark and soundproof room. Visual stimulation was provided by a 27-inch LCD display (Samsung S27A950D) with a screen resolution of 1920×1080 pixels, and a refresh rate of 120 Hz. The display was operated by a PC workstation (Mac pro, 3.2GHz Quad-Core Intel Xeon, ATI Radeon HD 5770 graphic card, 1GB GDDR5 memory) and placed in front of the subjects. The head position was fixed by a chin-rest and the viewing distance was 100 cm. A white cross of 2 cm length as a fixation point was displayed in center of the screen. Visual stimuli consisted of one or two white disks 3.2 cm in diameter on a black background. Visual angle of non-apparent motion condition was 2.8° and apparent motion condition was 5.6°. Sound stimuli were presented with mono sounds (65dB, 1,000Hz), which were delivered via two speakers (MM-SPWD3BK, Sanwa supply). These visual and auditory stimuli were developed and operated by computer program (Matlab & Psychtoolbox-3).

C. Design and Procedure

We consisted of audiovisual TOJ tasks in which the spatial location and duration of audiovisual stimuli was identical.

The participant sat on a chair in front of the stimulation and a constant head position was maintained by chin-rest. The audio-visual TOJ tasks were performed under two types of visual stimuli: apparent motion and non-apparent motion. The stimuli of apparent motion condition indicate long-range apparent motion. The long-range apparent motion is a phenomenon detected in much longer spatial displacements as several degrees of visual angle and temporal intervals more than 100ms of inter-stimulus intervals [10], [11]. Also, the stimulation of non-apparent motion condition was single flash as a conventional stimulation in TOJ task.

Apparent motion (Fig. 1(a)):

Each trial started with the fixation cross for 1.2 s, and a dark blank screen was followed for 800 ms. Next, one white circle of first visual stimulus showed up for 30 ms and after 137 ms as Stimulus onset asynchrony (SOA), the second stimulus was presented for 30 ms [36]. To assess the temporal discrimination of the auditory and visual stimuli pairs, one brief sound (30ms) as auditory stimuli was presented with the second visual stimulus. The subjects were instructed to conduct TOJ task between the second visual frame and the brief sound. The onset time of the auditory stimulus and the paired visual stimulus was randomly selected from the following stimulus onset asynchrony (SOA) values: -120, -90, -60, -30, 0, +30, +60, +90, and +120 ms (where the negative values indicate that the auditory stimulus preceded the visual stimulus). Then the participant made a forced-choice judgment with respect to the temporal discrimination of the order between auditory and visual stimulation by answering 'which one was first' as question mark. The answers consisted of 'light first' which was chosen by pressing the Z key and 'sound first' which corresponded to the X key. As a way to answer, 'light first' was selected when the light was ahead of the sound, and vice versa with 'sound first'.

Non-apparent motion (Fig. 1(B)):

The procedure of non-apparent motion was the same as in TOJ task of apparent motion. The fixation cross (1.5s) and the dark blank screen (800 ms) were presented respectively. After 137 ms (same as SOA of apparent motion), however, the second frame in apparent motion was only shown in this session, so the first visual frame was not presented. Then, the procedure for evaluating the temporal discrimination between auditory and visual stimulation, and the SOA values were the same as those used for apparent motion.

The total trials of experiment were 270 trials (2 visual conditions ×9 audiovisual SOAs ×15 repeats) with counterbalanced order. Participants performed 27 trials (9 audiovisual SOAs ×3 repeats) as one block for each condition. The practice of each session was conducted and the total performance took about an hour.

Prior to the session, we examined whether the participants perceived motion between two flashes and also confirmed that the motion was perceived during the experiment after the session.

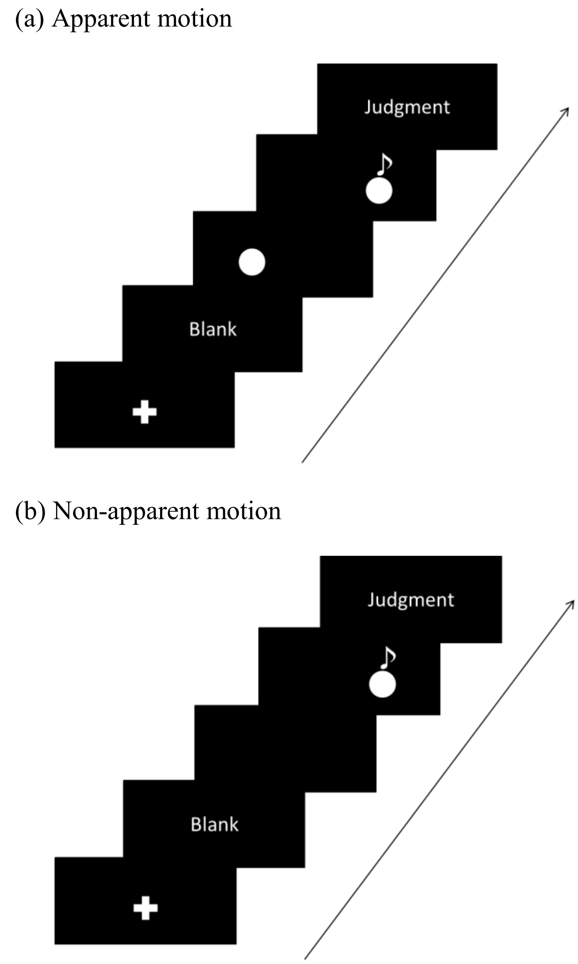


Fig. 1. Schematic illustration of the two conditions: Apparent motion (a) and Non-apparent motion (b).

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 [20]. The following equation was applied to the regression analysis:

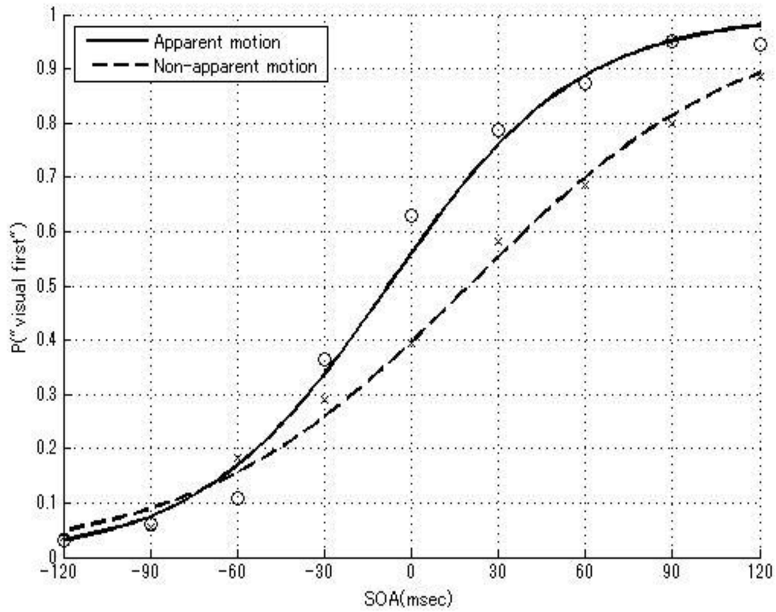
$$y = \frac{1}{1 + e^{\frac{(\alpha - x)}{\beta}}} \quad (1)$$

where α represents the estimated PSS, x denotes SOA, and β is related to *JND* as shown in the following:

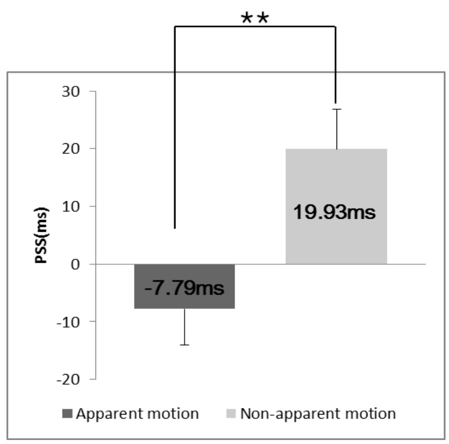
$$JND = \frac{X_{75} - X_{25}}{2} = \beta \log 3 \quad (2)$$

where X_p represents the SOA with p percent of 'auditory first' responses.

(a)



(b)



(c)

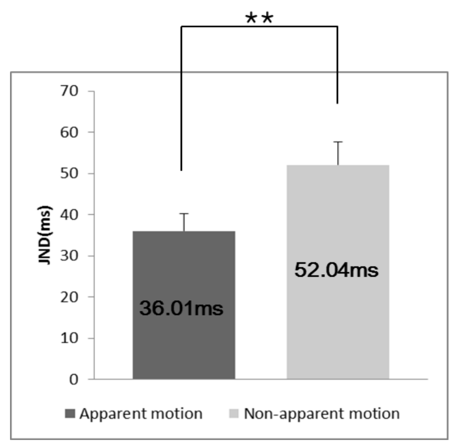


Fig. 2. (a) Psychometric curves fitted to the distribution of the mean TOJ data under the two conditions from all participants in experiment. (b) Mean PSSs under each condition. Error bars represent the standard error of the means. (c) Mean JNDs under each condition. Error bars represent the standard error of the means. $* : p < .05$, $** : p < .01$

Psychometric curves were fitted to the distribution of the mean TOJ data for each condition. We determined the JND and PSS values for each participant using regression analyses (Equation (1) and (2)) and processed the data statistically to obtain mean and standard error values.

III. Results

The results of one participant was excluded because the stimuli were not perceived as continuous motion. Fig. 2(a) presents the results of experiment. As shown in Fig. 2(b), the PSS under the condition of non-apparent motion was a positive value, 19.93 ms (SE = 6.91), but the PSS of apparent motion condition shifted to a negative value, -7.79 ms (SE = 6.31). The PSS of the negative value indicates that the audiovisual stimulus pairs were perceived as simultaneous when the auditory stimuli preceded the visual stimuli. Paired t-test on PSSs indicated significant deference between apparent motion and non-apparent motion condition (t(10) = -3.43, P < 0.01). Besides, the JND under apparent motion was smaller than the non-apparent motion (see Fig. 2(C)), and the JND sizes were 36.01 ms (SE = 4.19) and 52.04 ms (SE = 5.69) respectively. A significant difference in paired t-test was observed for the JNDs (t(10) = -3.83, P < 0.01).

IV. Discussion

This study confirmed that motion information influences temporal perception and we found new evidence that visual motion stimulation improves temporal perception. The results shows the PSS of non-apparent motion is similar to previous studies, which was usually shifted toward the visual-lead stimuli in TOJ task [1], but the PSSs of apparent motion was shifted to a negative value that is the opposite result of non-apparent motion. The negative value of PSS means that visual processing was faster. If physical transmission times through air and neural transmission times were considered it would be natural for auditory stimuli to be reached quickly in the brain [5], [1], but the apparent motion of the present study has resulted in faster visual processing. Also, JNDs was smaller in apparent motion condition. JND could also be considered as the degree of sensory integration representing temporal discrimination between the senses. We discuss our new findings to show that the visual apparent motion has specific characteristics in audiovisual temporal integration.

In this study, when the stimuli were apparent motion temporal processing on vision was faster. Previous studies have reported that PSS usually shifts toward a visual-lead stimulus, so simultaneity is maximally perceived if light comes slightly before sound [6], [12]–[14]. In addition, many studies have shown that hearing dominates vision and sound attracts or captures visual events in the time dimension [7], [15], [16]. Similarly, physical activity of the body is also susceptible to sounds other than sensory information in the study of sensorimotor synchronization [17], [18]. So, this finding suggests that motion stimulation accelerates the temporal perception.

With respect to temporal resolution, visual apparent motion resulted in the higher temporal resolution. Previous studies have reported that the different simultaneity or temporal resolution depends on the combination of sensory information and they can change according to a variety of factors such as stimulus intensity and attention [2], [4]. For example, the temporal ventriloquism effect has been shown that sound have an effect on JND and improve temporal resolution [7]. In this study, the visual apparent motion showed the higher temporal resolution. Therefore, we suggest that visual motion stimulation contributes to higher temporal discrimination.

It has not been reported that the quantitative evidence between visual motion and single flash on temporal perception. In previous researches, there is a robust illusory phenomenon called flash-lag effect examined the spatial alignment between motion and single flash [19]. When a moving target and single flash are aligned and participants were questioned about the same location between them. As a result, the moving target was perceived as more slightly shifted in the direction of motion relative to the flash. In other words, the flash was processed more slowly than the moving target. However, even if the moving object and the flash were presented with the same physical location, the moment for spatial alignment was determined after the flash. Therefore, there remains a need to examine the temporal perception between motion target and single flash respectively through a fixed spatial location. Thus, our results provide quantitatively new evidence that visual motion stimulation influences audiovisual temporal perception.

The relations between motion information and temporal perception have a significant effect on man-machine interface as well. Our findings will be important information to generate motion information by considering temporal perception on man-machine interface. Nevertheless, this study needs to be extended to topics clarifying predictive property of temporally constant interval and the effect of amount of stimulation because the timing of apparent motion was fixed and the amount of stimulation was different between apparent motion condition and non-apparent motion condition.

References

[1] J. Vroomen, M. Keetels, “Perception of intersensory synchrony: A tutorial review,” Attention, Perception, & Psychophysics 72 (4): 871-884, 2010.

[2] K. A. Schneider, D. Bavelier, “Components of visual prior entry. Cognitive Psychology,” 47, 333-366, 2003.

[3] Z. Shi, H. Zou, M. Rank, L. Chen, S. Hirche, H. J. Muller, “Effects of Packet Loss and Latency on the Temporal Discrimination of Visual-Haptic Events,” IEEE Transactions on haptics 3(1): 28-36, 2010.

[4] C. Spence, C. Parise, “Prior-entry: A review,” Consciousness and Cognition, 19, 364–379, 2010.

[5] M. D. Mauk, D. V. Buonomano, “The neural basis of temporal processing,” Annual Review of Neuroscience 27: 307–340, 2004.

[6] M. Zampini, D. I. Shore, C. Spence, “Audiovisual temporal order judgments,” Experimental Brain Research 152: 198-210, 2003.

[7] S. Morein-Zamir, S. Soto-Faraco, A. Kingstone, “Auditory capture of vision: Examining temporal ventriloquism,” Cognitive Brain Research 17: 154-163, 2003.

[8] V.S. Ramachandran, S. M. Anstis, “The perception of apparent motion,” Scientific American 254(6): 102-109, 1986.

[9] S. M. Anstis, D. M. Mackay, “The Perception of Apparent Movement,” Phil. Trans. R. Soc. Lond. B 290: 153-168, 1980.

[10] A. Larsen, J. E. Farrell, C. Bundesen, “Short- and Long-Range Processes in Visual Apparent Movement,” Psychol Res 45:11-18, 1983.

[11] O. J. Braddick, K. H. Ruddock, M. J. Morgan, D. Marr, “Low-Level and High-Level Processes in Apparent Motion,” Philosophical transactions Royal Society London B 290: 137-151, 1980.

[12] P. Jakowski, F. Jaroszyk, D. Hojan-Jezierska, “Temporal-order judgments and reaction time for stimuli of different modalities,” Psychol Res, 52:35 -38, 1990.

[13] M. Zampini, S. Guest, D. I. Shore, C. Spence, “Audio-visual simultaneity judgments,” Perception & Psychophysics, Vol. 67(3), 531-544, 2005.

[14] M. Kanabus, E. Szelg, E. Rojek, E. Pöppel, “Temporal order judgement for auditory and visual stimuli,” Acta Neurobiol. Exp., 62: 263-270, 2002.

[15] R. Fendrich, P. M. Corballis, “The temporal cross-capture of audition and vision”. Perception & Psychophysics 63(4): 719-725, 2001.

[16] P. Bertelson, G. Aschersleben, “Temporal ventriloquism: Crossmodal interaction on the time dimension. 1. Evidence from auditory–visual temporal order judgment,” International Journal of Psychophysiology 50: 147-155, 2003.

[17] B. H. Repp, A. Penel, “Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences,” Journal of Experimental Psychology: Human Perception and Performance 28 1085 – 1099, 2002.

[18] G. Aschersleben, P. Bertelson, “Temporal ventriloquism: crossmodal interaction on the time dimension 2. Evidence from sensorimotor synchronization,” International Journal of Psychophysiology 50, 157–163, 2003.

[19] R. Nijhawan, “Motion extrapolation in catching,” Nature 370: 256–257, 1994.

[20] D. J. Finney, Probit analysis a statistical treatment of the sigmoid response curve. Cambridge Univ. Press, 1952.