# Sounds can alter the perceived direction of a moving visual object

Wataru Teramoto	Department of Psychology, Graduate School of Arts and Letters, Tohoku University, Sendai, Miyagi, Japan, & Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan	
Souta Hidaka	Department of Psychology, Graduate School of Arts and Letters, Tohoku University, Sendai, Miyagi, Japan	$\bowtie$
Yoichi Sugita	Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan	$\bowtie$
Shuichi Sakamoto	Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan	$\bowtie$
Jiro Gyoba	Department of Psychology, Graduate School of Arts and Letters, Tohoku University, Sendai, Miyagi, Japan	$\bowtie$
Yukio lwaya	Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan	$\bowtie$
Yôiti Suzuki	Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan	$\bowtie$

Auditory temporal or semantic information often modulates visual motion events. However, the effects of auditory spatial information on visual motion perception were reported to be absent or of smaller size at perceptual level. This could be caused by a superiority of vision over hearing in reliability of motion information. Here, we manipulated the retinal eccentricity of visual motion and challenged the previous findings. Visual apparent motion stimuli were presented in conjunction with a sound delivered alternately from two horizontally or vertically aligned loudspeakers; the direction of visual apparent motion was always perpendicular to the direction in which the sound alternated. We found that the perceived direction of visual motion could be consistent with the direction in which the sound alternated or lay between this direction and that of actual visual motion. The deviation of the perceived direction of motion from the actual direction was more likely to occur at larger retinal eccentricities. These findings suggest that the auditory and visual modalities can mutually influence one another in motion processing so that the brain obtains the best estimates of external events.

Keywords: audiovisual interaction, motion perception, retinal eccentricity

Citation: Teramoto, W., Hidaka, S., Sugita, Y., Sakamoto, S., Gyoba, J., Iwaya, Y., & Suzuki, Y. (2012). Sounds can alter the perceived direction of a moving visual object. *Journal of Vision, 12*(3):11, 1–12, http://www.journalofvision.org/content/12/3/11, doi:10.1167/12.3.11.

# Introduction

Most objects and events in the external world generate concurrent inputs to several different sensory modalities. It has been assumed that each input is processed in the brain independently to some extent. However, we usually experience an integrated and unified percept of objects and events, suggesting that information from different sensory modalities is appropriately selected and bound together in the brain to represent a single object or event at several stages of perceptual processing. Indeed, recent studies on multisensory perception have revealed that different sensory modalities are closely related and mutually interplaying. In the domain of motion perception, several studies have suggested that visual information influences auditory motion perception (e.g., Kitagawa & Ichihara, 2002; Soto-Faraco, Lyons, Gazzaniga, Spence, & Kingstone, 2002; Soto-Faraco, Spence, & Kingstone, 2003; Soto-Faraco, Spence, & Kingstone, 2004). These

1

findings suggest that there are common neural substrates to motion perception between the visual and auditory modalities.

Auditory information has also been reported to affect visual motion perception (e.g., Freeman & Driver, 2008; Kim, Peters, & Shams, 2012; Maeda, Kanai, & Shimojo, 2004; Sekuler, Sekuler, & Lau, 1997; Watanabe & Shimojo, 2001). In contrast, several studies reported little or no influence of auditory information on visual motion information at perceptual level (Alais & Burr, 2004a; Meyer & Wuerger, 2001; Soto-Faraco et al., 2004; Wuerger, Hofbauer, & Meyer, 2003). A notable difference between these studies is a focus on the spatial characteristics or motion of auditory stimuli in the latter. For example, auditory effects on visual motion perception have been reported by studies manipulating the temporal relationship between a transient auditory stimulus and visual event (Freeman & Driver, 2008; Sekuler et al., 1997) or the semantic characteristics of an auditory stimulus (Maeda et al., 2004). Conversely, auditory effects on visual motion perception were reported to be absent or of smaller size at perceptual level by studies that manipulated the spatial aspects of auditory stimuli, including the interaural time difference (Alais & Burr, 2004a), interaural level differences by cross-fading a noise between two loudspeakers (Meyer & Wuerger, 2001; Wuerger et al., 2003), and direction of auditory apparent motion (Soto-Faraco et al., 2004).

It has recently been established that it is the precision/ reliability of different sensory inputs that determines their influence on the overall perceptual estimate (e.g., Ernst & Banks, 2002). This hypothesis could account for most of the observed influences of auditory information on visual motion perception (Alais & Burr, 2004b). Indeed, it has been demonstrated that changes in the location of a sound and auditory motion can induce or trigger visual motion perception of a static stimulus in far peripheral vision (Hidaka et al., 2009, 2011; Teramoto et al., 2010). Specifically, a blinking visual stimulus with a fixed location was perceived to be in lateral motion when its onset was synchronized to a sound with an alternating left-right source (Hidaka et al., 2009, 2011; Teramoto et al., 2010) or when it was presented, accompanied with a virtual stereo noise source smoothly shifting in a horizontal plane (Hidaka et al., 2011). This "Sound-Induced Visual Motion" (SIVM) phenomenon was more apparent when the blinking stimulus was located toward the periphery of the visual field.

The SIVM demonstrates that spatial shifts in sound alter the perception of a *static* visual stimulus. It remains unclear, however, whether auditory spatial information also affects the perception of a *moving* visual stimulus. In most studies using moving, rather than static, visual stimuli, the effect of auditory spatial information on visual motion perception has not been observed (Alais & Burr, 2004a; Meyer & Wuerger, 2001; Soto-Faraco et al., 2004). In the visual modality, motion signals are reported to be more salient than static signals (Dick, Ullman, & Sagi, 1987). Thus, auditory spatial information on visual motion perception might be limited to static visual stimuli. On the other hand, it could be notable that SIVM studies demonstrated the effect of auditory spatial information on visual stimuli especially in the peripheral visual field (Hidaka et al., 2009, 2011; Teramoto et al., 2010), whereas studies that did not find the auditory effect demonstrated visual stimuli primarily in the central visual field (Meyer & Wuerger, 2001; Soto-Faraco et al., 2002, 2003, 2004; Wuerger et al., 2003). While the spatial resolution of the visual system is excellent within the central visual field, it is dramatically reduced in regions located only a few degrees outside of this region. Recent studies suggest that the role of auditory spatial signals in cross-modal spatial localization depends on the spatial reliability of the visual signal (Alais & Burr, 2004b; Battaglia, Jacobs, & Aslin, 2003). Furthermore, Perrott, Costantino, and Cisneros (1993) reported that location discrimination performance at azimuth angles of 20° or larger was better for the auditory modality than for the visual modality. It can, therefore, be assumed that auditory spatial information can modulate visual motion perception when moving visual stimuli are presented in the peripheral visual field. If this would occur, it would be strong evidence that the auditory and visual systems influence each other in motion processing to a greater extent than what is previously believed.

The aim of the present study is to investigate the effect of auditory spatial information on the perception of moving visual stimuli. Our experiments demonstrate that the alternation of sound location can influence the perceived direction of visual motion within the peripheral visual field, even when visual stimuli are moving in a set direction (see demo in the Supplementary materials). In Experiment 1, we investigated whether the direction of visual apparent motion was modulated by a sound delivered alternately from two horizontally or vertically aligned loudspeakers (the direction of visual apparent motion was always perpendicular to the direction in which the sound alternated). We found clear effects of sounds on the perceived direction of visual motion. In Experiment 2, we investigated the effect of eye movements on the phenomenon observed in Experiment 1. In Experiment 3 as another control experiment, we investigated the discriminability of visual motion direction in peripheral vision in our experimental setup.

# **Experiment 1**

It is well known that localization cues based on binaural differences (binaural cues) as well as spectral changes in sound incident angle (spectral cues) can be used to localize sound sources and perceive sound movements (Asano, Suzuki, & Sone, 1990; Blauert, 1983). The relative importance of these cues depends on the location

or direction of motion of the sound. Sound localization and auditory motion perception in the horizontal plane are mediated mainly by binaural cues such as interaural time and level differences. Conversely, sound localization and auditory motion perception in the vertical plane are mediated by spectral change cues produced by the directional filtering associated with the pinnae, head, and shoulders. In Experiment 1-1, a sound was delivered alternately from the left and right loudspeakers. Therefore, binaural cues were likely to have been used to localize the sound and to perceive shifts in sound location. Experiment 1-2 was to investigate the contribution of spectral cues to the perceived direction of visual motion. We measured how the perceived direction of visual apparent motion is influenced by a sound delivered alternately from two loudspeakers aligned along the vertical meridian. Finding the perceived direction of motion to be influenced by sounds presented in the vertical as well as the horizontal plane may imply that the effect can be generated by both binaural and spectral cues and that it occurs throughout the entire two-dimensional audiovisual space.

## Methods

## Participants

There were ten participants in Experiment 1, including four of the authors (W.T., S.H., S.S., and Y.I.). All participants had normal or corrected-to-normal vision and normal hearing and, except for the authors, were naive to the purpose of the experiment. Informed consent was obtained from each participant before undergoing the procedures of the experiment, which were approved by the local ethics committee of Tohoku University.

### Apparatus

Visual stimuli were presented on a CRT monitor (Sony Trinitron GDM-F520, 21 inches,  $800 \times 600$  pixels) with a refresh rate of 60 Hz. The viewing distance was 76.4 cm. Red light-emitting diodes (LEDs) were used for fixation points at retinal eccentricities not possible on the monitor. Auditory stimuli were presented by two full-range loudspeakers (HOSIDEN, 0254-7N101, 30 mm  $\phi$ ) installed in small cylindrical plastic boxes (108 cm<sup>3</sup>) and positioned 76.4 cm to the left and right of the center of the CRT monitor (±45° in azimuth) in Experiment 1-1 and above and below the CRT monitor  $(\pm 45^{\circ} \text{ in elevation})$  in Experiment 1-2. Digital signals for the LEDs and auditory stimuli were converted to analog using audio interfaces (M-Audio ProFire Lightbridge and BEHRINGER ADA8000 8ch AD-DA converter). The experiment was controlled by a customized PC (Dell-XPS 710) running MATLAB (The Mathworks), with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and an open-source audio I/O library (Playrec, http://www.playrec.co.uk/).

A digital oscilloscope was used to confirm that the onset of visual and auditory stimuli was almost perfectly synchronized (to within  $\pm 10$  ms). The experiment was conducted in a dark anechoic room. Participants performed the experiment while seated and with head movements restrained by a chin rest.

## Stimuli

Experiment 1-1: The visual stimulus (target square) was a white square  $(1.0^{\circ} \times 1.0^{\circ}, 5.0 \text{ cd/m}^2)$  presented for 400 ms against a black background ( $0.2 \text{ cd/m}^2$ ). On each experimental trial, the square was presented six times, shifting in position by  $0.3^{\circ}$  up and down, with an interstimulus interval (ISI) of 100 ms so that apparent visual motion was induced along the vertical axis. The first square was presented to the upper side for half of the trials and to the lower side for the other half. The retinal eccentricity of the target square was varied by changing the position of a red fixation point such that the spatial relationship between the target square and the two loudspeakers was fixed, irrespective of retinal eccentricity of the target. The fixation point was positioned at a retinal eccentricity of  $2.5^{\circ}$ ,  $5^{\circ}$ ,  $10^{\circ}$ ,  $20^{\circ}$ , or  $40^{\circ}$  to the left of the center of the monitor so that the target square was presented in the dominant visual field (the right visual field was dominant for all participants). The fixation points at 20° and 40° of eccentricity were presented using LEDs; those at all other angles were presented on the monitor. The auditory stimulus was a 50-ms white noise burst with a cosine ramp of 5 ms at both onset and offset (sound pressure level: 74 dB; sampling frequency: 44.1 kHz; quantization: 16 bits).

Experiment 1-2: The stimuli were the same as in Experiment 1-1, with the following exceptions. The target square moved by  $0.3^{\circ}$  alternately to the left and right six times per trial. The fixation point was positioned at an angle of 4.88° to the left of the target square in order to test horizontal and vertical sound conditions at the same visual hemifield and at a retinal eccentricity along the vertical meridian of  $\pm 2.5^{\circ}$ ,  $\pm 5^{\circ}$ ,  $\pm 10^{\circ}$ ,  $\pm 20^{\circ}$ , or  $\pm 40^{\circ}$  (positive and negative values indicate the upper and lower visual fields, respectively). The upper and lower visual fields were tested in different sessions, the order of which was counterbalanced across participants.

#### Procedure

A fixation point was presented at the beginning of each trial, and participants pressed the start button upon seeing the fixation. After 500 ms of blank screen, the target square was presented six times with sound (sound condition) or without sound (no-sound condition). In the sound condition, sounds were presented alternately from the left and right loudspeakers in Experiment 1-1 and from the upper and lower loudspeakers in Experiment 1-2. The onset of each sound was synchronized with that of



Figure 1. (A) Schematic diagram of the experimental setup for Experiment 1 and (B) a presentation sequence of auditory and visual stimuli in a trial for the sound condition.

each target square (Figure 1). The first sound was delivered from one speaker for half of the trials and from the other speaker for the other half. These are randomly assigned in each trial for the sound condition. On each trial, a probe rod  $(3.0^{\circ} \times 0.1^{\circ}, 5.0 \text{ cd/m}^2)$  was presented 500 ms after the final disappearance of the target square. The initial orientation of the probe rod was randomized trial by trial. Participants had been instructed to report the direction in which they had perceived the target square to move by using buttons to rotate the probe rod  $(1^{\circ} \text{ with each button press})$ .

The experiment had two sessions consisting of 50 trials each. Each session contained two blocks of 25 trials: one block of sound condition trials and one block of no-sound condition trials. The order of sound and no-sound condition blocks was counterbalanced across the participants. Each of the five retinal eccentricities was presented 10 times in each condition; the order in which the different eccentricities were presented was randomized. Eye movements were not recorded.

## **Results and discussion**

Experiment 1-1: The average of the median deviation angles (difference between the directions of perceived and

actual motion of visual stimuli) across participants in each condition are shown in Figure 2 as a function of retinal eccentricity. A deviation angle of 0° corresponds to the direction in which the visual stimulus moved and a deviation angle of 90° to the direction in which the auditory stimulus alternated. A repeated-measures analysis of variance (ANOVA) with two within-participant factors, auditory condition (sound and no sound) and retinal eccentricity  $(2.5^{\circ}, 5^{\circ}, 10^{\circ}, 20^{\circ}, \text{ and } 40^{\circ})$ , revealed significant main effects of auditory condition ( $F_{1,9} = 11.40$ , p = 0.008) and eccentricity ( $F_{4,36} = 9.10, p < 0.001$ ). There was also a significant interaction between these factors ( $F_{4.36} = 7.25$ , p = 0.002) revealing the deviation angle as being significantly larger for the sound condition than for the no-sound condition at retinal eccentricities of 20° ( $F_{1,45} = 19.15$ , p < 0.001) and 40° ( $F_{1,45} = 23.43$ , p < 0.001) and also revealing that the deviation angle increased with increases in retinal eccentricity for the sound condition ( $F_{4,72} = 16.25, p < 0.001$ ) but not for the no-sound condition ( $F_{4,72} = 0.29$ , p = 0.883). These results suggest that auditory spatial information (alternation of sound source in the horizontal direction) can modulate the perceived direction of a moving visual stimulus, only when this is presented in the periphery of the visual field.

Experiment 1-2: The average of the median deviation angles across participants in each condition is shown in Figure 3 as a function of retinal eccentricity. A deviation



Figure 2. Results of Experiment 1-1. Mean deviation angles (difference between the directions of perceived and actual motion of visual stimuli) for the sound (black circles) and no-sound (white circles) conditions are shown as a function of retinal eccentricity (N = 10). A deviation angle of 0° corresponds to the direction in which the visual stimulus moved and a deviation angle of 90° to the direction in which the auditory stimulus alternated. Error bars denote the standard error of the mean. Asterisks denote significant differences between sound and no-sound conditions (\*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.005; ns, not significant).

angle of 90° corresponds to the direction in which the auditory stimulus alternated. The data were analyzed separately for each visual field (lower and upper) using ANOVA with two within-participant factors: auditory condition (sound and no sound) and retinal eccentricity  $(2.5^\circ, 5^\circ, 10^\circ, 20^\circ, \text{ and } 40^\circ)$ . The ANOVA for both visual fields revealed significant main effects of auditory condition (lower visual field:  $F_{1,9} = 21.75$ , p = 0.001; upper visual field:  $F_{1,9} = 9.28$ , p = 0.014) and eccentricity (lower visual field:  $F_{4,36} = 13.31$ , p < 0.001; upper visual field:  $F_{4.36} = 7.51$ , p < 0.001). There were also significant interactions between these factors (lower visual field:  $F_{4,36} = 14.83, p < 0.001$ ; upper visual field:  $F_{4,36} = 4.26$ , p = 0.006). The interactions revealed that the deviation angle was significantly larger for the sound condition than for the no-sound condition at retinal eccentricities of  $20^{\circ}$  ( $F_{1,45} = 12.29$ , p = 0.001) and  $40^{\circ}$  ( $F_{1,45} = 71.98$ , p < 0.001) for the lower visual field and  $10^{\circ}$  ( $F_{1.45} = 5.56$ , p = 0.023) and 40° ( $F_{1.45} = 22.18$ , p < 0.001) for the upper visual field. The interactions also revealed that the deviation angle increased with increases in retinal eccentricity for the sound condition (lower visual field:  $F_{4,72}$  = 26.41, p < 0.001; upper visual field:  $F_{4,72} = 11.64$ , p < 0.0010.001) but not for the no-sound condition (lower visual field:  $F_{4.72} = 1.08$ , p = 0.374; upper visual field:  $F_{4.72} =$ 1.10, p = 0.365). These results suggest that auditory spatial information (alternation of sound source in the horizontal direction) can modulate the perceived direction of a moving visual stimulus, only when this is presented in the periphery of the visual field.

To compare the data among the tested visual fields, an ANOVA was performed with three within-participant factors: visual field (right, lower, and upper visual fields), auditory condition (sound and no sound), and retinal eccentricity (2.5°, 5°, 10°, 20°, and 40°). The ANOVA revealed the significant main effects of auditory condition  $(F_{1,9} = 13.61, p = 0.005)$  and eccentricity  $(F_{4,36} = 18.41, p = 10.005)$ p < 0.001). Significant interactions were also found between auditory condition and eccentricity ( $F_{4,36}$  = 11.07, p < 0.001) and among the three factors ( $F_{8,72}$  = 2.48, p = 0.020). The two-way interaction revealed simple interactions between auditory condition and eccentricity  $(F_{4,108} > 3.82, ps < 0.006)$ ; the deviation angle was significantly larger for the sound condition than for the nosound condition at retinal eccentricities of 20° ( $F_{1,135} >$ 8.91, p < 0.003) and 40° ( $F_{1,135} = 27.43$ , ps < 0.001) for the right and lower visual fields and at 40° ( $F_{1,135}$  = 21.39, p < 0.001) for the upper visual field. We have thus found that a sound delivered across alternating locations can modulate the direction of perceived motion of a visual stimulus in either the vertical (Experiment 2) or horizontal (Experiment 1) plane. These results reveal that an effect of alternating sound location



Figure 3. Results of Experiment 1-2. Mean deviation angles (difference between the directions of perceived and actual motion of visual stimuli) for the sound (black circles) and no-sound (white circles) conditions are shown as a function of retinal eccentricity (N = 10). A deviation angle of 0° corresponds to the direction in which the visual stimulus moved and a deviation angle of 90° to the direction in which the auditory stimulus alternated. Negative and positive values in the horizontal axis indicate the retinal eccentricities for the lower and upper visual fields, respectively. Error bars denote the standard error of the mean. Asterisks denote significant differences between sound and no-sound conditions (\*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.005; ns, not significant).

on perceived direction of visual motion can be generated by both binaural and spectral cues, and it can thus be concluded that this phenomenon occurs throughout the entire two-dimensional audiovisual space.

# **Experiment 2**

It is possible that our findings of Experiment 1 could be accounted for by a confounding effect of eye movements induced by alternating sound sources. To test this possibility, we conducted a control experiment in which eye movements were recorded during the synchronized presentation of auditory and visual stimuli, the latter at either  $10^{\circ}$  or  $40^{\circ}$  of retinal eccentricity.

## Methods

The participants of Experiment 2 were the same ten people who participated in Experiment 1. Eye position was recorded from the left eye at a sampling rate of 60 Hz with an EMR-9 eye-tracking system (NAC Image Technology). Eyeglass shape prevented eye movements from being accurately recorded in three participants, and their data were excluded from the analysis. Experiment 2 consisted of three sessions, in each of which a different visual field (right, lower, and upper) was tested. The stimulus configuration for the right visual field session was the same as that used in Experiment 1-1: The visual stimulus moved alternately up and down in conjunction with a sound delivered alternately from the left and right loudspeakers. The upper and lower visual field sessions were consistent with those of Experiment 1-2: The visual stimulus moved alternately left and right in conjunction with a sound delivered alternately from the upper and lower loudspeakers. The visual stimulus was presented at either 10° or 40° of retinal eccentricity.

## **Results and discussion**

The averages of the median deviation angles across participants in each condition are shown in Figure 4 as a function of retinal eccentricity (Figure 4A for the right visual field and Figure 4B for the upper and lower visual fields). The trials during which the eye position deviated more than 1° (black symbols) are shown separately from those during which it was within 1° (white symbols). The mean ( $\pm SEM$ ) percentage of trials in which eye position deviated by more than 1° of visual angle in either the horizontal or vertical directions from the center of a fixation cross during stimulus presentation was 18.6  $\pm$  3.4% and 14.3  $\pm$  4.8% for retinal eccentricities of 10° and 40°, respectively, for the right visual field session. The percentage was 8.6  $\pm$  4.2% (10°) and 11.4  $\pm$  2.8% (40°)

6



Figure 4. Results of Experiment 2. Mean deviation angle (difference between the directions of perceived and actual motion of visual stimuli) for the sound condition is shown as a function of retinal eccentricity, (A) right visual field and (B) upper and lower visual fields. Trials during which eye position deviated by more than 1° from the center of a fixation point (black symbols) and those that did not deviate by more than 1° (white symbols) are separately shown (N = 7). A deviation angle of 0° corresponds to the direction in which the visual stimulus moved and a deviation angle of 90° to the direction in the vertical direction and an alternating left–right sound were synchronously presented. Asterisks denote significant differences between sound and no-sound conditions (\*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.005; ns, not significant).

for the lower visual field session and  $11.4 \pm 2.4\%$  (10°) and 20.7 ± 5.9% (40°) for the upper visual field session. An ANOVA for angle deviation was performed with three within-participant factors: eye deviation (more than 1° deviation trials and within 1° deviation trials), visual field (right, lower, and upper visual fields), and retinal eccentricity (10° and 40°). Only a significant main effect of retinal eccentricity was observed ( $F_{1,6} = 15.13$ , p =0.008). Further, significant correlations were not observed between the mean eye deviation and direction deviation for each trial (-0.157 < r < 0.076, p > 0.872) and between



7



Figure 5. Scatter plot of direction angle as a function of (A) mean eye deviation and (B) max eye deviation in each trial in Experiment 2. Regarding the vertical axis, a deviation angle of 0° corresponds to the direction in which the visual stimulus moved and a deviation angle of 90° to the direction in which the auditory stimulus alternated.

the maximum eye deviation and direction deviation for each trial (-0.146 < r < 0.134, p > 0.882; see Figure 5). The findings of Experiment 1 therefore could not be accounted for only by eye movements.

## **Experiment 3**

Previous work suggests that the positional information originating from the auditory modality would alter the perceived spatial position of visual stimuli, resulting in illusory visual motion perception (Alais & Burr, 2004b; Radeau & Bertelson, 1987). To test this possibility, we investigated whether the same effect could be observed with a longer ISI of visual stimuli (Kahneman & Wolman, 1970), in which visual apparent motion would be more difficult to perceive. If the positional capture of visual stimuli by the sounds are crucial, the same effect as that observed in Experiments 1 and 2 should be observed for the 1000-ms ISI condition. Additionally, to further single out the perceptual nature of the current phenomenon, signal detection theory (Macmillan & Creelman, 2004) was applied to the current experimental paradigm and was used to separately analyze the effects of the sounds on perceptual sensitivity to visual motion direction and criterion of visual motion discrimination.

## **Methods**

There were eight participants, including two of the authors (W.T. and S.H.). The visual stimulus (target square) was a white square  $(1.0^{\circ} \times 1.0^{\circ}, 5.0 \text{ cd/m}^2)$  presented for 400 ms against a black background (0.2 cd/m<sup>2</sup>). For each experimental trial, a fixation point

was presented at the beginning, and 500 ms after the participants pressed the start button, the square was presented six times, shifting in position by  $0.3^{\circ}$  either vertically or horizontally, with an ISI of either 100 ms or 1000 ms. Two retinal eccentricities (10° and 40°) were tested. For horizontal apparent motion trials, the first square was presented to the upper side for half of the trials and to the lower side for the other half. For vertical apparent motion trials, the first square was presented to the left for half of the trials and to the right for the other half. Further, a sound was delivered alternately from the left and right positions for half of the trials, while no sound was presented for the other half. The onset of each sound was synchronized with the presentation of each target square. The participants were asked to discriminate between horizontal and vertical apparent motion of visual stimuli for 320 trials: Visual motion (2; horizontal/vertical direction)  $\times$  Sound presentation (2; with sound/without sound)  $\times$  ISI of visual stimuli (2: 100 ms/1000 ms)  $\times$ Retinal eccentricity (2;  $10^{\circ}/40^{\circ}$ ) × Repetition (20). Each retinal eccentricity was tested in a different block. The order of the block was counterbalanced between the participants. Within each block, the order of the conditions was randomized. After completing all the trials, the participants were asked whether they perceived motion in the auditory stimuli. Almost all the participants reported that they perceived continuous or broken motion (a sound was heard to move from one side to the other, but the movement was rough or discontinuous) in the auditory stimuli for the 100-ms ISI condition and successive sounds for the 1000-ms ISI condition.

## **Results and discussion**

First, the proportions of correct responses were calculated for each condition. Then, as an index of sensitivity



Figure 6. Results of Experiment 3. (A) D-primes for motion direction discrimination, (B) response criterion ( $\beta$ ), (C) proportion correct of motion (or displacement) direction discrimination for horizontal visual motion, and (D) proportion correct of motion (or displacement) direction discrimination for vertical visual motion are shown. Error bars denote the standard error of the mean. Asterisks denote significant differences between sound and no-sound conditions (\*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.005; ns, not significant).

to the direction of visual motion, d-primes were computed on the basis of the signal detection theory (Macmillan & Creelman, 2004). This was because changes in the value of *d*-prime could be separated from those of the criterion, namely, response or decisional biases. The responses of vertical motion were regarded as "hits" in the vertical motion trials and as "false alarms" in the horizontal motion trials. The proportions of hits and false alarms with 0%or 100% values were corrected as 1/n or (n - 1)/n, respectively, where n was the total number of presentations (Anscombe, 1956; Sorkin, 1999). If the alternating left-right sound sources alter the perceived direction of visual motion, the values of *d*-prime should degrade for the sound condition, as compared with the no-sound condition.

The calculated values of *d*-prime are shown in Figure 6A. A three-way repeated-measures ANOVA was conducted with retinal eccentricity (2)  $\times$  sound presentation (2)  $\times$  ISI (2). This analysis revealed significant main effects of retinal eccentricity ( $F_{1,7} = 41.07, p < 0.001$ ) and ISI  $(F_{1,7} = 12.55, p < 0.001)$ , a one-way interaction effect between sound presentation and ISI ( $F_{1,7} = 7.52$ , p = 0.029), and a two-way interaction effect among the factors ( $F_{1,7} = 6.57$ , p = 0.049). The subsequent analysis revealed a significant simple-simple main effect of sound

presentation only for the 100-ms ISI condition at 40° of retinal eccentricity ( $F_{1,28} = 9.40, p = 0.005, \eta^2 = 0.010$ ). The value of *d*-prime for the 100-ms ISI condition decreased for the sound condition (d-prime = 1.63), as compared with the no-sound condition (*d*-prime = 2.01). According to the proportion data, this could be interpreted as being caused mainly by the vertical motion/ displacement being perceived as horizontal motion/ displacement (see Figures 6C and 6D). This result showed that the alternating left-right sounds altered the perceived direction of visual motion. In contrast, for the 1000-ms ISI condition at  $40^{\circ}$  of retinal eccentricity, the value of *d*-prime for the sound condition (*d*-prime = 1.12) was increased, as compared to that for the no-sound condition (d-prime =0.82), although it did not reach statistical significance  $(F_{1,7} = 4.13, p = 0.052, \eta^2 = 0.005)$ . That is, the sounds improved the sensitivity to visual motion/displacement discrimination for the 1000-ms ISI condition. This result might have occurred because the spatiotemporal uncertainty of visual stimuli at 40° of retinal eccentricity was reduced by sound presentation, due to a general alerting effect (e.g., Posner & Boies, 1971) or temporal capture effect (Aschersleben & Bertelson, 2003; Bertelson & Aschersleben, 2003; Fendrich & Corballis, 2001; Scheier, Nijhawan, & Shimojo, 1999). Thus, it can be concluded

that the current findings regarding the changes in the perception of visual motion direction by alternating left– right sounds cannot be accounted for by the positional capture of visual stimuli by sounds.

With regard to the analysis of the criterion of visual motion/displacement discrimination (see Figure 6B), the criterion was shifted to the horizontal motion/displacement responses for all the sound conditions ( $F_{1,28} > 9.58$ , ps < 0.005) except when the visual stimuli were presented with 100-ms ISI at 10° of retinal eccentricity ( $F_{1,28} = 0.006$ , p = 0.941). Thus, the present data showed that the alternating left–right sounds also altered the criterion of visual motion discrimination except for the 100-ms ISI condition at 10° of retinal eccentricity.

## **General discussion**

In the present study, we investigated whether and how auditory spatial information modulates the perceived direction of visual apparent motion. In Experiment 1-1, visual motion in the vertical direction was presented in the right visual field while a sound was alternately presented from two loudspeakers aligned in the horizontal plane. The results show that the perceived direction of visual motion can be modulated by the alternation of sound location, with participants reporting a perceived direction of visual motion that was inconsistent with visual inputs but consistent with the direction in which the sound alternated or lay between this direction and that of visual motion. The extent to which sound influenced the perceived direction of visual motion increased with increasing retinal eccentricity (from  $2.5^{\circ}$  to  $40^{\circ}$ ). In Experiment 1-2, we tested a situation where spectral cues were more important than binaural cues by presenting a sound that alternated between loudspeakers aligned in the vertical plane synchronized with visual motion in the horizontal direction. The results were almost identical to those of Experiment 1-1. Considering that spectral and binaural localization cues both play an essential role in the perception of sound in the diagonal plane (e.g., Grantham, Hornsby, & Erpenbeck, 2003), our findings can be generalized to the entire two-dimensional audiovisual space.

It has been reported that the temporal or semantic aspects of auditory information affect visual motion perception (Freeman & Driver, 2008; Maeda et al., 2004; Sekuler et al., 1997; Watanabe & Shimojo, 2001). In contrast, it has been shown that the spatial aspects of auditory stimuli, such as motion in space, have found less or no influence of auditory information on visual motion information at perceptual level (Alais & Burr, 2004a; Meyer & Wuerger, 2001; Soto-Faraco et al., 2002, 2003, 2004; Wuerger et al., 2003). The results of the present study suggest that the spatial aspects of a sound can alter visual motion perception in the peripheral visual field (where the spatial resolution of the visual system is low) even when visual stimuli are moving in a set direction.

Harrison, Wuerger, and Meyer (2010) and Meyer, Wuerger, Röhrbein, and Zetzsche (2005) investigated the effect of audiovisual cross-modal stimulation on motion perception and showed the audiovisual interaction in motion perception at the perceptual level. However, they did not independently measure the effect of auditory information on visual motion perception or the effect of visual information on auditory motion perception. The participants responded to any motion, irrespective of its source. Hence, our study is different from their studies.

Soto-Faraco et al. (2002, 2003, 2004) reported that the perceived direction of auditory motion could be profoundly affected by the direction of visual motion when these motion signals shared common motion paths and were presented at the same time. Specifically, when the direction of auditory and visual motion was in conflict, the auditory stimulus was perceived as moving in the same direction as the visual stimulus on 46% of trials. However, it was also reported that auditory motion had no effect on the perceived direction of visual motion. We found that the alternation of sound location did affect the perceived direction of visual motion. It should be noted that there was a difference in the auditory stimulus: alternation of sound location and auditory apparent motion. Furthermore, we tested in both the peripheral and central visual fields, but Soto-Faraco et al. (2002, 2003, 2004) tested only in the central visual field. The visual field is represented topographically in cortical areas involved in relatively low-level stages of visual processing, and the scale of the map varies depending on retinal location and cortical area (Daniel & Whitteridge, 1961; Rovamo & Virsu, 1979; Whitteridge & Daniel, 1961). Representations in the central visual field are larger and finer than those in the peripheral visual field, with the visual area represented by a single neuron increasing with increasing retinal eccentricity. The visual system might not be able to reliably distinguish the direction of motion of visual stimuli at larger retinal eccentricities. Conversely, though dependent on the type of sound, the spatial resolution of the auditory system (Mills, 1958) is almost the same or increases slightly up to 45° and dramatically at larger azimuth angles. Perrott et al. (1993) reported that location discrimination performance at azimuth angles of 20° or larger was better for the auditory modality than for the visual modality in a sequential discrimination task. These findings suggest that the relative importance of auditory spatial information might become greater at larger eccentricities, which is consistent with the finding of the current study that the size of the effect of auditory information on visual motion perception increased with increasing retinal eccentricity. Recently, Alais and Burr (2004b) reported that auditory stimuli modulated the perceived positions of visual stimuli when the reliability or intensity of these was decreased. A lower reliability of



Figure 7. Proportion of responses as a function of the range of deviation angle for Experiment 1. Error bars denote the standard error of the mean.

visual information would lead to a greater probability of an auditory effect at larger eccentricities. It is therefore very likely that audiovisual interactions are regulated in the brain in a compensatory manner for motion processing as well as for localization, depending on the reliability of visual and auditory inputs.

The difference between the directions of perceived and actual motion was less than 15° on 43% of trials and greater than 75° on 27% of trials at 40° of retinal eccentricity in Experiment 1-1. The difference was less than 15° on 28% of trials and greater than 75° on 40% of trials for the lower visual field, whereas it was less than  $15^{\circ}$  on 34% of trials and greater than  $75^{\circ}$  on 40%of trials for the upper visual field in Experiment 1-2 (see Figure 7). These results show that the effect of auditory information on visual motion perception can vary from trial to trial. This might happen because neither auditory nor visual information has a reliability and/or intensity sufficient to allow one modality to completely dominate the other and that they thus compete with one another at larger eccentricities. This interpretation is consistent with the idea of the maximum likelihood model of multisensory integration (e.g., Ernst & Banks, 2002; van Beers, Sittig, & Denier van der Gon, 1999) that multimodal interactions reflect the relative reliability of stimuli processed by different sensory modalities.

The current study has demonstrated that the spatial aspects of sound can alter the motion perception of visual stimuli moving within the peripheral visual field, suggesting that the idea that multimodal inputs interact depending on the accuracy and/or reliability of each input can be applied to motion perception. This can be considered as strong evidence that the auditory and visual modalities can mutually influence one another in motion processing so that the brain obtains the best estimates of external events.

# Acknowledgments

This research was supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Culture, Sports, Science and Technology (No. 19001004).

Commercial relationships: none. Corresponding author: Wataru Teramoto. Email: teraw@ais.riec.tohoku.ac.jp. Address: Tohoku University, 2-1-1 Katahira Aoba-ku, Sendai, Miyagi 980-8577, Japan.

## References

- Alais, D., & Burr, D. (2004a). No direction-specific bimodal facilitation for audiovisual motion detection. *Brain Research Cognitive Brain Research*, 19, 185–194.
- Alais, D., & Burr, D. (2004b). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.
- Anscombe, F. J. (1956). On estimating binomial response relations. *Biometrika*, 43, 461–464.
- Asano, F., Suzuki, Y., & Sone, T. (1990). Role of spectral cues in median plane localization. *Journal of the Acoustical Society of America*, 88, 159–168.
- Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism: Crossmodal interaction on the time dimension. 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology*, 50, 157–163.
- Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 20,* 1391–1397.

- Bertelson, P., & Aschersleben, G. (2003). Temporal ventriloquism: Crossmodal interaction on the time dimension. 1. Evidence from auditory–visual temporal order judgment. *International Journal of Psychophysiology*, 50, 147–155.
- Blauert, J. (1983). Spatial hearing: The psychophysics of human sound localization. Cambridge, MA: MIT Press.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of Physiology*, *159*, 203–221.
- Dick, M., Ullman, S., & Sagi, D. (1987). Parallel and serial processes in motion detection. *Science*, 237, 400–402.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.
- Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, 63, 719–725.
- Freeman, E., & Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Current Biology*, 18, 1262–1266.
- Grantham, D. W., Hornsby, B. W., & Erpenbeck, E. A. (2003). Auditory spatial resolution in horizontal, vertical, and diagonal planes. *Journal of the Acoustical Society of America*, *114*, 1009–1022.
- Harrison, N. R., Wuerger, S. M., & Meyer, G. F. (2010). Reaction time facilitation for horizontally moving auditory–visual stimuli. *Journal of Vision 10*(14):16, 1–21, http://www.journalofvision.org/content/10/14/16, doi:10.1167/10.14.16. [PubMed] [Article]
- Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., Suzuki, Y., & Iwaya, Y. (2009). The alternation of sound location Induces visual motion perception of a static object. *PLoS ONE*, 4, e8188.
- Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S., & Suzuki, Y. (2011). Auditory motion information drives visual motion perception. *PLoS ONE*, 6, e17499.
- Kahneman, D., & Wolman, R. (1970). Stroboscopic motion: Effects of duration and interval. *Perception* & *Psychophysics*, 8, 161–164.
- Kim, R., Peters, M. A., & Shams, L. (2012). 0 + 1 > 1: How adding noninformative sound improves performance on a visual task. *Psychological Science*, *23*, 6–12.
- Kitagawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, *416*, 172–174.

- Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A user's guide*. New Jersey: Lawrence Erlbaum Associates.
- Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, *14*, R990–R991.
- Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport*, *12*, 2557–2560.
- Meyer, G. F., Wuerger, S. M., Roehrbein, M., & Zetzsche, C. (2005). Low-level Integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, *166*, 538–547.
- Mills, A. W. (1958). On the minimum audible angle. Journal of the Acoustical Society of America, 30, 237–246.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Perrott, D. R., Costantino, B., & Cisneros, J. (1993). Auditory and visual localization performance in a sequential discrimination task. *Journal of the Acoustical Society of America*, 93, 2134–2138.
- Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78, 391–408.
- Radeau, M., & Bertelson, P. (1987). Auditory–visual interaction and the timing of inputs. Thomas (1941) revisited. *Psychological Research*, *49*, 17–22.
- Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*, 495–510.
- Scheier, C. R., Nijhawan, R., & Shimojo, S. (1999). Sound alters visual temporal resolution. *Investigative Ophthalmology & Visual Science*, 40, 4169.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308.
- Sorkin, R. D. (1999). Spreadsheet signal detection. Behavior Research Methods, Instruments, & Computers, 31, 46–54.
- Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., & Kingstone, A. (2002). The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Brain Research. Cognitive Brain Research*, 14, 139–146.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, 41, 1847–1862.
- Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Cross-modal dynamic capture: Congruency effects in the perception of motion across sensory modalities.

*Journal of Experimental Psychology: Human Perception and Performance, 30, 330–345.* 

- Teramoto, W., Manaka, Y., Hidaka, S., Sugita, Y., Miyauchi, R., Sakamoto, S., Gyoba, J., Iwaya, Y., & Suzuki, Y. (2010). Visual motion perception induced by sounds in vertical plane. *Neuroscience Letters*, 479, 221–225.
- van Beers, R. J., Sittig, A. C., & Denier van der Gon, J. J. (1999). Integration of proprioceptive and visual position information: An experimentally supported model. *Journal of Neurophysiology*, *81*, 1355–1364.
- Watanabe, K., & Shimojo, S. (2001). When sound affects vision: Effects of auditory grouping on visual motion perception. *Psychological Science*, 12, 109–116.

- Whitteridge, D., & Daniel, P. M. (1961). The representation of the visual field on the calcarine cortex. In R. Jung & H. Kornhuber (Eds.), *The visual system:* Neurophysiology and psychophysics (pp. 222–228). Berlin, Germany: Springer.
- Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics*, 65, 1188–1196.