Bulletin of the JSME Vol.11, No.6, 2024
Mechanical Engineering Journal

How close laterally should auditory cue be toward visual target to facilitate visual search under workload condition?

Kiichi NAKA* and Katsuya YAMAUCHI**
*School of Engineering, Kwansei Gakuin University
1 Gakuen-Uegahara, Sanda, Hyogo 669-1330, Japan
E-mail: k-naka@kwansei.ac.jp
**Faculty of Design, Kyushu University
4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan

Received: 1 July 2024; Revised: 18 August 2024; Accepted: 27 September 2024

Abstract
Auditory cues can draw individuals' spatial attention to visual targets and it enables individuals to find visual targets quickly. Some previous studies suggested that auditory cues should be in approximately the same functional field with visual targets (i.e., both events are presented within near field) to obtain quick responses even if in real situations such as driving. However, quantitative angular differences between auditory cues and visual targets are little investigated. This present study aimed to explore the angular differences that can elicit similar responses with auditory cues and visual targets presented from similar directions under workload conditions. Twenty-two participants were asked to perform visual search and tracking tasks simultaneously. The visual targets were at ±20°, ±40°, and ±60° in azimuth as 0° of frontal view with distance of 1.0 m from participants. The auditory cues were presented simultaneously with visual targets from 0°, ±20°, ±40°, and ±60°. The results of the response time analysis using a fitting approach revealed that the auditory cues within 40° from the visual targets could elicit similar responses similar with those when the auditory cues and visual targets were presented from the same direction even if the participants were deprived of their attentional resources by other tasks. When the angular difference between auditory cues and visual targets were larger than 60°, the participants' spatial attention was drawn to directions different from visual targets. It elicited a delay in finding visual targets because the participants should reallocate their spatial attention to visual targets. These results have meaningful implications for audio-visual user interface designs.

Keywords : Audio-visual, Cross-modal, Spatial cuing, Workload, Response time, Multimodal user interface

1. Introduction

Auditory cues presented in space facilitate the perception of a visual target in the vicinity, which is known as the audio-visual cross-modal spatial cuing effect (Driver and Spence, 1998; Lee and Spence, 2015, 2017; Santangelo et al., 2006; Spence and Driver, 2004; Spence et al., 1998, 2000). The audio-visual cross-modal spatial cuing effect has been studied for a long time in applied psychology and human factors (Spence and Soto-Faraco, 2020). In particular, many studies on vehicle driver assistance have been reported. Auditory cues, which contain directional information, draw the spatial attention of drivers and facilitate the visual search for visual information such as in-vehicle display, surrounding vehicles, and pedestrians. Therefore, drivers can elicit a faster response with lower subjective mental workload on driving manipulation (Chen et al., 2020; Ho and Spence, 2005; Ho et al., 2009; Lundqvist and Eriksson, 2019; Naka et al., 2021; Spence and Ho, 2008).

In previous studies, the auditory cues were presented from the same direction as the visual targets. However, it is not always possible to arrange audio-visual displays in the same direction owing to design constraints in actual vehicle interior designs. Therefore, determining how close the auditory cue should be arranged toward the visual target is beneficial to audio-visual multimodal user interface designs.

Paper No.24-00246
J-STAGE Advance Publication date : 9 October, 2024
[DOI: 10.1299/mej.24-00246]

The larger the distance between an auditory cue and a visual target, the weaker the audio-visual cross-modal spatial cuing effect (Mock et al., 2015; Mondor and Zatorre, 1995; Schmitt et al., 2001; Teder-Sälejärvi and Hillyard, 1998). By contrast, some studies argued that this effect can be obtained when the auditory cue and visual target are approximately in the same functional field, i.e., both events are presented from the near field (Driver and Spence, 1998; Ho et al., 2006; Lee and Spence, 2017; Schmitt et al., 2001; Spence and Driver, 2004). The auditory cue, namely, is not necessarily in the same direction as the visual target. For instance, when a visual target and an auditory cue were arranged within 18 degrees of angular differences, response times to visual targets were similar to those when the visual target and auditory cue were in the same direction (Gray et al., 2009).

Considering the application to actual driving situations, three issues should be explored, in addition to previous studies. First, responses to a visual target should be investigated at a wider angular difference between the visual target and auditory cue. The visual information in vehicles is often displayed in distal areas from the center of visual field such as left or right door mirrors. The visual and auditory displays can be displayed in distances larger than those examined in previous studies (e.g., Gray et al., 2009). Second, the influence of workload by a primary task (i.e., driving a vehicle) other than visual search should be considered. The drivers search for visual information while doing primary task that distract their attention. This fact could not be ignored because the less attentional resources mitigate the audio-visual cross-modal spatial cuing effect (Santangelo et al., 2007; Santangelo et al., 2008). Third, the effect of the auditory cue should be investigated under a condition where the auditory cue and visual target are presented simultaneously. In a previous study on audio-visual cross-modal spatial cueing effects, the auditory cue preceded the visual target by several-hundred milliseconds (e.g., Lee and Spence, 2015). This audio-visual inter-stimulus-interval represented the nature of an auditory spatial attention that it had a peak after several-hundred milliseconds from an arrival of an auditory stimulus. On the contrary, it is more plausible if the auditory and visual information are presented simultaneously for an actual audio-visual user interface design.

Therefore, this study investigated the angular differences between a visual target and an auditory cue, which can elicit responses similar to those when the visual target and auditory cue are presented simultaneously from the same direction, at a condition in which the participants are engaged in another task other than visual search. An experiment which comprised of two tasks—visual search task and tracking task—was conducted. Response times to visual targets were compared between auditory cue locations at each visual target location. Using the results of the response times analysis, this present study discusses the angular difference between visual target and auditory cue, which was required to elicit quick response under workload condition.

2. Method

2.1 Experimental set-up

The block diagram of the experimental setup is shown in Fig. 1, and the arrangement of LEDs, loudspeakers, one-handed joystick and keyboard, and an 8-inch display is shown in Fig. 2. This set-up consisted of two systems: a “visual search task system” and a “tracking task system”. The experimental program was built in Psychtoolbox-3 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) in MATLAB 2021a.

The PC used for the visual search task system was an Alienware m17 R4 (CPU: Intel core i9). Seven loudspeakers (Genelec 8010A) were placed in the azimuth plane at 0°, ±20°, ±40°, and ±60° from the participant's front at 0°. The distance between the participant’s head position and the loudspeakers was 1.1 m. The loudspeakers were controlled via an audio interface (MOTU 828ES). The LEDs were placed in front of each loudspeaker at a distance of 1.0 m from the participant's head position. The LED in front of the participants was red, whereas the other LEDs in pairs of the upper and lower were white. The distance between the upper and lower white LEDs was 1 cm. The luminance of the LEDs was measured using a luminance meter (Konica Minolta LS-110). The luminance was 145 cd/m^2 for the red LED and 245 cd/m^2 for the white LEDs. The LEDs were controlled using an Arduino Mega 2560. The loudspeakers and LEDs were positioned at the participant's eye level. The participant's input to the visual search task was obtained using a one-handed keyboard (Koolertron AE-AMAG09-RDB) with two keys arranged vertically next to each other. One auditory cue which had the lowest perceived urgency in the previous study of warning sounds was chosen (Naka and Yamauchi, 2023). The auditory cue was assumed that it did not evoke negative impressions that were not relevant to the task. The time-waveform and frequency spectrum of the auditory cue are shown in Fig. 3. The duration of the auditory cue was 500 ms, and contained multiple frequency components. The A-weighted sound pressure level of the auditory cues at the participants’ ear position was set at 70 dB.

[DOI: 10.1299/mej.24-00246]

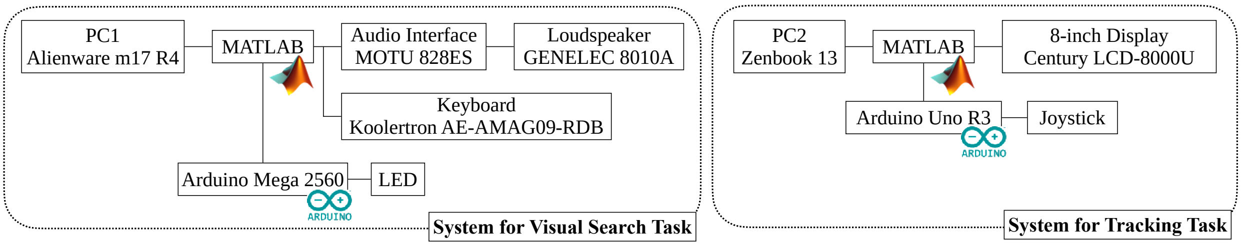

Fig. 1 Block diagram of the experimental system for visual search and tracking tasks.

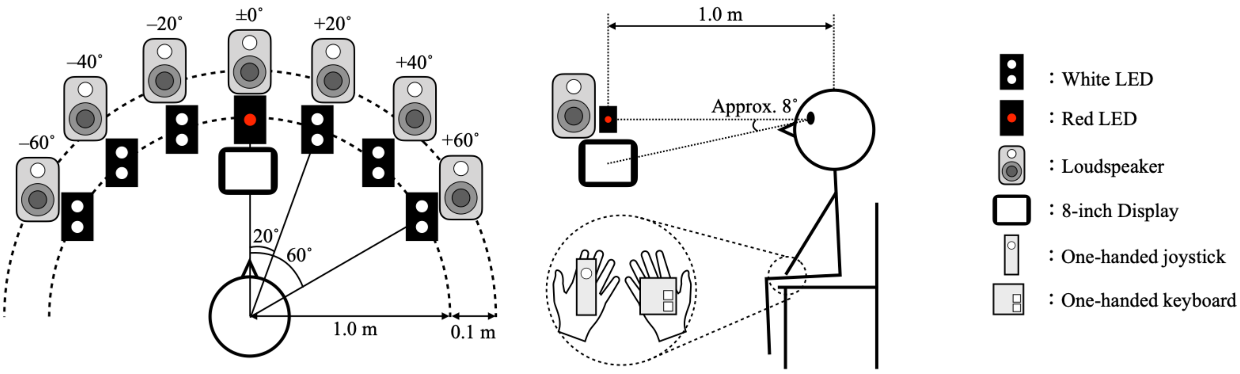

Fig. 2 Experimental set-up. Participants had one-handed joystick and keyboard in their left and right hand.

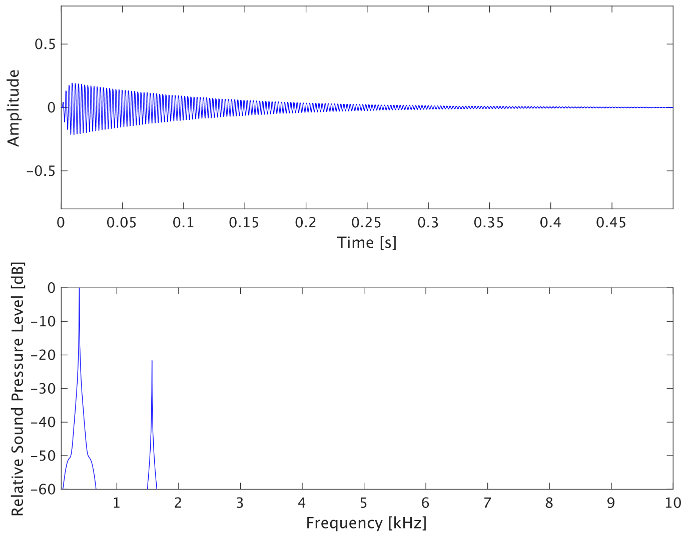

Fig. 3 Time-waveform and frequency spectrum of the auditory stimulus.

The PC used for the tracking task system was a Zenbook 13 (CPU: Intel core i7). An 8-inch display (Century LCD-8000U, resolution: 800 × 600 pixels) was placed underneath the LED in front of the participant (approximately 8° overhead angle). The participant's input to the tracking task was obtained using a joystick via Arduino Uno R3.

2.2 Participant

Twenty-two participants (15 males and seven females, ages ranging from 22 to 26 years) participated in the experiment. All the participants underwent audiometric testing using an audiometer (Rion AA-77A). They had normal hearing with their average hearing levels <10 dB at 500, 1000, 2000, and 4000 Hz. Written informed consent was obtained from each participant before the experiment. The study was approved by the Ethics Committee of Kyushu University and was based on the ethical standards of the Declaration of Helsinki.

2.3 Procedure

The accuracy of sound source localization was investigated at first. Subsequently, the participants were asked to perform the visual search and tracking tasks simultaneously.

2.3.1 Investigation of participants' accuracy of sound source localization

This study aimed to explore the arrangement of the visual target and auditory cue to facilitate visual search. Therefore, participants should have accuracy in sound source localization. To confirm the participants' accuracy for sound source localization, the auditory cue was presented ten times per loudspeaker, for a total of 70, in a random order, at 5 s intervals. The participants selected the loudspeaker which the auditory cue was presented.

2.3.2 Visual search and tracking tasks

The schematic diagram of a single trial in the visual search task is shown in Fig. 4. First, the red LED was turned on for 500 ms. Subsequently, one white LED was turned on as a visual target and an auditory cue was presented simultaneously from one loudspeaker. Before the experiment, the timing of the white LED and auditory cue presentation was measured by detecting a voltage on a lead of LED and by the microphone (Dayton Audio EMM-6) which was located at the participant's head position. These signal onsets were measured by an oscilloscope (Siglent SDS1104X-E), and they were calibrated to within ± 5 ms gap between onsets. Participants responded by pressing the up or down key as quickly and accurately as possible. This experimental design is known as the orthogonal spatial cuing paradigm (Spence and Driver, 1994, 1996, 1997; Ho et al., 2006), in which responses to horizontal search (i.e., left or right side) are intentionally obtained through elevation discrimination (i.e., upper or lower). The intertrial interval was random in the range of 2-4 s. The response times to white LED were collected in the visual search task.

While performing the visual search task, the participants were also instructed to always perform the tracking task. The screenshot of the tracking task display is shown in Fig. 5. In the tracking task, a white circle was displayed on an 8-inch display with a gray background. The white circle had a diameter of 1 cm. The circle moved at a constant speed at an interval of 5 s per revolution in a clockwise direction on a circumferential orbit. The diameter of the orbit was 8 cm, positioned at the center of the display. The experimental participants manipulated a black square icon, with a side of 0.6 cm, using a joystick and were asked to follow the white circle with the icon. The RGB values for each color were specified as (255, 255, 255) for white, (0, 0, 0) for black, and (170, 170, 170) for gray.

The conditions for the audio-visual stimuli were 42 combinations of 6 visual target positions (±20°, ±40°, ±60°) and 7 auditory cue positions (0°, ±20°, ±40°, ±60°). Each combination was presented ten times, for a total of 420 trials, in random order to each participant in the experiment. The experiment was conducted with breaks every 100 trials. The experiment was conducted in a soundproof room at Kyushu University. The background noise level of the soundproof room was 22.0 dB at the A-weighted sound pressure level, and the horizontal illuminance at the participants' eye level was approximately 500 lx.

3. Results

3.1 Accuracy of sound source localization

Two participants did not answer the correct directions more than 20%. The accuracy of sound source localization might affect the difficulty of the visual search (e.g., Naka et al., 2021). Hence, they were eliminated from the data because they did not perceive the correct auditory cue direction. The mean percentage of correct responses was 92% for the other 20 participants. Their response data were submitted for analysis.

3.2 Response time analysis of visual search task

For the tracking task, no participant lost joystick input for more than 1 s. Therefore, we analyzed the response times as the participants paid continuous attention to the tracking task while performing the visual search task.

The distribution of response times obtained for each audio-visual combination was analyzed by fitting an ex-Gaussian distribution. Balota and Yap (2011) suggested that the ex-Gaussian distribution should be applied to response time analysis because the ex-Gaussian distribution fits well with the response time data distribution. The ex-Gaussian distribution comprises three parameters: μ, σ, and τ $(\sigma > 0, \tau > 0)$ (details of the equation can be found in Haj et al.,

[DOI: 10.1299/mej.24-00246]

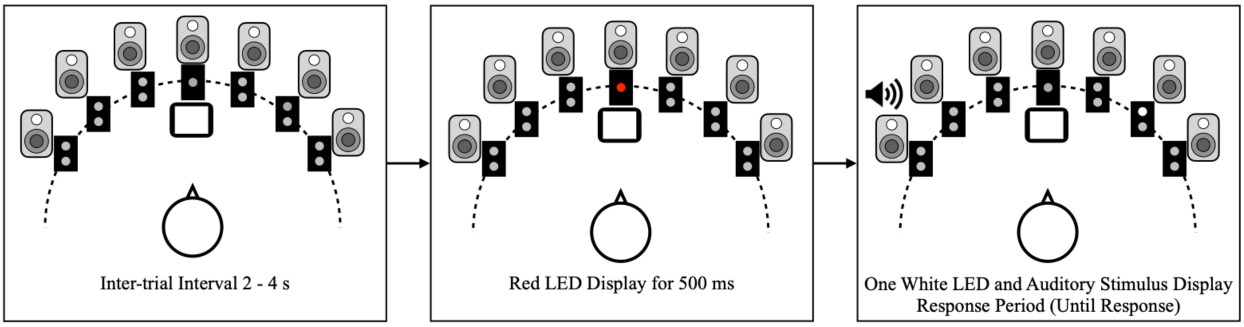

Fig. 4 Schematic diagram of a single trial of visual search task.

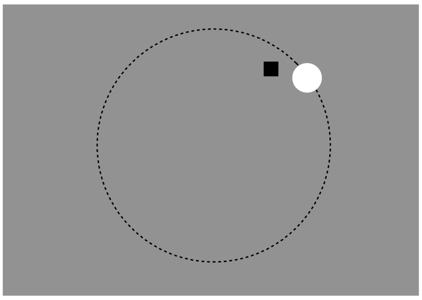

Fig. 5 Screenshot of a tracking task. The dashed line shows the orbit of white circle. Participants miniplate black square using a joystick.

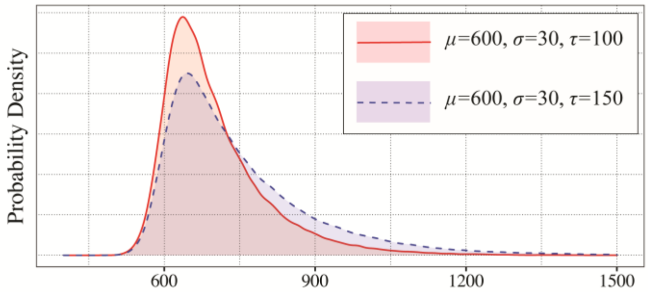

Fig. 6 Examples of ex-Gaussian distributions. The increase in τ stretches the tail of distribution.

2021). The mean of the ex-Gaussian distribution is defined as $\mu + \tau$, and τ influences the tail of the distribution, as shown in Fig. 6. Hence, the analysis using the ex-Gaussian allowed us to investigate the mean of response time and the shape of distribution, particularly focusing on the tail.

Response time analysis was conducted using R (ver. 4.0.3) and rstan package (ver. 2.21.2). Of the error responses, 2.2 % was excluded, and 8,216 response time data were submitted for analysis. We applied the uniform distribution as the prior distribution for each parameter. Markov Chain Monte Carlo (MCMC) method was used to estimate the posterior distribution of the parameters for each audio-visual combination. The model ran four chains. We obtained 11,000 samples per chain and discarded the first 1,000 samples as burn-in to minimize the effect of initial values on the posterior distribution. The model convergence was assessed using the potential scale reduction statistic $\hat{R}$ (< 1.1 for all parameters).

3.2.1 Mean response time

Figure 7 shows the posterior distribution of mean response time for each audio-visual combination. The annotation of same angle (SA) showed that the visual target and auditory cue were the same. The solid horizontal lines show the median of each SA condition. The boxplot shows the interquartile range of the posterior distribution. The area between upper and lower edge of the distribution and boxplot means 95% Bayesian credible interval (95% CI). For each comparison, the probability of direction ($p.d.$) between SA and each different angle (DA) condition was calculated. The $p.d.$ is the proportion of the difference distribution with the same sign (i.e., positive or negative) as the median of difference distribution (Makowski et al., 2019a, 2019b). In other words, the larger the difference of the values between SA and DA conditions, the greater the $p.d.$ is. The $p.d.$ of 0.950, 0.975, and 0.995 correspond to the frequentist p-value of 0.10, 0.05, and 0.01 (see also Wolpe et al., 2022).

When the visual target angle was -20°, no DA conditions yielded more than 0.950 of the probability of having a slower mean response time than the SA condition. When the visual target angle was +20°, the DA condition for -60° of the auditory cue had more than 0.950 of the probability of having a slower mean response time than the SA condition ($p.d. = 0.960, \text{median} = 28.6, 95\%\text{CI}[-6.3, 88.3]$).

[DOI: 10.1299/mej.24-00246]

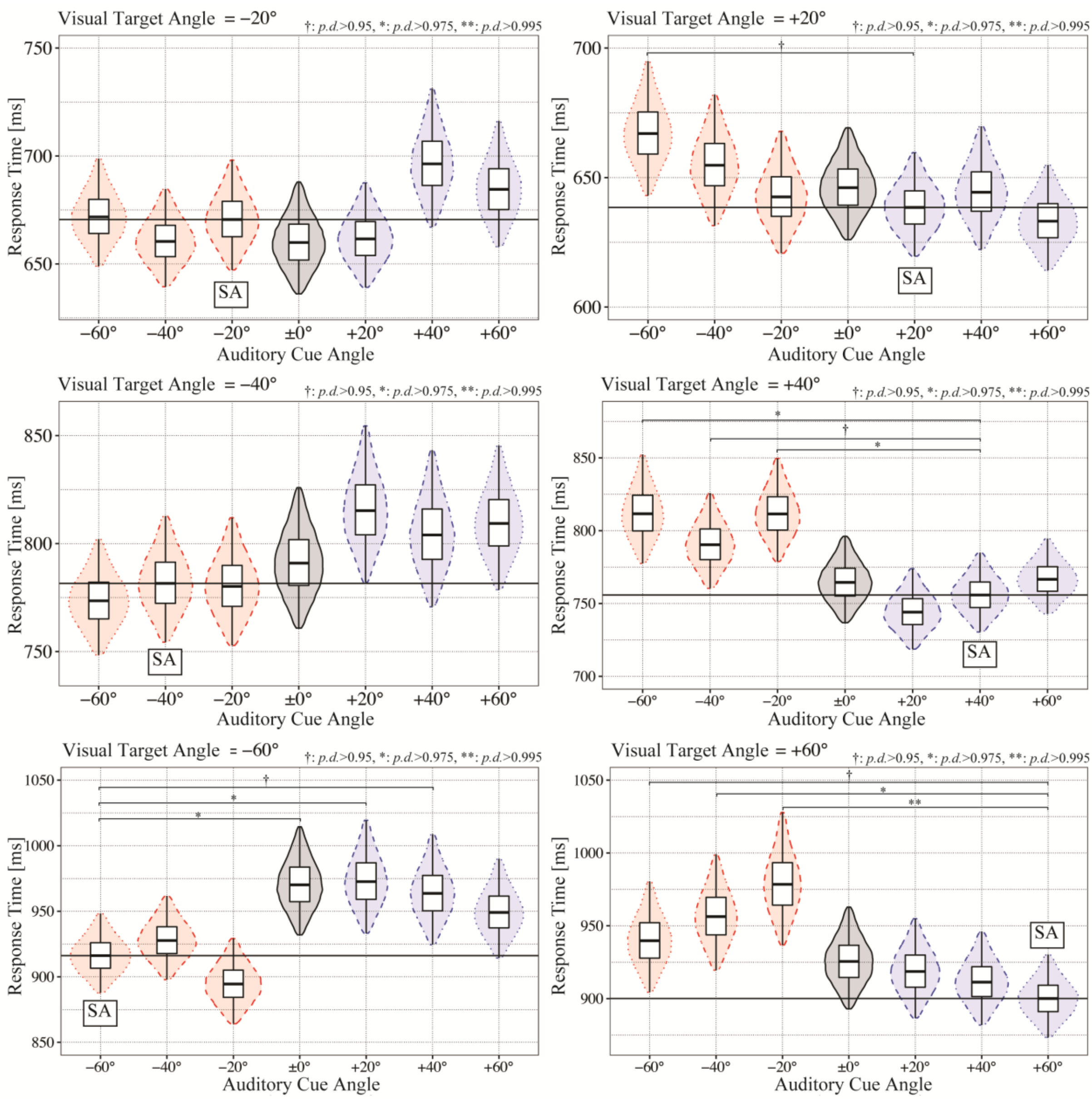

Fig. 7 Posterior distributions of mean response times for each audio-visual combination. SA indicates a condition in which the visual target and auditory cue were placed at the same angle direction. The combinations of SA condition and other conditions with more than 95% of $p.d.$ were annotated.

When the visual target angle was -40°, no DA conditions yielded more than 0.950 of the probability of having a slower mean response time than the SA condition. When the visual target angle was +40°, the DA conditions for -20°, -40°, and -60° of the auditory cue angles had more than 0.950 of the probability of having a slower mean response time than the SA condition (-20°: $p.d. = 0.994, \text{median} = 55.8, 95\%\text{CI}[12.2, 101.7]$; -40°: $p.d. = 0.950, \text{median} = 34.5, 95\%\text{CI}[-7.2, 77.9]$; -60°: $p.d. = 0.992, \text{median} = 55.8, 95\%\text{CI}[10.7, 104.1]$).

When the visual target angle was -60°, the DA conditions for 0°, +20°, and +40° of the auditory cue angles had more than 0.950 of the probability of having a slower mean response time than the SA condition (0°: $p.d. = 0.983, \text{median} = 54.0, 95\%\text{CI}[4.3, 106.4]$; +20°: $p.d. = 0.985, \text{median} = 56.4, 95\%\text{CI}[5.9, 111.2]$; +40°: $p.d. = 0.967, \text{Median} = 47.6, 95\%\text{CI}[-3.5, 100.3]$). When the visual target angle was +60°, the DA conditions for -20°, -40°, and -60° of the auditory cue angles had more than 0.950 of the probability of having a slower mean response time than the SA condition (-20°: $p.d. = 0.999, \text{median} = 78.4, 95\%\text{CI}[26.3, 134.7]$; -40°: $p.d. = 0.989, \text{median} = 56.3, 95\%\text{CI}[8.6, 106.8]$; -60°: $p.d. = 0.951, \text{median} = 36.7, 95\%\text{CI}[-6.3, 88.3]$).

[DOI: 10.1299/mej.24-00246]

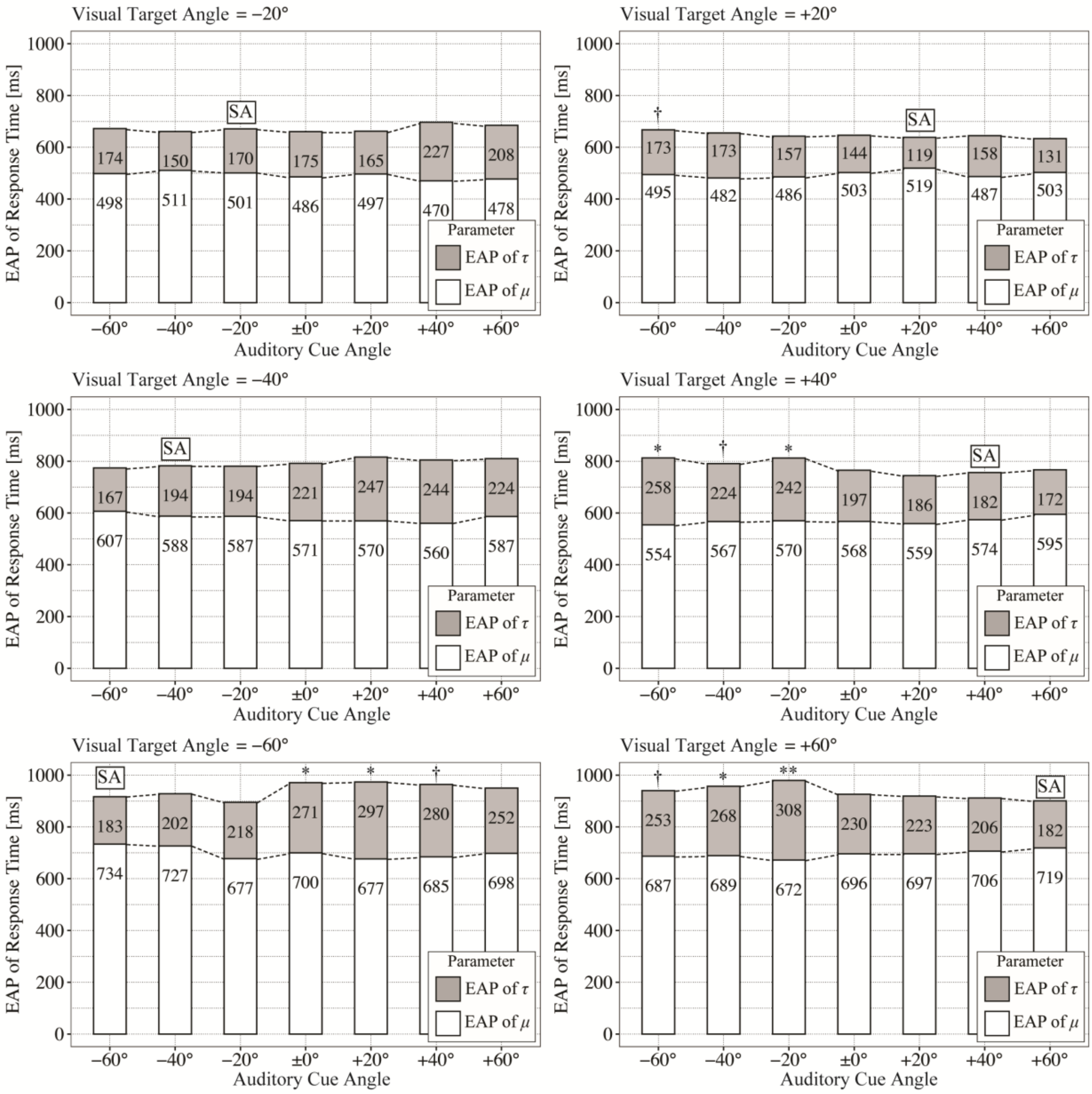

Fig. 8 Expected a posteriori of τ and μ for each audio-visual combination.

3.2.2 What parameter contributes to the increase in mean response time?

To explore the change in ex-Gaussian distribution with the increase in mean response time of the DA condition compared to the SA condition, we calculated the expected a posteriori (EAP) of μ and τ, i.e., the mean of the posterior distributions. Figure 8 shows the EAP of μ and τ for each audio-visual combination. The annotation shows the conditions at which the $p.d.$ of the difference distribution of the mean response time between SA and DA conditions were more than 0.950. For $p.d.$ greater than 0.950, the τ was increased compared to that of the SA condition.

3.3 Error rate analysis of visual search task

The repeated measures analysis of variance with the two within-participants factors of visual target and auditory cue did not yield any effects. The mean error rate in each condition was 2.3%, with a maximum error rate of 5.0% and a minimum error rate of 0.0%.

4. Discussion

When the auditory cues were presented within 40˚ of the visual targets, the participants could respond at speeds similar to those at which the auditory cue and visual target were presented from the same direction regardless of the

[DOI: 10.1299/mej.24-00246]

location of the visual targets. This result is consistent with the implications of the previous studies, suggesting that auditory cues can draw visual spatial attention when the auditory cues are approximately in the same functional field with the visual target, i.e., both events are not presented from completely same location but from near field (Driver and Spence, 1998; Ho et al., 2006; Lee and Spence, 2017; Schmitt et al., 2001; Spence and Driver, 2004). This study could demonstrate quantitatively angular differences between the auditory cue and visual target that elicit responses similar to those when the auditory cue and visual target were in the same direction, at a condition where the participants were continuously deprived of their attentional resources to frontal view.

When the angular differences were more than 60˚, the participants responded slower in several conditions than when the auditory cues and visual targets were in the same direction. In addition, the $\boldsymbol{\tau}$ increased when the participants' responses were slower. This indicates that an auditory cue presented from an angle largely different from the visual target drew the participants' spatial attention to a direction different from the visual target. The participants needed to re-allocate their spatial attention to the visual target. This process elicited larger response delays, which was indicated by an increase in $\boldsymbol{\tau}$.

This study has several limitations. At first, the effects of the ease of sound source localization need to be investigated in the visual search task. The accuracy of sound source localization for auditory cues can affect the performance of visual search (e.g., Naka et al., 2021). We used one auditory cue in this study, however, different results might be observed if we use other auditory cues which are easier (e.g., white noise) or more difficult (e.g., pure tone) to localize the sound source. Second, the difficulty of tasks should be considered. In dual-task conditions, the difficulty of one task would affect the performance of another task. Regarding this study, the region of auditory cues which elicit the faster response to visual targets might be modulated by the tracking task if it was more difficult.

This study showed the angular difference between the auditory cue and visual target that elicit responses similar to those when the auditory cue and visual target were presented from the same direction at a condition where the participants were deprived of their attentional resource by other tasks. These results have meaningful implication for multimodal user interface designs. Future study must validate this insight under real driving situations.

Acknowledgement

This work was partly supported by JSPS KAKENHI (grant number 22H03889) and SUZUKI FOUNDATION. We would like to thank Editage (www.editage.jp) for English language editing.

References

Balota, D. A. and Yap, M. J., Moving beyond the mean in studies of mental chronometry. Current Directions in Psychological Science, Vol.20, No.3 (2011), pp.160–166.

Brainard, D. H., The Psychophysics Toolbox. Spatial Vision. Vol.10, No.4 (1997), pp.433–436.

Chen, J., Šabić, E., Mishler, S., Parker, C. and Yamaguchi, M., Effectiveness of lateral auditory collision warnings: should warnings be toward danger or toward safety?, Human Factors: the Journal of Human Factors and Ergonomics Society, Vol.64, No.2 (2020), pp.418–435.

Driver, J. and Spence, C., Cross–modal links in spatial attention, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, Vol.353, (1998), pp.1319–1331.

Gray, R., Mohebbi, R. and Tan, H. Z. The spatial resolution of crossmodal attention: Implications for the design of multimodal interfaces. ACM Transactions on Applied Perception (TAP), Vol.6, No.1 (2009), pp.1–14.

Haj, A. E., Slaoui, Y., Solier, C. and Perret, C., Bayesian estimation of the ex-Gaussian distribution. Statistics, Optimization & Information Computing, Vol.9, No.4 (2021), pp.809–819.

Ho, C. and Spence, C., Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. Journal of Experimental Psychology: Applied, Vol.11, No.3 (2005), pp.157–174.

Ho, C., Tan, H. Z. and Spence, C., The differential effect of vibrotactile and auditory cues on visual spatial attention. Ergonomics. Vol.49, No.7 (2006), pp.724–738.

Ho, C., Santangelo, V. and Spence, C. Multisensory warning signals: when spatial correspondence matters, Experimental Brain Resarch, Vol.195, (2009), pp.261–272.

Kleiner, M., Brainard, D., Pelli, D. and Ingling, A., What's new in psychtoolbox-3. Perception. Vol.36, No.14 (2007),

pp.1–16.
Lee, J. and Spence, C., Audiovisual crossmodal cuing effects in front and rear space, Frontiers in Psychology, Vol.6, No.1086 (2015), pp.1–10.
Lee, J. and Spence, C., On the spatial specificity of audiovisual crossmodal exogenous cuing effects, Acta Psychologica, Vol.177, (2017), pp.78–88.
Lundqvist, L. M. and Eriksson, L. Age, cognitive load, and multimodal effects on driver response to directional warning, Applied Ergonomics, Vol.76, (2019), pp.147–154.
Makowski, D., Ben-Shachar, M. S., Chen, S. H. A. and Lüdecke, D., Indices of effect existence and significance in the Bayesian framework. Frontiers in Psychology, Vol.10, No.2767 (2019a), pp.1–14.
Makowski, D., Ben-Shachar, M. and Lüdecke, D., bayestestR: describing effects and their uncertainty, existence and significance within the Bayesian framework. The Journal of Open Source Software, Vol.4, No.40 (2019b), pp.1–8.
Mock, J. R., Seay, M. J., Charney, D. R., Holmes, J. L. and Golob, E. J., Rapid cortical dynamics associated with auditory spatial attention gradients, Frontiers in Neuroscience, Vol.9, No.179 (2015), pp.1–14.
Mondor, T. A. and Zatorre, R. J. Shifting and focusing auditory spatial attention, Journal of Experimental Psychology: Human Perception and Performance, Vol.21, No.2 (1995), pp.387–409.
Naka, K. and Yamauchi, K. Audio–visual cross-modal correspondences of perceived urgency: Examination through a speeded discrimination task. Multisensory Research, Vol.36, No.5 (2023), pp.413–428.
Naka, K., Yamauchi, K., Tanoue, N. and Kawata A., Effect of directional consistency of auditory sign and visual information on the drivers' information recognition, Journal of Acoustical Society of Japan, Vol.77, No.8 (2021), pp.491–499 (in Japanese).
Pelli, D. G., The VideoToolbox software for visual psychophysics: transforming numbers into movies, Spatial Vision, Vol.10, No.4 (1997), pp.437–442.
Santangelo, V., Belardinelli, M. O. and Spence, C. The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. Journal of Experimental Psychology: Human Perception and Performance, Vol.33, No.1 (2007), pp.137–148.
Santangelo, V., Finoia, P., Raffone, A., Belardinelli, M. O. and Spence, C. Perceptual load affects exogenous spatial orienting while working memory load does not. Experimental Brain Research, Vol.184, No.3 (2008), pp.371–382.
Santangelo, V., Lubbe, R. H. J. V. der, Belardinelli, M. O. and Postma, A., Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: a comparison of reflexive orienting mechanisms, Experimental Brain Research, Vol.173, No.1 (2006), pp.40–48.
Schmitt, M., Postma, A. and Haan, E. H. F. de, Cross-modal exogenous attention and distance effects in vision and hearing, European Journal of Cognitive Psychology, Vol.13, No.3 (2001), pp.343–368.
Spence, C. J. and Driver, J., Covert spatial orienting in audition: exogenous and endogenous mechanisms. Journal of Experimental Psychology: Human Perception and Performance, Vol.20, No.3 (1994), pp.555–574.
Spence, C. and Driver, J., Audiovisual links in endogenous covert spatial attention. Journal of Experimental Psychology: Human Perception and Performance. Vol.22, No.4 (1996), pp.1005–1030.
Spence, C. and Driver, J., Audiovisual links in exogenous covert spatial orienting. Perception & Psychophysics. Vol.59, No.1 (1997), pp.1–22.
Spence, C. and Driver, J., Crossmodal space and crossmodal attention (2004), Oxford University Press.
Spence, C. and Ho, C., Multisensory interface design for drivers: past, present and future, Ergonomics, Vol.51, No.1 (2008), pp.65–70.
Spence, C., Nicholls, M. E. R., Gillespie, N. and Driver, J., Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision, Perception & Psychophysics, Vol.60, No.4 (1998), pp.544–557.
Spence, C., Pavani, F. and Driver, J., Crossmodal links between vision and touch in covert endogenous spatial attention, Journal of Experimental Psychology: Human Perception and Performance, Vol.26, No.4 (2000), pp.1298–1319.
Spence, C. and Soto-Faraco, S., Crossmodal attention applied, Elements in Perception, (2020), pp.1–52, Cambridge University Press.
Teder-Sälejärvi, W. A. and Hillyard, S. A., The gradient of spatial auditory attention in free field: an event-related potential study, Perception & Psychophysics, Vol.60, No.7 (1998), pp.1228–1242.
Wolpe, N., Hezemans, F. H., Rae, C. L., Zhang, J. and Rowe, J. B. The pre-supplementary motor area achieves inhibitory control by modulating response thresholds. Cortex, Vol.152, (2022), pp.98–108.

[DOI: 10.1299/mej.24-00246]