Paper:

Auditory Virtual Reality for Insect Phonotaxis

Noriyasu Ando[*1,*2], **Hisashi Shidara**[*3], **Naoto Hommaru**[*4], **and Hiroto Ogawa**[*3]

[*1]Department of Systems Life Engineering, Maebashi Institute of Technology
460-1 Kamisadori-cho, Maebashi, Gunma 371-0816, Japan
E-mail: ando@maebashi-it.ac.jp
[*2]Research Center for Advanced Science and Technology, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
[*3]Department of Biological Sciences, Faculty of Science, Hokkaido University
Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan
[*4]Graduate School of Life Science, Hokkaido University
Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan

Insects have a sophisticated ability to navigate real environments. Virtual reality (VR) is a powerful tool for analyzing animal navigation in laboratory studies and is the most successful when used in the study of visually guided behaviors. However, the use of VR with non-visual sensory information, such as sound, on which nocturnal insects rely, for analyzing animal navigation has not been fully studied. We developed an auditory VR for the study of auditory navigation in crickets, *Gryllus bimaculatus*. The system consisted of a spherical treadmill on which a tethered female cricket walked. Sixteen speakers were placed around the cricket for auditory stimuli. The two optical mice attached to the treadmill measured the cricket's locomotion, and the sound pressure and direction of the auditory stimuli were controlled at 100 Hz based on the position and heading of the cricket relative to a sound source in a virtual arena. We demonstrated that tethered female crickets selectively responded to the conspecific male calling song and localized the sound source in a virtual arena, which was similar to the behavior of freely walking crickets. Further combinations of our system with neurophysiological techniques will help understand the neural mechanisms for insect auditory navigation.

1. Introduction

Insects have a sophisticated ability to navigate through challenging environments. Effective navigation is essential for insects to perform actions that allow them to survive and reproduce, such as finding suitable habitats, food, and mating partners. The behavior of several insect species has been intensively studied in neuroethology for many years, including pheromone-source localization of male moths, sound source localization of female crickets, and homing behaviors of desert ants and bees. Furthermore, behavioral algorithms or neural circuit models have been implemented on mobile robots and subsequently tested to evaluate biological hypotheses in real environments [1, 2]. These studies are the basis of bio-inspired or bio-mimetic autonomous robots that can behave in uncertain environments.

Among these model insects, cricket phonotaxis, in which female crickets recognize and localize the species-specific calling song of males, is a representative study in neuroethology. The biophysics of sound reception and the neural mechanisms of species-specific song recognition and steering behavior have been intensively studied [3–5] and that behavior has also been reproduced using mobile robots [6, 7]. Furthermore, laboratory studies on phonotactic behavior have produced sophisticated experimental tools for locomotion analysis, including a treadmill system that allows researchers to analyze insect locomotion in unlimited space [8]. Kramer's spherical treadmill (locomotion compensator), which allows an insect to walk freely on a sphere while locomotor data are measured, has been used to analyze cricket phonotaxis in a previous study [9]. In that study, a cricket with a reflective marker was made to walk on a sphere in response to a conspecific calling song, and the system controlled the sphere's rotation to keep the cricket's position at the top of the sphere. The other type of spherical treadmill, a trackball system in which a tethered insect walks on a ball, has also been used to analyze cricket phonotaxis. The rotation of the ball from insect locomotion is measured by optical sensors. This system is advantageous if the axes of the head and body of the insect are fixed as it allows researchers to present sensory stimuli (auditory [10], mechanical [11], visual, and combinations of the two of them [12, 13]) from an arbitrary direction relative to the body axis. Furthermore, fixation of the head and/or body allows the application of neurophysiological techniques to investigate neural activity during locomotion [14, 15]. The trackball system is relatively simple compared to Kramer's sphere that requires feedback control for positioning the sphere. Furthermore, the trackball system can be made using an op-

tical mouse with low costs has become a major tool for analyzing animal locomotion.

The trackball system has contributed to our understanding of the behavioral mechanisms of cricket phonotaxis [8]. However, the behavior observed on a spherical treadmill is only sound "tracking" behavior (steering toward the prescribed sound direction and walking toward it); this condition is different from true navigation in which the cricket tracks the sound and localizes the source. This difference is primarily due to the presence or absence of auditory sensory feedback. In a real-life situation, the direction and intensity of the sound will change depending on the position and direction of the cricket relative to the sound source.

To receive sensory feedback on a spherical treadmill, the loop between animal locomotion and sensory stimuli should be closed [16]. This so-called virtual reality (VR) system has been well developed for vision; visual stimuli are controlled by the movement of the trackball (i.e., insect locomotion) and are presented on a screen or on displays surrounding the trackball [17, 18]. Such VR systems, when combined with electrophysiology or imaging techniques, have become powerful tools to investigate visually guided animal navigation. In recent years, olfactory VR systems have also been reported [19–21].

However, auditory VR has not been fully established. To achieve sound source localization, a cricket must be able to move in a two-dimensional virtual space and receive auditory stimuli from any direction. Closed-loop control of left and right speakers for cricket phonotaxis has already been reported [22, 23]; however, the two speakers were installed in front of the cricket and could not reproduce situations where the cricket is pointing at the direction opposite to the sound source.

To achieve true sound source navigation of crickets, we developed an auditory VR system in which a trackball was surrounded by an array of speakers. This type of auditory VR was reported in a study of mammalian spatial navigation [24], but was limited to one-dimensional movement. In contrast, our system allowed a cricket to behave in a two-dimensional virtual space. We demonstrated that crickets are able to localize a virtual sound source.

2. Materials and Methods

2.1. Experimental Animal

The eighth (last) instar larvae of crickets (*Gryllus bimaculatus*) were purchased from Tsukiyono Farm in Gunma, Japan. The females were then isolated from the instar males and reared in an incubator at 27°C with a light : dark cycle of 12 : 12 h. Adult female crickets aged 2–3 weeks after imaginal molting were used for the phonotaxis experiment. All females were virgins and did not experience the conspecific male song. A small strip of paper tag was attached to the dorsal thorax of the crickets with an adhesive (G-17, Konishi, Osaka, Japan), and the tag was clipped to the attachment of the trackball to tether

the cricket on the ball. All experiments were performed during the dark phase in a dark room at room temperature (25°C–30°C).

2.2. Virtual Reality System

2.2.1. Overall Design

We designed an auditory VR system consisting of an array of 16 speakers surrounding a cricket tethered to a trackball (**Fig. 1**). The walking behavior of the tethered cricket was determined using two optical mice. The position and heading of the cricket relative to the sound source (male calling song) on a virtual space were calculated to set the parameters of sound stimulation (sound pressure level and direction). Our system allowed us to present auditory stimulus from all directions, which is a requirement for the study of virtual navigation in a fully two-dimensional space [a].

2.2.2. Trackball

The trackball used was a polystyrene ball 75 mm in diameter. Two optical mice (M-BGURSV, Elecom, Osaka, Japan) were placed at the back and the right side of the ball (**Fig. 2(a)**). The ball was made to float using a blower fan (MBDC12B4, Nidec Servo Corp., Gunma, Japan) placed at the bottom of the setup. The "back mouse" (master) measured the pitch (fore/aft) and yaw (right/left) rotation while the "side mouse" (slave) measured the roll (right/left) rotation of the ball at a sampling rate of 100 Hz. The ball movement was measured by the two mice, and was sent to two microcontroller boards (Arduino Due) via universal serial bus (USB); the two controllers were connected by I^2C to calculate the cricket's position (x_t, y_t) and heading relative to the y-axis $(\theta_t, -\pi \sim \pi)$ at each sampling time (t) in a virtual space. The calculated locomotion data were then sent to a personal computer via serial connection for display and storage. The programs were developed with Arduino IDE (ver. 1.8) for the microcontrollers and Processing (ver. 3) for data acquisition with a PC. Detailed calculations of the position of the cricket in a virtual arena from the trackball measurements are shown below.

The ball movement (i.e., the cricket's movement on the trackball) in each direction, which was measured as the number of counts at each sampling time (10 ms), was converted to real units ($\Delta x_{m,t}$, $\Delta y_{m,t}$, $\Delta y_{s,t}$; unit [mm]; **Fig. 2(b)**) by multiplying with the constants obtained from the calibration process (circumference of the ball divided by the number of counts required for the ball to go around). Since crickets do not actively walk laterally like crabs, we considered left-right rotation ($\Delta x_{m,t}$) and fore-aft movement ($\Delta y_{m,t}$) to be the main elements of movement, and the lateral element ($\Delta y_{s,t}$) to be a deviation.

The calculation of the main elements was based on two-wheeled robots controlled by an insect-driven trackball [25, 26]. Excluding lateral deviation, ball rotation ($\Delta x_{m,t}$, $\Delta y_{m,t}$) within a single sampling duration (t to $t+1$) represented movement from A to B′ in the virtual space
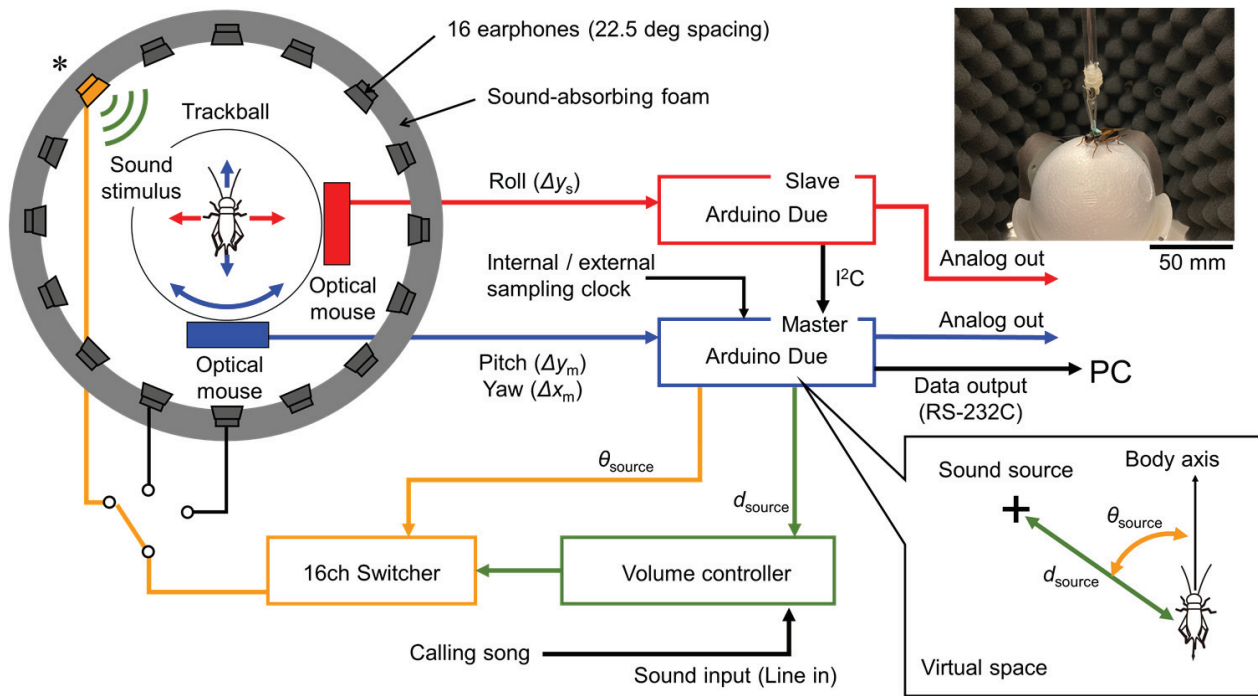
Fig. 1. Diagram of the auditory virtual reality (VR) system. The cricket was placed on a trackball surrounded by earphones that were embedded in sound-absorbing foam (inset). When the sound source is at $\theta_{source} = -\pi/4$ (front left, $-45°$) in a virtual space, one earphone at the corresponding angle (earphone indicated by an asterisk, covering $\theta_{source} = -45 \pm 11.25°$) outputs the sound. The sampling clock of the system can also be supplied externally via an I/O port on Arduino to synchronize other equipment, such as a data logger for physiological recording and video capturing. The two digital-analog output ports of Arduino can also be used for recording with an external data logger to correct data from multiple devices.
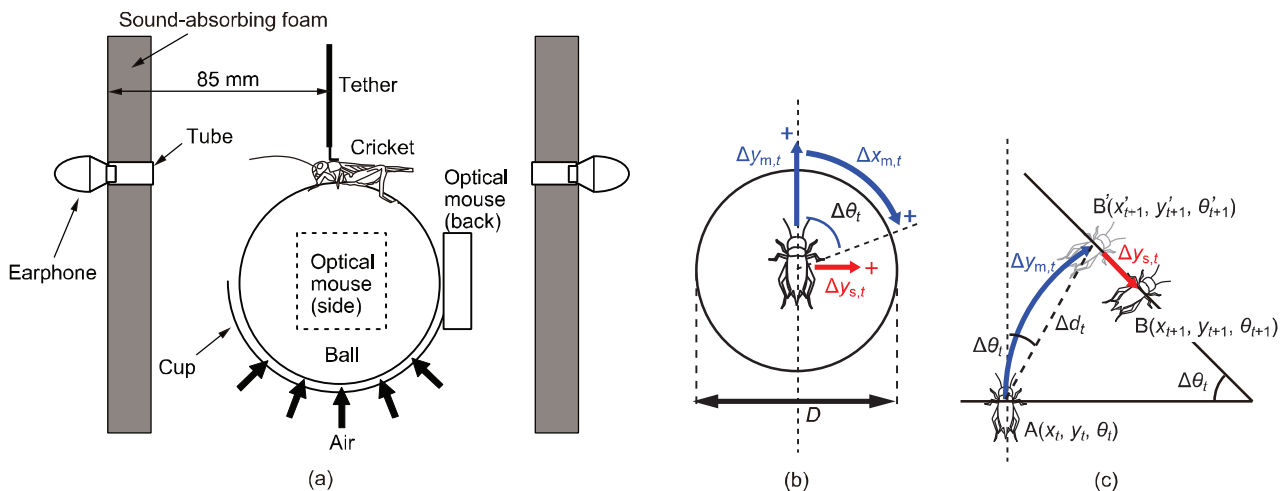

Fig. 2. Trackball system. (a) Side view of the trackball. (b) Definitions of variables for trackball movements. $\Delta x_{m,t}$, $\Delta y_{m,t}$, $\Delta y_{s,t}$, and $\Delta \theta_t$ indicate left-right rotation, fore-aft movement, left-right deviation, and change in cricket's heading, respectively. The arrows with plus signs indicate positive directions. D indicates the diameter of the ball. (c) Definition of variables for cricket locomotion in a virtual space.

(**Fig. 2(c)**). The change in the cricket's heading ($\Delta \theta_t$) was calculated as

$$\Delta \theta_t = \frac{2\Delta x_{m,t}}{D}, \quad \ldots \ldots \ldots \ldots (1)$$

where D is the diameter of the ball (75 mm). Subsequently, the position and heading at B′ were calculated

as follows:

$$x'_{t+1} = x_t + \Delta d_t \sin \Delta \theta_t, \quad \ldots \ldots \ldots (2)$$

$$y'_{t+1} = y_t + \Delta d_t \cos \Delta \theta_t, \quad \ldots \ldots \ldots (3)$$

$$\theta'_{t+1} = \theta_t + \Delta \theta_t, \quad \ldots \ldots \ldots \ldots (4)$$

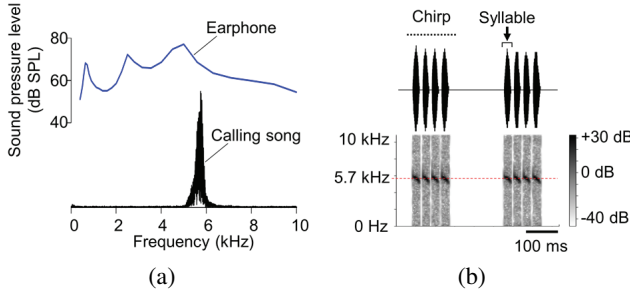where Δd_t is the direct distance between A and B′, and is

Fig. 3. Male calling song for auditory stimulus. (a) Frequency response of the earphone (upper, dB SPL) and the fast Fourier transform spectrum of the calling song (lower). (b) Waveform and a sonogram of the calling song.

calculated as follows:

$$\Delta d_t = 2\frac{\Delta y_{m,t}}{\Delta \theta_t}\sin\frac{\Delta\theta_t}{2} \quad (\Delta x_{m,t} \neq 0), \quad . \quad . \quad . \quad (5)$$

$$\Delta d_t = \Delta y_{m,t} \quad\quad\quad (\Delta x_{m,t} = 0). \quad . \quad . \quad . \quad (6)$$

Finally, lateral deviation was added to move the position from B′ to B (**Fig. 2(c)**) as follows:

$$x_{t+1} = x'_{t+1} + \Delta y_{s,t}\cos\Delta\theta_t, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

$$y_{t+1} = y'_{t+1} + \Delta y_{s,t}\sin\Delta\theta_t, \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$\theta_{t+1} = \theta'_{t+1}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

2.2.3. Auditory Stimulator

The speaker array consisted of eight stereo earphones (5 mm in diameter; H310, Monotaro, Osaka, Japan; the frequency response was shown in **Fig. 3(a)**), which allowed the production of auditory stimulus from 16 different directions (spacing 22.5°, each side of the stereo earphones was operated independently). Each earphone was fixed through a hole on an acrylic sheet 0.5 mm thick, and a sound-absorbing foam was attached to the sheet to reduce undesirable sound reflection. A plastic tube (6 mm in inner diameter) was then attached to the earphone to guide sound through the foam. Finally, the sheet was rolled into a cylinder (170 mm in diameter, the distance between a cricket (the center) and each earphone was 85 mm) and covered with a trackball.

The auditory stimulus was the conspecific calling song of *G. bimaculatus*. The calling song, which consisted of four syllables (peak frequency at 5.7 kHz, each 30 ms in duration), repeated every 275 ms (**Fig. 3(b)**). We also prepared an artificial calling song with a peak syllable frequency at 4.8 kHz to confirm whether female crickets in VR selectively responded to the calling song at the syllable frequency of 5.7 kHz. The 4.8 kHz syllable is attractive for European *G. bimaculatus*, while it is not attractive for Japanese ones [27]. These songs were prepared as Waveform Audio File format and played on a personal computer. The playback sound was input to a volume controller (PGA2311, Texas Instruments, Dallas, TX, USA).

The relationship between sound pressure levels (dB SPL) and the distance from a sound source was imple-

mented in the microcontroller. The sound pressure level declines with increasing distance from a male cricket [28]. We approximated the attenuation of sound pressure depending on the distance by a linear relationship and set the sound pressure level of 75 dB at 0 mm and 45 dB at 500 mm. A similar attenuation was obtained from the measurement of male calling song ($N = 5$) with a noise dosimeter (SD-2200, FUSO, Tokyo, Japan) in a sound-attenuated room. The volume controller adjusted the sound pressure of the calling song depending on the distance between the cricket (x_t, y_t) and the sound source $(0,0)$ in the virtual space. The distance from the source ($d_{source,t}$) was calculated as follows:

$$d_{source,t} = \sqrt{x_t{}^2 + y_t{}^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Then, the sound pressure level $P_{sound,t}$ at $d_{source,t}$, from the sound source was determined. We set the size of the sound source because crickets rarely experienced the maximum sound pressure levels of 75 dB if only a single point $(0,0)$ was the sound source. The size of the source at 10 mm in diameter was set, and $P_{sound,t}$ was calculated as follows:

$$P_{sound,t} = -0.0612d_{source,t} + 75.6$$
$$(10 < d_{source,t} \leq 500), \quad . \quad . \quad . \quad (11)$$

$$P_{sound,t} = 75.0 \quad (0 \leq d_{source,t} \leq 10). \quad . \quad . \quad . \quad (12)$$

The output of the volume controller was then sent to a 16-channel switcher consisting of eight 2-channel phototriac couplers (AQW210, Panasonic, Osaka, Japan). Once the heading was determined relative to the sound source ($\theta_{source,t}, -\pi \sim \pi$), the calling song was made to play from one of the 16 earphones at the corresponding angle. The $\theta_{source,t}$ was calculated as follows:

$$\theta_{source,t} = -\frac{x_t\sin\theta_t + y_t\sin\theta_t}{\sqrt{x_t{}^2 + y_t{}^2}}. \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$\theta_{source,t}$ was multiplied by -1 when the sound source was on the left side relative to the body axis. In the case of **Fig. 1**, where the sound source is at $\theta_{source} = -\pi/4$ (front left, $-45°$) in a virtual space, one earphone covering $\theta_{source} = -45\pm11.25°$ outputs the sound. The closed-loop control of the sound stimulus was performed at the same frequency as the sampling rate (100 Hz) with a time delay of a single sampling time (10 ms).

2.2.4. Sound Calibration

Since the sound propagates through the air between each earphone and the cricket, we calibrated the relationship between the volume controller input and sound pressure level at the center of the ball (85 mm away from earphones). We entered a pure tone of 5 kHz to the volume controller as a test tone, and the sound at the center of the ball was sampled and amplified with a microphone (type 4939, Brüel & Kjær, Nærum, Denmark) and microphone amplifier (type 2670, Brüel & Kjær, Nærum, Denmark). The microphone was mounted downward 5 mm above the ball. The sound pressure level at 5 kHz was then analyzed using an FFT analyzer (CF-5210, Ono Sokki, Yokohama, Japan).
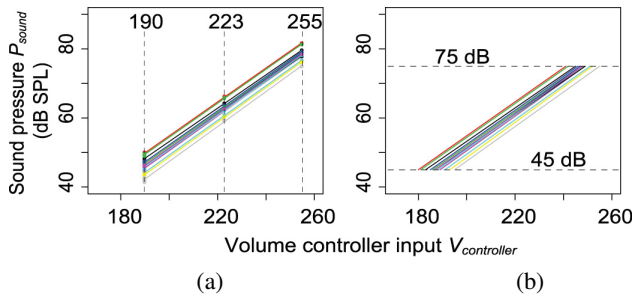
(a)   (b)

Fig. 4. Calibration of sound volume. (a) Measurements of sound pressure levels of all 16 earphones relative to the volume controller input (190, 223, and 255). (b) Relationship between the range of sound pressure levels (45–75 dB) for auditory stimulus and the volume controller input.

The volume controller could adjust the sound volume through 256 steps (volume controller input: 0–255). The sound pressure of the tone from each earphone was measured. The volume controller inputs were set at 190, 223, and 255. The measurement indicated that the relationship between the volume controller input and sound pressure level was linear (**Fig. 4(a)**), and the individual difference of earphones could be calibrated by fitting individual lines. We determined the volume controller input ($V_{controller,t}$) to reproduce the target sound pressure level in i-th earphone as follows:

$$V_{controller,t} = \frac{P_{sound,t} - \beta_i}{\alpha_i} \quad (45 \leq P_{sound,t} \leq 75), \quad . \ (14)$$

where α_i and β_i are the slope and intercept of the fitting line in i-th earphone.

Based on the linear relationship of the sound attenuation, α_i and β_i relative to the calling song were estimated and the controller volume range was individually determined to limit the actual sound pressure at the cricket's position to a range of 45–75 dB (**Fig. 4(b)**).

2.3. Experimental Condition

The virtual arena was an unlimited two-dimensional space (**Fig. 5(a)**). The sound source was set at $(0,0)$ and the maximum sound pressure level (75 dB) was presented when the cricket's position was within 10 mm from the source. The starting position was 500 mm away from the source $(0,-500)$, and the initial heading was toward the source ($\theta_0 = \theta_{source,0} = 0$).

Two experiments were conducted using the VR. One was conducted to evaluate the capability of the crickets to localize the sound source in a VR environment ($N = 16$, each cricket performed a single trial for each condition). The conditions were as follows: **Fig. 6(a)** no sound stimuli, **(b)** conspecific calling song (syllable frequency, 5.7 kHz), **(c)** artificial calling song (4.8 kHz), and **(d)–(f)** conspecific calling song at constant sound pressure levels of 45, 60, and 75 dB, respectively (**Figs. 6** and **7**). The no-sound stimuli in **Fig. 6(a)** were recorded approximately 1 min after the cricket was attached to the VR system; therefore, locomotor activities were different among
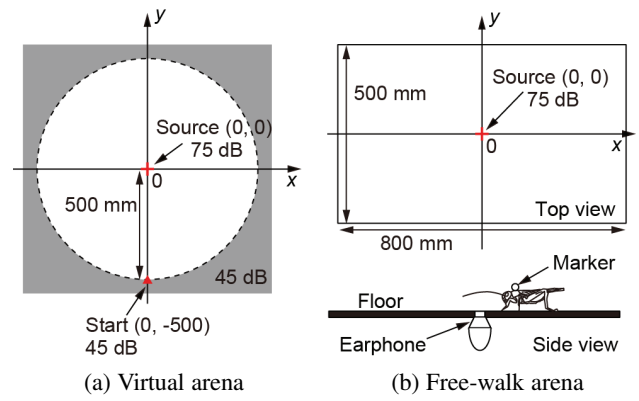

(a) Virtual arena   (b) Free-walk arena

Fig. 5. Experimental arenas. (a) Arena for virtual reality. The triangle and cross signs indicate the cricket starting position and the sound source position, respectively. The sound pressure level increased linearly as the cricket approached the sound source, while it remained constant at 45 dB when the cricket was more than 500 mm away (gray area). (b) Arena for the free-walk experiment. The start position was not determined because crickets were allowed to walk around the wall continuously.

individuals. The sound pressure level increased linearly from 45 dB to 75 dB as the cricket approached 500 mm to 10 mm from the sound source (Eqs. (11) and (12)), while it remained constant at 45 dB when the cricket was more than 500 mm away in **(b)** and **(c)**. The conspecific calling song (syllable frequency, 5.7 kHz) at a constant sound pressure level in all positions was used as the auditory stimulus in the experiments **(d)–(f)**.

The other experiment was conducted to demonstrate the capability of the VR system to present arbitrary sound environments ($N = 17$, each cricket performed a single trial for each condition). In addition to **Fig. 8(a)** no sound stimuli and **(b)** the conspecific calling song (syllable frequency, 5.7 kHz) with increasing sound pressure levels as the cricket approached the sound source (Eqs. (11) and (12)), we conducted **(c)** calling song (syllable frequency, 5.7 kHz) with decreasing sound pressure levels as the cricket approached the sound source (**Fig. 8**).

We analyzed the trajectory data for 120 s from the beginning of the auditory stimulus. Individuals that did not move for 120 s were omitted from the analysis (the number of analyzed data for each condition is shown in the corresponding figure legends).

We also conducted phonotaxis experiments under the free-walk condition to compare with phonotaxis in VR. We focused on the behavior around a sound source without physical cues, because crickets could not experience physical contact at the sound source in VR. The free-walk arena was a plastic box sized 500 mm (W) × 800 mm (L) × 200 (H) mm (**Fig. 5(b)**). The auditory stimulus was given through a single earphone embedded in the floor in the center of the box to eliminate the physical cues of a sound source. The sound pressure was adjusted to 75 dB at the center of the box. The behavior was captured by a USB camera with infrared illumination. A white
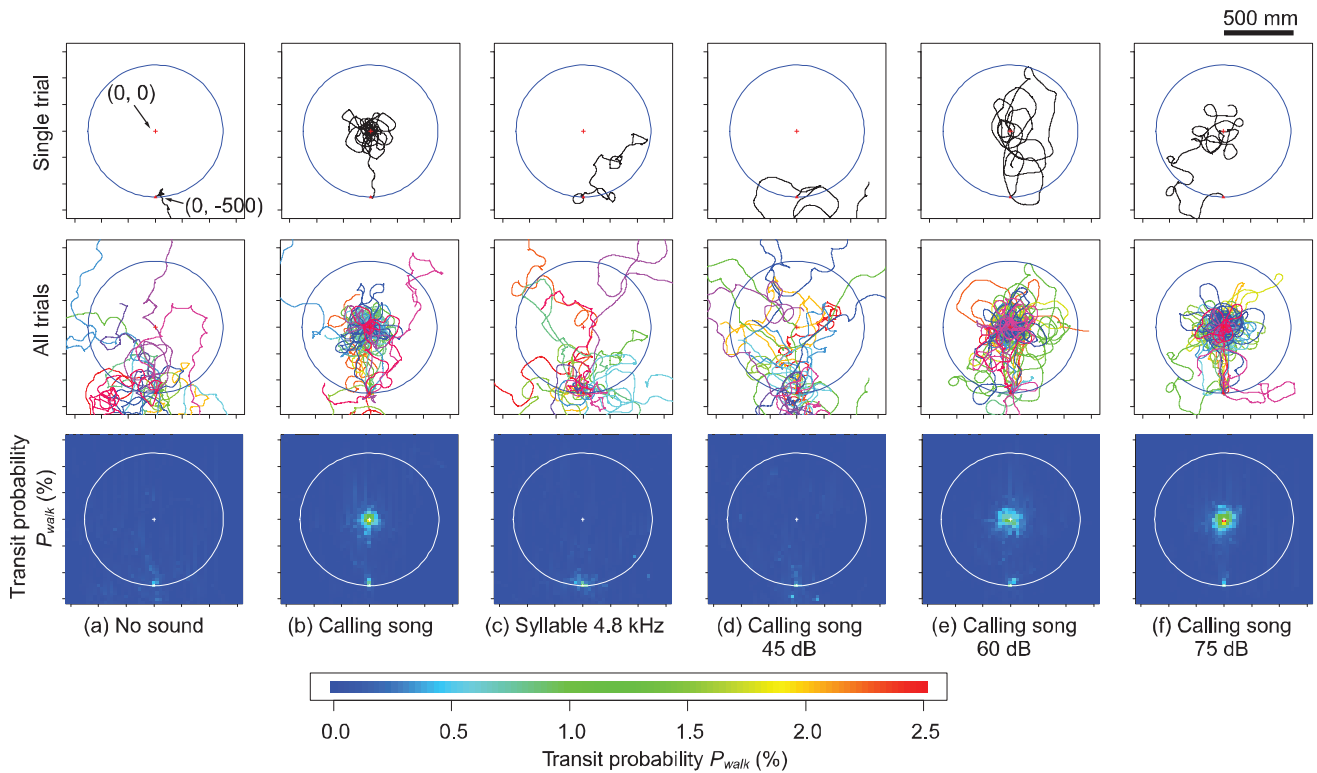
500 mm

(a) No sound (b) Calling song (c) Syllable 4.8 kHz (d) Calling song 45 dB (e) Calling song 60 dB (f) Calling song 75 dB

Transit probability P_{walk} (%)

0.0   0.5   1.0   1.5   2.0   2.5

Transit probability P_{walk} (%)

Fig. 6. Trajectories and density plots of cricket locomotion in the virtual space (color online). For each experimental condition, the top panel shows a representative trajectory, the middle panel shows all the trajectories individually colored, and the bottom panel shows a density map of the mean transit probability P_{walk} of all trials. The start position was $(0, -500)$ and the source was the center $(0, 0)$ of the arena. Experimental conditions were (a) no stimulus ($N = 13$), (b) conspecific calling song ($N = 14$), (c) artificial calling song with a peak syllable frequency of 4.8 kHz ($N = 15$). The sound pressure levels of the sound stimuli were modulated based on the position of the cricket (Eqs. (11) and (12)). (d)–(f) Conspecific calling song at constant sound pressure levels of (d) 45 dB, (e) 60 dB, and (f) 75 dB ($N = 15$ each).
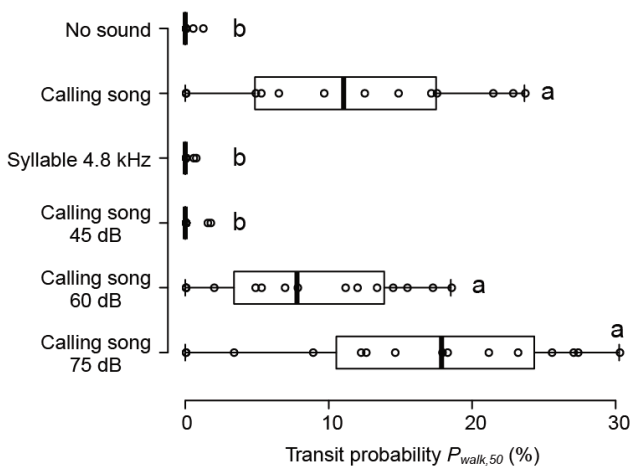

Fig. 7. Comparison of transit probability within 50 mm from the source. The left and right sides of the box indicate the first and third quartiles, and the bar represents the median. The whiskers indicate the $1.5 \times$ interquartile range. Plots with different letters indicate significant differences between them ($p < 0.05$, Steel-Dwass test).

polystyrene ball (4 mm in diameter) was attached to the dorsal thorax of the crickets to track their position with software (Dipp Motion V, Ditect, Tokyo, Japan). We used 12 females and analyzed each trajectory for 60 s from the beginning of the auditory stimulus.

2.4. Data Analysis

To visualize the trajectory characteristics under each experimental condition, we calculated the transit probability of crickets during walking (P_{walk}) every 25×25 mm mesh as follows:

$$P_{walk} [\%] = \frac{N_{walk}}{N_{sample}} \times 100, \quad \ldots \ldots \ldots (15)$$

where N_{walk} is the number of samples with crickets that were walking in each mesh and N_{sample} is the number of samples for analysis (12,000 samples for 120 s). We calculated data during walking because it was difficult to evaluate whether crickets responded to sound stimuli from the data when they stopped. The "walking" was defined as moving at a speed > 10 mm/s, which included the walking speed of female crickets during phonotaxis on a trackball (2–9 cm/s) [29]. The transit probability P_{walk} is visualized as a density map in **Figs. 6**, **8**, and **9**. Source localization was defined when crickets arrived within 50 mm from
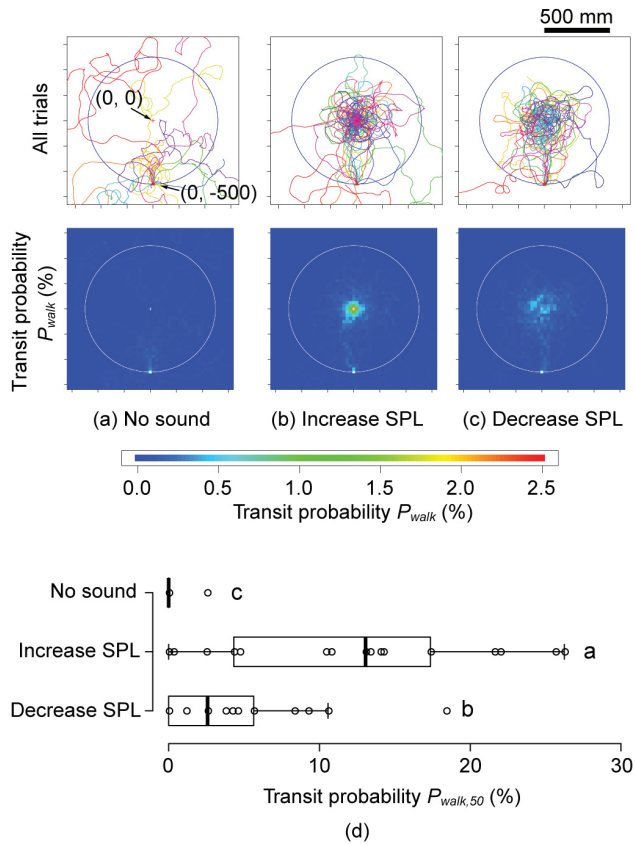
Fig. 8. Manipulation of the sound pressure gradient (color online). For each experimental condition, the top panel represents all the trajectories colored individually while the bottom panel is the density map of the mean transit probability P_{walk} of all trials. Experimental conditions were (a) no stimulus ($N = 17$), (b) conspecific calling song with increasing sound pressure as the insect approached ($N = 17$), and (c) calling song with decreasing gradient as the insect approached ($N = 17$). In (c), the sound pressure level decreased linearly from 75 dB to 45 dB as the cricket approached 500 mm to 10 mm from the sound source, while it remained constant at 75 dB when the cricket was more than 500 mm away. (d) Comparison of transit probability within 50 mm from the source. Plots with different letters indicate significant differences between them ($p < 0.05$, Steel-Dwass test).

the sound source, and then quantified the performance of source localization as the transit probability within 50 mm from the sound source ($P_{walk,50}$):

$$P_{walk,50}\ [\%] = \frac{N_{walk,50}}{N_{sample}} \times 100, \quad \ldots \ldots \quad (16)$$

where $N_{walk,50}$ is the number of samples in which the cricket was within 50 mm from the source while walking. Statistical tests for comparing $P_{walk,50}$ between experimental conditions were performed using the Steel-Dwass test for multiple groups or the Wilcoxon signed-rank test for two groups at a significance level of $p < 0.05$.
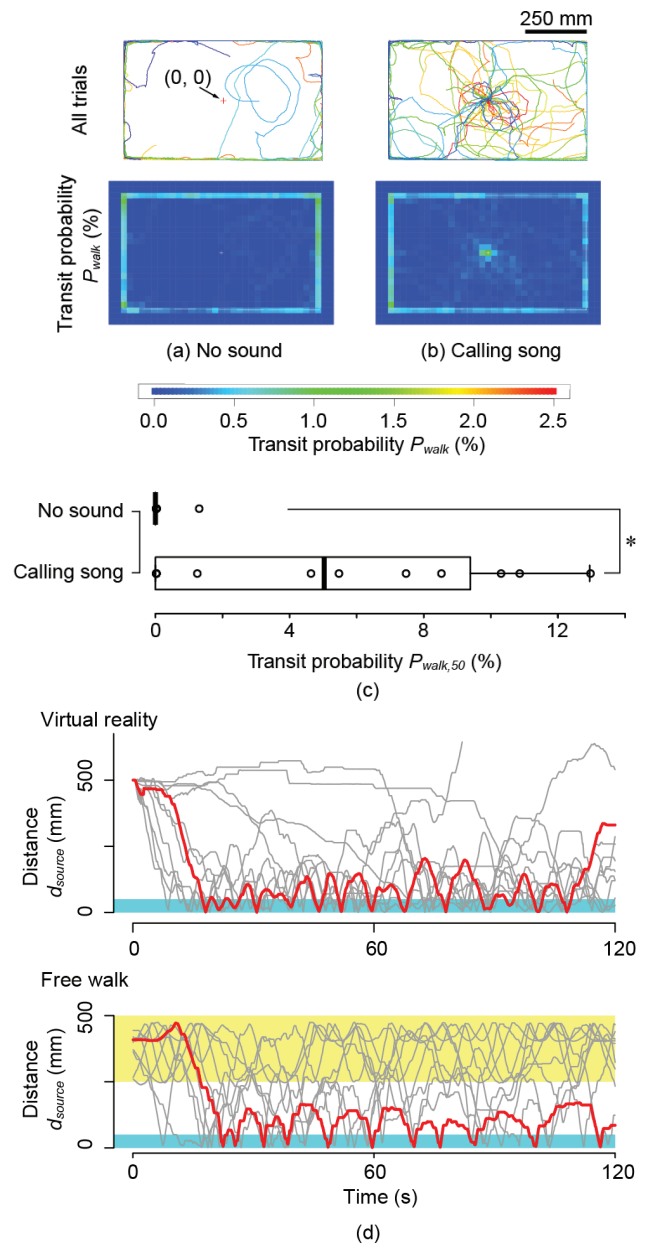

Fig. 9. Phonotaxis during free-walk (color online). For each experimental condition, the top panel indicates all trajectories colored individually, while the bottom panel shows the density map of the mean transit probability P_{walk} of all trials. Experimental conditions were (a) no stimulus ($N = 12$) and (b) conspecific calling song ($N = 12$). (c) Comparison of transit probability within 50 mm from the source. An asterisk indicates a significant difference ($p < 0.05$, Wilcoxon signed-rank test). (d) Time course of the distance from the source d_{source} in VR (upper, $N = 11$) and free-walk (lower, $N = 10$) for 120 s during calling song stimulus. Trials that approached within 50 mm from the source (lower shaded area) were displayed. Representative "back and forth" behavior is indicated by bold lines. The upper shaded area ($d_{source} = 250$–500 mm) in the free-walk setup indicates the range of d_{source} where crickets followed the wall.

3. Results

3.1. Localization of the Sound Source

Without auditory stimuli, the crickets walked around the start position in random directions (**Fig. 6(a)**). When the conspecific calling song was played, 11 of the 14 crickets walked within 50 mm of the sound source (**Fig. 6(b)**). Furthermore, crickets overran the source, went around the source, and approached the source again. This "back and forth" pattern was specific to the conspecific calling song and increased transit probability near the source. The crickets did not show this behavior when the artificial calling song with different syllable frequencies was presented (**Fig. 6(c)**). The transit probability $P_{walk,50}$ under the conspecific calling song was significantly higher than those with no stimulus and those presented with 4.8 kHz syllables ($p < 0.05$, Steel-Dwass test; **Fig. 7**). These results indicated that the tethered crickets responded to the conspecific calling song and localized the sound source in VR.

We next presented the conspecific calling song at three constant sound pressures (45, 60, and 75 dB) to investigate the behavioral threshold of sound pressure for successful phonotaxis (**Figs. 6(d)–(f)**). The crickets did not show "back and forth" behavior around the source at 45 dB, although they tended to walk toward the source. The crickets approached the source and performed the "back and forth" behavior when the calling song was presented at constant sound pressures of 60 dB and 75 dB. The median of the transit probability $P_{walk,50}$ increased as the sound pressure level increased (45 dB, 0 [0, 1.88]; 60 dB, 19.2 [11.5, 38.0]; 75 dB, 35.5 [26.8, 45.2]; median [interquartile range]), and the values at 60 dB and 75 dB were significantly higher than those at 45 dB ($p < 0.05$, Steel-Dwass test; **Fig. 7**), which were not significantly different from those with no stimulus and those presented with 4.8 kHz syllables ($p > 0.05$). These results indicate that presenting the calling song at high sound pressures is an essential sensory cue for accurate orientation.

The accuracy of sound source localization depending on sound pressure levels was also tested using a reversed gradient of pressure levels (**Fig. 8**). In this setup, the pressure level of the calling song was set at 75 dB at the start position, decreasing to 45 dB as the cricket approached the source. The crickets performed sound localization in both normal and reversed pressure gradients (**Figs. 8(a)–(c)**). The trajectories and transit probability around the source in the reversed condition were sparse compared to those in the normal condition. The transit probabilities $P_{walk,50}$ of the normal gradient were significantly higher than those of the reversed gradient ($p < 0.05$, Steel-Dwass test; **Fig. 8(d)**), although both were significantly higher than those without auditory stimulus ($p < 0.05$). These results also indicate that a sufficient sound pressure level is required for accurate localization of the source.

3.2. Comparison with Free-Walk Phonotaxis

The "back and forth" behavior observed around the sound source can be regarded as a sign of successful localization as it was only observed when the conspecific calling song was presented at sufficient sound pressure levels. However, such behavior was not observed in the free-walk condition as each trial was terminated when a freely walking cricket reached the sound source (such as a speaker, [30]) and made physical contact with it. To confirm whether the "back and forth" behavior can be observed in the free-walk condition, we conducted a free-walk experiment with a sound source embedded on the floor.

The crickets followed the wall of the arena and rarely went through the center without sound stimulus (**Fig. 9(a)**). When the calling song was presented, crickets approached the sound source located at the center of the arena (**Fig. 9(b)**) [a]; $P_{walk,50}$ was significantly higher with than without stimulus ($p < 0.05$, Wilcoxon signed-rank test; **Fig. 9(c)**). Furthermore, the crickets performed "back and forth" behavior around the source, which can be expressed as the continuous d_{source} oscillation near the source (**Fig. 9(d)**).

4. Discussion

4.1. Auditory Navigation in VR

The crickets performed phonotaxis to the conspecific calling song and localized the sound source in VR. The performance of localization, as quantified by $P_{walk,50}$ increased with increasing sound pressure levels. This result was consistent with the trackball (open-loop) and free-walk studies of cricket phonotaxis [29–31], indicating that high sound pressure provides a cue for the exact direction of the sound source. In contrast, the crickets were rarely able to localize the source at a constant sound pressure level of 45 dB. The sound pressure of 45 dB is closed to the thresholds for activating the auditory interneurons in crickets [32, 33]. However, the sound pressure level of 45 dB was that at the starting point ($d_{source} = 500$) of a normal sound pressure gradient setup in which most crickets localized the sound source (**Figs. 6(b)** and **8(b)**). Therefore, a sound pressure level at 45 dB was not below the behavioral threshold for phonotaxis but instead approximated the threshold. Although a sound pressure of 45 dB is insufficient to localize the sound source, it might prompt the cricket to find a preferable orientation or location where it can receive higher sound pressures.

The relationship between the distance from the source and sound pressure levels differs among environments [28, 34]. We set the minimum sound pressure level of 45 dB at 500 mm from the source based on the measurement in the sound-attenuated room, however, it has been reported that the sound pressure level was above 50 dB at 2 m from a male cricket in an area of scrubland [28]. The proposed VR system allows us to apply arbitrary profiles of the sound pressure levels, therefore,

our system will be helpful in understanding how female crickets utilize the instantaneous change of sound pressure level for phonotaxis.

The "back and forth" behavior observed in VR was also observed in the free-walk condition (**Fig. 9**), though the comparisons were limited to be qualitative. Therefore, this behavior could occur when there are no additional sensory cues at the sound source. The "back and forth" pattern would be a repetition of targeted steering driven by direct auditory information rather than search behavior [35]. Although it is unnatural that there is no male cricket at the sound source, the phonotaxis observed in tethered crickets in auditory VR would be comparable to those in free-walk conditions.

Taken together, we established an experimental procedure to elicit cricket phonotaxis and sound source localization in two-dimensional VR.

4.2. Limitations and Improvements

There are two ways to present auditory stimuli to a tethered animal on a spherical treadmill from any direction. One is to install multiple speakers, and the other is to rotate one speaker around the animal. We adopted the former to achieve a high-speed response without excess mechanical noise. For example, the direction of the sound must change up to $180°$ within a single sampling time when the cricket goes through the sound source. With the installation of multiple speakers, achieving a fast angular velocity is only possible by selecting a speaker with a predetermined angle.

The limitations of using multiple speakers include the constraints of angular resolution, the need for calibration of the sound pressure levels of all speakers, and the temporal gap during speaker switching. The angular acuity of hearing in *G. bimaculatus* is high enough to discriminate an angular change of $1°$ in the frontal region ($±30°$) [36], while the inter-speaker angle of our system was $22.5°$. We have not evaluated the effects of the limited angular resolution on phonotaxis; this should be done by increasing or decreasing the number of speakers. In a previous VR setup using a two-speaker system, the volume of the left and right speakers was adjusted to simulate the angle of incidence of sound between the two speakers [23]. This stereo audition is a plausible alternative to using an array of speakers; however, the presentation of the sound from a single source is preferable to analyze the directional control of phonotaxis as crickets can detect phase differences between sounds [37]. Therefore, adjusting the volume of two adjacent speakers on an array [24] would be a solution to interpolate the inter-speaker angle. The calibration was manually performed in this study, which can be automated by implementing feedback control on the sound pressure levels. The temporal gap in sound stimulation during switching was not noticeable when the sound played was a calling song consisting of intermittent chirps and syllables; however, it can be smoothened by delaying the end of the sound that is played on the selected speakers.

4.3. Future Directions

Virtual reality allows us to investigate the neural mechanisms behind spatial navigation in animals. The mechanism of locomotion in cricket phonotaxis consists of simple reactive steering [38]. However, a recent behavioral study indicated that their behavior consists of multiple strategies that are executed depending on the insect's internal state and presence of external stimuli [30]. This implies that the higher center for motor control in the brain is involved in phonotaxis. Therefore, we targeted the central complex, which is the brain region responsible for navigational tasks in insects, to apply neurophysiological techniques during phonotaxis. By combining auditory VR with intracellular recording, we successfully identified several neurons encoding locomotor commands [39].

In contrast, studies on mouse navigation have reported that different behaviors were observed between the real and virtual worlds. One reason for this discrepancy is believed to be the limited number of sensory modalities used for sensory feedback [40]. Animals are surrounded by a rich source of sensory inputs in nature; however, multisensory integration is important for selecting behaviors that lead to efficient navigation. Therefore, multisensory VR is required to bridge the gap between the behaviors observed in the virtual and real worlds [19, 21]. In our preliminary study, auditory VR was combined with four flat-panel displays surrounding the trackball to add visual feedback [41]. Multimodal VR combined with neurophysiological techniques is expected to further elucidate the neural mechanisms of insect navigation in more natural environments.

References:

[1] B. Webb, "Robots in invertebrate neuroscience," Nature, Vol.417, No.6886, pp. 359-363, 2002.

[2] B. Webb, "Robots with insect brains," Science, Vol.368, No.6488, 244, 2020.

[3] F. Huber, T. E. Moore et al., "Cricket Behavior and Neurobiology," Cornell University Press, 1989.

[4] H. H. Wilson, T. Mito et al., "The Cricket as a Model Organism," Springer Japan, 2017.

[5] B. Hedwig, "Pulses, patterns and paths: neurobiology of acoustic behaviour in crickets," J. Comp. Physiol. A, Vol.192, No.7, pp. 677-689, 2006.

[6] B. Webb, "Using robots to model animals: A cricket test," Robot. Auton. Syst., Vol.16, Nos.2-4, pp. 117-134, 1995.

[7] B. Webb, R. R. Harrison et al., "Sensorimotor control of navigation in arthropod and artificial systems," Arthropod Struct. Dev., Vol.33, No.3, pp. 301-329, 2004.

[8] B. Hedwig, "Trackball Systems for Analysing Cricket Phonotaxis," H. H. Wilson, T. Mito et al. (Eds.), "The Cricket as a Model Organism," Springer Japan, pp. 303-312, 2017.

[9] T. Weber, J. Thorson et al., "Auditory-behavior of the cricket. 1. Dynamics of compensated walking and discrimination paradigms on the Kramer treadmill," J. Comp. Physiol., Vol.141, No.2, pp. 215-232, 1981.

[10] B. Hedwig, "A highly sensitive opto-electronic system for the measurement of movements," J. Neurosci. Meth., Vol.100, Nos.1-2, pp. 165-171, 2000.

[11] M. Oe and H. Ogawa, "Neural basis of stimulus-angle-dependent motor control of wind-elicited walking behavior in the cricket *Gryllus bimaculatus*," PLOS ONE, Vol.8, No.11, e80184, 2013.

[12] H. Bohm, K. Schildberger et al., "Visual and acoustic course control in the cricket *Gryllus bimaculatus*," J. Exp. Biol., Vol.159, pp. 235-248, 1991.

[13] H. Haberkern and B. Hedwig, "Behavioural integration of auditory and antennal stimulation during phonotaxis in the field cricket *Gryllus bimaculatus*," J. Exp. Biol., Vol.219, Issue 22, pp. 3575-3586, 2016.

[14] E. Staudacher and K. Schildberger, "Gating of sensory responses of descending brain neurones during walking in crickets," J. Exp. Biol., Vol.201, No.4, pp. 559-572, 1998.

[15] K. Schildberger, "Behavioral and neuronal mechanisms of cricket phonotaxis," Experientia, Vol.44, No.5, pp. 408-415, 1988.

[16] E. Roth, S. Sponberg et al., "A comparative approach to closed-loop computation," Curr. Opin. Neurobiol., Vol.25, pp. 54-62, 2014.

[17] D. A. Dombeck and M. B. Reiser, "Real neuroscience in virtual worlds," Curr. Opin. Neurobiol., Vol.22, No.1, pp. 3-10, 2012.

[18] H. Naik, R. Bastien et al., "Animals in virtual environments," IEEE Trans. Vis. Comput. Graph., Vol.26, No.5, pp. 2073-2083, 2020.

[19] B. A. Radvansky and D. A. Dombeck, "An olfactory virtual reality system for mice," Nat. Commun., Vol.9, No.1, 839, 2018.

[20] C. A. Hernandez-Reyes, S. Fukushima et al., "Identification of exploration and exploitation balance in the silkmoth olfactory search behavior by information-Theoretic modeling," Front. Comput. Neurosci., Vol.15, 629380, 2021.

[21] P. K. Kaushik, M. Renz et al., "Characterizing long-range search behavior in Diptera using complex 3D virtual environments," Proc. Natl. Acad. Sci., Vol.117, No.22, pp. 12201-12207, 2020.

[22] A. Yamashita, N. Ando et al., "Closed-loop locomotion analyzer for investigating context-dependent collision avoidance systems in insects," J. Robot. Soc., Vol.27, No.7, pp. 704-710, 2011.

[23] D. Santos-Pata, A. Escuredo et al., "Insect behavioral evidence of spatial memories during environmental reconfiguration," V. Vouloutsi, J. Halloy et al. (Eds.), "Biomimetic and Biohybrid Systems," Springer International Publishing, pp. 415-427, 2018.

[24] A. Funamizu, B. Kuhn et al., "Neural substrate of dynamic Bayesian inference in the cerebral cortex," Nat. Neurosci., Vol.19, 1682, 2016.

[25] S. Emoto, N. Ando et al., "Insect-controlled robot – evaluation of adaptation ability," J. Robot. Mechatron., Vol.19, No.4, pp. 436-443, 2007.

[26] N. Ando, S. Emoto et al., "Insect-controlled robot: a mobile robot platform to evaluate the odor-tracking capability of an insect," J. Vis. Exp., Vol.118, No.118, e54802, 2016.

[27] A. Miyashita, H. Kizaki et al., "No effect of body size on the frequency of calling and courtship song in the two-spotted cricket, *Gryllus bimaculatus*," PLOS ONE, Vol.11, No.1, e0146999, 2016.

[28] L. W. Simmons, "The calling song of the field cricket, *Gryllus bimaculatus* (deGeer) – Constraints on transmission and its role in intermale competition and female choice," Anim. Behav., Vol.36, pp. 380-394, 1988.

[29] B. Hedwig and J. E. A. Poulet, "Mechanisms underlying phonotactic steering in the cricket *Gryllus bimaculatus* revealed with a fast trackball system," J. Exp. Biol., Vol.208, No.5, pp. 915-927, 2005.

[30] N. Hommaru, H. Shidara et al., "Internal state transition to switch behavioral strategies in cricket phonotaxis," J. Exp. Biol., Vol.223, Issue 22, jeb229732, 2020.

[31] J. F. Stout, C. H. Dehaan et al., "Attractiveness of the male *Acheta domesticus* calling song to females. 1. Dependence on each of the calling song features," J. Comp. Physiol., Vol.153, No.4, pp. 509-521, 1983.

[32] G. Horseman and F. Huber, "Sound localisation in crickets," J. Comp. Physiol. A, Vol.175, No.4, pp. 399-413, 1994.

[33] K. Kostarakos, M. Hartbauer et al., "Matched filters, mate choice and the evolution of sexually selected traits," PLOS ONE, Vol.3, No.8, e3005, 2008.

[34] S. Hirtenlehner and H. Romer, "Selective phonotaxis of female crickets under natural outdoor conditions," J. Comp. Physiol. A, Vol.200, No.3, pp. 239-250, 2014.

[35] F. Steinbeck, A. Adden et al., "Connecting brain to behaviour: a role for general purpose steering circuits in insect orientation?," J. Exp. Biol., Vol.223, Issue 5, jeb212332, 2020.

[36] S. Schoneich and B. Hedwig, "Hyperacute directional hearing and phonotactic steering in the cricket (*Gryllus bimaculatus* deGeer)," PLOS ONE, Vol.5, No.12, e15141, 2010.

[37] K. M. Seagraves and B. Hedwig, "Phase shifts in binaural stimuli provide directional cues for sound localisation in the field cricket *Gryllus bimaculatus*," J. Exp. Biol., Vol.217, Issue 13, pp. 2390-2398, 2014.

[38] B. Hedwig and J. F. A. Poulet, "Complex auditory behaviour emerges from simple reactive steering," Nature, Vol.430, No.7001, pp. 781-785, 2004.

[39] K. Kai, H. Shidara et al., "Neural activity in the central complex of the crickets during phonotaxis," Abstr. Soc. Neurosci., 694.604, 2019.

[40] M. Minderer, C. D. Harvey et al., "Neuroscience: Virtual reality explored," Nature, Vol.533, No.7603, pp. 324-325, 2016.

[41] K. Makino, N. Ando et al., "Auditory-visual virtual reality for the study of multisensory integration in insect navigation," U. Martinez-Hernandez, V. Vouloutsi et al. (Eds.), "Biomimetic and Biohybrid Systems," Springer, pp. 325-328, 2019.

Supporting Online Materials:

[a] Experiment video of the VR system. https://sites.google.com/view/andolab-e/research/movies [Accessed March 18, 2021]

Name:
Noriyasu Ando

Affiliation:
Associate Professor, Department of Systems Life Engineering, Faculty of Engineering, Maebashi Institute of Technology

Address:
460-1 Kamisadori-cho, Maebashi, Gunma 371-0816, Japan
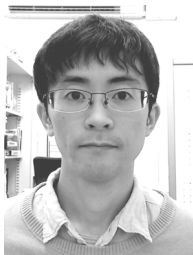**Brief Biographical History:**
2006-2016 Assistant Professor, The University of Tokyo
2016-2019 Project Lecturer, The University of Tokyo
2019- Associate Professor, Maebashi Institute of Technology
Main Works:
● "Insect-machine hybrid robot: closing loops with mobile robots," Current Opinion in Insect Science, Vol.42, pp. 61-69, 2020.
● "Toward robotic application of insect sex pheromone-searching behavior: lessons from silkmoths," Y. Ishikawa (Ed.), "Insect sex pheromone research and beyond – from molecules to robots –," Springer, pp. 287-313, 2020.
Membership in Academic Societies:
● The Robotics Society of Japan (RSJ)
● The Zoological Society of Japan (ZSJ)
● The Japanese Society for Comparative Physiology and Biochemistry (JSCPB)

Name:
Hisashi Shidara

Affiliation:
Assistant Professor, Department of Biochemistry, Graduate School of Medicine, Mie University

Address:
2-174 Edobashi, Tsu, Mie 514-8507, Japan
Brief Biographical History:
2017-2018 Postdoctoral Fellow, Hokkaido University
2018-2021 Research Fellow (PD), Japan Society for the Promotion of Science (JSPS)
2021- Assistant Professor, Mie University
Main Works:
• "Internal state transition to switch behavioral strategies in cricket phonotaxis behavior," J. of Experimental Biology, Vol.223, Issue 22, jeb229732, 2020.
• "Compartmentalized cGMP Responses of Olfactory Sensory Neurons in Caenorhabditis elegans," J. of Neuroscience, Vol.37, pp. 3753-3763, 2017.
Membership in Academic Societies:
• Society for Neuroscience (SfN)
• The Japan Neuroscience Society (JNS)
• The Biophysical Society of Japan (BSJ)

Name:
Naoto Hommaru

Affiliation:
Graduate Student, Graduate School of Life Science, Hokkaido University

Address:
Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan
Brief Biographical History:
2018-2020 Graduate Student, Hokkaido University
Main Works:
• "Internal state transition to switch behavioral strategies in cricket phonotaxis behavior," J. of Experimental Biology, Vol.223, Issue 22, jeb229732, 2020.

Name:
Hiroto Ogawa

Affiliation:
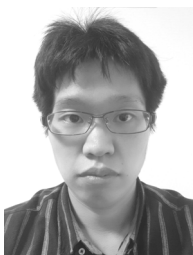Professor, Department of Biological Sciences, Faculty of Science, Hokkaido University

Address:
Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810, Japan
Brief Biographical History:
1997-2008 Assistant Professor, Saitama Medical School
2008-2016 Associate Professor, Hokkaido University
2016- Professor, Hokkaido University
Main Works:
• "Internal state transition to switch behavioral strategies in cricket phonotaxis behavior," J. of Experimental Biology, Vol.223, Issue 22, jeb229732, 2020.
• "Trade-off between motor performance and behavioural flexibility in the action selection of cricket escape behaviour," Sci. Rep., Vol.9, 18112, 2019.
• "Direction-specific adaptation in neuronal and behavioral responses of an insect mechanosensory system," J. Neurosci., Vol.35, Issue 33, pp. 11644-11655, 2015.
Membership in Academic Societies:
• Society for Neuroscience (SfN)
• The Japan Neuroscience Society (JNS)
• The Zoological Society of Japan (ZSJ)