

Field observations of ecoacoustic dynamics of a Japanese bush warbler using an open-source software for robot audition HARK

Original paper

Article history:

Received: 6 October 2017

Accepted: 30 April 2018

Published: 27 June 2018



Check for updates

*Correspondence:

RS: reiji@nagoya-u.jp

Peer review:

Double blind

Copyright:

© 2018 Suzuki et al. © This is an open access article distributed under the Creative Commons Attribution License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited and its authors credited.

Keywords:

sound source localization; microphone array; bird songs; HARK; virtual experience of soundscape; japanese bush warbler

Citation:

Suzuki R., Sumitani S., N., Matsubayashi S., Arita T., Nakadai K., and Okuno H. G. (2018). Field observations of ecoacoustic dynamics of a Japanese bush warbler using an open-source software for robot audition HARK. *Journal of Ecoacoustics*. 2: #EYAJ46.
<https://doi.org/10.22261/JEA.EYAJ46>

Reiji Suzuki^{1*}, Shinji Sumitani¹, Naren², Shiho Matsubayashi³, Takaya Arita¹, Kazuhiro Nakadai^{4,5}, Hiroshi G. Okuno⁶

¹Graduate School of Informatics, Nagoya University Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

²School of Informatics and Sciences, Nagoya University Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

³Graduate School of Engineering, Osaka University, 2-1, Yamadaoka, Suita, Osaka 565-0871, Japan

⁴Department of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology 11-20, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan

⁵Honda Research Institute Japan Co., Ltd. 8-1 Honcho, Wako-shi, Saitama 351-0188, Japan

⁶Graduate School of Creative Science and Engineering, Faculty of Science and Engineering, Waseda University 3-4-1, Okubo, Shinjyukuku, Tokyo 169-8555, Japan

Abstract

We report on a simple and practical application of HARK, an easily available and portable system for bird song localization using an open-source software for robot audition HARK, to a deeper understanding of ecoacoustic dynamics of bird songs, focusing on a fine-scaled temporal analysis of song movement — song type dynamics in playback experiments. We extended HARKBird and constructed a system that enables us to conduct automatic playback and interactive experiments with different conditions, with a real-time recording and localization of sound sources. We investigate how playback of conspecific songs and playback patterns can affect vocalization of two types of songs and spatial movement of an individual of Japanese bush-warbler, showing quantitatively that there exist strong relationships between song type and spatial movement. We also simulated the ecoacoustic dynamics of the singing behavior of the focal individual using a software, termed Bird song explorer, which provides users a virtual experience of acoustic dynamics of bird songs using a 3D game platform Unity. Based on experimental results, we discuss how our approach can contribute to ecoacoustics in terms of two different roles of sounds: sounds as tools and subjects.

Introduction

Acoustic interactions are important for understanding intra- and inter-specific communication in songbird communities. Recently, it is also pointed out that ornithological community can make significant contributions to the nascent field of soundscape ecology (Pijanowski et al., 2011) in terms of measurement, process and application of soundscape (Gasc et al., 2017). We believe that sound source localization using microphone arrays can greatly contribute to both research fields by providing fine-scale observation of spatially grounded acoustic events to gain fundamental knowledge to understand biological — environmental relationships in an acoustic space.

Songbirds are one of the driver species of such ecoacoustic dynamics, and males produce long vocalizations called songs to advertise their territory or attract females in a breeding season (Catchpole and Slater, 2008). For a deeper understanding of ecological functions and semantics (Daimon et al., 2017) of their songs, it is important to clarify the fine-scaled and detailed relationships among their characteristics of songs (e.g., song types) and behavioral contexts (e.g., movement, with/without neighboring rivals) of them.

Using microphone arrays is a promising approach to acoustically monitor wildlife that produce sounds (Blumstein et al., 2011) for this purpose, because it can provide directional or spatial information of recorded vocalization from recordings. There have been several empirical studies to spatially localize or estimating the direction of arrival (DOA) bird songs using multiple microphones for playback experiments (Mennill et al., 2006, 2012; Araya-Salas, et al., 2017; Hedley et al., 2017) and localization of songs of wild birds (Collier et al., 2010; Harlow, et al., 2013). However, microphone arrays are still not widely used by field researchers because of the limited availability of both software and hardware.

Suzuki et al. (2017) are developing an easily available and portable system for bird song localization called HARKBird. It automatically extracts sound sources (i.e., bird songs) and the DOA of each localized sound, both of which are useful to grasp the soundscape around the microphone array. HARKBird consists of a standard laptop PC with an open source software for robot audition HARK (Honda Research Institute Japan Audition for Robots with Kyoto University) (Nakadai et al., 2017) combined with a low-cost and commercially available microphone array. They showed the existence of temporal overlap avoidance in the singing behaviors of some forest species (Suzuki et al., 2017) and successful spatial localization of song posts of the great reed warblers by using multiple microphone arrays (Matsubayashi et al., 2017). Suzuki et al. (2018) further used multiple and self-developed 16-ch microphone arrays and evaluated both the spatial and the temporal localization accuracy of songs of these great reed warblers, by comparing the position and duration of localized songs around the song posts with those annotated by human observers, and found significant temporal overlap avoidance and an asymmetric relationship between songs of the two singing individuals.

In this paper, we report on an application of HARKBird to a deeper understanding of ecoacoustic dynamics of bird songs, focusing on temporal patterns of song characteristics and their behavioral contexts, which is a fine-scaled temporal analysis of song movement (i.e., changes in DOA of songs) — song type dynamics in playback experiments.

For this purpose, we focus on Japanese bush-warbler (*Cettia diphone*), which is one of the most popular songbird species in Japan and widely distributed in our experimental field. Males of Japanese bush warblers sing two types of songs: type-H and type-L, that were similar but slightly different (Hamao, 2007). A type-H song sounds as “Hoh-hokekyo” and its frequency is relatively high (Figure 1a). A type-L song has intermittent whistles and sounds as “Hoh-hohohokekyo” (Figure 1b) with relatively lower frequency. The type-L song is known as a threat to rivals in the vicinity because territory owners frequently use this type in the periphery of their territory. Also, they return this type of song to replayed songs of other males (Hamao, 2007). Momose (1986) reported they tend to sing type-L frequently while reducing the number of type-H songs and tended to explore other individuals when the songs of conspecific individuals were replayed. It was also recently reported that detailed song structures (e.g., the portion and inflation of frequency modulation) reflect ecological differences in their habitats (e.g., mainland or islands) (Hamao, 2013). We believe that detailed

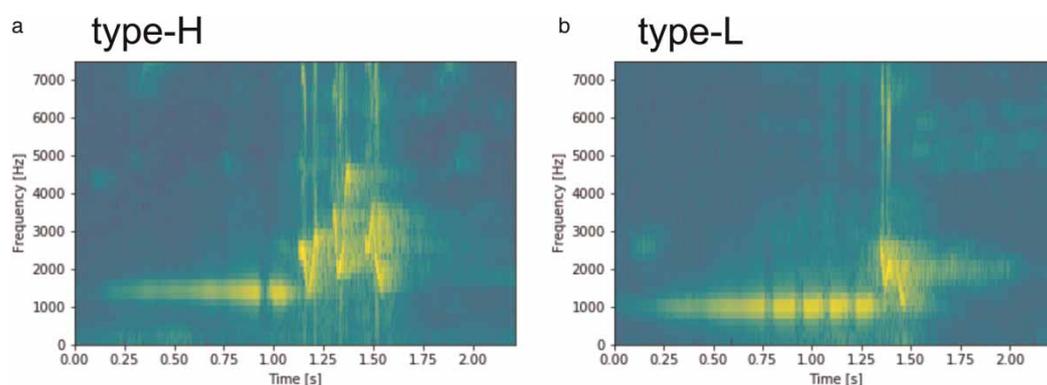


Figure 1. An example spectrogram of a type-H song (a) and a type-L song (b).

These songs were used for playback experiments (SH and SL, respectively).

analyses on the temporal relationship between song movement and song type reveal further understanding of ecological roles of their songs.

By extending the framework of HARKBird, we construct a system that enables us to conduct automatic playback and interactive experiments with different conditions, with a real-time recording and localization of sound sources. We investigate how playback of conspecific songs and playback patterns can affect vocalization of these two types of songs and spatial movement of an individual of JBWA. The results showed quantitatively that there exist strong relationships between song type and spatial movement, and some of the ecoacoustic dynamics of songs were simulated using a software, termed Bird song explorer, which provides users a virtual experience of acoustic dynamics of bird songs using a 3D game platform Unity (Naren et al., 2017).

Based on experimental results, we also discuss how our system can contribute to ecoacoustics in terms of two different roles of sounds: sounds as tools and subjects (Farina and Gage, 2017; Pavan, 2017).

Materials and methods

HARKBird and the developed system

We used HARKBird¹ to estimate the DOA of the sound sources acquired from a microphone array. The sound source localization algorithm of HARK is based on the MULTiple SIGNAL Classification (MUSIC) method (Schmidt, 1986) using multiple spectrograms with the short time Fourier transformation. Furthermore, we can extract separated sounds as wave files for each localized sound using GHDSS (Geometric High order Decorrelation based Source Separation) method. See Suzuki et al. (2017) for additional details of HARKBird and Nakadai et al. (2010, 2017) for HARK. We adjusted the parameters to localize songs of the JBWA as much as possible while suppressing localizations of other sound sources (e.g., water flows).²

We used a laptop PC (TOUGHBOOK CF-C2; Panasonic) and a 8-channel microphone array (TAMAGO; System in Frontier Inc., Tokyo, Japan (<http://www.sifi.co.jp/en/>)) placed on a tripod. The TAMAGO has 8 microphones that are horizontally arranged 45 degrees apart around its egg-shaped body, which enables us to conduct 24 bit, 16 KHz recording with a PC connected with a USB cable. We adopted Ubuntu Linux 12.04 in which HARK and hark-python were installed to execute sound source localization processes using HARKBird.

Because the current HARKBird does not support realtime recording with automatic and interactive playback, we modified the network of HARK in HARKBird so that the system can record and localize in realtime while performing automatic or interactive playback of a specified sound file using a loudspeaker (MM-SPBTBK; Sanwa Supply) that was connected to the laptop with Bluetooth. Specifically, we added a Python script that will be executed every time step (0.5 seconds) of the real-time localization process, and set up the script so that a song file will be replayed at the appropriate timing based on experimental setups.

Playback and interactive experiments

We conducted automatic playback and interactive experiments on an individual of JBWA around his territory at the Inabu field, the experimental forest of Field Science Center, Graduate School of Bioagricultural Sciences, Nagoya University, in central Japan (35°21'N, 137°57'E), in May 21st and 22nd, 2016. The forest is mainly composed of conifer plantation (Japanese cedar, Japanese cypress, and red pine), with small patches of broadleaf trees (*Quercus*, *Acer*, *Carpinus*, etc.). In this forest, common bird species are known to vocalize actively during a breeding season. JBWAs in particular breed in a bamboo thicket.

Specifically, we placed a system on an opened space (a parking space) in the field (Figure 2). The space was surrounded by trees or a bank. The focal individual usually sang on several trees around the loudspeaker and the microphone array, and thus we can estimate the spatial change in his song location as the directional change in his song location.

We used four different songs for playback experiments: the type-H song (SH) of the focal individual which was recorded before the experiment, the type-L song (SL) of this individual, the type-H song of another JBWA male (OH), and the type-L song of this male (OL). This another individual inhabited approximately 200 m away from the system. We conducted noise reduction and normalization on the recordings of these songs. We

1 The website of HARKBird (2017). Last updated 25 November 2017. <http://www.alife.cs.is.nagoya-u.ac.jp/~reiji/HARKBird/>.

2 We used (1) the expected number of sound sources for the MUSIC method: 3 sources, (2) the lower bound frequency for the MUSIC method: 2200 Hz to detect each vocalization by localizing the high frequency part of both types of songs while reducing low frequency noise, and (3) the threshold for source tracking: 28.5. See Suzuki et al. (2017) or the website of HARKBird for the details of each parameter.

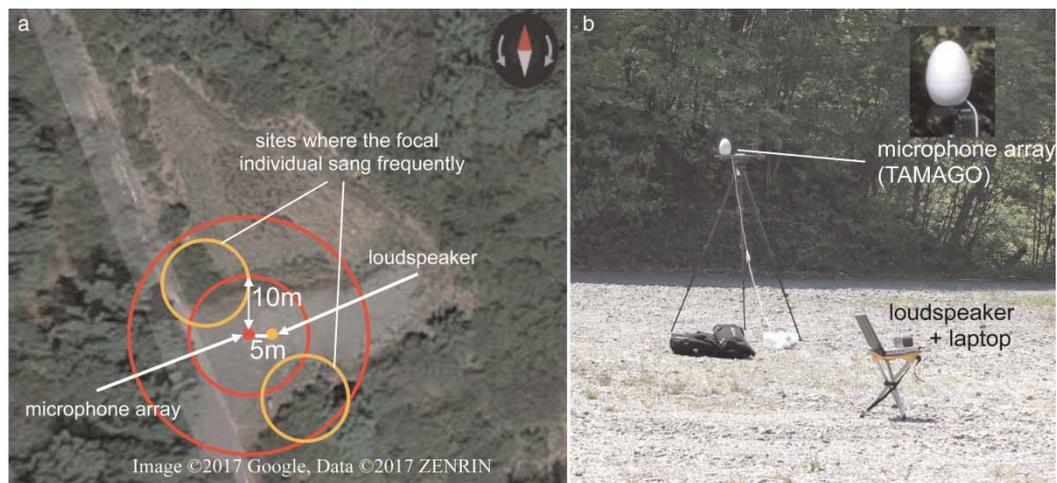


Figure 2. (a) The experimental field and the set up for playback experiments. (b) A snapshot of the system.

replayed these songs using three types of playing intervals. The two of them are the fixed time interval settings in which each song was replayed every 8 or 12 seconds (F8 and F12). In addition, we also interactively replayed about every 3 seconds after the song of the focal individual was localized by the system (I3). As a result, there are 4 (SH, SL, OH and OL) \times 3 (F8, F12 and I3) = 12 types of experimental settings. We also conducted two no-playback experiments (i.e., observation only) (N1 and N2). We started recordings of these no-playback experiments when the focal individual was singing around the field.

For each experimental setting, we conducted a playback experiment for a few minutes. We picked up the duration of 500 seconds during which the focal individual was singing repeatedly in the field. As for the case of OH/F12, we could not obtain such 500 seconds recording because the focal individual once went away from the field for a while. Thus, we extended the duration to 750 seconds to make the net duration during which the focal individual sang repeatedly 500 seconds. Note that we confirmed that the focal individual only sang during experiments from the field observation. We also confirmed that the localized songs were of the focal individual by checking the patterns of each song because it is reported that individual identification of male JBWA was possible by a combination of song patterns (Hamao, 1993).

After experiments, we manually classified the localized songs of the focal individual into two types (L or H). We could finally obtain the timing, DOA and type of each song of the focal individual during the experiments.

Results

Playback experiments

Figure 3 shows an example of the experiment (I3), showing that both songs of the focal individual and the replayed songs were successfully localized at different directions. The web-based versions of Bird song explorer, which will be explained later in Discussion, with the two experimental results (N1 and SH/F8) are available online,³ which provides users a virtual experience of acoustic dynamics of bird songs with these conditions (Figure 4). A user is represented as an avatar agent,⁴ and can move around the virtual field in which the focal individual was singing by using a keyboard interface (forward (up), backward (down), turn left (left), turn right (right), jump (space)).

We visualized the directional distribution of each song type (Figure 5) in all experiments.⁵ We observed the typical effects of song playback on the behavior of the focal individual. For example, in the cases without song playback (N1 and N2), the focal individual sang both type-H and type-L songs frequently at one or two specific locations on trees. On the other hand, in the other cases with song playback, the number of the type-H songs decreased while that of the type-L songs remained relatively consistent.

³ The website of Bird Song Explorer (2017). Last updated 26 October 2017. http://www.alife.cs.i.nagoya-u.ac.jp/~naran/bird_song_explorer/JBWA/.

⁴ We used Unity-chan (© Unity Technologies Japan/UCL) for the representation of the avatar.

⁵ The data of playback and interactive experiments on a Japanese bush warbler (2018). Last updated 7 June 2018. <http://www.alife.cs.i.nagoya-u.ac.jp/~reiji/jea2018/>.

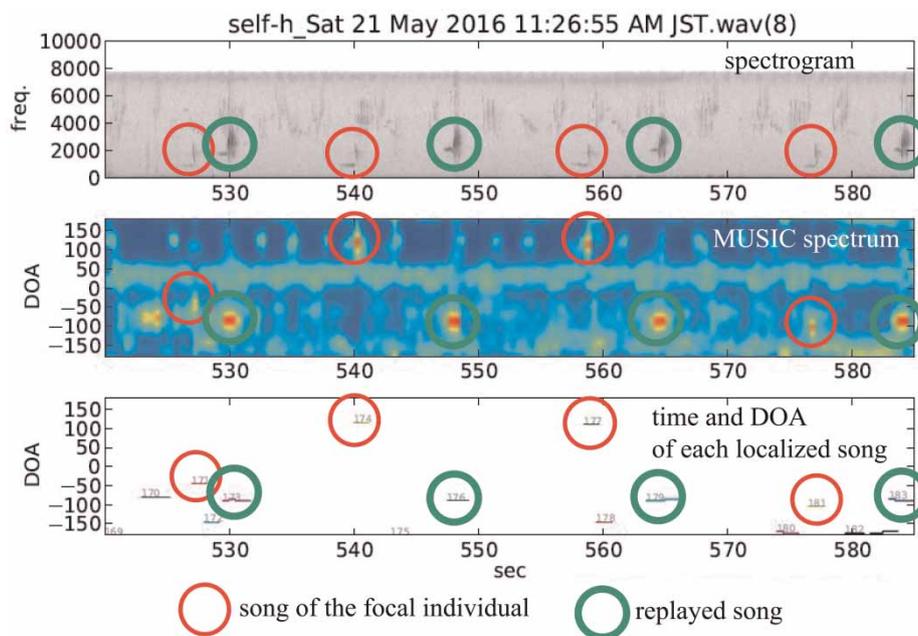


Figure 3. An example of the experiment (I3), showing that both songs of the focal individual and the replayed songs were successfully localized at different directions from the microphone array.

Top: the spectrogram of the recording of a channel. Middle: MUSIC spectrum, which represents the likelihood of the sound existence, calculated for each time and direction. Bottom: the time and the direction of the localized sources by extracting high local peaks (red) in the middle panel.

We focused on the effects of experimental conditions on the average amount of change in the DOA between consecutive songs and the proportion of type-H songs (Figure 6). We regarded the change in the DOA as the rough estimate of the change in the spatial location, because the distance between peripheral trees on which the focal individuals sang frequently and the microphone array was similar (approximately 10 m) (Figure 2). The result clearly shows that the focal individual tended to sing the type-H songs and tend not to change its DOA largely when N1 and N2. On the other hand, he tended to sing the type-L songs and moved frequently and far, flying over the loudspeaker, in the other cases with song playback. This tendency was more significant in the cases with 8 second interval of playback than the cases with 12 second, implying that more frequent playback had stronger effects on the behavior of the focal individual. However, we observed strong effects in the case of interactive playback while the average playback interval was 19.2 seconds, which was longer than 12 seconds (F12). This implies that interactive playback had different effects on the behavior of the focal individual. We also observe a weak tendency that these effects



Figure 4. A screenshot of the Bird song explorer (Naren et al., 2017) which was customized for simulating playback experiments.

A user can experience the soundscape in which the focal songs of JBWA and playback songs by a loudspeaker come from their estimated locations.

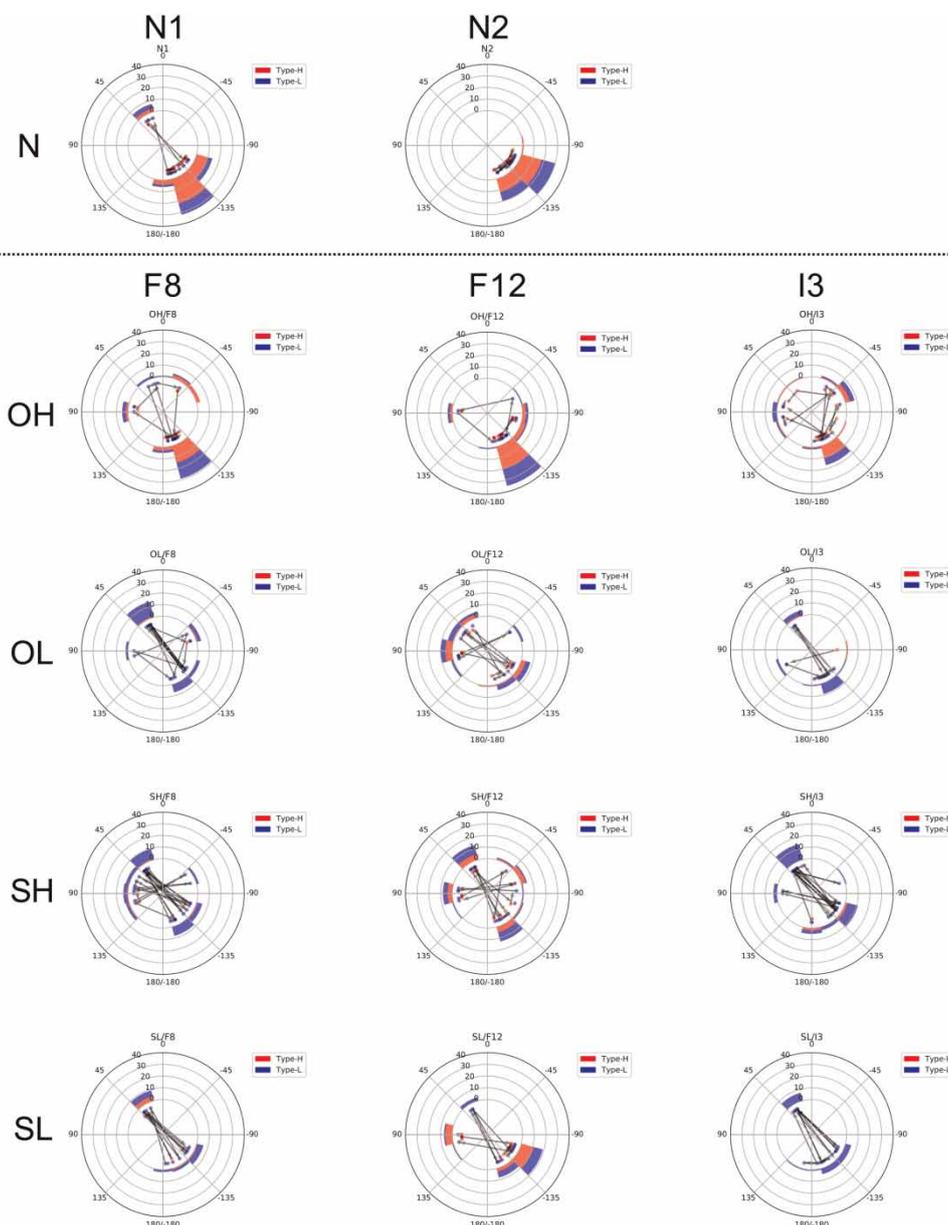


Figure 5. The directional distribution of each song type in all experiments.

The circular bar graph represents the histogram of the direction from the microphone array to the location at which the focal individual sang in 500 seconds. The arrows in the histogram represents the trajectory of movement in each experiment. Red: type-H, blue: type-L. N1, N2: no playback (observation only). OH (OL): the type-H (type-L) song of another individual. SH (SL): the type-H (type-L) song of the focal individual. F8 (F12): the fixed playback interval of 8 (12) seconds, I3: interactive playback.

were stronger when the songs of the focal individual were replayed than the cases of the songs of other individual and when the type-L songs were replayed than the cases of the type-H songs. This also implies the type of playback songs might affect responses from the focal individual. There were bimodal peaks in the histogram of the directional changes and the song type of the subsequent songs (i.e., the type of song that was sung after the movement) in the DOA (Figure 7). The large changes in the direction correspond to the movement between distant trees, while the small changes correspond to the movement on a tree or neighboring trees. It should be noticed that the focal individual tended to sing type-L songs after large directional changes, because the peak on the right consists of the type-L songs. This implies the song movement — song type relationship in his behavioral dynamics.

Statistical analysis

To clarify such behavioral tendency, we conducted a randomization test of the proportion the vocalization of type-L songs after a large change in the DOA (Figure 8). Specifically, we discretized the time series of directional change by classifying the amount of directional changes into large ($>T$ degree) and small ($\leq T$ degree) ($T = 45$ for

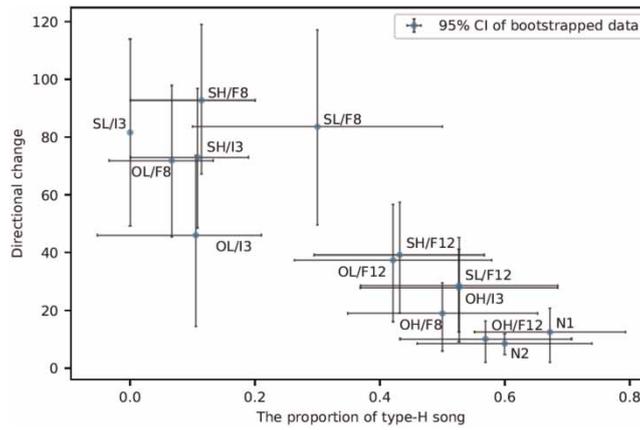


Figure 6. The proportion of type-H songs in all vocalizations and the amount of changes in the DOA between songs in all experiments.

The line represents 95% bootstrap percentile confidence interval based on 100,000 resamplings.

N2, OH/I3, SH/F12, SL/I3; $T=30$ for the other settings), and calculated the estimate of the proportion of type-L songs after a large directional change when the order of the directional changes was randomized, repeating 100,000 times of a random series generation. The observed proportion of the type-L was almost 1 and this was

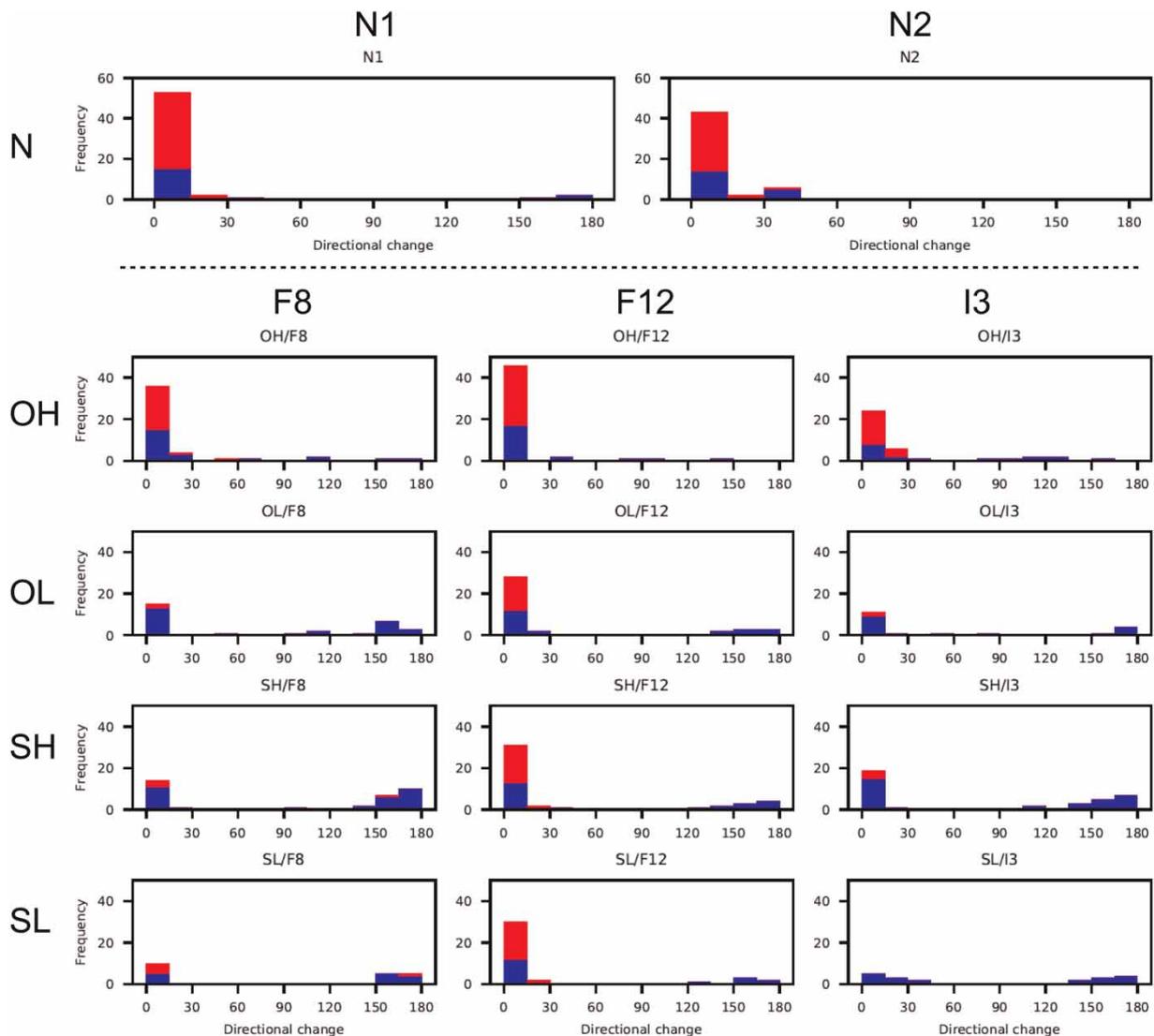


Figure 7. Histograms of absolute changes in the DOA between consecutive songs (red: type-H, blue: type-L).

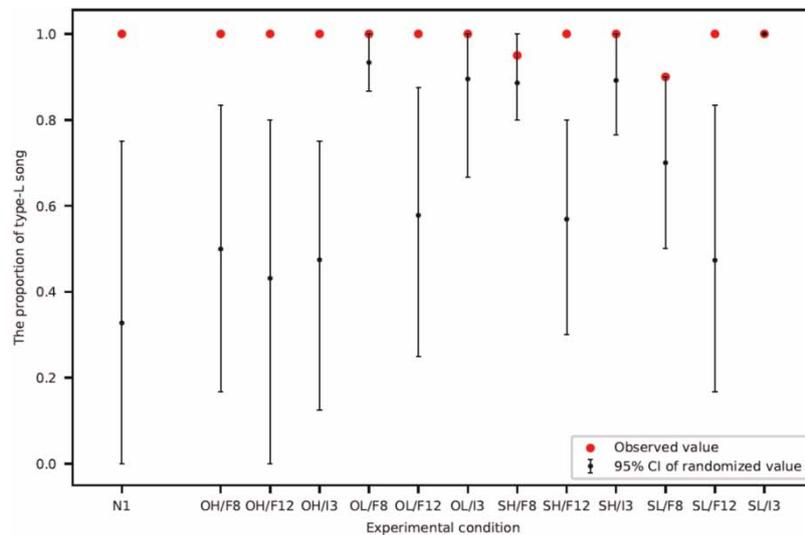


Figure 8. A randomization test of the proportion the vocalization of type-L songs after a large change in the DOA. Each red point represents the observed proportion of vocalization of the type-H song. The black dot and the error bar represents the average value and the 95% bootstrap percentile confidence interval of 100,000 resamplings.

significantly larger than those with randomized cases in almost all experiments, showing this behavioral tendency is significant.

Also, we measured transfer entropy between the discretized directional changes and the subsequent song type (Figure 9), which has been recently used to analyzing information flows in complex systems (Bossomaier et al., 2016). Specifically, this measure quantifies the expected amount of directional information flow from one time series to another; the transfer entropy $TE_{Y \rightarrow X}$ from a discrete time series $Y = \{y_t\}_{t=1,2,\dots}$ (source) to another discrete time series $X = \{x_t\}_{t=1,2,\dots}$ (destination). Given the past one value of X , the amount of reduction in the uncertainty about the future value of X (i.e., the reduced entropy of the transition probability of X) by knowing the past one value of Y is calculated as follows:

$$TE_{Y \rightarrow X} = \sum_{x_{t+1}, x_t, y_t} p(x_{t+1}, x_t, y_t) \log \left(\frac{p(x_{t+1} | x_t, y_t)}{p(x_{t+1} | x_t)} \right) \tag{1}$$

In our case, X and Y correspond to the time series of the change in the DOA and the type of subsequent song.

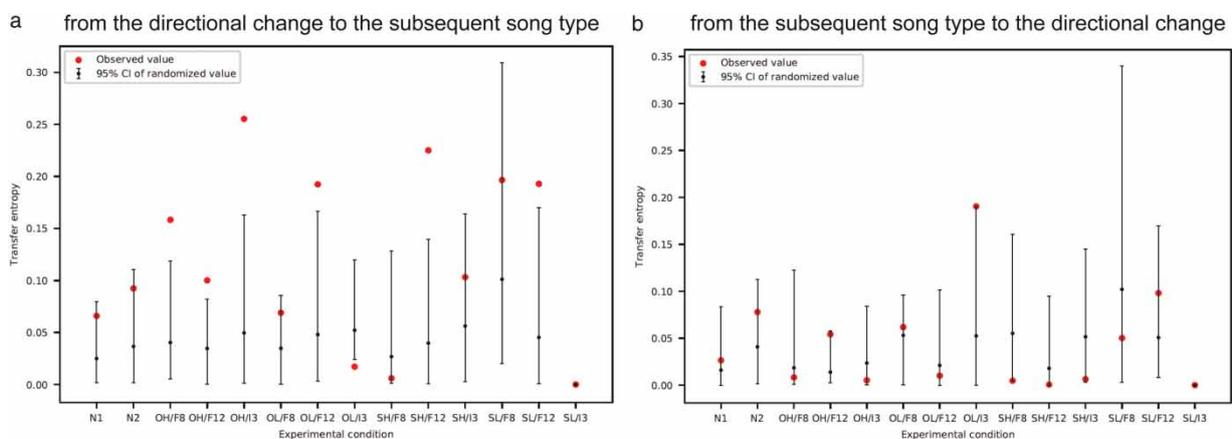


Figure 9. The transfer entropy between the discretized directional changes and the subsequent song type. a: from the directional change to the subsequent song type. b: from the subsequent song type to the directional change. Each red point represents the observed value and the black dot and the error bar represents the average value and the 95% bootstrap percentile confidence interval of 100,000 resamplings of the source (former) time series.

We compared the observed value of the transfer entropy with its estimated value of randomized cases in which the source time series was randomly shuffled. The observed information flow from the directional change to the subsequent song type is basically higher than the one in the randomized case, and significantly high in 6 cases (Figure 9a). On the other hand, there was no significant information flow from the subsequent song type to the directional change (Figure 9b). This means that the song movement clearly affects the type of the subsequent song.

Discussion

We conducted automatic playback and interactive experiments to understand the dynamics between song characteristics and behavioral contexts in various playback conditions by extending HARKBird and using a real-time localization feature of HARK.

Our system successfully observed effects of playback on behavioral patterns of an individual of JBWA in different environmental conditions. While decreasing the number of songs, the focal individual altered his behaviors by a large change in his position measured by DOAs in playback experiments. This observed behavioral tendency well fitted with the field observation in previous studies (Momose 1986).

While the results were from the observation of a single individual, experimental results with various combinations of different playback songs and fixed playback intervals showed a general tendency that the shorter playback interval had stronger effects among fixed playback experiments. On the other hand, the interactive experiments had also strong effects while the average song interval was 19.2 seconds. Maynard et al. (2012) showed duetting male long-tailed manakins avoid overlapping neighbors but do not avoid overlapping playback-simulated rivals. Our results may imply the differences in effects of fixed interval playback and interactive playback. We could also observe a weak tendency that type-L songs and songs of focal individual had stronger effects on the behavior of the focal individual, which might be because type-L songs had more aggressive roles and replaying songs of the focal individual himself could be more unnatural situations for the focal individual. Our system can contribute to analyze detailed behavioral changes in different playback conditions because the settings of conditions can be modified easily by modifying a script.

Furthermore, the fine-scaled data enabled us to further clarify the song movement — song type relationship that the focal individual tended to sing the type-L song after a large directional change in the DOA from the microphone array, using a randomization test. We also clarified this tendency using transfer entropy analyses, which showed the significant information flow from the spatial behavior to the singing behavior. This tendency appears to be plausible when we consider that the type-L song is used as a threat to rivals in the vicinity (Momose, 1986; Hamao, 2007), because there is a high possibility of the existence of rivals when the focal individual moves to another new place.

To obtain these data, we were necessary to manually remove unnecessary sound sources and classified songs of the focal individual into two classes (type-H and type-L). However, the DOA of each song was estimated from the system, which is the most important information that could not be easily and consistently obtained by human observers. We believe that techniques of machine learning will enable us to discriminate these sounds automatically according to the research purposes.

Natural sounds can be both the tool (i.e., an ecological attribute that can be utilized to investigate a broad array of applications such as the indirect measurement of biodiversity or habitat quality) and the subject (to understand the properties of sound, its evolution, and its function in the environment) of ecological research in the context of ecoacoustics (Farina and Gage, 2017; Pavan, 2017). We believe that our system could use biological sounds for both cases.

As for the former case, aiming at acoustic monitoring of bird behaviors in their habitat space, we could successfully observe a strong directionality of a JBWA's songs without playback. This not only showed the presence of the focal individual during the experimental periods but also potentially reflected the spatial structure and the acoustic quality of the habitat space (i.e., trees in an open space). The playback experiments further showed how such directional distributions of songs were modified by external biological factors (i.e., simulated intrusions by a rival). It should be noted that the system could quantitatively measure the behavioral changes with and without external factors, by the average change in the DOA of the sound source. This implies that these data can be used for long-term monitoring of bird behaviors in their habitat space.

As for the latter case, aiming at understanding functions of their songs, observed differences showed that external biological factors affected both spatial distributions and the song type patterns of conspecifics, while manual classification of song types is still required. We also observed the movement — song type relationship of the focal

individual. This is also a clear example of the strong environmental and behavioral relationships, which cannot be easily quantitatively measured without the sound source localization systems.

In addition, the obtained data are rich enough to recreate the soundscapes of bird songs in the field where we can virtually experience such acoustic environments. We developed a software, termed Bird song explorer, which provides us a virtual experience of acoustic dynamics of bird songs using a 3D game platform Unity as a new way of representation of acoustic behaviors with spatial information (Naren et al., 2017). As illustrated in Figure 4, a user, represented as an avatar, can explore a virtual forest based on the extracted songs from a real recording. In this virtual forest of which spatial structure is roughly simulated, an avatar hears each separated sound (song) replayed at the position which reflects its DOA at its localized timing, assuming the microphone array was placed in the center of the forest. The user can experience the soundscape composed of these sounds by using an OS dependent software with a surround sound system or a VR interface (e.g., Oculus Rift), or a web-based stereo system (as introduced above). We believe that this kind of representation of acoustic data with spatial information will be an important technique to better represent ecoacoustic dynamics of bird behaviors.

Future work includes more detailed analyses of bird behaviors using multiple microphone arrays and the development of automatic sound source classification based on machine learning techniques.

Acknowledgements

We thank Charles E. Taylor and Martin L. Cody (UCLA) for constructive comments and suggestions for experiments and analyses, and Naoki Hijii, Naoki Takabe and Norio Yamaguchi (Nagoya University) for supporting field experiments.

Funding sources

This work was supported in part by MEXT/JSPS KAKENHI Grant Number JP15K00335, JP16K00294, JP24220006, JP17H06841, and JP17H06383 in #4903.

Competing interests

All authors declare that they have no conflict of interest.

References

- Araya-Salas M., Wojczulanis-Jakubas K., Phillips E. M., Mennill D. J. and Wright T. F. (2017). To overlap or not to overlap: context-dependent coordinated singing in lekking long-billed hermits. *Animal Behaviour*. 124: 57–64. <https://doi.org/10.1016/j.anbehav.2016.12.003>
- Blumstein D., Mennill D. J., Clemins P., Girod L., Yao K., et al. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*. 48: 758–767. <https://doi.org/10.1111/j.1365-2664.2011.01993.x>
- Bossomaier T., Barnett L., Harré M. and Lizier J. T. (2016). An introduction to transfer entropy - information flow in complex systems. Berlin, Germany: Springer. <https://doi.org/10.1007/978-3-319-43222-9>
- Catchpole C. K. and Slater P. J. B. (2008). *Bird Song: Biological Themes and Variations*. Cambridge, UK: Cambridge University Press.
- Collier T. C., Kirschel A. N. G. and Taylor C. E. (2010). Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network. *The Journal of the Acoustical Society of America*. 128: 182–189. <https://doi.org/10.1121/1.3425729>
- Daimon K., Hedley R. W. and Taylor C. E. (2017). Semantic Inference of Bird Songs Using Dynamic Bayesian Network. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, eds. Singh S. P. and Markovitch S., AAAI Press. 4911–4912.
- Gasc A., Francomano D., Dunning J. B., and Pijanowski B. C. (2017). Future directions for soundscape ecology: the importance of ornithological contributions. *The Auk*. 134: 215–228. <https://doi.org/10.1642/AUK-16-124.1>
- Farina A. and Gage S. H. (2017). *Ecoacoustics A new science*. In *Ecoacoustics: The Ecological Role of Sounds*. eds. Farina A. and Gage S. H., Oxford, UK: Wiley-Blackwell. 1–11.
- Hamao S. (2013). Acoustic structure of songs in island populations of the Japanese bush warbler, *Cettia diphone*, in relation to sexual selection. *Journal of Ethology*. 31: 9–15. <https://doi.org/10.1007/s10164-012-0341-1>
- Hamao S. (1993). Individual identification of male Japanese bush warbler *Cettia diphone* by song. *Japanese Journal of Ornithology*. 41: 1–7. <https://doi.org/10.3838/jjo.41.1>
- Hamao S. (2007). Japanese Bush Warbler. *Bird Research News*, 4, 2. Last updated 3 February 2014. http://www.bird-research.jp/1_shiryoseitai/uguisu.pdf
- Harlow Z., Collier T., Burkholder V., and Taylor C. E. (2013). Acoustic 3D localization of a tropical songbird. In *Proceedings of the 2013 IEEE China Summit and International Conference on Signal and Information Processing (China SIP 2013)*, IEEE. 220–224.
- Hedley R. W., Huang Y. and Yao K. (2017). Direction-of-arrival estimation of animal vocalizations for monitoring animal behavior and improving estimates of abundance. *Avian Conservation and Ecology*. 12: 6. <https://doi.org/10.5751/ACE-00963-120106>

- Maynard D. F., Ward K. A., Doucet S. M. and Mennill D. J. (2012). Calling in an acoustically competitive environment: duetting male long-tailed manakins avoid overlapping neighbours but not playback-simulated rivals. *Animal Behaviour*. 84: 563–573. <https://doi.org/10.1016/j.anbehav.2012.06.008>
- Matsubayashi S., Suzuki R., Saito F., Murate T., Masuda T., et al. (2017). Acoustic monitoring of the great reed warbler using multiple microphone arrays and robot audition. *Journal of Robotics and Mechatronics*. 27: 224–235. <https://doi.org/10.20965/jrm.2017.p0224>
- Mennill D. J., Burt J. M., Fristrup K. M. and Vehrencamp S. L. (2006). Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *The Journal of the Acoustical Society of America*. 119: 2832–2839. <https://doi.org/10.1121/1.2184988>
- Mennill D. J., Battiston M. and Wilson D. R. (2012). Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. *Methods in Ecology and Evolution*. 3: 704–712. <https://doi.org/10.1111/j.2041-210X.2012.00209.x>
- Momose H. (1986). Mechanism of maintaining territories by acoustic communication. In: *Reproductive Strategies of Birds*, edited by Yamagishi S. Tokyo, Japan: Toukaidaigaku-shuppankai. 127–157.
- Nakadai K., Takahashi T., Okuno H. G., Nakajima H., Hasegawa Y. and Tsujino H. (2010). Design and implementation of robot audition system ‘HARK’ |open source software for listening to three simultaneous speakers. *Advanced Robotics*. 24: 739–761. <https://doi.org/10.1163/016918610X493561>
- Nakadai K., Okuno H. G. and Mizumoto T. (2017). Development, deployment and applications of robot audition open source software HARK. *Journal of Robotics and Mechatronics*. 27: 16–25. <https://doi.org/10.20965/jrm.2017.p0016>
- Naren, Suzuki R., Arita T., Nakadai K., and Okuno H. G. (2017). Bird song explorer: a virtual forest application based on a surround sound for experiencing singing behavior of birds. *Proceedings of the 79th National Convention of Information Processing Society of Japan*, 4, 239–240.
- Pavan G. (2017). Fundamentals of soundscape conservation. In *Ecoacoustics: The Ecological Role of Sounds*, eds. Farina A. and Gage S. H., Oxford, UK: Wiley-Blackwell. 235–258.
- Pijanowski B. C., Villanueva-Rivera L. J., Dumyahn S. L., Farina A., Krause B. L., et al. (2011). Soundscape ecology: the science of sound in the landscape. *BioScience*. 61 (3): 203–216. <https://doi.org/10.1525/bio.2011.61.3.6>
- Schmidt R. (1986). Bayesian nonparametrics for microphone array processing. *IEEE Transactions on Antennas and Propagation (TAP)*. 34: 276–280. <https://doi.org/10.1109/TAP.1986.1143830>
- Suzuki R., Matsubayashi S., Nakadai K. and Okuno H. G. (2017). HARKBird: exploring acoustic interactions in bird communities using a microphone array. *Journal of Robotics and Mechatronics*. 27(1): 213–223. <https://doi.org/10.20965/jrm.2017.p0213>
- Suzuki R., Matsubayashi S., Saito F., Murate T., Masuda T., et al. (2018). A spatiotemporal analysis of acoustic interactions between great reed warblers (*Acrocephalus arundinaceus*) using microphone arrays and robot audition software HARK. *Ecology and Evolution*. 8 (1): 812–825. <https://doi.org/10.1002/ece3.3645>