1T-02

Visualizing soundscapes and quantifying interspecific interactions in forest animal vocalizations using robot audition technology

Hao Zhao[1] Reiji Suzuki[1] Shinji Sumitani[1] Shiho Matsubayashi[2]
Takaya Arita[1] Kazuhiro Nakadai[3, 4] Hiroshi G. Okuno[5]

[1] Nagoya University [2] Osaka University [3] Tokyo Institute of Technology

[4] Honda Research Institute Japan [5] Kyoto University

1. Introduction

We have been proposing and discussing novel applications of robot audition techniques to visualize and analyze soundscape dynamics of forest animal vocalizations by using HARKBird [1], a bird song localization software based on a robot audition software HARK, and methods in ecoacousics [2]. So far, we visualized directional soundscapes of forest birds and cicadas using a false-color spectrogram of which color reflected the direction of arrival of their vocalizations and showed a possibility of inter- and intra-specific interactions among them using a simple classification method of separated sounds into bird and cicada vocalizations [3], but still needs further investigation.

This paper reports on further analyses on the applicability of our methods to discuss such inter- and intra- specific interactions, focusing on quantifications of vocal activities of birds and cicadas. We first visualize example patterns of their vocalizations in forests using directional soundscapes. Then, we describe an updated method to classify their vocalizations by using three ecoacoustic indices. We further analyze their vocal durations in playback experiments of cicada vocalizations, which implies there exist interspecific interactions between birds and cicadas.

2. HARKBird

HARKBird [1] is a collection of Python scripts that enable us to conduct a field recording using microphone arrays connected to a laptop PC and to analyze the recording using a network of HARK which are designed to localize and separate bird songs in fields. The HARKBird can estimate the existence and the direction of arrival (DOA) of each sound source by using the MUSIC method, and further separated the localized sound using the GHDSS method. It also provides an interactive annotation tool (Fig. 1A) used in this study. The detailed description of HARKBird and the scripts are available from our website[1].

3. Experiments and soundscapes

We conducted playback experiments on two consecutive days (June 24 and 25, 2021) to observe vocalizations of cicadas (*Terpnosia nigricosta*) and birds (e.g., Siberian blue robin (Luscinia cyane), Red-billed leiothrix (*Leiothrix lutea*)) in response to playback sounds of cicadas replayed around their territories at our field site in the Inabu field, the experimental forest of Field Science Center, Nagoya University. We used a 10-minute recording on June 27th, 2020, in which multiple cicadas (*Terpnosia nigricosta*) vocalized through the recording. We used the HARKBird to export the information about localized sound sources (i.e., the begin and end time, DOA, and its separated sound (wave file)), which were used in the visualization and analyses.

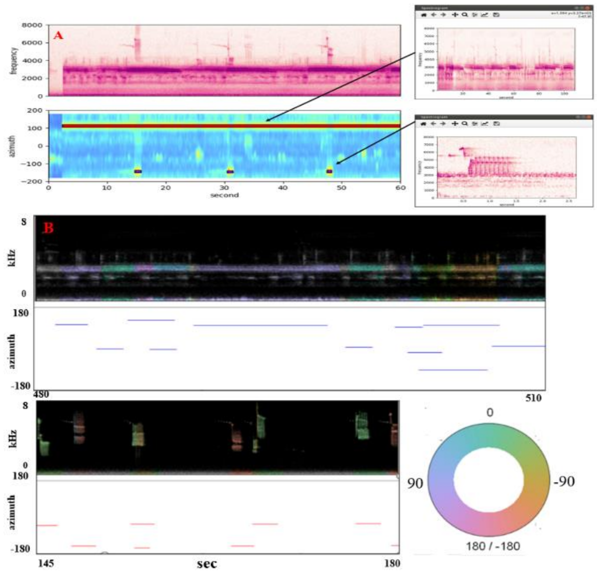

Fig. 1: A) The HARKBird annotation tool showing the information about localized sound sources; B) the directional false-color spectrum.

Fig. 1B shows example patterns of their vocalizations visualized with directional soundscapes based on false-color spectrograms [3]. Bars in the bottom graph represent the duration and DOA of localized sounds, and the colored regions in the spectrogram represent the spectral features of the localized sounds and their color orresponds to their DOAs. In the top spectrogram, cicada songs were indicated with different colors, which means that multiple individuals were singing in different directions alternately in this recording. In the bottom, two or three birds were vocalizing in different directions. These results show that the localization results reflect the vocal pattern of birds and cicadas well.

4. Classification of vocalizations

Acoustic indices can reflect distinct attributes of

1 http://www.alife.cs.is.nagoya-u.ac.jp/~reiji/HARKBird/

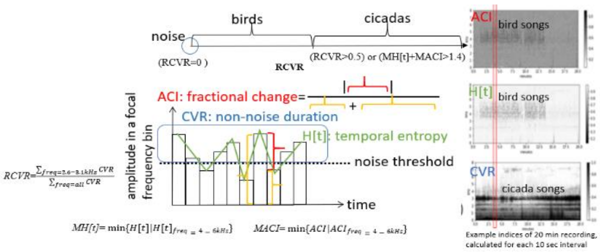

Fig. 2: The automatic classification method based on acoustic indices.

soundscape. We used the following three acoustic indices (Fig. 2) adopted in a visualization of soundscapes with a false-color spectrogram [4]:

- Acoustic complexity index (ACI) [Pieretti et al. 2011]: The average absolute fractional change in the amplitude values over time (complex animal vocalization (e.g. birds)),
- Temporal entropy (H[t]): The normalized entropy of the all amplitude values (short repetitions of vocalizations (e.g. birds)).
- Acoustic cover (CVR): The fraction of amplitude values where it exceeds a noise threshold (constant vocalizations (e.g., insects)).

We classified the localized sound sources into three classes (birds, cicadas, and noise) as follows:

- Calculate the values of the three indices of all frequency bins (from 0.0 to 8.0 kHz) for each separated sound file, and normalize these values.
- Calculate the RCVR, the sum of the CVR values corresponding to the frequency range from 2.6 to 3.1 kHz, which is further divided by the sum of the entire values. Then, calculate the minimum values of ACI and H[t] corresponding to the frequency range from 4.0 to 6.0 kHz, which are called MACI and MH[t], respectively.
- Each sound source is classified the following criteria: 0) If RCVR=0 then noise, 1) else if RCVR>0.5 then cicadas (long and broad vocalizations), 2) else if MH[t]+MACI>1.4 then cicadas (short and sharp vocalizations), 3) otherwise birds.

We tried to classify the localized sound sources in a 10-minute recording using the above classification method, and manually checked the classification results. The classification accuracy for localized sounds was 84％ (not shown).

5. Vocal activity analysis

In order to investigate inter- and intra-specific effects between birds and cicadas, we analyzed a six hour-recording at 9:50-15:40 on June 24th, 2021, where we replayed a 10-minute recording of cicada songs with an interval of 10 minutes to see the effects of song playback on their vocal activities and their mutual effects.

We classified the localized sound sources in the recording using our automatic classification method to extract the vocal activities of cicadas and birds. Their activity in each 150-second time segment was calculated as the total duration of localized sounds in the segment.

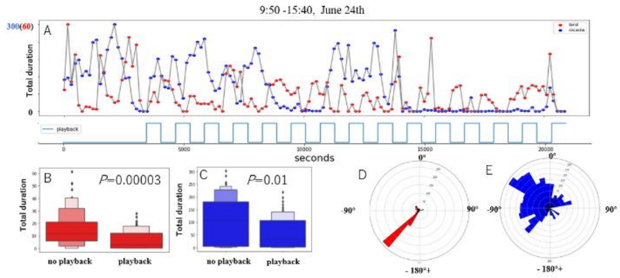

Fig. 3: The vocal activity of birds (red) and cicadas (blue) based on the proposed classification for about 6 hours.

Then, we conducted a Kolmogorov-Smirnov test to see if there were significant differences among vocal activities in playback and non-playback conditions.

Fig. 3 shows (A) the changes in the vocal activity of cicadas (blue) and birds (red) with and without playback conditions through the recording, the boxplot of their vocal activity in each condition ((B) and (C)), and their directional distribution of vocalizations ((D) and (E)). (A) shows that birds appeared to vocalize more actively when cicada vocalized less actively, and vice versa. (B) and (C) show that the vocal activities of both species were significantly active in the playback conditions than those in the non-playback conditions. This supports the inter-specific effects above, and further supports the existence of intra-specific effects among cicadas. (D) and (E) show that this soundscape of birds was dominated by a single individual singing at around -140 degrees, while those of cicadas were composed of many cicadas around the microphones.

6. Conclusion

This paper discussed an application of robot audition techniques to a soundscape analysis of a complex situation of vocalizing birds and cicadas in early summer. We showed that directional soundscapes can illustrate the patterns of their vocalizations and proposed an updated method to classify their vocalizations by using three acoustic indices. The preliminary quantitative analysis of their vocal activities implied that there exist inter- and intra-specific interactions between birds and cicadas.

Acknowledgements

We thank Naoki Takabe (Nagoya University) for supporting experimental recordings. This work was supported in part by JSPS KAKENHI JP21K12058, JP20H00475, JP19KK0260, JP17H06383 (#4903).

References

[1] R. Suzuki, S. Matsubayashi, R. Hedley, K. Nakadai, H. G. Okuno: HARKBird: Exploring acoustic interactions in bird communities using a microphone array. *Journal of Robotics and Mechatronics*, 27 (1): 213-223, 2017.

[2] A. Farina and S.H. Gage, Ecoacoustics: *The Ecological Role of Sounds*, John Wiley and Sons, 2017.

[3] H. Zhao, R. Suzuki, S. Sumitani, S. Matsubayashi, T. Arita, K. Nakadai, H. G. Okuno: Visualizing soundscape of animal vocalizations in forests using robot audition techniques, *JSAI Technical Report*, SIG-Challenge-057-04, pp. 15-20, 2020.

[4] M. Towsey, L. Zhang, M. Cottman-Fields, J. Wimmer, J. Zhang, and P. Roe: Visualization of long-duration acoustic recordings of the environment, *Procedia Computer Science*, 29:703-712, 2014.