Proceedings of ACML-TAB2019

Position estimation using multi-channel audio signals

Kazuki Fujimori KAZUKI-FUJIMORI@HIROSHIMA-U.AC.JP
Bisser Raytchev BISSER@HIROSHIMA-U.AC.JP
Kazufumi Kaneda KIN@HIROSHIMA-U.AC.JP
Hiroshima University

Emyo Fujioka EFUJIOKA@MAIL.DOSHISHA.AC.JP
Shizuko Hiryu SHIRYU@MAIL.DOSHISHA.AC.JP
Doshisya University

Toru Tamaki TAMAKI@HIROSHIMA-U.AC.JP
Hiroshima University

Abstract

In this paper, we describe a method for estimating the flight position of flying bats by using audio signals. This paper proposes a deep network architecture that takes audio signals obtained from an 18ch microphone array as input and outputs a bat presence map. Experimental Results the data which recorded the sound of one flying bat at 500kHz are shown.

Keywords: Position estimation, audio signals, CNN

1. Introduction

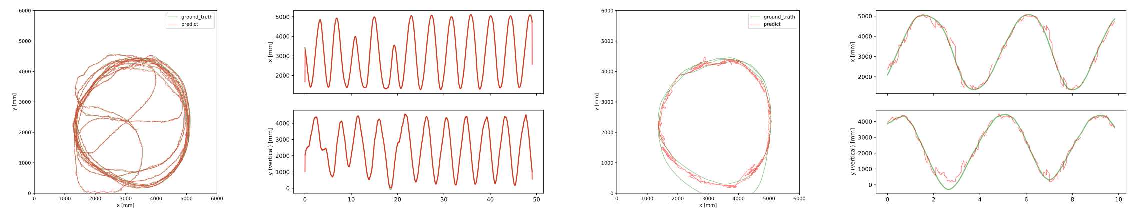
The analysis of the three-dimensional position of bats is an important issue in ecology, and research is being conducted to measure the position using techniques such as the difference in arrival time with a microphone array and the measurement of a stereo with a camera(Fujioka et al. (2014), Kloepper and Bentley (2017)). In this study, we investigate the feasibility of an approach that uses heat maps from 18ch microphone array signals.

2. Proposed method

When the coordinates are obtained by direct regression, the network structure depends on the number of flying bats. Therefore, we use maps to estimate multiple bat locations. A map is an image that shows higher values in places where bats seem to be flying. Even if the number of bats changes, maps can be expressed in the same format, so a network structure does not depend on the number of bats.

2.1. N-sheet simultaneous estimation

Bats are always moving with emitting pulse sounds. In the sound data, there is a silent interval between bat pulses. If one map is estimated by inputting a sound clip(a sample) with a short time width, the moving distance of the bat within the sample becomes small. However, silent samples are more often to be used. On the other hand, a longer sound clip is used as a sample to estimate a map, silent clips are less used, but the distance of bat



(a) training data

(b) test data

Figure 1: Results

movement within the sample will be large. Therefore, we use a long clip sample to estimate multiple maps of multiple time steps. As a result, the different of bat positions within a sample is small while reducing the use of silent samples.

3. Result

Fig.1 shows the locations estimated by the proposed network. Detected and peak positions are connected in chronological order and plotted as a path. Fig.1a is the result of the training data, and Fig.1b is of the test data. The left figure plots the detected peak position in the estimated map. The right figure is plotted with horizontal axis of time and vertical axis of x, y coordinates.

For the plots of the training data, it can be seen that most of the estimation is close to the true value, and the RMSE value is 21 mm. The results of the test data is worse a little, RMSE of 222 mm, but in an acceptable range. Further improvement is a future work.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers JP16H06540, JP16H06542.

References

Emyo Fujioka, Ikkyu Aihara, Shotaro Watanabe, Miwa Sumiya, Shizuko Hiryu, James A. Simmons, Hiroshi Riquimaroux, and Yoshiaki Watanabe. Rapid shifts of sonar attention by pipistrellus abramus during natural hunting for multiple prey. *The Journal of the Acoustical Society of America*, 136(6):3389–3400, 2014. doi: 10.1121/1.4898428. URL https://doi.org/10.1121/1.4898428.

Laura N. Kloepper and Ian Bentley. Stereotypy of group flight in brazilian free-tailed bats. *Animal Behaviour*, 131:123 – 130, 2017. ISSN 0003-3472. doi: https://doi.org/10.1016/j.anbehav.2017.07.019. URL http://www.sciencedirect.com/science/article/pii/S0003347217302312.