Indoor localization based on analysis of environmental ultrasound

Yudai Nagama[1] and Takeshi Umezawa[2] and Noritaka Osawa[2]

[1] Graduate School of Science and Engineering Chiba University, Chiba, Japan
[2] Graduate School of Engineering Chiba University, Chiba, Japan

Abstract. This study proposes and evaluates a method to estimate indoor position by environmental ultrasound as latent information in a target environment. The method converts environmental ultrasounds into spectrograms and builds a non-linear regression model based on the spectrograms using a convolutional neural network to estimate the indoor position and direction. This study evaluated the proposed method's estimation accuracy of the position and direction in a room. In a position estimation experiment using ultrasounds recorded at 20 measuring point in the target room, a position was estimated based on a regression model trained with training data. The model has a root mean squared error (RMSE) of 1.01 m for test data. This result demonstrates that the proposed method is effective for indoor position estimation. In a direction estimation experiment, the estimation error was 78.81° for test data. This result shows that it is difficult to estimate the direction of an ultrasonic microphone used in the experiment.

Keywords: Environmental Ultrasound, Spectrogram, Convolutional Neural Network, Deep Learning

1 Introduction

Indoor position estimation techniques are being studied at present, but many of them require installation of devices such as beacons. We think that information which already exists in an indoor environment makes it possible to perform indoor position estimation which keeps cost down. In this research, we focused on environmental ultrasound as latent information in an indoor environment. Environmental ultrasound refers to a mixture of ultrasounds emitted from indoor devices such as air conditioning equipment and switching power supply.

Tsuchiya et al. have shown that it is possible to identify multiple rooms by using environmental ultrasound [1]. They converted the ultrasound into a spectrogram and analyzed it using a convolutional neural network (CNN). We extend the area classification method to a regression method in an area and propose a new method to estimate a position and direction based on environmental ultrasounds. We performed experiments to evaluate the proposed method and obtained the results that the proposed method is effective for position estimation but not for direction estimation.

2 Related Work

In the case of a method using radio waves, a method based on triangulation is often used if the position of the transmitting source is known. A method of triangulation mainly uses received signal strength or time of flight between a transmitting source and a receiving terminal. When the position of the transmitting source is unknown, many methods based on position fingerprinting are used. In the case of using environmental ultrasounds, estimation based on position fingerprinting is performed.

There are many methods that use Wi-Fi, which is already installed in many indoor environments [2], [3], [4], [5], [6]. Methods using Bluetooth [3], [7], [8] and using a special wireless signal [9] require installment of the beacons and terminals, and such installment may cost much.

Tsuchiya et al. showed that environmental ultrasounds in rooms have unique characteristics and demonstrated that analysis of ultrasounds could distinguish multiple rooms with an accuracy of 97.5% [1]. Their work also reported that ultrasounds at different locations in the same room had different characteristics. Therefore, we think it is possible to perform indoor position estimation in one room by exploiting environmental ultrasounds.

3 Proposed Method

We propose a method to estimate the position and direction within a room using environmental ultrasound. The method is partially based on the work of Tsuchiya et al. but differs in that we estimate the coordinates and azimuth of a recoding ultrasonic microphone by using non-linear regression models. The regression models are built by CNN with spectrograms of ultrasound. Since our method is a type of fingerprinting, we need to collect ultrasounds at various positions and directions to build non-linear regression models for coordinates and azimuth.

3.1 Spectrogram

An example of a spectrogram is shown in Fig.1. A spectrogram is a heat map representing relationships among time, frequency, and intensity. In the spectrogram, the horizontal axis represents time, the vertical axis represents a frequency band, and the gray scale represents intensity. Short-term changes such as beats of fan noises of an air conditioner can be important features of ultrasound fingerprinting. The spectrogram can include such importance features which cannot be expressed by simple acoustic features such as Fourier coefficients for a period.

3.2 Convolutional Neural Network (CNN)

CNN is one of deep learning algorithms that show high accuracy in image recognition. CNN performs feature extraction in the image by the convolution layer and the pooling layer that allows misalignment.

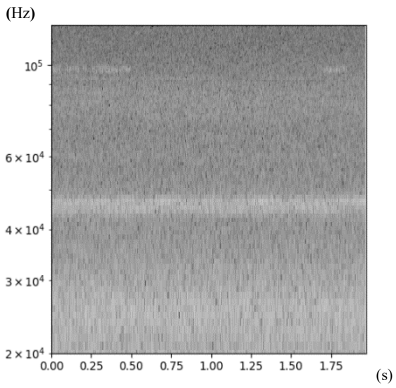


Fig.1. Example of environmental ultrasound spectrogram

4 Position Estimation

We experimented to evaluate the effectiveness of position and direction estimation by the proposed method. This section describes the experiment and its results.

4.1 Experimental Procedure

We collected ultrasounds in one room. The recorded environmental ultrasound is converted into a spectrogram. Pairs of the spectrogram and recording position of ultrasound were accumulated to a dataset. We divide the created data set into training data and test data. Non-linear regression models of the positions were built by CNN. The features of the spectrogram for each position of training data were learned. Recording positions of test data are estimated using the created models, and the estimation error is evaluated.

4.2 Experimental Environment

Ultrasounds were collected in a room whose dimensions are 5.93 m long × 4.73 m wide × 2.60 m high. The room has 0.3 m × 0.4 m pillars at its four corners, which are indicated by black rectangles in Fig.2(a).

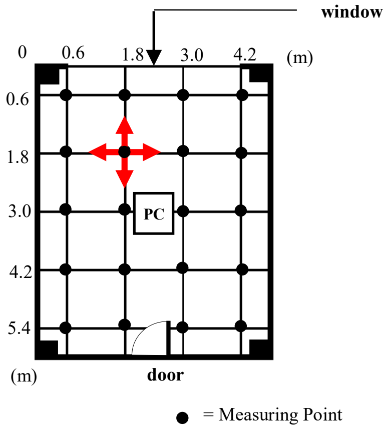
A Dodotronic Ultramic 250k ultrasonic microphone was used as the recording device. Since the maximum sampling rate of the microphone is 250 kHz, ultrasounds can be recorded up to 125 kHz. The microphone was connected to a notebook-type personal computer (PC) which captured and saved ultrasounds. The PC was fixed at the center of the room to reduce the influence of any ultrasound generated by the PC.

4.3 Experimental Conditions

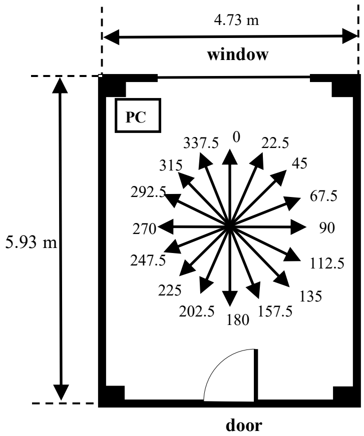
We placed 20 measuring points indicated by dots in Fig.2(a), which were points at the intersections on a 1.2 × 1.2 m grid, within a room, and performed recording at the 20 points. Also, to evaluate the influence of recording direction, recordings in four directions were collected at each measuring point as shown in Fig.2(a). This dataset which contains four direction data is called dataset R-all.

The length of one recording was two seconds, and we repeatedly recorded ultrasounds at one-second intervals. In this experiment, we only used sound data of 20k to 125 kHz because a spectrum of the audible range is affected by sounds caused by daily human activities (e.g., human voices).

To investigate the influence of recording direction on position estimation, we compared the estimation errors of the datasets for each direction and all directions. We refer to the set of all data recorded with direction 0° as dataset R0, which contains 10,000 recordings. Similarly, sets of data recorded with directions 90°, 180°, and 270° are referred to respectively as datasets R1, R2, and R3. For dataset R-all, 200 data were collected in each direction, 800 data in each position, 16,000 data for all. We used 80 % of the data in each direction at each point as training data and the remaining 20% as test data.



(a).Position Estimation



(b).Direction Estimation

Fig.2. Recording Condition

4.4 Structure and Parameters of CNN

Table 1 shows the structure of CNN used in this study. In Table 1, the shape of the input array to each layer and the shape of the output array from each layer is expressed by (X, Y, Z) where the size of the input array is X × Y and Z is the number of input or output channels.

CNN was trained with a batch size of 32 for 100 epochs. The learning rate started from 10^{-5} and was attenuated by 10^{-6} at each update. Momentum was 0.9. The stochastic gradient descent (SGD) method was used to optimize the learning parameters.

Table.1 Structure of CNN

Layer	Input array	Output array
Input		(512,512,1)
Convolutional 1	(512,512,1)	(512,512,32)
Pooling 1	(512,512,32)	(255,255,32)
Convolutional 2	(255,255,32)	(255,255,32)
Pooling 2	(255,255,32)	(127,127,32)
Convolutional 3	(127,127,32)	(127,127,64)
Pooling 3	(127,127,64)	(63,63,64)
Convolutional 4	(63,63,64)	(63,63,64)
Pooling 4	(63,63,64)	(31,31,64)
Flatten	(31,31,64)	61504
Fully Connected 1	61504	512
Fully Connected 2	512	1
Output	1	

4.5 Evaluation

Root mean squared error (RMSE) is used as a quantitative measure of estimation error. RMSE shows an average of distance between coordinates of a ground truth and an estimation. The unit of distance is meter in this research.

In this experiment, we configure two non-linear regression model which estimate X and Y coordinate. We defined RMSE of X and Y coordinates independently by (1), (2) as RMSE_x and RMSE_y where a both of data *i* is respectively (x_i, y_i) and $(\widehat{x_i}, \widehat{y_i})$ and the number of data is n. RMSE between ground truth positions and estimated positions is defined (3).

$$\mathrm{RMSE}_x = \sqrt{\Sigma_{i=1}^{n}(\widehat{x_i} - x_i)/n} \quad (1)$$

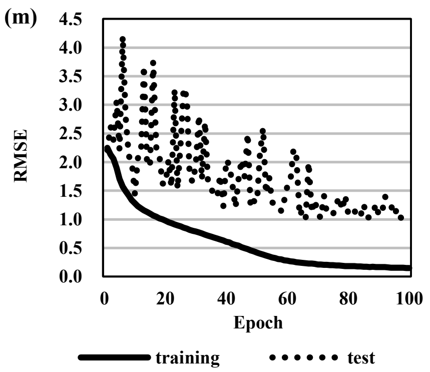
$$\mathrm{RMSE}_y = \sqrt{\Sigma_{i=1}^{n}(\widehat{y_i} - y_i)/n} \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\mathrm{RMSE}_x{}^2 + \mathrm{RMSE}_y{}^2} \quad (3)$$

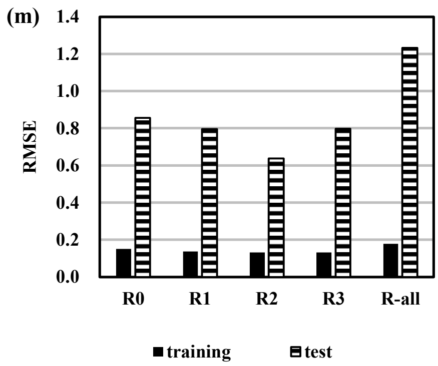
4.6 Result

Fig.3(a) shows the transition of RMSE of training data and test data for R-all. The horizontal and vertical axes of the graph represent respectively the number of epochs of learning and RMSE. The solid and dotted lines in Fig.3(a) indicate RMSE for the training data and the test data, respectively. At epoch 100, the RMSE of the training data was 0.15 m, and that of the test data was 1.01 m.

Fig.3(b) shows the RMSE values dataset R0, R1, R2, R3 and R-all. The RMSE values of dataset R0, R1, R2, and R3, each of which includes data for one direction, is lower than the RMSE values of dataset R-all, which includes all of four direction data.



(a). Transition of Position Estimation Error



(b).Comparison of Position Estimation Error in Each Dataset

Fig.3. Position Estimation Result

4.7 Discussion

In the experiment using the dataset R-all that mixed the recording data of four directions, RMSE for training data was about 0.15 m. This result shows that learning epoch is sufficient and the indoor environment ultrasounds had different characteristics for each position.

However, RMSE for test data fluctuated between 1 m and 1.5 m, as shown in Fig.3(a). Moreover, the estimation performance for training data was higher than that for test data. Those suggest that the non-linear regression model fell into overfitting.

The RMSE of R0, R1, R2, and R3 using only unidirectional data is lower than that of R-all. This result suggests that recording directions affect position estimation accuracy. If we get orientation information, for example, from a magnetic field sensor in a smartphone, we could improve the accuracy and robustness to microphone direction of the proposed method.

5 Direction Estimation

From result of position Estimation, recording direction of microphone effect position estimation accuracy. In this section, we performed direction estimation experiment by proposed method.

5.1 Experimental procedure

We followed the same procedure as the experiment described in Section 4. The same microphone and room as the position estimation experiment were used for the recording. In this experiment, we estimate indoor direction by non-linear regression model.

5.2 Experimental Conditions

Ultrasounds were recorded in 16 directions from the center of the room, as shown in Fig.2(b). An ultrasonic microphone was placed on a desk at the center of the room. The recording angle was varied from 0 to 337.5° in increments of 22.5°.

Dataset contains 1,000 items of data in each direction, for a total of 16,000 items. We used 80% of the data in each direction as training data and the remaining 20% as test data.

5.3 Result

Fig. 4. shows the transition of the estimation error. The horizontal axis of the graph represents the number of epochs of learning and the vertical axis represents estimation error. The unit of estimation error is a degree. The solid and dotted lines in the graph indicate the estimation error for the training data and test data of datasets, respectively. At epoch 100, the estimation error for the test data was respectively 78.81°.

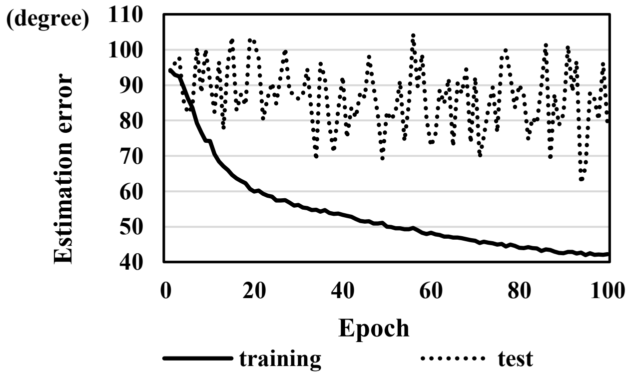


Fig.4. Transition of Direction Estimation Error

5.4 Discussion

As a result of direction estimation error for test data is about 80°. In the position experiment, Fig.3(b) shows that environmental ultrasound recorded in 90 ° units has different features. If the direction estimation result is used to specify the recording direction of the microphone in position estimation, an error of about 90 degrees is insufficient in accuracy from the R-all result. Therefore, further investigation is needed to clarify the cause.

6 Conclusion

This paper proposed and evaluated a method to estimate indoor position and direction within a room based on environmental ultrasounds. In a position estimation, the results suggest that environmental ultrasound can be used to estimate a position in one room and that estimation accuracy can be improved by utilizing orientation information together. In a direction estimation, it was difficult to estimate the direction of an ultrasonic microphone used in the experiment.

References

1. Tatsuro Tsuchiya, Takeshi Umezawa, and Noritaka Osawa, "An Indoor Area Estimation Method Analyzing Spectrograms of Environmental Ultrasounds by Convolutional Neural Network", 2018 Ubiquitous Positioning, Indoor Navigation and Location and Location-Based Services (UPINLBS), DOI: 10.1109/UPINLBS.2018.8559701
2. Duc V. Le and Paul J.M. Havinga, "SoLoc: Self-organizing indoor localization for unstructured and dynamic environments", 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2017.8115900
3. Viet-Cuong Ta, Trung-Kien Dao, Dominique Vaufreydaz, and Eric Castelli, "Smartphone-Based User Positioning in a Multiple-User Context with Wi-Fi and Bluetooth", 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2018.8533809
4. Brieuc Berruet, Oumaya Baala, Alexandre Caminada, and Valery Guillet, "A Deep Learning Based CSI Fingerprinting Indoor Localization in IoT Context", 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2018.8533777.
5. Xingli Gan, Baoguo Yu, Lu Huang, and Yaning Li, "Deep learning for weights training and indoor positioning using multi-sensor fingerprint", 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2017.8115923.
6. Xiaoyun Zhou, Jie Wei, Fang Zhao, Haiyong Luo, and Langlang Ye, "A shop-level location algorithm based on CNN for crowdsourcing fingerprint", 2018 Ubiquitous Positioning, Indoor Navigation and Location and Location-Based Services (UPINLBS), DOI: 10.1109/UPINLBS.2018.8559873
7. Maani Ghaffari Jadidi, Mitesh Patel, Jaime Valls Miro, Gamini Dissanayake Jacob Biehl, and Andreas Girgensohn, "A Radio-Inertial Localization and Tracking System with BLE Beacons Prior Maps", 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2018.8533827
8. Tomofumi Takayama, Takeshi Umezawa, Nobuyoshi Komuro, and Noritaka Osawa, "An Indoor Positioning Method Based on Regression Models with Compound Location Fingerprints", 2018 Ubiquitous Positioning, Indoor Navigation and Location and Location-Based Services (UPINLBS), DOI: 10.1109/UPINLBS.2018.8559728
9. A. De-La-Llana-Calvo, J. L. L´azaro-Galilea, A. Gardel-Vicente, D. Rodr´ıguez-Navarro, and I. Bravo-Munoz, "Characterization of Multipath Effects in Indoor Positioning Systems Based on Infrared Signals", 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), DOI: 10.1109/IPIN.2018.8533816