Reverberation-Robust Acoustic Indoor Localization

Jae Choi, Jeunghun Kim, Shin Jae Kang, Nam Soo Kim

Department of Electrical and Computer Engineering and INMC,
Seoul National University, Republic of Korea
{sjchoi,jhkim,sjkang}@hi.snu.ac.kr, nkim@snu.ac.kr

Abstract

We propose an effective approach to acoustic indoor localization that works in reverberant environments. The proposed localization system has been implemented to operate in an inaudible frequency range for practical applications. In order to achieve high performance, we propose an effective acoustic source data structure with a synchronization algorithm. To evaluate the proposed system, series of experiments were conducted in simulated reverberant environments and have shown good performance.

Index Terms: Acoustic indoor localization, OFDMA-CDM, TDOA.

1. Introduction

For the last several years, a pervasive use of smart devices such as smartphones or tablets has brought a growth in location based services (LBS). Most of LBS, such as navigation systems, traffic alerts, and etc., depend on the global positioning system (GPS) which only works outside. Indoor localization has also become important for applications such as virtual tour guide at museums, tracking personal possessions, or guide at shopping malls. Conventionally, RFID, Wi-Fi and camera have been deployed as indoor positioning techniques. However, performance of these techniques has not been satisfactory due to various indoor interferences or insufficient line of sight (LOS) by obstacles.

Recently, studies on indoor localization using acoustic sensors has been increasing due to their several advantages. Unlike other techniques, they can detect signals through obstacles, i.e., they get less influence from LOS. Loudspeakers and microphones are already available in most indoor environments; loudspeakers are installed in buildings for background music or notification announcement and people carry various mobile devices with microphones.

Even though acoustic indoor localization (AIL) is likely to be categorized as sound source localization (SSL), there are differences in their goal and system setting. In SSL, the goal is to find the positions or directions of target sources by using multiple microphones at a known position, while in AIL, we are aiming at locating a target microphone with reference to the source positions with fixed known positions.

The basic idea of AIL is to estimate the relative position of the target microphone based on its incoming acoustic signal. Since the received signal is a superposition of source signals played from a number of loudspeakers, the source signals have to be carefully designed. One of the most important issues in AIL is how to design the source signals. In the traditional approaches, specific sequences such as the pseudo-random sequences or Gold codes [1, 2] are utilized, which are then modulated by means of the code division multiple access (CDMA) [1], or direct sequence spread spectrum techniques [1, 3]. In order to estimate the position of the target receiver, the time of arrival (TOA) and time difference of arrival (TDOA) are measured based on time delays. In a TOA-based system, position is estimated by calculating distances between the receiver and the sources while considering the relation between the time delay and the speed of sound in a synchronous manner. Whereas the TDOA-based methods assumes the receiver and the sources to be asynchronous and compute the position through multilateration or iterative estimation [3, 4].

Previous studies have focused on accurate indoor localization, however they lacked in considering actual environmental issues. In real environments, there usually exist acoustic reverberation and the near-far effects which implies the situation that the target signal is masked by nearby strong signals making it difficult to retrieve the desired signal.

In developing our approach to AIL, we have focused on several principles for a practical application: First, we aim to influence the least on human hearing by utilizing the inaudible frequency band which can be processed by off-the-shelf audio devices. Second, we target the system to operate in large reverberant environments with sub-meter accuracy. Based on these principles, we propose the AIL system with an efficient source data structure which is designed to deal with environmental issues and sychronization algorithm to cope with the multipath effects. A series of simulated experiments is conducted to evaluate the performance of the proposed scheme.

2. Acoustic indoor localization using TDOA

In this section, we explain the general procedure of acoustic indoor localization. The target receiver is assumed to be surrounded by a set of loudspeakers with known positions placed in a reverberant environment as shown in Figure 1. Suppose that M sound sources are simultaneously emitted from the loudspeakers. The room impulse response (RIR) from the i-th loudspeaker to the receiver can be described in a simplified form with Dirac delta function $\delta(t)$ as

$$h_i(t) = \sum_k a_{i,k}\delta(t - b_{i,k}) \quad (1)$$

where $a_{i,k}$ represents the amplitude of the RIR measured at time $b_{i,k}$. Based on (1), the received signal $y(t)$ at time t is given by

$$y(t) = \sum_{i=1}^{M} [h_i(t) * x_i(t)] + n(t) \quad (2)$$

where $*$ indicates convolution, $x_i(t)$ represents the i-th sound source and $n(t)$ is the background noise at time t.

In this scenario, our goal is to estimate the receiver position from received signal $y(t)$. To compute the location from

$y(t)$, we need precise estimation of time delays $\{\tau_i\}$ between the receiver and each source. In designing $x_i(t)$, the characteristics of the cross-correlation function between sources and auto-correlation function of each source are important in order to provide high accuracy in estimating $\{\tau_i\}$. The cross-correlation between different sources should be kept as small as possible to avoid interferences from other sources, whereas the auto-correlation function of each source needs to have a salient peak for a successful estimation of τ_i. Among many kinds of pseudo-random sequences, we apply the Gold sequence due to its good auto- and cross-correlation properties [5].

Given the received signal $y(t)$ and the source signals $\{x_i(t)\}$, the time delay τ_i is estimated by following equation

$$\tau_i = \underset{\tau}{\operatorname{argmax}}\, R_i(\tau) \tag{3}$$

where $R_i(\tau)$ represents the following cross-correlation function

$$R_i(\tau) = \sum_t y(t+\tau)x_i(t). \tag{4}$$

If we assume the speed of sound c is constant, measuring TOAs implies measuring the distances $\{l_i\}$ between the source and receivers using the relation $l_i = c\tau_i$. This is possible if the receiver and sources are synchronized, however, this TOA-based system is often not the case in reality.

In the TDOA-based system, the time delays $\{\tau_i\}$ cannot be directly utilized to calculate the distance because the receiver and sources are assumed to be asynchronous. Instead, the position of the receiver is estimated by the time differences $\{\tau_{ij}\}$. If the positions of the receiver and each source are denoted by $\boldsymbol{r}$ and $\boldsymbol{s}_i$, respectively, the time delay τ_i in (3) can be represented as

$$\tau_i = \eta_i + \frac{||\boldsymbol{s}_i - \boldsymbol{r}||}{c} - \gamma \tag{5}$$

where $\|\cdot\|$ indicates the Euclidean norm, and γ and η_i imply the signal capturing time and emission time of the i-th source, respectively. The time difference τ_{ij} is now given by

$$\begin{aligned}\tau_{ij} &= \tau_i - \tau_j \\ &= \eta_i - \eta_j + \frac{\|\boldsymbol{s}_i - \boldsymbol{r}\|}{c} - \frac{\|\boldsymbol{s}_j - \boldsymbol{r}\|}{c} \\ &= \frac{\|\boldsymbol{s}_i - \boldsymbol{r}\|}{c} - \frac{\|\boldsymbol{s}_j - \boldsymbol{r}\|}{c} \end{aligned} \tag{6}$$

where η_i and η_j cancel out because we assume that all the sources emit the signals simultaneously. Since the target position $\boldsymbol{r}$ can not be directly obtained from the time differences $\{\tau_{ij}\}$, we apply the iterative TDOA algorithm.

3. Acoustic source design and synchronization

In this section, we propose a source data structure and synchronization algorithm for an efficient localization in actual environments. For this we have focused on the following issues: First, we need to consider the reverberations which cause multipath effects in designing the acoustic source signals. Second, acoustic source data structure should be designed carefully to mitigate the near-far effects. Third, we need to deal with the case in which the power of indirect signal component becomes stronger than the direct signal component in reverberant environments.

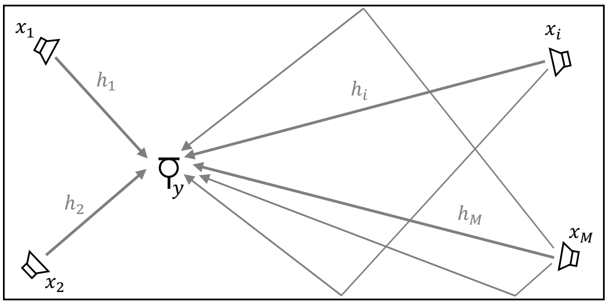


Figure 1: *Relation of the receiver and the sources. x_i, y, and h_i represent the i-th source, the received signal, and their room impulse response, respectively.*

3.1. RMS delay spread in multipath environments

The root mean square (RMS) delay spread is one of the important measures in understanding a multipath channel. The RMS delay spread σ_{rms} is defined as the standard deviation of the power of $h(t)$ in (1), as given by

$$\sigma_{rms} = \sqrt{\frac{\sum_k a_k^2 (b_k - \bar{b})^2}{\sum_k a_k^2}} \tag{7}$$

where we omit the sub-script i for simplicity which denotes the source index and $\bar{b} = (\sum_k a_k^2 b_k)/\sum_k a_k^2$ represents the mean delay. If the RMS delay spread σ_{rms} is relatively short compared to the symbol duration, the intersymbol interference is prevented, i.e., the channel can be considered flat [5]. For this reason, when designing the source signal, the symbol length should be set long enough to deal with the multipath effects.

3.2. Source data structure for AIL

The near-far effects occur when the desired source and the interfering sources overlap in time slots and frequency bands. In wireless communications, the system increases the transmitting power so that it becomes strong enough to retrieve the desired signal regardless of the interference or attenuation, which is called the power control. Unfortunately, it is impractical to apply the technique to AIL, since it is an one-way communication so there is no way to feedback the channel information or signal power. An alternative method to deal with the near-far effects is to design the source signals such that they reside in non-overlapping frequency regions.

In order to achieve this, we borrow an idea from the orthogonal frequency division multiple access-code division multiplexing (OFDMA-CDM) approach [6] for the acoustic data structure design. OFDMA is a multiple access scheme that divides the bandwidth into closely spaced multiple subcarriers assigned to each source. OFDMA-CDM employs OFDMA for multiple access and additionally applies CDM for the diversity of each source.

Figure 2 shows how the proposed scheme generates the i-th source signal $x_i(t)$. The vector $\boldsymbol{d}_i$ represents the data symbol assigned to the i-th source. Each data symbol $\boldsymbol{d}_i$ is spread by pseudo-random sequence which we chose as the Gold sequence $\boldsymbol{g}$. The resulting transmission vector $\boldsymbol{p}_i$ is interleaved evenly onto its subcarriers by the source-specific frequency mapper producing the source symbol $X_i(f)$. The source symbol $X_i(f)$ described in the frequency domain is then transformed into the corresponding time domain source $x_i(t)$ by inverse Fourier

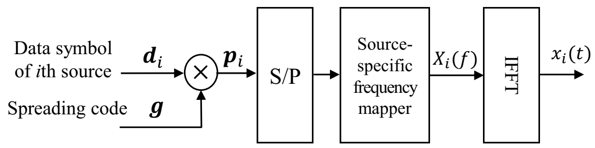


Figure 2: *Schematics of generating the source data. Data symbol $\boldsymbol{d}_i$ of each source is spread by a unique sequence $\boldsymbol{g}$ followed by source-specific frequency mapping.*

transform. Interested readers are encouraged to refer to [5, 6] for the basic structure and process of OFDMA-CDM.

In this method, the source-specific frequency mapper assigns each source to different subcarriers in a specified order, which has the following advantages: First, by mapping each source in non-overlapping frequency regions, it enables to mitigate the near-far effects. Second, each source signal becomes robust due to the combined effect of interleaving and diversity from structural characteristics of OFDMA-CDM, which is advantageous on the commonly used microphones and loudspeakers because their frequency responses cannot be regarded flat.

3.3. Direct path detection in synchronization

Another important issue of multipath environments comes up in the synchronization process. Due to the reverberation, the power of the first reflection component of a source signal sometimes becomes similar to or even stronger than that of the direct path component [7]. This becomes a problem because indirect path components are emphasized in the cross-correlation function. Most of the conventional AIL systems focus on detecting or enhancing the highest peak in the cross-correlation function. Only few studies have attempted to consider the preceding signals [8]. Therefore, when processing the correlation, we need to consider not only the prominent peaks but also the precedence of those peaks in estimating the time delays.

We propose a simple but effective method by refining the synchronization process in (3) and (4). In this method, we derive the modified peak time indices $\{\tau'_i\}$ in the following way:

$$\tau'_i = \begin{cases} \hat{\tau}_i, & \text{if } R_i(\hat{\tau}_i) > \lambda R_i(\tau_i), \\ \tau_i, & \text{otherwise} \end{cases} \tag{8}$$

where τ_i indicates the time delay obtained from (3) and $\hat{\tau}_i$ denotes the preceding peak as following

$$\hat{\tau}_i = \underset{\tau_i - W_{pre}+1 \leq \tau < \tau_i}{\text{argmax}} R_i(\tau). \tag{9}$$

W_{pre} and λ denote the length of the searching window and the threshold ratio of the peak, respectively. In our experiments, W_{pre} and λ were determined empirically based on the room acoustics [7].

4. Performance evaluation

In this section, we evaluated the performance of the proposed acoustic source data structure and the synchronization algorithm in simulated reverberant environments. As a performance measure, the root mean square error (RMSE) between the measured and the estimated positions was calculated over the entire data symbols for each target position.

Table 1: *Simulation environment.*

Parameters	Value
Dimension (m)	8×10
Loudspeaker position (m)	(1,1), (7,1), (1,9), (7,9)
RT_{60} (sec)	0.5, 1.1

Table 2: *System configuration.*

Parameters	Value
Sampling rate (kHz)	48
Frequency band (kHz)	18 - 21.5
Symbol length (samples)	4096, 8192
W_{pre} (samples)	500
λ	0.65

4.1. Experimental setup and system configuration

To evaluate the performance of the proposed approach, we conducted a series of experiments in simulated two-dimensional environments. The environmental parameters are listed on Table 1. Four sound sources were placed at each corner of the simulated room. When simulating the reverberation of the room, the RIRs were generated by Allen and Berkley's image method [9]. Acoustic signal was generated using frequency band above 18 kHz for inaudibility. The symbol length of the source data was arranged to be sufficiently longer than the average of RMS delay spread σ_{rms}. The synchronization parameters λ and W_{pre} was set empirically based on the room acoustics. The configuration of the AIL system for evaluation is given on Table 2.

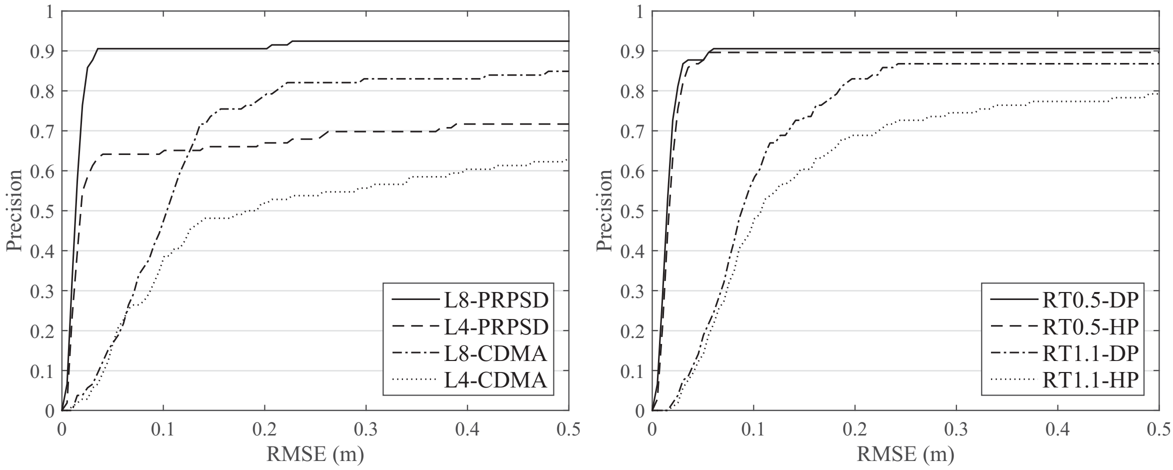
4.2. Evaluation of acoustic data structure

The first simulation was conducted to confirm the structure of the acoustic source data. The simulation result in Figure 3(a) shows the comparison between the proposed OFDMA-CDM data scheme (PRPSD) and a conventional CDMA-based scheme for different data symbol lengths of 8192 (L8) and 4096 (L4) samples. The CDMA-based system was implemented as a conventional structure from [1] with some modification for fair comparison. The RT60 was fixed to 0.5 sec for the reverberant environment simulation and the average RMS delay spread calculated from the generated RIRs was 1120 samples. All the signals used in this experiment had the same sampling rate, frequency band and symbol length as shown in Table 2.

We can see that longer data symbol length cases (L8) are more robust than the shorter ones (L4). This is because the data symbol length in this experiment was sufficiently longer than the RMS delay spread which the channel can be regarded flat. For the structural difference, the PRPSD outperformed the CDMA-based source data structure. This can be accounted for by the fact that the CDMA-based technique separates the source data in the code domain resulting overlaps in the frequency domain, which is not suitable for mitigating the near-far effects.

4.3. Performance of synchronization technique in reverberant environments

The second experiment was conducted to evaluate the effect of the direct path detection algorithm in various reverberant environments. Using the parameters of the previous experiment, the data symbol length was fixed to 8192 samples (L8) and performance was measured by the PRPSD. For simulation of severe



(a) (b)

Figure 3: *Performance evaluation of the proposed system: (a) comparison of the conventional structure with the proposed structure in different data symbol lengths, and (b) performance of proposed synchronization method in various reverberant conditions.*

reverberant environments, the RIR with RT60 of 0.5 and 1.1 sec cases were generated and denoted as RT0.5 and RT1.1, respectively. The case of direct path detection algorithm and conventional synchronization algorithm in (3) are denoted as DP and HP, respectively. From the result shown in Figure 3(b), we can see that DP outperformed HP in all environments. One notable observation from this experiment is that the performance of DP was significantly better than HP in severe reverberation (RT1.1). This result indicates that there exists more portion of stronger indirect signals in reverberant environments. With the PRPSD and DP, more than 80% of the points show RMSE less than 0.2 m which means that it is applicable in reverberant environments.

5. Conclusion

In this paper, the indoor localization system using inaudible acoustic signals that operates in real environments is proposed. We have focused on designing a structure of acoustic source to deal with the reverberation and near-far effects. The direct path detecting algorithm in synchronization process is proposed to cope with multipaths in reverberant environment. The proposed scheme have shown good performance in a series of simulated reverberant room. In the future, we plan to further study potential interference on the AIL systems such as ambient, white or transient noises.

6. Acknowledgements

This research was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A2A2A01045874), and by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1016) supervised by the IITP(Institute for Information & communications Technology Promotion).

7. References

[1] C. Sertatıl, M. A. Altınkaya, and K. Raoof, "A novel acoustic indoor localization system employing cdma," *Digital Signal Processing*, vol. 22, no. 3, pp. 506–517, May 2012.

[2] N. Aloui, K. Raoof, A. Bouallegue, S. Letourneur, and S. Zaibi, "Performance evaluation of an acoustic indoor localization system based on a fingerprinting technique," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–16, Jan. 2014.

[3] P. Lazik and A. Rowe, "Indoor pseudo-ranging of mobile devices using ultrasonic chirps," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, 2012, pp. 99–112.

[4] F. Hoflinger, R. Zhang, J. Hoppe, A. Bannoura, L. Reindl, J. Wendeberg, M. Buhrer, and C. Schindelhauer, "Acoustic self-calibrating system for indoor smartphone tracking (assist)," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012, pp. 1–9.

[5] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX.* Wiley, 2008.

[6] S. Kaiser and W. A. Krzymien, "Performance effects of the uplink asynchronism in a spread spectrum multi-carrier multiple access system," *European transactions on telecommunications*, vol. 10, no. 4, pp. 399–406, Jul. 1999.

[7] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality.* Springer Berlin Heidelberg, 2007.

[8] K. W. Wilson and T. Darrell, "Learning a precedence effect-like weighting function for the generalized cross-correlation framework," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 2156–2164, Nov. 2006.

[9] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, Apr. 1979.