Poster Abstract: Enhancing Fingerprint-based Smartphone Localization Using Acoustic Time-of-Flight for Complex Indoors

Yukiya Mita, Hiroaki Murakami, Takuya Sasatani, Matthew Ishige, Yoshihiro Kawahara
The University of Tokyo Japan
{mita,murakami,sasatani,mishige,kawahara}@akg.t.u-tokyo.ac.jp

ABSTRACT

Robust indoor positioning systems provide stable location-aware applications, enhancing our daily experiences. Fingerprint-based positioning techniques enable estimation of a user's position in complex indoor environments. While previous studies have used the received signal strength indicator or power spectral density as fingerprints, they typically achieved only submeter accuracy. This paper presents Geometric Sound Profile (GSP) as a novel location fingerprint to elevate the performance ceiling of fingerprint-based positioning. GSP is derived from the cross-correlation of transmitted and received signals based on transmission time, and a user's position is computed using weighted k-nearest neighbors. Our experiments demonstrate a median error of 0.66 m, marking a significant advancement over previous fingerprinting techniques.

KEYWORDS

Acoustic Sensing, fingerprint, indoor localization, time of flight

ACM Reference Format:
Yukiya Mita, Hiroaki Murakami, Takuya Sasatani, Matthew Ishige, Yoshihiro Kawahara. 2023. Poster Abstract: Enhancing Fingerprint-based Smartphone Localization Using Acoustic Time-of-Flight for Complex Indoors. In *The 21st ACM Conference on Embedded Networked Sensor Systems (SenSys '23), November 12–17, 2023, Istanbul, Turkiye.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3625687.3629535

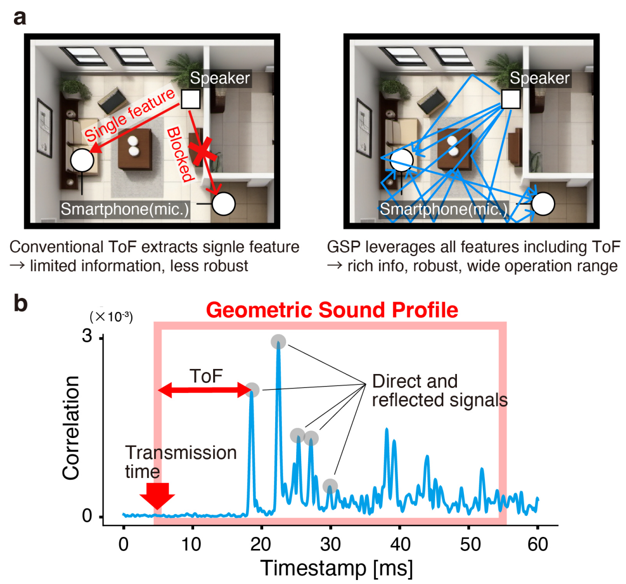

Figure 1: (a) The distinction between the conventional ToF approach and our GSP approach, (b) Details of GSP.

1 INTRODUCTION

With modern individuals predominantly spending their day indoors, 'indoor location information' is now critical for facilitating indoor activities. Robust positioning systems are paramount in aiding emergency evacuations and supporting visually impaired individuals, especially within intricate indoor environments. With the complexities of using the global navigation satellite system indoors, numerous alternative indoor positioning methodologies have surfaced. Among them, our work is based on acoustic localization using speakers and microphones, which has two major advantages: (i) Speakers exist ubiquitously in indoor public spaces, and most smartphones have microphones. (ii) Acoustic signal processing makes it possible to compute arrival times of direct and reflected signals at a smartphone's application layer.

SenSys '23, November 12–17, 2023, Istanbul, Turkiye

ACM ISBN 979-8-4007-0414-7/23/11…$15.00
https://doi.org/10.1145/3625687.3629535

Typically, acoustic positioning employs trilateration, leveraging signals from at least three speakers to calculate a microphone's location [1]. However, complex and enclosed real-world indoor environments often pose challenges, including limited speaker availability and obstructions leading to non-line of sight (NLoS) conditions. To navigate these challenges, fingerprint-based positioning methods, which utilize signal characteristics as location fingerprints, have been introduced. The most prevalent among these methods leverage the received signal strength indicator (RSSI) of Wi-Fi or Bluetooth, albeit with positioning errors spanning several meters due to the variability of RSSI [2]. Wang *et al.* used the power spectrum density (PSD) of acoustic signals from four speakers as a location fingerprint [4]. While this method surpasses the accuracy of using RSSI, its performance lags considering the number of speakers utilized. In indoor environments, wireless or acoustic signals consist not only of a direct signal but also of reflected signals originating from walls, furniture and so on, as shown in Fig. 1a. When metrics such as RSSI or PSD are employed, these direct and reflected signals are combined and represented as a single value, leading to a loss of the inherent spatial characteristics and information of the signal.

In this paper, we present Geometric Sound Profile (GSP) as a novel location fingerprint to enhance fingerprint-based positioning performance. GSP, derived from the cross-correlation of received

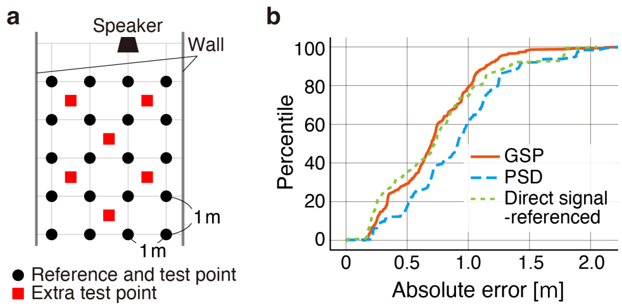

Figure 2: (a) Experimental setup and (b) cumulative distribution function of the positioning error

and transmitted signals combined with the transmission time, captures richer geometric features compared to RSSI or PSD (detailed explanation in §2). We employ GSP as our location fingerprint and use weighted k-nearest neighbors (WKNN) for localization.

2 GEOMETRIC SOUND PROFILE

The Geometric Sound Profile (GSP) introduces two novel aspects: it pivots on the cross-correlation of the transmitted signal from a speaker and received data via a microphone, and segments this cross-correlation by transmission time (see Fig. 1b). Existing methods such as RSSI and PSD, generate fingerprints from superimposed signals, preventing the full utilization of information from multiple paths. Our approach, employing cross-correlation, captures characteristics of multiple propagation paths as distinct peaks. Moreover, the use of absolute timestamps - based on the signal transmission time - allows us to determine the time-of-flight (ToF) for multiple signals. This provides a direct link to the spatial distance information of the smartphone's location. Thus, GSP shows a stronger correlation between location and profile than RSSI or PSD, suggesting that GSP offers advanced features beyond existing profiles.

In acoustic positioning with independent speakers and smartphones, estimating transmission time is tough. Despite this, recent advancements have made the estimation of transmission time feasible [3], laying the groundwork for our approach. We employ chirp signals for transmission, which aid in separating reflected signals. It's also assumed that the necessary parameters of the transmission signal for cross-correlation processing on the smartphone are received via nearby Wi-Fi.

3 POSITIONING EVALUATION

Matching Algorithm. Building upon previous work [4], we employ a WKNN approach for position estimation. GSP is derived by extracting a 50 ms segment from the cross-correlation function, starting at the transmission time and fingerprints are computed as the averages of GSPs obtained from receiving 20 chirp signals at each reference point. During the localization phase, we use the Euclidean distance between the calculated GSP from received signals and the fingerprints at each reference point. We select k ($k > 2$) reference points successively with smaller Euclidean distances and apply weighting based on the Euclidean distance, and calculate position.

Experiment Setup. For the transmission system, we used a PC (MacBook Pro) to generate audio signals, paired with an audio interface (ROLAND RUBIX24), an amplifier (Fostex AP20d), and speakers (Fostex FT28D). The speaker outputs a chirp signal in the range of 4 kHz to 10 kHz with a duration of 20 ms, and a transmission interval of 250 ms. On the receiving end, we used a Google Pixel 5 as the microphone, setting its sampling rate to 48 kHz. All data analysis was conducted offline. The experiment was conducted in an environment with a ceiling height of 4.0 m. The z coordinate of the speaker and microphone were set to 2.3 m and 1.0 m, respectively. As shown in Fig. 2a, 20 reference points served as location fingerprints. We also introduced six test points, leading to a total of 26 points for evaluating positioning performance.

Result. To evaluate the effectiveness of GSP, we performed comparative experiments against PSD and cross-correlation referenced to the direct signal. Fig. 2b shows the cumulative distribution functions of the positioning error for these three fingerprint types, and the meddian errors were 0.663 (GSP), 0.913 (PSD), and 0.731 m (direct signal-referenced), respectively. Our findings underscore that GSP, which references the transmission time in its cross-correlation, outperforms the direct signal-referenced cross-correlation. Moreover, the accuracy of the GSP-based method improved by 27.4 % compared to the PSD-based method.

4 CONCLUSION AND FUTURE WORK

This paper introduces Geometric Sound Profile (GSP), a novel location fingerprint designed to enhance the accuracy of fingerprint-based localization. Our approach achieved a median error of 0.66 m, demonstrating a substantial improvement over conventional fingerprinting techniques. Future work will evaluate the efficacy of GSP as a location fingerprint under various environments (*e.g.,* NLoS environments and large spaces), diverse conditions (*e.g.,* speaker position and dynamic environment) and different sound sources (*e.g.,* types of signals, frequency bands). Additionally, we intend to investigate deep learning models tailored for extracting and learning from ToF information. Our overarching goal is to develop methods that effectively leverage ToF information, enhancing the utility and accuracy of GSP-based positioning across diverse scenarios.

Acknowledgements. This work was supported by a joint research program with Daikin Industries, Ltd. (Research Institute for an Inclusive Society through Engineering, Space Design for Co-creation).

REFERENCES

[1] Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, and Paul Webster. 1999. The anatomy of a context-aware application. In *Proc. ACM/IEEE MobiCom'99*. 59–68.

[2] Suining He and S.-H. Gary Chan. 2016. Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons. *IEEE Communications Surveys & Tutorials* 18, 1 (2016), 466–490.

[3] Takumi Suzaki, Hiroaki Murakami, Masanari Nakamura, Hiroki Watanabe, Hiromichi Hashizume, and Masanori Sugimoto. 2022. PT-Sync: COTS Speaker-based Pseudo Time Synchronization for Acoustic Indoor Positioning. In *Proc. IEEE IPIN'22*. 1–8.

[4] Zexing Wang, Ruizhi Chen, Shihao Xu, Zuoya Liu, Guangyi Guo, and Liang Chen. 2021. A novel method locating pedestrian with smartphone indoors using acoustic fingerprints. *IEEE Sensors Journal* 21, 24 (2021), 27887–27896.