

鏡像スピーカを用いた スマートフォン高精度3次元測位手法

村上 弘晃^{1,a)} 中村 将成¹ 橋爪 宏達² 杉本 雅則¹

受付日 2019年3月4日, 採録日 2019年9月11日

概要: 本稿では, スピーカ 2 台から送信される音響信号の直接波と反射波を用いたスマートフォンのための新たな高精度 3 次元測位手法について述べる. 我々は, 壁と床からの 1 次反射波を, 壁と床に対するスピーカの鏡像体からの信号ととらえる, 鏡像スピーカの新概念を新たに提案する. また, スマートフォンを水平にもち, 2 つの搭載マイクロフォンを利用することで実在スピーカおよび鏡像スピーカからの信号を特定する. 2.5 m の間隔でスピーカを 2 台設置し, フロアがリフレクティブな廊下とアブソプティブな廊下にて各々 12 点での測位実験を行った. 評価実験を通して, 90th-percentile での測位誤差が 37.7 mm 未満であることと, 最大偶然誤差が 5.22 mm 未満であることを確認した. さらに, Dilution of Precision (DoP) やシミュレーションを用いた考察を行い, 鏡像スピーカの有用性を明らかにした.

キーワード: 音響測位, スマートフォン, マルチパス, 鏡像スピーカ, フロアプラン

Smartphone Accurate 3-D Localization Using Mirror Image Speakers

HIROAKI MURAKAMI^{1,a)} MASANARI NAKAMURA¹ HIROMICHI HASHIZUME² MASANORI SUGIMOTO¹

Received: March 4, 2019, Accepted: September 11, 2019

Abstract: In this paper, we describe a novel high-accuracy 3-D positioning method for a smartphone using direct and reflected acoustic signals from two existing installed speakers. We propose a new concept called *Mirror image speaker* regarding primary reflected signals from a wall and a floor as signals transmitted from a virtual speaker at a mirror image position of its original one. We hold a smartphone horizontally, and specify a real speaker and mirror image speakers by using two microphones built-in a smartphone. In our experiments, we installed speakers at intervals of 2.5 m along a corridor and estimated the smartphone position at several places on a reflective and an absorptive floor. From these results, we obtained 90th-percentile errors of less than 37.7 mm and maximum random errors of less than 5.22 mm for 3-D positioning. Furthermore, we clarified the usefulness of mirror image speakers using Dilution of Precision (DoP) measures and confirmed it through computer simulations.

Keywords: acoustic localization, smartphone, multi-path, mirror image speakers, floor plan

1. はじめに

IoT の発展にともない, 様々な場面において位置情報を

求められる動きが強まっている. 広く使われている位置推定システムに GPS (Global Positioning System) があるが, 屋内では GPS の信号が届かず, ショッピングモールや地下街, 美術館でのガイドはいまだにパンフレットや看板に頼らざるを得ない場面が多い. そのため, 屋内位置認識技術に対する要望は高い.

屋内位置認識手法の 1 つに音響測位がある. スマートフォンを用いる音響測位では, 一般的に伝搬時間ではなく, 到来時間差を利用する TDoA (Time Difference of Arrival) 方式を用いる. これは, 位置計測で必要な送信側

¹ 北海道大学大学院情報科学研究科情報理工学専攻
Department of Computer Science and Information Technology, Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan

² 国立情報学研究所
National Institute of Informatics, Chiyoda, Tokyo 101-8430, Japan

a) hmurakami@ist.hokudai.ac.jp

のスピーカと受信側のデバイスのマイクロ秒精度での時刻同期が困難なためである。

TDoAによる音響測位は、電波強度を用いるRSSI (Received Signal Strength Indication) に比べ、位置の推定精度が高いという利点がある。また、設置済みの既存スピーカを使用することができれば、追加の設備なしに測位が可能となる。しかし、到来時間差を利用するTDoA方式では、3次元測位の場合最低4台のスピーカを必要とする。建物内には複数のスピーカが設置されているが、廊下ではスピーカが一行に配置されている場合が多い。受信機であるスマートフォンから遠距離にあるスピーカの音響信号は、スマートフォンの最短距離にあるスピーカからの音響信号と比較し減衰が大きく、通常のスピーカ配置で検出可能なスピーカの数は3台未満である場合が多い。そのため、できる限り少数のスピーカによる位置推定手法の開発は重要な課題である。

そこで我々はスピーカ2台から送信される直接波と、反射波を利用したスマートフォンのための新たな高精度3次元位置推定手法を提案する。本手法では、測位誤差の要因となる壁や床からの反射波を、壁や床に対する鏡像スピーカからの信号ととらえることで、仮想的にスピーカの数を6台に増加させる。鏡像スピーカの位置は既知のフロアプランに基づき設定され、設置スピーカ2台による高精度3次元測位を可能とする。また、スマートフォンを水平にもち、2つの搭載マイクロフォンを利用することで、判別の難しい実在スピーカおよび鏡像スピーカからの信号を特定する。評価実験では、実際の環境を想定し、反射率の高いリフレクティブなフロアと反射率の低いアブソプティブなフロアの2種類にて測位実験を行った。各々12点での実験に基づき、90th-percentileでの測位誤差が37.7mm未満であることを確認した。加えて、スマートフォンの方角に依存する測位誤差を実測し、十分実用的であることを確認した。さらに、提案手法の高さ方向に対する測位精度や、障害物に対する影響をDoP (Dilution of Precision) とシミュレーションを用いて議論し、鏡像スピーカの有効性を明らかにした。

本稿の貢献は下記の3点である。

- 鏡像スピーカの提案と有効性の検証
- スマートフォンに内蔵された2つのマイクロフォンを利用しての3次元測位手法の提案
- 各マイクロフォンでの高精度な3次元測位の実現

2. 関連研究

これまでに、音響信号を利用した様々な測位手法が提案されている。たとえば、Active Bats [1] は、送信機と受信機間での伝搬時間を利用するToA (Time of Arrival) 方式を用いた超音波位置認識システムである。屋内に多数

の超音波受信機を取り付け、Batタグと呼ばれる超音波送信機の位置を推定する。文献 [1] では広帯域超音波を使用しており、送信機と受信機を時刻同期した3次元測位において、95th-percentileで測位誤差0.09m未満を達成している。Cricket [2] では、屋内にRF信号と超音波信号を送信するビーコンを設置し、2つの信号の受信時刻からToA方式により受信機の位置を推定する。このシステムは、AoA (Angle of Arrival) によりターゲットの角度を計測するCricket Compass [3] に拡張されている。5つの超音波受信素子を受信機にL字型に実装することで、角度計測における 2π 周期の曖昧性を回避している。上記のようなToA方式による位置計算では、非常に高精度な測位を実現できる。

しかし、スマートフォンを含む市販のモバイルデバイスでは、ToAで必要となる送受信機間の高速かつ正確な時刻同期の実現は難しく、たとえばスマートフォンカメラと照明を用いた時刻同期の例 [4] などを除いて、同分野の報告はなされていない。そのため、時刻同期を必要としないTDoA方式による位置認識が広く行われている。ASSIST [5] は、スマートフォンからチャープ信号を送信し、TDoA方式を用いて2次元位置を計算する。屋内に設置したマイクロフォンが利用できない時は、スマートフォン内蔵の加速度センサとカルマンフィルタを用いて位置計測を行う。また、Nakamuraら [6] は狭帯域信号で高精度な受信時刻検出が可能な位相一致法 (PAM) [7] を拡張したFDM-PAMを提案し、TDoAによる高正確度な3次元測位を実現している。文献 [6] では、“sync pattern” と呼ばれる信号を周波数分割多重化にて送信することで毎秒7.8回の測位を可能とし、偶然誤差2.8cm以下を達成している。TDoA方式では、到来時間差を利用するため、3次元測位を行うために4台以上のスピーカを使用する。そして、ホールのような大きな空間については、そのような数のスピーカ設置が可能である。しかし、廊下のようにスピーカ数が限られる環境では必要数のスピーカ設置は困難で、TDoAによる3次元での音響測位の実現をできない場合がある。

近年、音響信号の反響音を利用したレーダーシステムが多く提案されている。BatMapper [8] では、スマートフォンから送信する音響信号の反響音を利用し、建築物のマップを生成する。上部と下部のマイクロフォンごとにチャープ信号をベースとした異なる最適な音響信号を作成しており、マップ作成においてドアの検出も可能としている。SAMS [9] では、BatMapper [8] よりも帯域幅を大きくとったFMCW (Frequency Modulated Continuous Wave) を用いた、より高精度なマップ生成手法を提案している。同様にスマートフォンから送信する音響信号の反響音を利用し、BatTracker [10] では、スマートフォンの相対位置のトラッキングを可能としている。一方、Oliverら [11] は部屋の中央にtweeterを設置し、高周波チャープ信号を送信

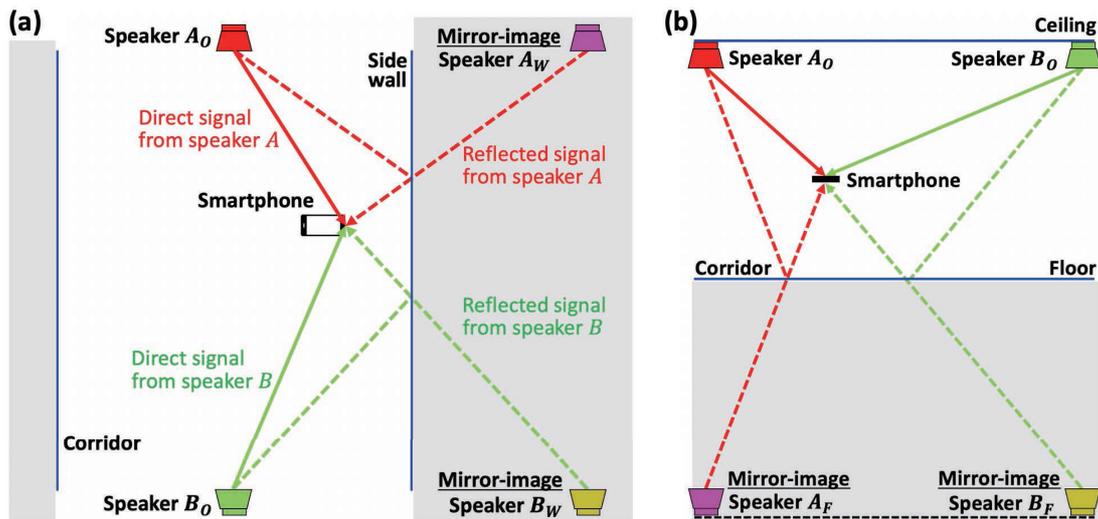


図 1 提案する鏡像スピーカの概要
 Fig. 1 The overview of our proposed mirror image speakers.

する．反響音を含む受信信号の周波数応答の変化から部屋内の占有率を計測する手法を提案している．ただこれらは形状測定システムであり，位置計測を目的としたものではない．

音響測位では，設置スピーカからの直接波を利用することが一般的で，反響音は測位誤差の要因となるマルチパスとして知られている．本稿では，マルチパスを壁や床に対する鏡像スピーカからの信号ととらえ利用することで，スピーカ 2 台による高精度な 3 次元測位を実現する．音響測位ではマルチパスを利用することはこれまで試みられていないが，電波測位では近年マルチパスを利用する研究が注目されている．SALMA [12] は，マルチパスとフロアプランを用いた UWB (Ultra Wide Band) の測位手法で，送信機 1 台での 2 次元測位を行った．送信機には 4 つの指向性アンテナを設置し，1 つの無指向性アンテナを搭載したデバイスで部屋の 4 つの側壁からのマルチパスを受信する．90th-percentile で 20 cm 以下の測位誤差を達成している．

3. 提案手法

3.1 鏡像スピーカの概要

廊下のような狭い環境では，壁や床からの反響音（マルチパス）が発生し，測位誤差の要因となることが知られている．我々は，このマルチパスを積極的に使用し，壁や床による鏡像スピーカからの信号（図 1）ととらえることで，スピーカの数を仮想的に補強する手法を試みる．

提案手法では信号強度を確保するため，物体に 2 回以上反射した 2 次以上の反射波は利用せず，壁・床に 1 回反射した 1 次反射波のみを利用する．このとき，スピーカは天井に設置されていると想定されるため，天井に対する 1 次反射波は生じない．スマートフォンに内蔵される 2 つのマイクロフォンを利用し，自身の座標を計算する．スマート

フォンは図 1 のように水平に持つことを想定しており，上部のマイクロフォンを左マイクロフォン，下部のマイクロフォンを右マイクロフォンとして扱う（以降，左マイクロフォンおよび右マイクロフォンと呼称）．このとき，右マイクロフォンでは，マイクロフォンの指向性により左壁に対する 1 次反射波を安定して受信することは難しい．同様に左マイクロフォンでは，右壁に対する 1 次反射波を安定して受信することが困難である．ゆえに，右マイクロフォンでは右壁に対する 1 次反射波を，左マイクロフォンでは左壁に対する 1 次反射波を利用する．

1 つのマイクロフォンで，実在するスピーカ A_0 と，壁に対する鏡像スピーカ A_W （図 1(a)）と床に対する鏡像スピーカ A_F （図 1(b)）からの信号を受信可能である．同様に実在スピーカ B_0 についても B_W と B_F からの信号を受信可能である．このとき，各鏡像スピーカの 3 次元座標はフロアプランより計算できる．以降は実在スピーカおよび鏡像スピーカをスピーカ S_K ($S \in \{A, B\}$, $K \in \{O, W, F\}$) と総称する．

3.2 位置計算手法

送受信機間での音響信号の伝搬時間から距離を計算し位置を算出する ToA 方式では，最低 3 台のスピーカを用いることで 3 次元測位を行うことが可能である．しかしすでに述べたように，送受信機間での時刻同期を必要とするため，スマートフォンへの応用は難しい．一方の TDoA 方式は，時刻同期が不要な位置計算手法である．そのため，本稿では TDoA 方式を用いて位置計算を行う．TDoA 方式では複数の送受信機間の到来時間差（式 (1)）から受信機の位置を計算する．送信機 i と受信機間の距離を r_i ($i \in \{1, \dots, k, \dots, n\}$) とすると，

$$r_k - r_i = c(t_k - t_i), \quad (1)$$

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$$

このとき、送信機 i の位置を (x_i, y_i, z_i) 、受信機の位置を (x, y, z) 、送信機 i からの信号受信時刻を t_i 、音速を c とする。上述のとおり、3次元座標を計算するためには最低4台 ($n \geq 4$) のスピーカを必要とする。

提案手法では、実在スピーカおよび鏡像スピーカの合計6台での TDoA 測位を行う。そのため、式 (2) に最小二乗法を適用することで3次元座標を算出する。

$$J_k = \sum_{i=1}^6 d_{ki}(x, y, z)^2, \quad (2)$$

$$d_{ki}(x, y, z) = r_k - r_i - c(t_k - t_i)$$

このとき、 J_k は送信機 k を参照する残差平方和であり、 d_{ki} は送信機 k と送信機 i の座標と到来時間差から得られる関数である。

3.3 信号特定手法

スマートフォンの左右のマイクロフォン ($M \in \{L, R\}$) の位置を計算するために、スピーカ S_K から送信される信号の到来時刻 $t_{S_K}^M$ を特定する必要がある。本節では、廊下のフロアプランと左右のマイクロフォンの基線長に着目した信号特定手法を提案し、左右のマイクロフォンの座標を推定する。提案手法での処理の流れを図2に示す。

3.3.1 送信信号

送信信号にはソナーやレーダーで使用されるチャープ信号を用いる。以下が使用する線形チャープ $a(t)$ である。

$$a(t) = \sin \left(\phi_0 + 2\pi \left(f_0 t + \frac{k}{2} t^2 \right) \right), \quad (3)$$

$$k = \frac{f_1 - f_0}{T}$$

このとき、 f_0 を開始周波数、 f_1 を終了周波数とし、 ϕ_0 を初期位相 (本稿では $\phi_0 = 0$)、 T を f_0 から f_1 までのスweep時間とする。生活雑音や人間の声で生じる音響の周波数帯域を避けるため、12–18 kHz の周波数帯域を用い、 $T = 10$ ms とする。本稿では、使用する2つのスピーカを判別するために、スピーカ A_O ではダウンチャープ ($f_0 = 18$ kHz, $f_1 = 12$ kHz)、スピーカ B_O ではアップチャープ ($f_0 = 12$ kHz, $f_1 = 18$ kHz) を用いる。また、1秒間に2回の頻度で測位を更新することを想定し、2つの信号を250 ms ごとに交互に送信する。

3.3.2 事前処理

左右のマイクロフォン $M \in \{L, R\}$ で得られる受信信号の時系列データと、スピーカ $S \in \{A, B\}$ からの送信信号との相互相関値を各々算出し、包絡線 $E_S^M(t)$ を求める。これにより、直接波ならびに反射波の到来時刻を包絡線のピークとして算出することができる。また、外れ値を除去する

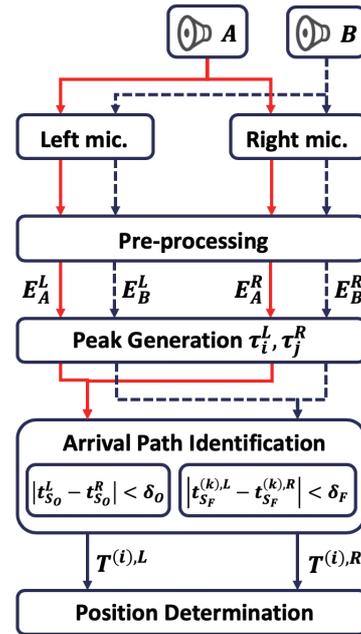


図2 提案手法の処理の流れ

Fig. 2 Processing flow of the proposed method.

ため、 m 次平滑化を行う。本稿では、サンプリングレート48 kHz に対し $m = 10$ とした。

3.3.3 ピークの生成

スピーカ S からの到来時刻 $t_{S_K}^M$ を推定するために、時間窓を設定する。2次以上の反射波では大きく減衰することが多く、得られる包絡線 $E_S^M(t)$ の最大値は、実在スピーカ S_O および鏡像スピーカ S_W, S_F のいずれかからの送信信号であることを実験より確認している。特に、スマートフォンとスピーカが近接している場合には、 S_O からの信号が最大となり (図3(a))、スマートフォンとスピーカが離れている場合には、 S_W や S_F からの信号が最大となる (図3(b))。また、実在スピーカ S_O からの信号が最も早く到来するため、フロアプランに応じて、各マイクロフォンでの直接波と反射波の到来時間差 $|t_{S_W}^M - t_{S_O}^M|$ または $|t_{S_F}^M - t_{S_O}^M|$ の最大値 η を算出することが可能である。これより、 $E_S^M(t)$ が最大となる到来時刻 $t_{max,S}^M$ を基準とし、 $[t_{max,S}^M - 1.5\eta, t_{max,S}^M + 1.5\eta]$ で時間窓をとる (図3)。時間窓において、 η ではなく 1.5η としているのは、反射波の重畳により2次以降の反射波の振幅が大きくなる場合を想定し、マージンを大きくとるためである。

次に $t_{S_K}^M$ の候補となるピークの到来時刻 τ_i^M ($i \in \{1, 2, \dots\}$) を算出する。このとき、 i は得られるピークの時系列での順序を示している。 $[t_{max,S}^M - 1.5\eta, t_{max,S}^M + 1.5\eta]$ において、下記の条件を満たす t を τ_i^M とする。

- $[t - \alpha, t + \alpha]$ で $E_S^M(t)$ が最大
- $E_S^M(t) > 2\mu_S^M$

このとき、 μ_S^M は $[t_{max,S}^M - 1.5\eta, t_{max,S}^M + 1.5\eta]$ における $E_S^M(t)$ の平均値である。また、 α はピークを検出するための

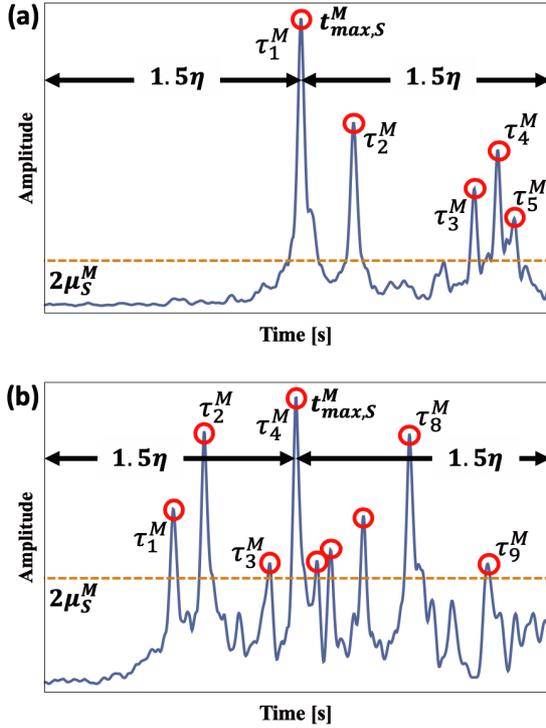


図 3 ピークの生成例：スマートフォンとスピーカが (a) 近接している場合、(b) 離れている場合

Fig. 3 Example of peak generation: cases where a smartphone microphone and an installed speaker are (a) close, (b) remote.

時間窓長のマージンである。実験では送信信号の周波数帯域幅 6 kHz に基づく時間分解能が $1/(6 \times 10^3) = 1.67 \times 10^{-4}$ (s) であることを考慮したうえで、自己相関により生じるサイドローブの誤検出を防ぐため、2 倍の $\alpha = 3.33 \times 10^{-4}$ (s) と設定した。図 3 は、条件を満たす τ_i^M を示している。なお、 η の値は実験によって設定される (4.1.2 項)。

3.3.4 到来パスの識別

$E_S^L(t)$ と $E_S^R(t)$ で観測されるピークの到来時間差に着目することで、 $t_{S_O}^M$ および $t_{S_F}^M$ を推定する。また、 $t_{S_W}^M$ の推定については、マイクロフォンの反対側にある壁からの反射波を安定して取得することが難しいため、到来時間差による推定を実施しない。これは、先にも述べたが、マイクロフォンの指向性が強いことに起因する。

$t_{S_O}^M$ の推定。 得られた到来時刻の候補 τ_i^L, τ_j^R ($i, j \in \{1, 2, \dots\}$) を用いて $t_{S_O}^M$ 、すなわち直接波の到来時刻の推定を行う。各マイクロフォン間での直接波の到来時間差はマイクロフォンの基線長に依存し、マイクロフォンの基線長から得られる $|t_{S_O}^L - t_{S_O}^R|$ の最大値を δ_O と定義する。このとき、式 (4) を満たす最初の τ_i^L と τ_j^R の組み合わせ (図 4(a)) を、各マイクロフォンでの直接波の到来時刻とし、 $t_{S_O}^L, t_{S_O}^R$ とする。

$$|\tau_i^L - \tau_j^R| \leq \delta_O \quad (4)$$

$t_{S_F}^M$ の候補の算出。 同様に床からの反射波の推定を行い、

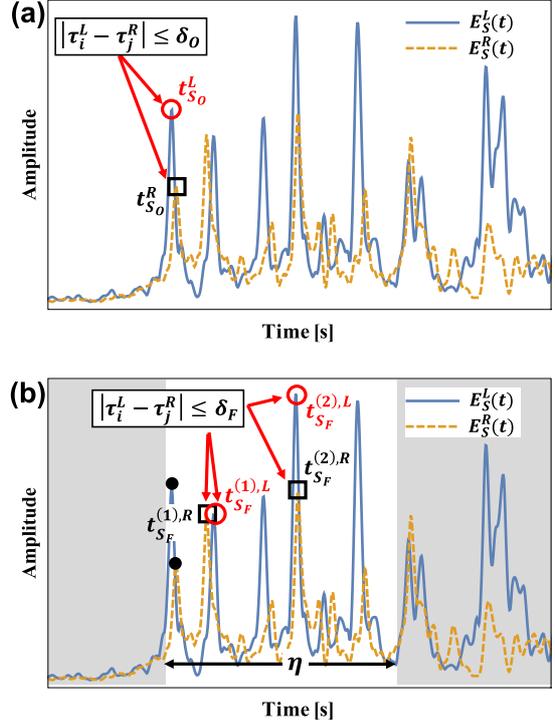


図 4 $E_S^L(t)$ と $E_S^R(t)$ で観測されるピークの到来時間差：(a) 実在スピーカ S_O での例、(b) 鏡像スピーカ S_F での例

Fig. 4 Time difference of arrival of peaks in $E_S^L(t)$ and $E_S^R(t)$: based on (a) real speaker S_O , (b) mirror image speaker S_F .

式 (5) を満たす τ_i^L と τ_j^R のすべての組み合わせ (図 4(b)) を $t_{S_F}^{(k),L}, t_{S_F}^{(k),R}$ ($k \in \{1, 2, \dots\}$) とする。

$$|\tau_i^L - \tau_j^R| \leq \delta_F \quad (t_{S_O}^M < \tau_i^M < t_{S_O}^M + \eta) \quad (5)$$

このとき、 δ_F はフロアプランとマイクの基線長に基づく $|t_{S_F}^L - t_{S_F}^R|$ の最大値である。

$t_{S_W}^M$ の候補の算出。 壁からの反射波では、左右のマイクロフォンの到来時間差に着目した候補の算出ができない。そのため、推定した $t_{S_O}^M$ を除くすべての τ_i^M ($t_{S_O}^M < \tau_i^M < t_{S_O}^M + \eta$) を $t_{S_W}^M$ の候補 $t_{S_W}^{(l),M}$ ($l \in \{1, 2, \dots\}$) とする。

$t_{S_K}^M$ の候補の組合せ。 上記を満たす、 $t_{S_K}^M$ の候補となる組合せの集合 $\mathbf{C}^{(i)}$ ($i \in \{1, 2, \dots\}$) を下記のように表現する。

$$\mathbf{C}^{(i)} = \{T^{(i),L}, T^{(i),R}\}, \quad (6)$$

$$T^{(i),L} = \{t_{A_O}^L, t_{A_F}^{(i),L}, t_{A_W}^{(i),L}, t_{B_O}^L, t_{B_F}^{(i),L}, t_{B_W}^{(i),L}\}$$

$$T^{(i),R} = \{t_{A_O}^R, t_{A_F}^{(i),R}, t_{A_W}^{(i),R}, t_{B_O}^R, t_{B_F}^{(i),R}, t_{B_W}^{(i),R}\}$$

このとき、 $t_{S_O}^M$ は一意に定まっており、 $t_{S_F}^{(i),M}$ は (5) 式、すなわち $|t_{S_F}^{(i),L} - t_{S_F}^{(i),R}| \leq \delta_F$ を満たしていることに注意する。なお、 δ_O, δ_F の値は実験によって設定される (4.1.2 項)。

3.3.5 位置の決定

各マイクロフォンの 3 次元位置 P_e^L および P_e^R を求める。到来時刻の候補 $\mathbf{C}^{(i)}$ を用いて、式 (2) の最小二

乗法により, $J_k^{(i,L)}, J_k^{(i,R)}$ を求め, 各マイクロフォンの3次元位置 $\{P_e^{(i,L)}, P_e^{(i,R)}\}$ を推定する. このとき, 以下の (7), (8), (9) 式を満たす $\{P_e^{(i,L)}, P_e^{(i,R)}\}$ を最終候補 $\mathbf{D}^{(j)} = \{P_e^{(j,L)}, P_e^{(j,R)}\}$ とする.

$$J_k^{(i,L)} \leq \beta \quad (7)$$

$$J_k^{(i,R)} \leq \beta \quad (8)$$

$$|P_e^{(i,L)} - P_e^{(i,R)}| \leq l_s + \gamma \quad (9)$$

式 (2) に関して, 6つの信号から得られる到来時間差は5つであり, 到来時間差から導出される5つの双曲面の交点として各マイクロフォンの座標を求めることができる. 求める3つの変数に対し5つの方程式 (TDoA) があるため, 求めるスマートフォンの座標に近いほど式 (2) における残差平方和 J の値は0に近づき, 一方で J が0に収束せず大きな値を取る場合は外れ値と考えられる. 本稿では本実験のスケールで得られた計測結果を基に明らかに外れ値となる残差平方和の閾値が0.01以上であることを確認し, 閾値 β を0.01と設定した. ゆえに, 式 (7) および (8) の条件を加えている. 本実験では, (1) 式における t_k を $t_{A_0}^M$ としている. また, 式 (9) では, マイクロフォンの基線長 l_s を利用することで推定位置の外れ値を除外している. このとき, γ はマージンであり, 実験に基づき $\gamma = 0.1$ とした.

すでに述べたように, 本手法ではマイクロフォンの指向性を考慮し, マイクロフォンの向きとは反対側の壁からの反射波は利用しない. しかし, その反射波が大きな振幅を持っていた場合, 図5のように壁に平行でスピーカ A_0, B_0 を結ぶ基線を通る面に対称な座標が推定されることがある. これを避けるために, $\mathbf{D}^{(j)}$ における4つの S_W からの信号の振幅 $E_S^M(t_{S_W}^M)$ に着目する.

$$\sum_{M \in \{L,R\}} \sum_{S \in \{A,B\}} E_S^M(t_{S_W}^M) \quad (10)$$

最終的に, 式 (10) が最大となる $\mathbf{D}^{(j)}$ をマイクロフォンの3次元位置 P_e^L および P_e^R の組合せとして決定する. これにより, 反対側の壁からの反射波を含む測位座標の組み合わせ $\mathbf{D}^{(j)}$ をつねに除外できることを実験により確認している.

4. 評価実験

4.1 実験設定

4.1.1 実験環境

本実験では, 図6(a)のとおり幅2mの廊下に2.5mの間隔で2台のスピーカを配置した. また, スピーカのz座標は1.9mとした.

送信スピーカに Fostex FT200D, スマートフォンに SONY Xperia XZ1, 信号発生器に NF WF1948, スピーカドライバとして独自のアンプを使用した. また, オーディオ録音のサンプリングレートを48kHzとした. この

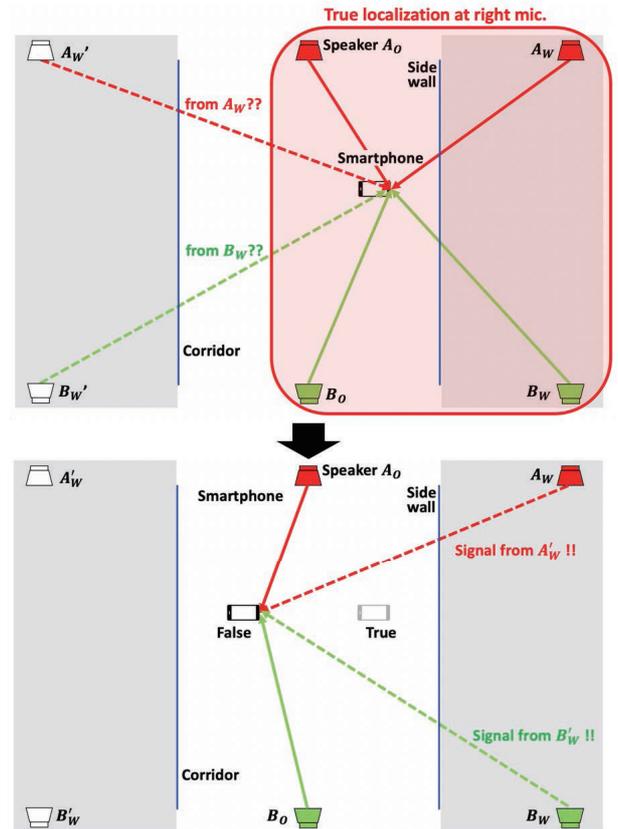


図5 対称な座標が推定されるケース
Fig. 5 Case where bilaterally symmetric coordinates are estimated.

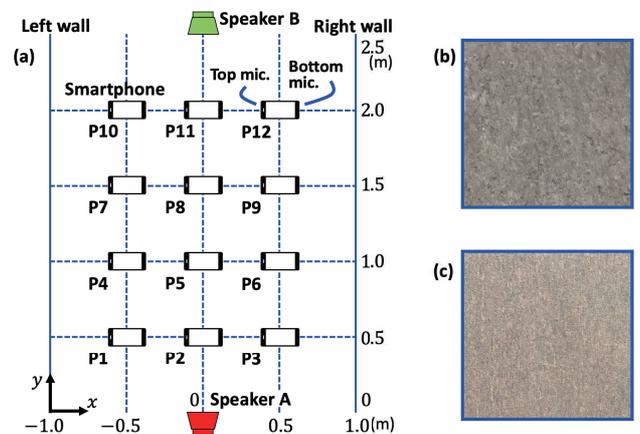


図6 実験環境: (a) xy 平面におけるスピーカの配置とスマートフォンの測位座標, (b) リフレクティブなフロアの材質, (c) アブソルプティブなフロアの材質
Fig. 6 Experiment environment: (a) speakers' placement and smartphones' positioning coordinates in xy -plane, (b) reflective floor material, (c) absorptive floor material.

とき, 2台のスピーカは信号発生器により同期しているものとする.

実験では, 環境音を含む実験下においてスマートフォンの各マイクロフォンで録音されたデータに対し, オフラインにて提案手法を適用し測位を行っている.

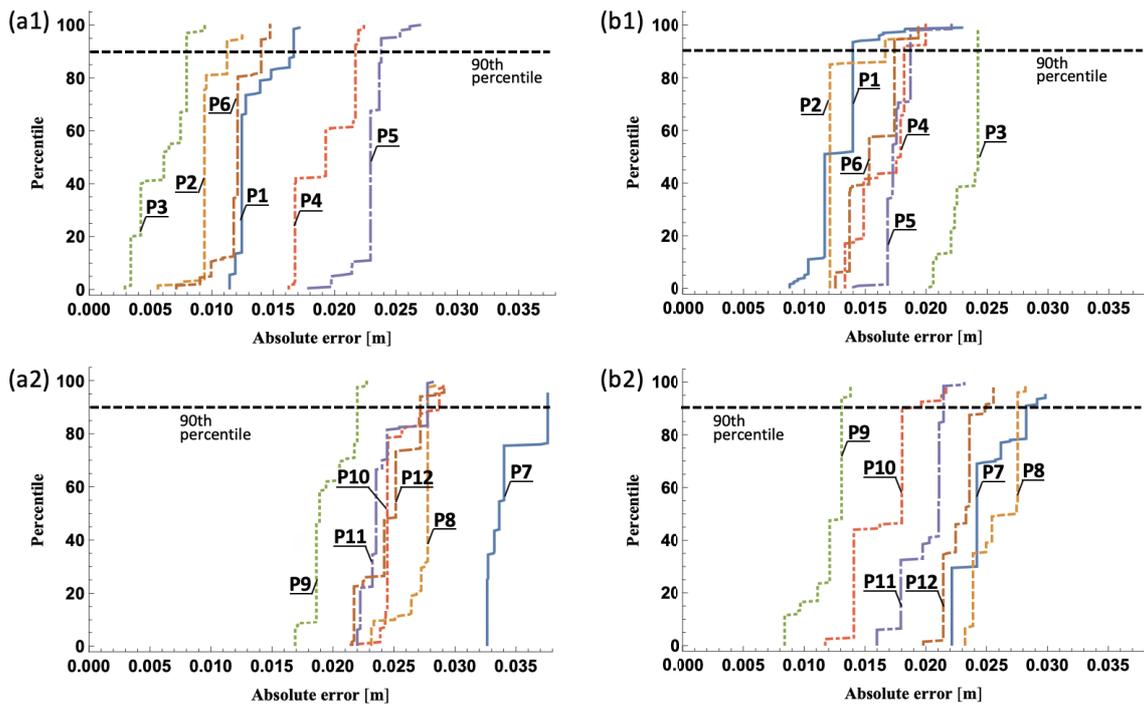


図 7 リフレクティブなフロアにおける累積誤差関数：(a) 左マイクロフォン, (b) 右マイクロフォン

Fig. 7 Cumulative error function on a reflective floor : (a) left microphone, (b) right microphone.

4.1.2 フloorプランとマイクロフォンの基線長に基づくパラメータの設定

提案手法で使用するフロアプランとマイクロフォンの基線長に基づくパラメータの設定方法について説明する。下記のパラメータの設定ではスピーカ A を想定する。これは、スピーカ B においても同様のことがいえるためである。また、音速は $c = 343.5 \text{ m/s}$ (気温 20 度) とする。

まず、 $|t_{S_F}^M - t_{S_O}^M|$ または $|t_{S_W}^M - t_{S_O}^M|$ の最大値 η を求める。本実験環境において $A_O = \{0, 0, 1.9\}$, $A_F = \{0, 0, -1.9\}$ であり、 $|t_{S_F}^M - t_{S_O}^M|$ が最大となるのは、スマートフォンに内蔵されたマイクロフォンの座標 p_i^M が A_O と A_F の基線長にある場合、すなわち $p_i^M = \{0, 0, z_c\}$ ($0 < z_c < 1.9$) となる場合である。本稿では、スマートフォンを手を持った場合の z 座標はたかだか 1.5 m と考え、 $z_c = 1.5$ とした。このとき、 $|t_{S_F}^M - t_{S_O}^M| \approx 8.73 \times 10^{-3}$ (s) である。同様に、 $A_O = \{0, 0, 1.9\}$, $A_W = \{\pm 2.0, 0, 1.9\}$ において $|t_{S_W}^M - t_{S_O}^M|$ が最大となるのは、 $p_i^M = \{\mp 1.0, 0, 1.5\}$ のときであり、 5.67×10^{-3} (s) となる。ゆえに、 $\eta \approx 8.73 \times 10^{-3}$ (s) と設定した。

次に、 $|t_{S_O}^L - t_{S_O}^R|$ が最大値 δ_O をとるのは $p_i^L = \{1.0, 0, 1.5\}$, $p_i^R = \{1.0 - l_s, 0, 1.5\}$ または、 $p_i^L = \{-1.0 + l_s, 0, 1.5\}$, $p_i^R = \{-1.0, 0, 1.5\}$ のときであり、 $\delta_O \approx 3.90 \times 10^{-4}$ (s) である。このとき、 l_s はスマートフォンに内蔵されたマイクロフォン間の基線長であり、 $l_s = 0.146$ である。

同様に、 $|t_{S_F}^L - t_{S_F}^R|$ が最大値 δ_F をとるのは $p_i^L =$

$\{1.0, 0, 0\}$, $p_i^R = \{1.0 - l_s, 0, 0\}$ または、 $p_i^L = \{-1.0 + l_s, 0, 0\}$, $p_i^R = \{-1.0, 0, 0\}$ のときであり、 $\delta_F \approx 1.86 \times 10^{-4}$ (s) が得られる。

4.2 反射率が異なるフロアでの測位実験

提案手法では、壁と床に対する反射波を利用するため、壁と床の材質による測位結果への影響が考えられる。特に、床ではカーペットのように信号を吸収する材質も使用されている。そこで、提案手法の有効性を確認するために、床の反射率が異なる 2 つの廊下にて測位実験を行った。計測回数は各々 200 回である。

4.2.1 リフレクティブなフロア

理想的な環境として、信号の反射率が高いリフレクティブなフロアにて測位実験を行った。このとき、床の材質は図 6(b) に示す硬質材である。また、スマートフォンを手を持った測位を想定し、P1–P12 におけるスマートフォンの高さは $z = 1.1 \text{ m}$ とした。

図 7 にリフレクティブなフロアにおける 3 次元での測位誤差の累積密度分布を示す。また、各マイクロフォンでの 90th-percentile での測位誤差は、各々 37.7, 28.3 mm 未満であった。

4.2.2 アブソープティブなフロア

次に、反射波の検出が難しいと考えられる、信号の反射率が低いアブソープティブなフロアにて測位実験を行った。床の材質は図 6(c) に示すカーペットである。実験では、

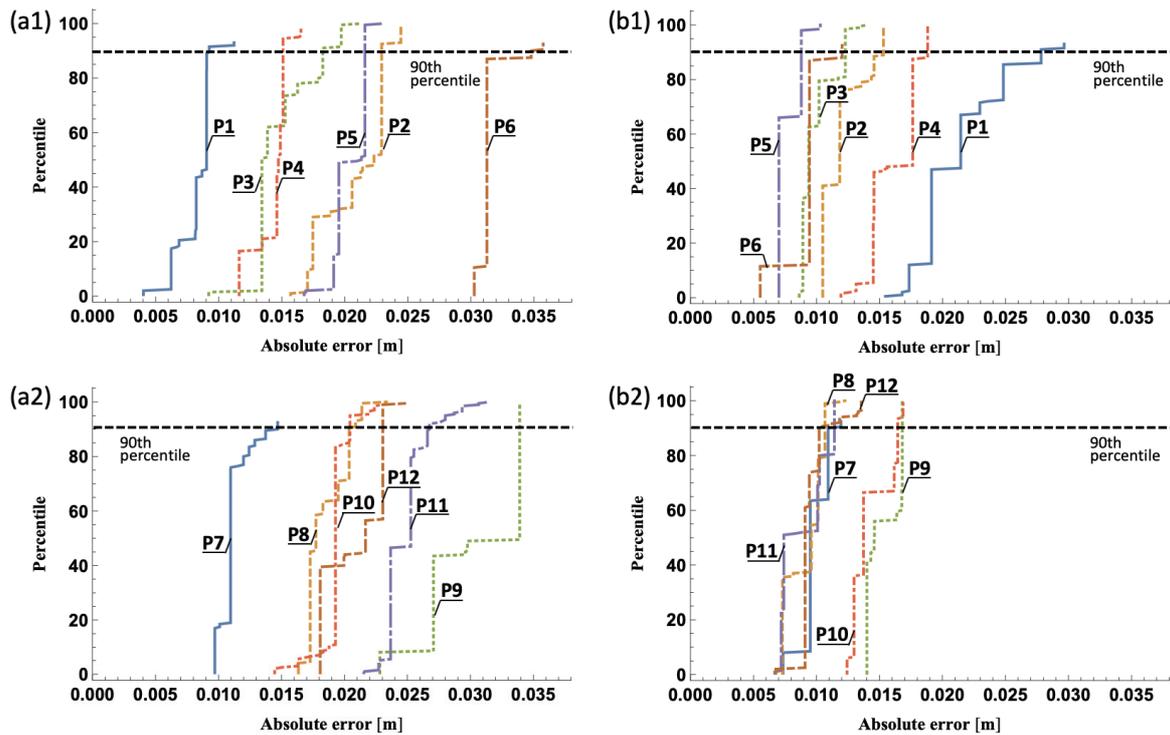


図 8 アブソーブティブなフロアにおける累積誤差関数 ($z = 1.1$ m) : (a) 左マイクロフォン, (b) 右マイクロフォン

Fig. 8 Cumulative error function on an absorptive floor at $z = 1.1$ m: (a) left microphone, (b) right microphone.

前項と同様に P1–P12 におけるスマートフォンの高さを $z = 1.1$ m とした。図 8 は、 $z = 1.1$ m でのアブソーブティブなフロアにおける 3 次元での測位誤差の累積密度分布である。このとき、各マイクロフォンでの 90th-percentile での測位誤差は、各々 34.8、27.9 mm 未満であった。

さらに、高さ方向に対する提案手法の有効性を確認するため、スマートフォンの高さを $z = 1.4$ m に設置し測位実験を行った。 $z = 1.4$ m における実験では、実験環境 (図 6 (a)) より P1–P6 と P7–P12 に対称性があることを考慮し、P1–P6 にて測位を行った。図 9 は、 $z = 1.4$ m でのアブソーブティブなフロアにおける 3 次元での測位誤差の累積密度分布であり、各マイクロフォンでの 90th-percentile での測位誤差は、各々 31.6、32.9 mm 未満である。

4.3 異なる角度での測位実験

提案手法では、スマートフォンを水平に持つことを想定している。そこで、スマートフォンを xy 平面上で回転させての測位実験を実施し、その限界を調査した。実験では、壁に対してスマートフォンが垂直な位置を 0° とし、 -75° から 15° ずつ 75° までスマートフォンを回転させての測位実験をアブソーブティブなフロアにて行った。また、計測座標は図 6 (a) の P5 ($z = 1.1$ m) とし、計測回数は 200 回とした。

図 10 は、スマートフォンによる測位誤差の箱

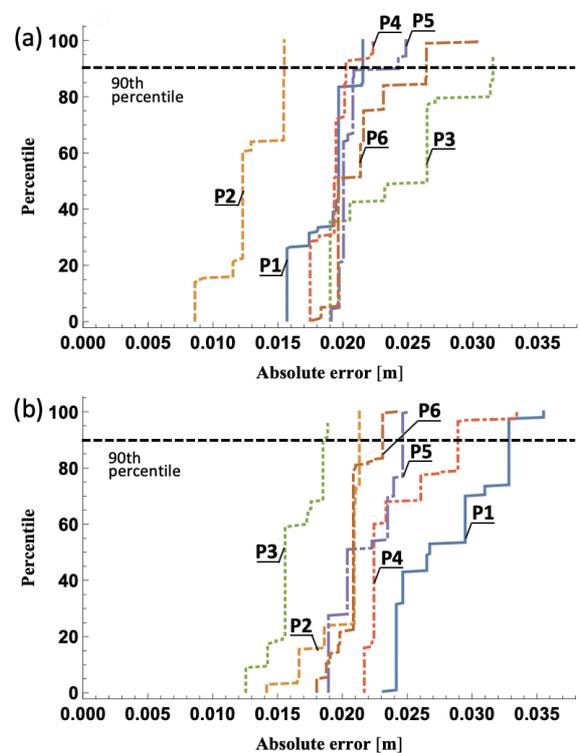


図 9 アブソーブティブなフロアにおける累積誤差関数 ($z = 1.4$ m) : (a) 左マイクロフォン, (b) 右マイクロフォン

Fig. 9 Cumulative error function on an absorptive floor at $z = 1.4$ m: (a) left microphone, (b) right microphone.

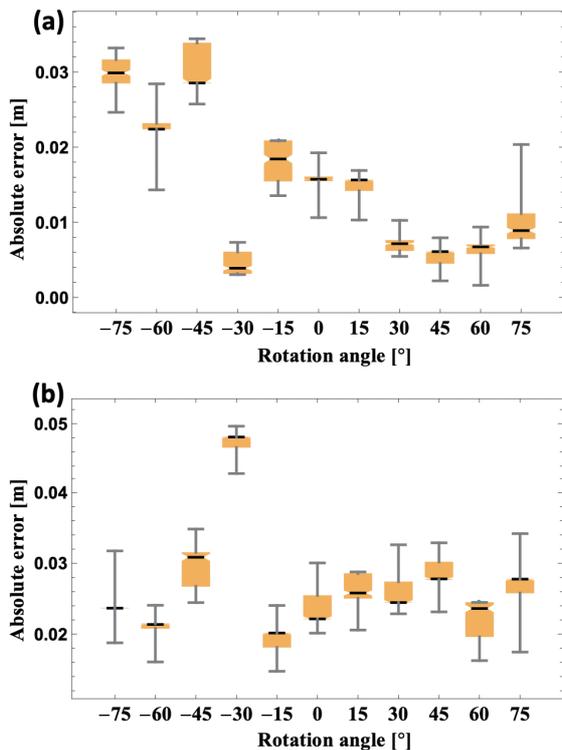


図 10 スマートフォンの角度による絶対測位誤差の箱ひげ図 (黒線は中央値を表している): (a) 左マイクロフォン, (b) 右マイクロフォン

Fig. 10 Box-and-whisker plot of absolute positioning error at different smartphone angles (black line represents median value): (a) left microphone, (b) right microphone.

ひげ図である。図 10 より、スマートフォンの xy 平面上での回転に対し、誤差は 1.6~49.7mm であり、十分実用的であることが確認できる。

5. 考察

5.1 測位誤差の評価

4.2 節における計測結果を系統誤差と偶然誤差を用いて評価する。ある計測位置 p での 200 個のデータを \mathbf{x}_i^p ($i \in \{1, 2, \dots, 200\}$)、その平均値を $\bar{\mathbf{x}}^p = (1/200) \sum_i \mathbf{x}_i^p$ 、真値を \mathbf{X}^p とおくと、

- 系統誤差: $|\mathbf{X}^p - \bar{\mathbf{x}}^p|$
- 偶然誤差: $\sqrt{(1/200) \sum_i |\mathbf{x}_i^p - \bar{\mathbf{x}}^p|^2}$

4.2 節における実験での最大系統誤差は 42.1 mm であり、位置認識として十分実用的な値である。誤差の要因としては、実験機材やスマートフォン設置時の測定誤差が影響しているものと考えられ、より厳密な機器を用いての測定により改善が見込めると考えられる。

特筆すべきは、最大偶然誤差の小ささであり、5.22 mm 未満を達成している。提案手法により非常に高精度な測位が実現できることを確認できた。

5.2 高さ方向における測位精度

提案する実在スピーカ 2 台を用いた鏡像スピーカでの 3 次元測位では、床に対する鏡像スピーカからの信号を取得することができる。そのため、従来の天井に設置されたスピーカ 4 台での 3 次元測位と比較し高さ方向 (z 座標) に対する測位精度の向上が見込まれる。3 次元位置認識における誤差は受信時刻計算の誤差によって生じるが、この誤差の位置計算に対する影響度は送信機と受信機の位置関係によって異なり、DoP を用いて評価できる。そこで、DoP を用いて、提案手法の高さ方向に対する測位精度の有効性を示す。TDoA における DoP の算出には Bingao ら [13] の論文を参照する。

天井に設置されたスピーカ 4 台での 3 次元測位について考える。本稿での実験環境を想定し受信機の座標を $P4 (z = 1.1)$ における左マイクロフォンの座標 $(-0.573, 1.0, 1.1)$ とし、天井に設置されたスピーカ 4 台のスピーカの座標を $A_1 = \{-1, 0, 1.9\}$, $A_2 = \{1, 0, 1.9\}$, $A_3 = \{-1, 0, 2.51.9\}$, $A_4 = \{1, 2.5, 1.9\}$ とする。このとき、高さ方向に対する Vertical DoP は 67.78 であり、床に対し平行な Horizontal DoP は 11.66 である。ゆえに、従来の天井に設置されたスピーカ 4 台での 3 次元測位では、高さ方向に対する測位精度が低くなりやすく、3 次元での測位が難しいことが確認できる。一方で、提案手法での鏡像スピーカ 4 台を含む 6 台 ($A_O, A_W, A_F, B_O, B_W, B_F$) での 3 次元測位では、Vertical DoP が 0.68, Horizontal DoP が 1.18 と、きわめて理想的な DoP をとることが確認でき、鏡像スピーカを用いた 3 次元測位での有用性が明らかである。

5.3 障害物を含む測位

提案手法では、信号の反射を利用するため、障害物による測位精度の影響が考えられる。特にスマートフォンを手にとって測位する場合、図 11 のように使用者自身が障害物となって、いくつかの信号が受信できないケースが考えられる。提案手法では、測位に必要な 6 台の信号をすべて受信できることを前提としているが、今後の課題として障害物を考慮する必要がある。そこで、3 次元測位に最低限必要なスピーカ 4 台からの信号のみが取得できた場合を想定し、DoP と実測値によるシミュレーションを用いて提案する鏡像スピーカの可能性を議論する。このとき、受信機であるスマートフォンの座標は前節と同様に、 $P4 (z = 1.1)$ における左マイクロフォンの座標 $(-0.573, 1.0, 1.1)$ を仮定する。また、実測値として、リフレクティブなフロアにおける $P4 (z = 1.1)$ での左マイクロフォンの受信信号を用いる。

5.3.1 3 次元測位が困難なケース

前節において、天井のみにスピーカが設置された場合での測位では、 z 座標に対する測位精度が低下することを述

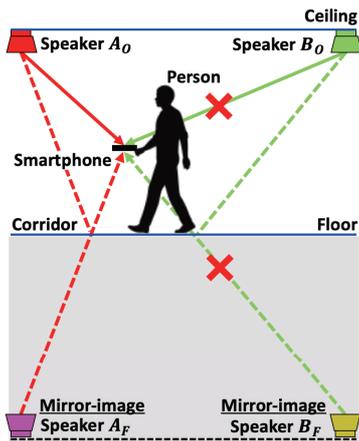


図 11 スマートフォンを手に保持した場合の信号の遮断
 Fig. 11 Received signal block when a person has smartphone in his hand.

べた。これは 4 台のスピーカが同一平面上にあることに起因する。そのため、 z 座標が同一である A_O, A_W, B_O, B_W では、位置に対する Position DoP は 89.48、実測値によるシミュレーションでは系統誤差が 0.60 m と 3 次元での測位は難しい。また、 x 座標が同一である A_O, A_F, B_O, B_F での 3 次元測位では、Position DoP が 89.56、実測値によるシミュレーションでは系統誤差が 0.57 m であり、同様に 3 次元での測位が難しいことを確認した。そのため、同一平面上にあるスピーカに対しては 3 次元での測位は難しいと考えられる。

しかし、スピーカ A_O, A_W, B_O, B_W における xy 平面での系統誤差は 0.15 m、スピーカ A_O, A_F, B_O, B_F における yz 平面での系統誤差は 0.04 m であり、得られたスピーカの組み合わせを考慮することで 2 次元測位として活用することが期待できる。

5.3.2 3次元測位が可能なケース

同一平面上に位置していない 4 台のスピーカを用いれば高精度の 3 次元測位が可能である。図 11 は、ユーザ自身によりスピーカ B_O, B_F からの信号が遮断され、 A_O, A_W, A_F, B_W からの信号は受信可能である場合を示している。このとき、6 台のスピーカすべてを用いた場合での Position DoP が 1.36 であるのに対し、スピーカ A_O, A_W, A_F, B_W での Position DoP が 1.93 である。また、実測値でのシミュレーションによる測位誤差の累積密度分布を図 12 に示す。図 12 では、6 台すべてを受信できた場合 (All) の測位誤差と、3 次元測位が可能な 2 つの例を取り上げている。DoP や図 12 より、提案手法は想定される障害物の影響に対しても十分実用的であることが期待できる。

5.4 反射波の重畳

提案手法では、直接波に加え反射波を利用するため、同一の信号を複数回取得することになる。このとき、受信機

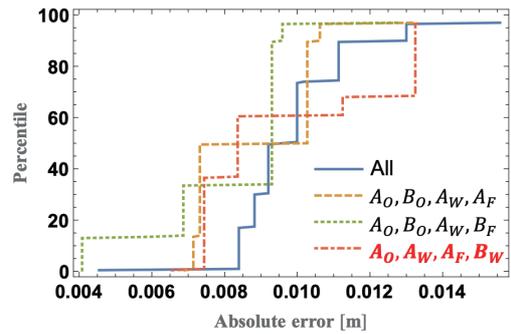


図 12 3次元測位が可能なスピーカの組合せによる測位精度の違い

Fig. 12 Positioning accuracy comparison by combination of 3-D positionable speakers.

の位置によっては直接波ならびに反射波が時間的に重なる場合がある。 S_O と S_W や S_O と S_F からの信号が重畳する受信機の位置は、壁や床の境界面であるため実用上問題は小さい。しかし、 S_W と S_F については重畳する可能性がある。そのため、提案手法では $t_{S_W}^M$ と $t_{S_F}^M$ が同時刻に到来することを想定し、受信信号から得られる t_S^M の候補 τ_i^M を $t_{S_W}^M$ と $t_{S_F}^M$ の 2 つの候補として重複することを許している。しかし、重畳によって本来の到来時刻との差異が生じるため測位精度の悪化につながることを考えられる。また、提案手法では式 (7)、(8) を利用しているため、本来の到来時刻との差異が大きいと測位することができない。今後の課題として、移動を想定する測位では、慣性センサとの併用やパーティクルフィルタを用いることで、式 (7)、(8) を利用しない提案手法の改善を検討中である。

6. むすび

本稿では、設置スピーカ 2 台から送信される音響信号の直接波に加えて、測位誤差の要因として知られるマルチパスを利用した新たな測位手法を提案した。提案手法では、壁と床からの 1 次反射波を壁と床に対する鏡像スピーカからの信号ととらえることで、スピーカの数を仮想的に補強し高精度な 3 次元測位を可能とした。また、スマートフォンに内蔵された 2 つのマイクロフォンを用いることで、本来判別が困難とされるマルチパスの判別を行い、実在および鏡像スピーカからの信号を特定する手法を提案した。評価実験により、90th-percentile での測位誤差が 37.7 mm 未満であることと、最大偶然誤差が 5.22 mm 未満であることを確認し、提案手法の有用性を確認した。さらに、DoP と実測値でのシミュレーションによる考察を行い、鏡像スピーカの有効性を明らかにした。

提案手法では、左右のマイクロフォンの座標を高精度に測位できるため、スマートフォンの方向を推定できることが期待される。また、マルチパスを活用するため、歩行時に背中が妨げとなり本来得られない信号の補強が可能であ

ることも期待される。今後は、スマートフォンによる歩行方向の絶対角度の推定とともに、障害物に対して頑健な移動時の測位手法を検討する。

本提案手法は位置推定の基礎技術と位置付けられる。そこで、複数スピーカの識別、Wi-Fi 測位、PDR (Pedestrian Dead Reckoning) などの手法を統合することにより、建物内を移動するユーザに対する実用的な屋内測位の実現を目指す。

謝辞 本研究は JSPS 科研費 JP19H04222 の助成を受けたものです。

参考文献

- [1] Harter, A., Hopper, A., Steggle, P., Ward, A. and Webster, P.: The anatomy of a context-aware application, *Wireless Networks*, Vol.8, No.2/3, pp.187–197 (2002).
- [2] Priyantha, N.B., Chakraborty, A. and Balakrishnan, H.: The cricket location-support system, *Proc. 6th Annual International Conference on Mobile Computing and Networking*, pp.32–43, ACM (2000).
- [3] Priyantha, N.B., Miu, A.K., Balakrishnan, H. and Teller, S.: The cricket compass for context-aware mobile applications, *Proc. 7th Annual International Conference on Mobile Computing and Networking*, pp.1–14, ACM (2001).
- [4] Akiyama, T., Sugimoto, M. and Hashizume, H.: Time-of-arrival-based indoor smartphone localization using light-synchronized acoustic waves, *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, Vol.100, No.9, pp.2001–2012 (2017).
- [5] Höflinger, F., Zhang, R., Hoppe, J., Bannoura, A., Reindl, L.M., Wendeberg, J., Bühner, M. and Schindelbauer, C.: Acoustic self-calibrating system for indoor smartphone tracking (assist), *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp.1–9, IEEE (2012).
- [6] Nakamura, M., Akiyama, T., Sugimoto, M. and Hashizume, H.: 3D FDM-PAM: Rapid and precise indoor 3D localization using acoustic signal for smartphone, *Proc. 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pp.123–126, ACM (2014).
- [7] Hashizume, H., Kaneko, A., Sugano, Y., Yatani, K. and Sugimoto, M.: Fast and accurate positioning technique using ultrasonic phase accordance method, *TENCON 2005: 2005 IEEE Region 10*, pp.1–6, IEEE (2005).
- [8] Zhou, B., Elbadry, M., Gao, R. and Ye, F.: BatMapper: Acoustic Sensing Based Indoor Floor Plan Construction Using Smartphones, *Proc. 15th Annual International Conference on Mobile Systems, Applications, and Services*, pp.42–55, ACM (2017).
- [9] Pradhan, S., Baig, G., Mao, W., Qiu, L., Chen, G. and Yang, B.: Smartphone-based Acoustic Indoor Space Mapping, *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol.2, No.2, p.75 (2018).
- [10] Zhou, B., Elbadry, M., Gao, R. and Ye, F.: BatTracker: high precision infrastructure-free mobile device tracking in indoor environments, *Proc. 15th ACM Conference on Embedded Network Sensor Systems*, p.13, ACM (2017).
- [11] Shih, O. and Rowe, A.: Occupancy estimation using ultrasonic chirps, *Proc. ACM/IEEE Sixth International Conference on Cyber-Physical Systems*, pp.149–158, ACM (2015).

- [12] Großwindhager, B., Rath, M., Kulmer, J., Bakr, M.S., Boano, C.A., Witrisal, K. and Römer, K.: SALMA: UWB-based Single-Anchor Localization System using Multipath Assistance, *Proc. 16th ACM Conference on Embedded Networked Sensor Systems*, pp.132–144, ACM (2018).
- [13] Li, B., Dempster, A.G. and Wang, J.: 3D DOPs for positioning applications using range measurements, *Wireless sensor network*, Vol.3, No.10, p.334 (2011).



村上 弘晃 (学生会員)

2016年北海道大学工学部情報エレクトロニクス学科卒業。現在、同大学大学院情報科学研究科情報理工学専攻博士課程在学中。屋内音響測位システムに関する研究に従事。



中村 将成 (正会員)

2014年北海道大学工学部情報エレクトロニクス学科卒業。2016年北海道大学大学院情報科学研究科情報理工学専攻修士課程修了。2018年北海道大学大学院情報科学研究科情報理工学専攻博士課程修了。博士(情報科学)。2016年より三菱電機株式会社情報技術総合研究所に勤務、現在に至る。



橋爪 宏達 (正会員)

1984年東京大学大学院工学系研究科博士課程修了。同年東京大学助手(文献情報センター)。1986年学術情報センター助教授。1989年～1990年ペンシルバニア大学客員助教授。1990年文部省学術調査官(併任)。1998年学術情報センター教授。2000年国立情報学研究所教授。2002年より総合研究大学院大学教授を併任し、現在に至る。



杉本 雅則 (正会員)

1990年東京大学工学部航空学科卒業。1995年同大学院工学系研究科博士課程修了。博士(工学)。同年学術情報センター(現、国立情報学研究所)助手。1999年東京大学助教授。2012年より北海道大学教授となり、現在に至る。