



HOKKAIDO UNIVERSITY

Title	音響を用いた屋内測位およびエコー埋め込みに関する研究
Author(s)	須崎, 太久弥
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第16500号
Issue Date	2025-03-25
DOI	https://doi.org/10.14943/doctoral.k16500
Doc URL	https://hdl.handle.net/2115/94921
Type	doctoral thesis
File Information	Takumi_Suzaki.pdf



令和 6 年度 博士論文

音響を用いた屋内測位および エコー埋め込みに関する研究

北海道大学 大学院情報科学院 情報理工学コース
数理科学分野 知能情報学研究室
須崎 太久弥

令和 6 年 12 月 2 日

Research on Acoustic Indoor Positioning and Echo Embedding

Takumi Suzuki

Laboratory of Intelligent Information Systems
Research Group of Mathematical Science
Course of Computer Science and Information Technology
Graduate School of Information Science and Technology
Hokkaido University

December 2 2024

概要

位置情報を利用したサービスの増加に伴い、測位技術の需要が高まっている。特に近年では、政府や企業による DX (デジタルトランスフォーメーション) の推進により、生産性の向上や業務効率化を目的とした位置情報の活用も期待されている。広く知られている測位技術として GPS を代表とする GNSS (Global Navigation Satellite System) が挙げられるが、屋内環境では人工衛星からの信号が外壁によって遮断される、もしくは大幅に減衰するため、測位精度が著しく低下する。今後、ショッピングモールや駅構内など、屋内でも位置情報を活用したサービスの展開が予想されるため、GNSS に代わる測位技術の開発が急がれている。

本論文では、屋内での高精度な測位技術を実現するため、音響を利用した測位手法を提案する。既存の屋内測位技術として Wi-Fi や Bluetooth を利用した電波ベースの手法、LED による可視光を利用した手法などが提案されているが、音響は測位精度とコストの面で優位性を持つ。これは音速は電波に比べて遅く、信号の受信時刻を高精度に検出できることに起因する。また、モバイルデバイスには標準でマイクロフォンが搭載されており、専用ハードウェアを必要としないため、低コストでの実現が可能である。これらの利点により、音響を用いた手法は屋内測位分野における有力な技術として注目されている。

音響を用いた測位手法は、測位精度や実装コストの面で優位性を持つ一方、いくつか課題も存在する。1 点目の課題として送受信機間の時刻同期が困難であることが指摘されている。音響測位では、信号が空気中を伝搬する時間を基に位置を推定するため、送信機 (スピーカ) と受信機 (マイクロフォン) の精密な時刻同期が一般的に必要とされる。時刻同期を不要とする TDoA (Time Difference of Arrival) アルゴリズムも存在するが、これを採用する場合、時刻同期済みの手法と比較して、測位計算に必要なスピーカの数が増加する。

2 点目の課題として高精度測位が可能な範囲 (カバレッジ) が限定的であることが指摘されている。モバイルデバイスに搭載されているマイクロフォンは指向性を持ち、その周波数特性によって測位精度が変化する。そのため、スピーカとモバイルデバイスの配置関係や環境の影響により、測位精度が低下することがある。

3 点目の課題として送信信号の帯域が可聴域と重複することが指摘されている。多くのモバイルデバイスでサポートされているサンプリング周波数は最大 48kHz であり、受信可能な周波数は 24kHz までに限定される。一方で、人の可聴域は 18kHz 程度であるため、送信

信号として利用可能な帯域は 18kHz から 24kHz に限られる．この帯域幅は測位用の信号として十分ではないため，測位精度の低下につながる可能性がある．

本論文では，これらの課題を解決するための改善手法を提案する．1 点目の課題に対しては，床からの反射波に着目し，送受信機間の擬似的な時刻同期を可能にする手法を提案する．この手法では，鏡像法を用いて床からの反射波を仮想スピーカからの直接波とモデル化することで，擬似的にスピーカ数を増加させる．これにより，受信側で幾何学的に送信時刻を推定することが可能となる．また，送信時刻の推定には，モバイルデバイスの高さ情報が必要となるが，これは音響センシングによる測距によって推定する．

2 点目の課題に対しては，TDM (時間分割多重) と FDM (周波数分割多重) を組み合わせた新たな送信方式を提案し，測位領域の拡張を試みる．この手法では，各タイムスロットで周波数が異なる信号を送信することで，周波数特性による測位精度の影響を抑制する．複数のタイムスロットを用いることで，送信信号の帯域を広げ，測位精度の向上が期待できる．

3 点目の課題に対しては，音響電子透かしの分野に着目し，エコーを測位用信号として BGM に埋め込む手法を提案する．この手法では，人間の聴覚が反響音に対して鈍感である特性を利用し，広帯域の信号を BGM に埋め込む．帯域と時間分解能にはトレードオフの関係があるが，広帯域化することで時間分解能を向上させ，高精度な測位が期待できる．

本論文は以上の 3 つの課題に対する解決手法を提案し，その有効性を示したものである．

目次

第 1 章	序論	9
1.1	研究背景	9
1.2	研究目的	10
1.3	研究課題	10
1.4	本論文の構成	12
第 2 章	関連研究	13
2.1	測距と測位	13
2.1.1	測距方式	13
2.1.2	測位方式	16
2.2	測位と定位	17
2.2.1	アクティブ型の測位	18
2.2.2	パッシブ型の測位	19
2.2.3	アクティブ型の定位	19
2.2.4	パッシブ型の定位	20
2.3	既存の位置推定技術	21
2.4	音響を用いた位置推定技術	24
2.5	音響センシング技術	25
2.6	音響情報ハイディング技術	26
第 3 章	鏡像法による送受信機間の擬似的な時刻同期手法	29
3.1	はじめに	29
3.2	基本原理	29
3.3	アクティブセンシングによる床からの高さ推定	30
3.3.1	送信信号の設計	30
3.3.2	高さ z の推定	32
3.4	送信時刻推定	33
3.4.1	送信時刻の特徴	33

3.4.2	ユーザのアクションに基づく選択	34
3.4.3	ToA ベースの測距・測位	34
3.5	予備実験：スマートフォンの周波数特性	36
3.6	時刻同期手法	38
3.6.1	実験設定	38
3.6.2	時刻同期精度の評価	39
3.6.3	測距精度の評価	39
3.6.4	測位精度の評価	42
3.7	考察	43
3.7.1	測距および測位誤差の原因	43
3.7.2	測位結果の曖昧性	45
3.8	まとめ	45
第 4 章	FDM と TDM を組み合わせた測位領域の拡張手法	47
4.1	はじめに	47
4.2	送信方式	47
4.3	受信信号の検出	49
4.4	コヒーレント積分	49
4.5	TDoA による測位	50
4.6	群遅延の補正	51
4.7	評価実験	51
4.7.1	事前実験 1	51
4.7.2	事前実験 2	51
4.7.3	測位実験の環境	53
4.7.4	測位実験	54
4.7.5	実験結果	57
4.8	考察	61
4.8.1	群遅延の補正	61
4.8.2	コヒーレント積分	62
4.8.3	タイムスロット数	62
4.8.4	FDM による測位	62
4.8.5	TDM による測位	63
4.9	限界と展望	63
4.10	まとめ	64
第 5 章	エコー埋め込みを用いた BGM による屋内測位	65

5.1	はじめに	65
5.2	提案手法	65
5.2.1	埋め込み方法	66
5.2.2	検出方法	67
5.2.3	測位アルゴリズム	70
5.3	評価実験	71
5.3.1	設定と実験環境	72
5.3.2	音質	73
5.3.3	音源のジャンルと測位精度	73
5.3.4	埋め込み強度と測位精度	74
5.3.5	PN 系列の長さや測位精度	74
5.3.6	計測位置と測位精度	74
5.4	考察	74
5.5	限界と展望	75
5.6	まとめ	77
第 6 章	結論と今後の展望	79
6.1	結論	79
6.2	今後の展望	79
参考文献		83

目次

2.1	測距の概略図	14
2.2	測距方式の比較	14
2.3	測位方式の比較	16
2.4	音響を用いた測位と定位の比較	18
3.1	時刻同期手法の概要	31
3.2	アクティブセンシングの概要	32
3.3	インコヒーレント積分の回数 N による波形の比較	33
3.4	各フレームで推定される送信時刻の候補	35
3.5	推定された送信時刻に基づく距離の組合せ p の集合	35
3.6	推定された送信時刻候補のクラスター	36
3.7	スマートフォンのスピーカとマイクロフォンの周波数特性：(左上) 上部のスピーカ, (右上) 下部のスピーカ, (左下) 上部のマイクロフォン, (右下) 下部のマイクロフォン	37
3.8	実験環境	38
3.9	送信信号の波形 (上) とその周波数 (下)	39
3.10	参照スピーカとマイクロフォンの位置関係	39
3.11	アクティブセンシングによる高さ推定の精度	41
3.12	距離による同期精度	42
3.13	各計測地点での測位結果	43
3.14	移動時の測位結果	43
3.15	測位誤差の累積分布関数 (CDF)	44
3.16	高さの推定誤差 Δz で生じるスピーカとスマートフォン間の距離 d_{xy} と測位誤差の関係	45
4.1	提案送信方式の概要	48
4.2	コヒーレント積分による包絡線のシミュレーション	50
4.3	実験設定と結果	52

4.4	機種ごとの群遅延の比較	52
4.5	実験環境	53
4.6	スピーカとスマートフォンの配置	53
4.7	各実験の測位結果 (□, ◇: 真の位置, ×, *: 平均推定位置)	58
4.8	各実験の誤差の累積分布関数 (CDF)	59
4.9	各測位点での測位計算に失敗した割合	60
4.10	各測位点での誤差が 1.0m 以下になった割合	60
4.11	FDM 方式により P13 で得られた包絡線と拡大図	61
5.1	提案システムの概要	66
5.2	検出手順の概要	67
5.3	到来する合成エコーカーネル $k_{rec}(t)$	68
5.4	(a) 受信信号より得られるケプストラム, (b) 各スピーカに割り当てた PN と の相関, (c)(b) の相関を (a) のケプストラムを用いてフィルタリングした結果	70
5.5	各種パラメータによる音質	71
5.6	音源のジャンル	71
5.7	埋め込み強度 α	71
5.8	PN の長さ L_{PN}	71
5.9	測位実験の結果	72

表目次

2.1	測位と定位の分類	18
2.2	各手法の特徴の比較	21
4.1	実験条件	55
4.2	測位実験の結果	55
5.1	音響処理に関するパラメータ	71
5.2	MOS: Mean Opinion Scores	72

第 1 章

序論

1.1 研究背景

位置情報を利用したサービスの増加に伴い、測位技術の需要が高まっている。特に近年では、政府や企業による DX（デジタルトランスフォーメーション）の推進により、生産性の向上や業務効率化を目的とした位置情報の活用が期待されている [1]。広く知られている測位技術として、GPS を代表とする GNSS (Global Navigation Satellite System) が挙げられるが、屋内環境では人工衛星からの信号が外壁によって遮断される、もしくは大幅に減衰するため、測位精度が著しく低下する [2]。今後、屋内でも位置情報を活用したサービスの展開が予想されるため、GNSS に代わる測位技術の開発が急がれている。

屋内測位技術の具体的な応用例として、以下のようなサービスが挙げられる。

- 大型商業施設、駅構内、空港、博物館などでのナビゲーション
- 災害時の避難誘導システム
- 人流の把握および行動分析
- 自律ドローンやロボットの自己位置推定
- 医療施設での位置把握

このように、産業、防災、医療、エンターテインメントなど、様々な分野で屋内測位技術の需要が高まっており、その経済効果は 2030 年までに世界で 1325 億米ドルに達すると予測されている [3]。

屋内測位技術の実現には、「デバイスフリー」と「デバイスベース」の2種類のアプローチが考えられる。「デバイスフリー」の手法では、屋内に設置されたカメラや人感センサ類を用いてターゲットの位置を推定する。ターゲット側では測位のための専用機材が不要であるため、煩しき等の負担が少ない利点がある反面、外部の機材設置コストやプライバシーの面で課題がある。他方で「デバイスベース」の手法では、人が保有するモバイルデバイスを

用いてその位置を推定する。アクセスポイントやビーコンを用いることで、既存のインフラを活用できるため、測位システムを普及させやすいという利点がある。本論文では、モバイルデバイスをターゲットにした「デバイスベース」のアプローチで、音響に着目した屋内測位技術についてまとめる。

1.2 研究目的

本論文では、モバイルデバイスをターゲットにした測位誤差 10 cm 以下になるような測位技術の開発を目的とする。10 cm 以下の誤差は様々なサービスの要求を満たすために必要な精度であり、例えば、隣り合う人との区別が十分可能な距離である。近接学 (proxemics) とよばれるパーソナルスペースに関する研究分野では、対人距離を密接距離、個体距離、社会距離、公衆距離に分類し、距離とコミュニケーションの関係性について論じている [4, 5]。日常生活において対人距離が最短である密接距離は 15~45 cm であるため、10 cm の誤差は人の位置を識別する上で十分な精度である。

また近年では、スーパーマーケットでの購買行動に関する分析も積極的に行われているが、測位誤差が数メートル単位であるため、商品カテゴリーの推定にとどまっている [6, 7]。測位誤差が数 cm レベルに達することで、商品単位レベルかつリアルタイムでの位置情報を推定できるようになり、推薦システムの実現や購買行動の分析において新たな知見を得られる可能性がある。

筆者の知る限り、屋内測位技術で最も古いものは IR (InfraRed; 赤外線) を利用した Active Badge [8] であり、部屋単位での人の位置を推定する手法である。その後も、電波を利用した RADAR [9] など、毎年新しい技術が提案されてきた。しかし、実際に社会に展開されている技術は測位誤差が数 m 以上のものにとどまっている。高精度な位置情報を利用するサービスの需要も踏まえると、測位誤差が 10 cm 以下で設備コストが安価な技術の開発が不可欠である。

1.3 研究課題

本論文では、GNSS のようにモバイルデバイスで受信可能であり、測位誤差が 10 cm 以下の要求を満たすことが期待できる音響に着目する。音響測位は GNSS と同様に、波が空間を伝搬する遅延時間を利用して位置計算を行う。音速 (340 m/s) は、光速 (3×10^8 m/s) と比較して空気中での伝搬速度が遅く、送受信時刻の検出が容易であるため測位精度が高いことが知られている。

現在までに、音響を利用した様々な手法が提案されているが課題も指摘されている。次に本論文で解決を目指す課題とその解決方針を述べる。

送受信機間の時刻同期

1 つ目の課題として、送受信機間の時刻同期が困難であることが挙げられる。音響測位では、信号が空気中を伝搬する時間を基に位置を推定するため、送信機（スピーカ）と受信機（マイクロフォン）のマイクロ秒単位の精密な時刻同期が求められる [10, 11, 12]。時刻同期を不要とする TDoA (Time Difference of Arrival) アルゴリズムも存在するが、測位計算に必要なスピーカの数が増加する制約が生じる。

本論文では、床からの反射波に着目し、送受信機間の擬似的な時刻同期を可能にする手法を提案する。この手法では、鏡像法 [13] によって床からの反射波を仮想スピーカからの直接波と捉える。これにより、擬似的にスピーカ数を増加させ、受信側で幾何学的に送信時刻を推定することが可能となる。送信時刻の推定には、モバイルデバイスの高さ情報が必要だが、この情報も音響センシングによる測距を用いて推定する。

測位領域の拡張

2 つ目の課題として、高精度測位が可能な範囲（カバレッジ）が限定的であることが挙げられる。モバイルデバイスに搭載されているマイクロフォンは指向性を持つため、その周波数特性によって測位精度が変化する。このため、スピーカとモバイルデバイスの配置関係や環境の影響により、測位精度が低下する場合がある。

本論文では、多重化方式である TDM (Time Division Multiplexing; 時間分割多重) と FDM (Frequency Division Multiplexing; 周波数分割多重) を組み合わせた新たな送信方式を提案し、測位領域の拡張を試みる。この手法では、各タイムスロットで周波数が異なる信号を送信することで、周波数特性による測位精度への影響を抑制する。複数のタイムスロットを用いることで、送信信号の帯域を広げることが可能となり、測位精度の向上が期待できる。

可聴域での測位

3 つ目の課題として、送信信号の帯域が可聴域と干渉してしまう点が挙げられる。多くのモバイルデバイスでサポートされているサンプリング周波数は最大 48 kHz であり、受信可能な周波数は高々 24 kHz までに制限される。一方で、人の可聴域は 18 kHz 程度であるため、送信信号として利用可能な帯域は 18 kHz から 24 kHz に限られる。測位が目的の場合、この帯域幅は十分ではないため、測位精度の低下につながる可能性がある。

本論文では、音響情報ハイディングの分野に着目し、エコーを測位用信号として BGM に埋め込む手法を提案する。この手法では、人の知覚が反響音に対して鈍感であることを利用して、広帯域の信号を BGM に埋め込む。帯域と時間分解能にはトレードオフの関係があ

り，広帯域であるほど時間分解能が向上するため，高精度な測位が期待できる。

1.4 本論文の構成

以下に本論文を構成する各章の概要を示す．第1章では，屋内測位技術の基礎概念および背景を説明し，本論文の概要と貢献を述べた．第2章では，位置推定を行う研究および音響に関する研究をまとめ，本論文の位置づけを示す．音響以外の既存の屋内測位技術にも言及し，特徴と性能を比較する．第3章では，床からの反射波に着目した疑似的な時刻同期手法を提案し，その基礎概念と詳細を説明する．また，実験による時刻同期の誤差および測位精度の評価を行い，提案手法の有効性を示す．第4章では，送信方式の工夫による測位領域の拡張手法を提案し，その基礎概念と詳細を説明する．また，実験による従来の送信方式との比較を行い，提案手法の有効性を示す．第5章では，エコーを BGM に埋め込むことによる測位手法を提案し，その基礎概念と詳細を述べる．また，実験による音源のジャンルや埋め込みパラメータによる測位精度の比較を行い，提案手法の有効性を示す．第6章では，本論文で提案した手法について得られた成果を総括し，今後の課題および展望について記述する．

第 2 章

関連研究

本章では、本論文の位置づけを明確にするために、関連研究について詳しく述べる。はじめに、測位と密接に関連する測距技術について説明し、その違いと特徴を述べる。ここでは測位アルゴリズムについても言及する。次に、測位と概念的に近い定位技術について説明し、分類を行いながらその違いについて述べる。続いて、音響以外の既存の位置推定技術について比較を行いながらまとめる。その後、音響を用いた既存の測位技術について説明する。続いて、分野的に近い音響を利用したセンシング技術について触れ、最後に 5 章で扱う音響情報ハイディング技術に関連する研究を説明する。

2.1 測距と測位

測位は測距を多次元に拡張した概念と言える。測距はターゲットとの距離を推定する技術であり、古くから軍事分野を中心に発展してきた歴史がある [14, 15]。対して測位は測距技術を応用して、ターゲットの位置を推定する技術と言える。測距の代表的な技術として、電波を利用した RADAR (レーダ; Radio Detection And Ranging), レーザ光を利用した LiDAR (Light Detection And Ranging), 音波を利用した SONAR (SOund Navigation And Ranging) などが代表的な測距技術である。

2.1.1 測距方式

測距方式は ToF (Time of Flight) 方式と RSS (Received Signal Strength) 方式に分類できる。ToF 方式は波が空間を伝搬する時間を利用するのに対し、RSS 方式は波の強度が距離によって減衰する性質を利用する。

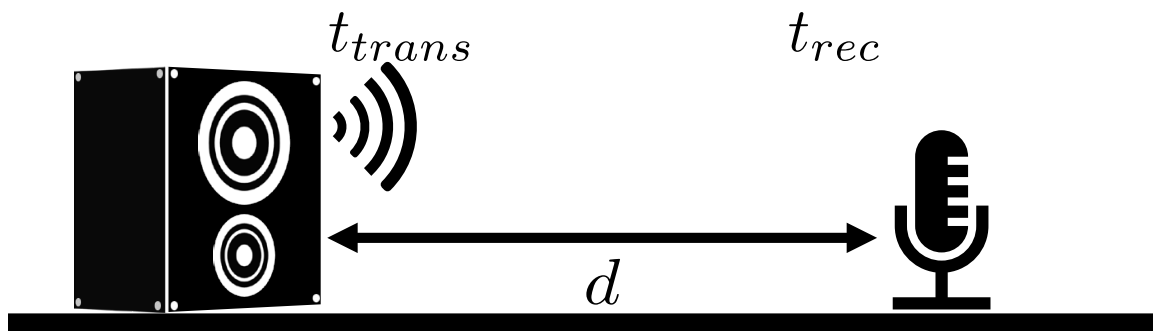
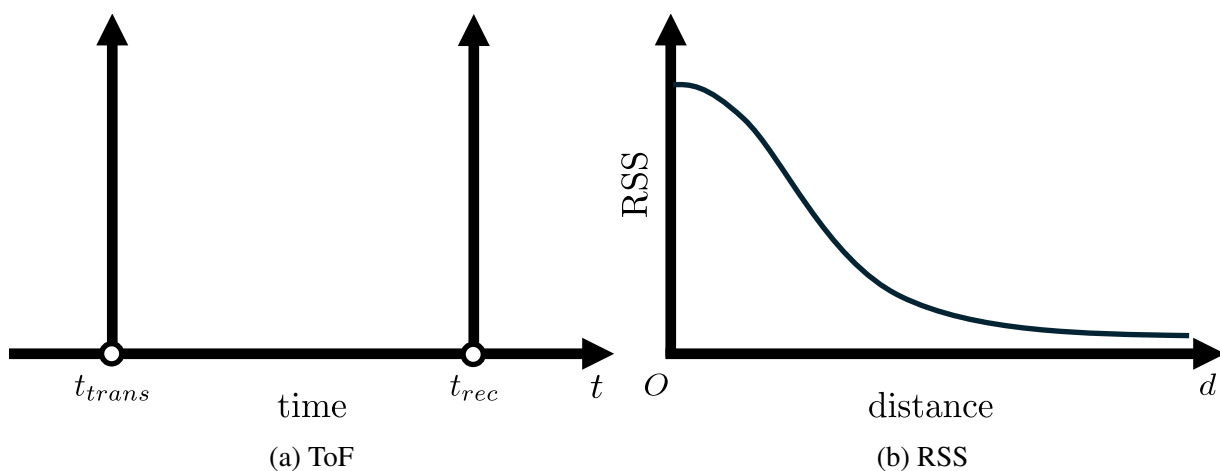


図 2.1: 測距の概略図



送信と受信時刻の差が距離に対応する。

RSS の減衰具合が距離に対応する。

図 2.2: 測距方式の比較

ToF 方式

まず図 2.1 のようにスピーカとマイクロフォン間の距離を推定することを考える。音波の場合、音速はおよそ 340 m/s であることが知られているため、この値を既知として図 2.2a の時間差から、以下の式で推定できる。ここで、 d は距離、 c は音速、 t_{rec} は受信時刻、 t_{trans} は送信時刻、 Δt_{sync} は同期誤差である。

$$d = c \cdot (t_{rec} - t_{trans} + \Delta t_{sync}) \quad (2.1)$$

レーダや LiDAR のように、ターゲットに反射させて測距を行う場合、往復分を考慮して以下の式で距離を推定する。

$$d = \frac{c}{2} \cdot (t_{rec} - t_{trans} + \Delta t_{sync}) \quad (2.2)$$

ToF 方式は高精度な測距が期待できる一方で、送信および受信機間の精密な時刻同期が必

要である。式 (2.1) にて受信機側の視点では、送信された時刻 t_{trans} が不明であるため、実際に距離を推定することはできない。また、同期誤差 Δt_{sync} の分だけ測距誤差が生じることになり、例えば音響測距の場合、1 ms (0.001 秒) の同期誤差でも 0.340 m の測距誤差が生じる。そのため、音響を利用する場合はマイクロ秒単位の同期精度が必要になる。時刻同期を行う手法としては、ネットワークの NTP (Network Time Protocol) や PTP (Precise Time Protocol) などの利用が考えられるが、NTP は誤差が大きいことで知られているため測距や測位には向かない [16]。PTP は NTP よりも高精度同期が可能だが、専用ハードウェアが必要になる、もしくは通信容量などの環境に依存するため実用的ではない [17]。事前に LED を用いて可視光通信による時刻同期を行い、音響によって測距を行う手法 [18] も提案されているが、実用的な状況は限定的である。一般に、音響を用いて無線環境での測距を行う場合、マイクロ秒単位の同期精度が要求されるが実現は困難である。一方で、式 2.2 のようにターゲットからの反射を利用して測距を行う場合、送受信機間是有線による電気信号によって容易に同期がとれる。

無線環境での時刻同期が不要な測距手法として BeepBeep [19] が提案されている。この手法ではモバイルデバイス間で互いに音響信号を送り合い、同間隔で送信するようにスケジューリングすることで、時刻同期の誤差が打ち消され、時刻同期が不要な測距方式を提案している。このように互いに信号を送り合い、その往復時間である RTT (Round Trip Time) から測距を行う手法は TWR (Two Way Ranging) と呼ばれ、主に WiFi [20] や UWB [21] などの電波を用いる場合に利用される。TWR は扱いやすい一方で、ターゲット側およびアンカ側の両者に送信および受信モジュールが必要であり、エネルギー消費、空間での時間もしくは周波数の占有効率、通信オーバーヘッドなどのリソース面で課題が残る [22]。

式 (2.2) のようにターゲットからの反射波を利用して測距を行う代表的な測距手法として FMCW (Frequency Modulated Continuous Wave) がある [23]。FMCW では、送信信号として周波数が連続的に増加もしくは減少する信号を送信し、受信信号との周波数差が遅延に対応する性質を利用して距離を推定する。FMCW は一種のパルス圧縮技術であり、周波数帯域を広くとることで距離分解能の改善が見込める。主にレーダ分野で利用されてきた歴史があるが、LiDAR やミリ波センサなどの自動運転を目的とした車載センサ分野でも注目されている [24]。

RSS 方式

RSS 方式は、図 2.2b のように波の強度が距離によって減衰する性質を利用する。自然界における信号強度が距離の二乗に反比例する逆二乗の法則を利用し、この減衰モデルから逆算することで送信機からの距離を推定する [25]。RSS 方式は ToF 方式と比べて時刻同期が不要であるため、実装が容易である反面、測距精度は劣化する傾向がある。特にターゲットまでの距離が離れるほど、裾が広がるような減衰モデルになるためノイズの影響を受けやす

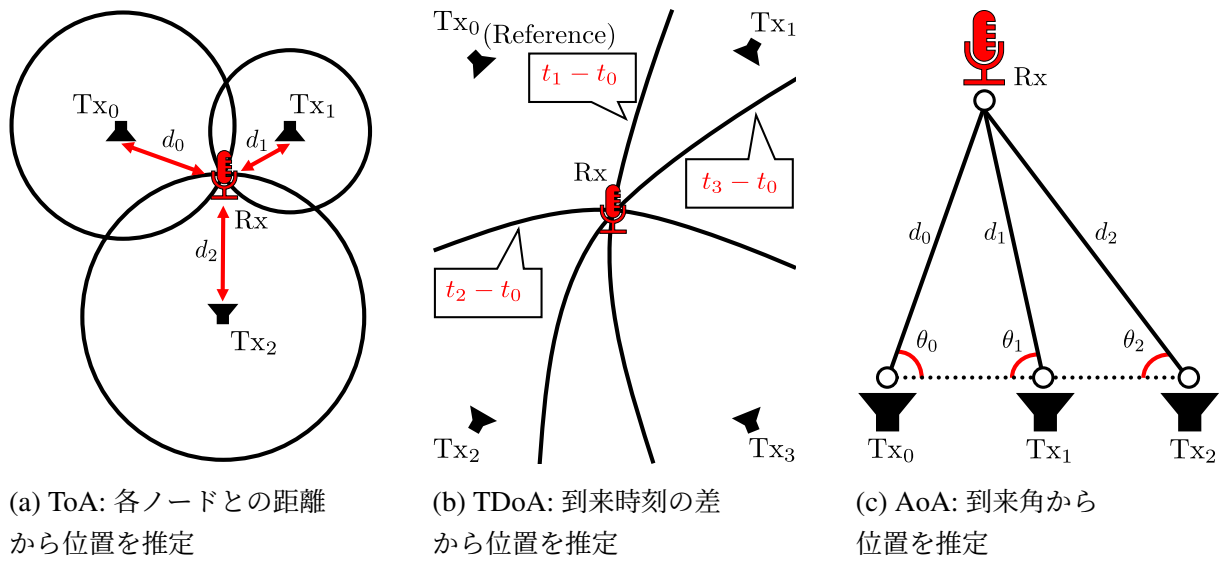


図 2.3: 測位方式の比較

くなる [26]. 可視光, Wi-Fi, BLE は光速で伝搬するため ToF の利用が難しく, 代わりに RSS を用いた手法が主流である. 逆に, ToF が利用可能な音響でも, アプリケーションによっては RSS を利用する場合がある [27, 28].

実際の屋内環境での減衰モデルは, 家具などのオブジェクトの影響によって理論モデルに従わない. これを逆に利用し, 屋内各所で事前に計測した RSS パターンを用いて測距・測位を行う手法も提案されている. このような RSS パターンとの照合によって測位を行う手法はフィンガープリント方式と呼ばれ, Wi-Fi や BLE などの電波を用いた手法で広く利用されている [29, 30]. 一方で, フィンガープリント方式は過学習させるような手法なので汎化性能が低く, 新たな環境に適用するには再度計測が必要となる. また, 人が密集する場所やオブジェクトが頻繁に移動する場所では, RSS パターンが常に変動するため工夫が必要となる [31].

2.1.2 測位方式

ターゲットの位置推定を行う基本原理は図 2.3 に示すように ToA (Time of Arrival), TDoA (Time Difference of Arrival), AoA (Angle of Arrival) に大きく分類できる. ToA および TDoA は, 信号が空間を伝搬する時間を利用するのに対し, AoA は信号の到来角を利用する.

ToA

ToA は図 2.3a のようにターゲットとなる受信機と複数の送信機との距離を用いて位置を推定する手法である。複数の送信機との ToF を利用した測距を行うため、時刻同期が必要になる。基本原理は各ノードとの距離から位置を推定するという面で三辺測量と等しい。

ToA の代表例は GPS をはじめとする GNSS である。人工衛星からの信号には送信時の時刻が含まれているため、受信機側でデコードを行い、受信時刻との差から距離を推定することができる。これを最低 4 つの人工衛星で処理を繰り返すことで位置を推定する。

音響を用いた ToA による測位手法として PTP などネットワーク経由で時刻同期をとる方法が考えられる。しかし、スピーカおよびマイクロフォンの両方で専用ネットワークカードを搭載する必要がある。その他にも ZigBee モジュールを用いて送受信機間での同期をとる手法が提案されている [32, 33]。LED による可視光通信で事前に同期を取り、音響を用いた ToA による測位を行う手法も提案されている [18, 34]。

TDoA

TDoA は図 2.3b のように複数の送信機と受信機との到達時刻の差を利用して位置を推定する手法である。ToA 方式と比較して送受信機間の時刻同期が不要であるため実装が容易である反面、測位計算に必要な送信機数が増加する欠点がある。これは計算時には送信時刻を未知として受信時刻のみを使用するため、ToA よりも余分に式を立てる必要があるためである。解は非線形最小二乗法による最適化問題として解くことが可能で、音響を用いた測位分野では時刻同期が困難であるため TDoA 方式が主流である。

AoA

AoA は図 2.3c のようにターゲットに到来した信号の到来角を用いて、幾何学的な角度関係を推定する手法である。各ノードの角度を求め、その交点から位置を推定する。AoA 方式は、複数の受信機を等間隔になるように配置してアレイを構築することで、遅延差や位相から信号の到来角を推定することができる [35]。

2.2 測位と定位

測位と定位は、広義には位置を推定するという共通の意味を持つが、研究分野に応じて使い分けられることがある。本論文では、本研究の位置づけを示すために便宜上これらを区別することにする。

本論文では、測位をモバイル端末などの自身の位置を推定する技術、定位を位置を知りたいターゲットとの位置関係を推定する技術として扱う。位置を推定する興味の対象が自身

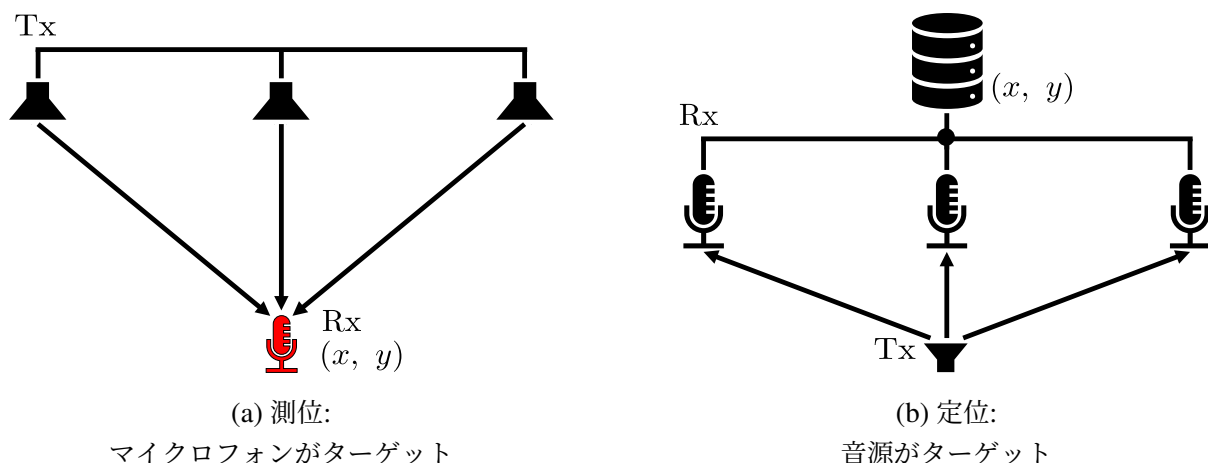


図 2.4: 音響を用いた測位と定位の比較

表 2.1: 測位と定位の分類

種類	測位	定位
アクティブ	アクティブ測位 : 自身から信号を発信して自己位置推定 (例: コウモリやイルカのエコーロケーション, SLAM)	アクティブ定位 : 信号を発信したターゲットの位置推定 (例: 震源地推定, 話者の位置推定)
パッシブ	パッシブ測位 : 外部インフラからの信号を受信して自己位置推定 (例: GPS/GNSS, 天測航法)	パッシブ定位 : 何も発信しないターゲットの位置推定 (例: 戦闘機の探知, RFID タグ)

の場合は測位, 他者の場合は定位と捉えることもできる. 特に音響を用いた研究分野では, ターゲットがマイクロフォン側 (受信機) の場合に測位, 音源側 (送信機) の場合に定位として扱われる傾向がある.

また測位と定位は表 2.1 に示すように, アクティブ型とパッシブ型に大別できる. アクティブ型は能動的に自身から信号を発信するのに対し, パッシブ型は信号を発信しない. 本論文では GPS/GNSS のように受信のみで自己位置推定を行うパッシブ型の測位手法を提案する.

2.2.1 アクティブ型の測位

アクティブ型の測位では, 自身から信号を発信して, その反射波から自己位置推定を行う. 代表例として, コウモリやイルカのエコーロケーションが挙げられる. これらの

動物は自身で信号を出し、その反射波から周囲の環境や獲物の位置を認識している [36]. 近年ではドローンや走行ロボットによる自律ナビゲーション技術の需要が高まっているが、未知の環境下でも地形を把握するため、自己位置推定と空間マッピングを繰り返す SLAM (Simultaneous Localization and Mapping) に関する研究も活発に行われている [37]. SLAM の自己位置推定には LiDAR などの複数のセンサ類をカルマンフィルタやパーティクルフィルタなどで統計的に統合して推定を行う [38]. カメラによる映像ベースの vSLAM (visual SLAM) は AR や VR, 機械学習による画像認識の発展により近年の研究トレンドになっている [39].

アクティブ測位は環境適応能力が高く、外部に専用のインフラが不要であり、自己完結できる利点がある。一方でエネルギー消費や計算コストが大きく、アルゴリズムが複雑になる欠点がある。

2.2.2 パッシブ型の測位

パッシブ型の測位では、受信のみで自己位置推定を行う。代表例の GPS/GNSS は革新的な発明であり、世界中どこにいても GPS の受信モジュールのみでの自己位置推定が可能になった。従来は天測航法により、天体と水平線の角度を測定して位置を推定していたが、GPS の登場で天気や時間帯に左右されずに推定することが可能になった。

パッシブ測位は、発信を行わないためエネルギーやプライバシーの面で利点がある。また、受信機のみで位置推定ができるため普及させやすい。一方で、GPS の場合は人工衛星のような外部の専用インフラが必要であり、設置コストや維持コストがかかる欠点がある。本論文の目的は、屋内測位技術の確立であるため、普及させやすい利点を活かしたパッシブ型の測位に基づいた手法を提案する。

パッシブ型の測位の難しさとして、特に音響の場合、各スピーカから送信された信号が空間で合成されるため、受信側ではこれを識別する必要がある点が挙げられる。これはチャネル間の多重アクセス干渉 MAI (Multiple Access Interference) と呼ばれ、多重化などの手法で避ける工夫が必要である。多重化の手法としては TDM, FDM, CDM が考えられるが、音響測位の分野では FDM が採用されることが多い。

2.2.3 アクティブ型の定位

アクティブ型の定位では、信号を発信したターゲットの位置を推定する。代表例として震源地推定が挙げられる。地震の発生時には、各地に配置された地震計にて波形データが記録され、P 波と S 波の速度の違いから三辺測量の原理で震源地の推定を行う [40].

また、アクティブ型の定位は、受信機の空間的な配置によって集中型と分散型に分類できる。集中型は、複数の受信機を半波長以下になるように等間隔に配置してビームフォーミン

グにより DoA (Direction of Arrival) を推定する。遅延和法が代表的な手法だが、空間分解能が低い欠点がある [41, 42]。相関行列の固有ベクトルがもつ性質を用いて、高い分解能で定位を行う MUSIC 法 [43] や ESPRIT 法 [44] などのサブスペース法も代表的な手法である。近年ではドローンを代表とする UAV (Unmanned Aerial Vehicle) の位置推定に関する研究も行われており、MUSIC 法を改良した Root-MUSIC 法による手法も提案されている [45]。

対して分散型は、複数の受信機を空間に広く分散されて配置し、到来時刻の差を利用して位置を推定する。TDoA 方式による測位での送受信機の入れ替えを行った形で位置推定が可能である。チャンネル間の遅延時間差の検出には相互相関法が一般的に利用されるが、位相による重み付けで修正を加えた GCC-PHAT や GCC-SCOT など提案されている [46]。近年では、定位精度の向上を目的とした、補助関数法によるサブサンプル推定に関する研究も行われている [47]。

AR や VR の分野では HRTF (Head Related Transfer Function; 頭部伝達関数) による音像定位に関する研究も活発に行われている [48]。ヘッドフォンやイヤホンには左右で合計 2 つのチャンネルがあるが、出力される時間および振幅を変化させることで、ユーザに音源の位置を知覚させる。HRTF を利用したナビゲーションの研究も行われており、3D オーディオを利用してユーザに目的地までのガイドを提供する研究も行われている [49]。

測位や定位の精度は一般に送信信号の帯域幅に依存するが、アクティブ型の定位は、動物の鳴き声や会話など変調を前提としないため、精度が不安定になることが指摘されている [50]。また、定位のターゲット数が増加すると分離が困難になる課題もある。一方でパッシブ型の測位では、送信機から測位専用に変調された信号を使用できるため位置推定には優位であると言える。

さらにモバイルデバイスから信号を送信し、分散型の受信機によって位置推定を行う手法では、発信のためのエネルギー消費が大きい課題もある。また、外部の受信機で常に録音することになるのでプライバシーやセキュリティの問題もある。

2.2.4 パッシブ型の定位

パッシブ型の定位では、デバイスフリー型とデバイスベース型に分類可能で、その違いは外部インフラからの信号を積極的に反射させるか否かである。デバイスフリー型の定位では、何も発信しないターゲットの位置を推定する。代表例として、レーダによる戦闘機の位置推定が挙げられる。戦闘機の探知は防衛において重要だが、探知するには戦闘機からの反射波を受信する必要がある。この反射波の強度は戦闘機の RCS (Radar Cross Section; レーダ反射断面積) に依存し、RCS が大きいほど探知しやすい。探知を避けるためのステルス戦闘機では、RCS を小さくする工夫がされている [51]。

RFID (Radio Frequency IDentification) は小型チップを内蔵した ID タグと、それを読み取

表 2.2: 各手法の特徴の比較

	音響	GPS/GNSS	Wi-Fi/BLE	UWB	可視光	PDR
信号	音	電波	電波	電波	光	慣性センサ
測距	ToF	ToF	RSS	ToF	RSS	積分
誤差	<0.3m	<10m	<3m	<1m	<1m	累積
範囲	△	◎	○	○	△	◎
設備コスト	○	×	○	×	○	○
専用受信機	◎	◎	◎	×	◎	◎
消費電力	◎	◎	△	△	△	△
セキュリティ	◎	◎	×	△	○	◎

るリーダーで構成される。RFID では設置インフラをタグとしてリーダーをターゲットとするアクティブ型と、逆に設置インフラをリーダーとしてタグをターゲットとするパッシブ型が考えられる。後者は、リーダーからの電波をアンテナで反射させることで電力不要かつ低コストの通信が行えるため、物流や自動レジなどの分野で近年注目されている [52]。RFID を用いた位置推定手法も検討されているが、反射される電波が微弱で通信可能な距離が数 m 程度であるため、展開可能なアプリケーションは限定的である [53]。

パッシブ型の定位はターゲットから発信を行わない、もしくは電力を必要としないためエネルギー消費の面で利点がある。一方で、反射波の強度が弱いため通信距離が短く、大規模な屋内環境での位置推定には向かない。

2.3 既存の位置推定技術

今日、屋内測位の実現を目標として、様々な技術が提案されている。表 2.2 のように電波、音波、可視光を利用した手法に分類できるが、それぞれに特徴があり利点と欠点がある。

GPS/GNSS

GPS/GNSS は屋内での利用が困難であることが知られているが、この課題の解決を試みた研究も存在する。GPSMirror[54] はバックスキッタ (Backscatter; 後方散乱) を用いて GPS からの信号を再反射させることで屋内環境での測位を実現している。タグは低コストであり壁に容易に設置できるが、中央誤差が 3.7 m であり未だに精度が低い。COIN-GPS[55] では、10 inch × 10 inch のアンテナを用いてビームフォーミングにより GPS 信号の利得の改善を目指しているが、アンテナが大きいことや測位精度が 10 m 程度であり実用的ではない。

Wi-Fi

Wi-Fi は日常生活に欠かせない通信手段であり、規格化された送受信機であるアクセスポイントが今日屋内の至る所に設置されている。これをそのまま測位に活用することで、測位のインフラ設置コストを削減できるメリットがある [56]。また、Wi-Fi 信号は物体を貫通する性質を持つため、障害物が存在する NLoS (Non Line of Sight; 見通し外) 環境に強い利点がある。

Wi-Fi ベースのアプローチは距離分解能が低く、ToF の利用が困難であるため RSS によるフィンガープリントに基づく測位が主流 [29] だが、アンテナアレイを活用した AoA に基づく測位手法も提案されている。

ArrayTrack[57] はアンテナアレイを持つアクセスポイントで MUSIC 法による DoA および AoA 方式でターゲットの位置推定を行う手法である。発信したターゲットの位置を推定するため、アクティブ型の定位に分類できる。プロトタイプを用いた評価実験では、3 台で中央誤差 0.75 m、6 台で全てを用いた中央誤差 0.26 m を達成しているが、アンテナアレイが多く必要になる。

SpotFi [58] は、超解像アルゴリズムを用いてアンテナ 3 本を用いた汎用アクセスポイントでの AoA による測位を可能にした手法である。CSI (Channel State Information; チャネル状態情報) の取得が課題となっていたが、各アンテナを組み合わせることでこれを解決した。評価実験でのマルチパスを多く含む環境において、中央誤差 0.4 m になることを確認している。LoS (Line of Sight; 見通し内) 環境でのアクセスポイントが 2 つの環境下では中央誤差 1.6 m になることが報告されている。

Wi-Fi ベースの位置推定手法は、既存のインフラが利用できる一方で、精度の分散が大きく安定しない課題がある。また通信の応用であるため、セキュリティ面での懸念がある。

BLE

Bluetooth の低消費電力版である BLE (Bluetooth Low Energy) は IoT 機器間で近距離通信に利用される通信規格であり、屋内測位にも応用されている [59]。今日ではワイヤレスイヤホンの普及でモバイルデバイスでは標準で利用可能であるため、専用ハードウェアを必要としない利点がある。測位手法は Wi-Fi と同様に RSS を利用したフィンガープリントが主流だが、通信可能距離が短いため、普及させるには多くのビーコンが必要になりコストがかかる。DABIL [60] はノイズ除去を目的としてオートエンコーダを組み入れたフィンガープリントによる測位手法で、評価実験によって平均誤差が 1.09 m になることを報告している。オートエンコーダによって環境による RSS の変動に強くなり、ビーコンをロストした場合でも平均誤差が 1.60 m に抑えられる利点がある。BLE は Wi-Fi と同様に普及させやすいも

の、測位精度の面で課題が多い。

UWB

UWB (Ultra Wide Band; 超広帯域無線) は電波を利用するが、超広帯域であるため良好な距離分解能を持つ通信規格であり、位置推定やセンシングを目的としたアプリケーションに応用されている。帯域幅と測距精度についての調査では、電波を用いる場合でも帯域幅が広いほど測距精度が改善することを確認している [61]。UWB では送受信機間での時刻同期は困難なため、TWR によるタグ間で双方向通信により位置推定を行う方式が一般的ではあるが、受信のみでのパッシブ型やアクティブ型で AoA を利用する手法も検討されている [62, 63]。

Uloc[64] では、従来の TWR の冗長性による電力消費を削減するため、一方向通信での AoA による定位手法を提案している。ターゲットであるタグからの信号を設置タグで受信し、MUSIC 法ベースのアルゴリズムに基づく定位を行う。評価実験によって、中央誤差が 3.6 cm になることを確認している。Chorus[65] でも同様に TWR の冗長性を解決するため、受信のみでパッシブ測位を実現するための手法を提案している。複数ノードからの信号をタグで受信し、TDoA による位置推定により、デシメートルレベルの精度を達成している。

UWB は音響と同様にセンシングの分野でも利用されている。研究例としては、透過率を利用した水分の成分推定 [66] や、人の呼吸数推定などの生体センシング [67, 68] などがある。

UWB は位置推定精度の面では期待できるが、送受信機の両側で専用ハードウェアが必要であるため設備コストがかかるため、普及させるには課題が多い。

可視光

可視光測位では、PD (Photo Diode) や内蔵カメラで天井の LED 変調光を検出し、距離や角度などからターゲットの位置を計算する [69]。PD の低消費電力版の ALS (Ambient Light Sensor; 環境光センサ) を用いた測位手法も検討されている [70]。ALS による手法は、画面の明るさ自動制御を目的としてモバイルデバイスに標準で搭載されているため、普及させやすい利点がある。一方で LED からは変調光を送信する必要があるため、設置インフラに変更を加える必要がある点や、障害物、自然光による影響を受けやすい欠点がある。

PDR

PDR (Pedestrian Dead Reckoning; 歩行者自律航法) は、モバイルデバイスに内蔵されたセンサ類を用いて歩行者の移動経路を推定する技術である [71]。主に加速度、ジャイロ、地磁

気、気圧、などのセンサ類を用いて、歩行者の歩幅と方向を推定する．相対的な移動経路が推定できるが、絶対位置の推定はできず、誤差が蓄積するため、他の測位技術と組み合わせた定期的なキャリブレーションが必要である．機械学習ベースでセンサ類の値を学習させて移動経路を推定する試みもある．RIDI[72]では加速度センサの値を二重積分することで移動経路の推定を行っている．NILoc[73]では、IMU データをトランスフォーマーで学習し、高精度な移動経路推定を実現している．PDR はモバイルデバイスで完結するため、設備コストがかからず、今後の普及が期待される．

ミリ波

ミリ波を利用した位置推定技術も Beyond 5G 技術として注目されている．Millimetro[74]は、主に道路標識や障害物検知を目的とした FMCW による相対的な自己位置推定を行う手法である．事前に障害物に設置するタグ側では Van Atta Array[?] を用いて、到来角と同じ方向に反射波を返す仕組みを取り入れている．反射された信号を MIMO アンテナで受信し、ドップラ FFT による位置推定を行う．屋外での評価実験にて、中央誤差が 22 cm になることを確認している．Van Atta Array のように反射を利用する技術は、メタサーフェスの分野で近年再注目されている．ミリ波レーダはアクティブ型の測位になるため、本研究の目的には適していないが自動運転の分野など応用先は広い．

2.4 音響を用いた位置推定技術

音響を利用した初の位置推定手法は、筆者の知る限り AT&T ケンブリッジ研究所の Active Bat [75] である．この手法では送信機からの音波を送り、格子状に設置されたマイクロフォンで受信することで送信機の位置推定を行う．本論文の分類ではアクティブ型の定位に分類される．同じようなシステムとしては ASSIST [76] や反射波も利用した手法 [77] が提案されている．パッシブ型測位の初の試みは、MIT の Cricket [78] である．Cricket ではビーコンから電波と超音波を同時に送信し、リスナー (受信機) では受信した電波と超音波の遅延時間差を利用して測距および測位を行う．Cricket は Cricket Compass [79] に拡張され、2つのマイクロフォンを利用して AoA による方位の機能を追加した．国内からは DOLPHIN [80] とよばれる、電波により同期されたノード間で音波を逐次繰り返し送信することで、ノードの位置推定を行う手法が提案されている．

モバイルデバイスから音響信号を送信し、その反射波を検出して周囲環境を把握するシステムも提案されている．BatMapper [81] および SAMS [82] は主に壁やドアの検出を目的として、フロアプランの構築を行っている．BatTracker [83] では反射波の振幅やドップラシフトを加味した MHT (Multi-Hypothesis Tracking) によるトラッキングを行っている．

2010 年頃からはパッシブ型の測位に関する研究が活発に行われている。Yang らは車内でドライバーがスマートフォンを操作しているか判定するための測位システムを提案している [84]。車内にある 4 つのスピーカから信号を送信し、その到来時刻差から位置推定を行う手法だが、座席の分類にとどまっている。Lazik らはこれを拡張し、屋内でスマートフォンをターゲットとした高精度測位を実現している [85]。高精度測位のために距離分解能が良好なチャープ信号を採用し、受信のみで測位を行う方式は現在に至るまで広く利用されている。Lazik らはこれを拡張した ALPS [86, 87] も提案している。

現在に至るまで、受信のみで位置推定を行うパッシブ型測位に関する研究は、送信方式を工夫したもの [33, 88, 12, 89, 90, 91, 92, 93], PDR と組み合わせたもの [94, 95], 反射波を利用したもの [96, 97] などが提案されている。

2.5 音響センシング技術

音響を用いたセンシング技術も、モバイルデバイスやスマートスピーカの普及によって近年活発に研究が行われている。呼吸数測定やジェスチャー認識など、生体情報のセンシングに関する研究が多いが、位置の変化を検知するという面で測距や測位と共通する点が多い。

音響の測距精度の高さを利用して、デバイスフリー型のジェスチャー認識や、呼吸数モニタリングを行う手法が提案されている。FM-Track [98] は、主にジェスチャー認識を目的としたトラッキング手法であり、FMCW とマイクロホンアレイ技術によって 0.11 cm 以下の誤差になることを報告している。LASense [99] は、測距技術の応用による、胸部の時間による変位を計測した呼吸数を測定する手法である。FMCW を利用することで 6 m ほどの長距離でも精度の高い推定が可能であることを報告している。DF-Sense [100] でも同様の技術およびマイクロホンアレイを採用することで複数ターゲットのモニタリングを実現している。CW-AcouLen[101] では、ホルムヘルツ素子を利用したメタサーフェスを設計し、NLoS 環境下におけるセンシング技術を提案している。

また、デバイスベース型のトラッキングに関する研究も行われている。CAT [102] では、VR や AR への展開を目的としたトラッキング手法を提案し、FMCW と IMU センサを統合することで推定精度の改善を図っている。Swadloon [103] では、モバイル端末を移動させた際に生じるドップラシフトを利用したトラッキング手法を提案している。Vernier [104] では、位相差を利用した手法を提案し、トラッキング誤差が 4 mm 以下になることを報告している。これらのトラッキング技術の精度は高いが、基本的には変位の検出のみに対応しているため、測位のような絶対位置の推定はできない。また、手法によっては初期位置を与える必要がある欠点がある。

音響の物理的な性質を利用したセンシングに関する研究も存在する。VECTOR [105] では、音速が温度によって変化する特徴を利用して、空間の温度分布を推定する手法を提

案している．ToF の検出が重要であることから，分解能が高い系列信号を利用する工夫を行っている．Meta-speaker[106] では，空気の非線形性を利用して，任意の位置に音源を生成する技術を提案している．送信機（パラメトリックスピーカ）からは非可聴音の信号が送信されるが，ビームが交差した位置では非線形性によって可聴音が再生される技術である．MuDIS[107] でも同様にパラメトリックスピーカを用いてターゲット方向にビームを生成し，対象の人のみにアナウンスを提供する技術を提案している．同様に非線形性を利用した手法として，Backdoor [108] と呼ばれる通常のマイクロフォンでは受信できない 40 kHz 程度の高い周波数の信号を，録音可能にする技術も存在する．このような非線形音響学の KZK 方程式 [109] に基づく研究や，音響のメタサーフェスを利用した研究は近年増加傾向にあり注目を集めている．

2.6 音響情報ハイディング技術

5 章では，BGM などの音源に測位用のコードを埋め込んだ手法について説明する．本節ではこの技術に関連する音響情報ハイディングに関する研究をまとめる．

音響情報ハイディング技術は文献 [110] によれば，偽造防止（違法コピーや違法配布など），音声改ざん検出，なりすまし防止などを目的として，メディアコンテンツ内に意図的に情報を隠す技術の総称だとされている．類似の分野として「音響電子透かし」と「音響ステガノグラフィ」が挙げられるが，文献内ではこれらの違いと分類についても詳細に議論されている．音響情報ハイディング技術の主な課題は，音質の劣化を最小限に抑えつつ，必要な情報を確実に埋め込むことであり，これが研究を進める上で主要な目標となっている．

違法コピーの検出を目的とする場合，DVD などの媒体に直接接続することになるので，透かしの検出は比較的容易に実現可能である．一方，空間を伝搬させる無線通信を目的とする場合，環境ノイズなどによりその難易度は大幅に高くなる．空間を伝搬させるハイディング技術には，本論文で扱うエコー拡散法 [111] の他に，群遅延や蝸牛遅延特性を利用した手法 [112, 113]，オクターブ類似性の効果を利用した手法 [114]，周波数マスキング特性を利用した手法 [115] などが提案されている．また，通信の高速化を目的として OFDM 技術を音響に適応した手法 [116] や，近年では位相変調を利用した手法 [117] など提案されている．

続いて本論文で使用するエコー拡散法に関連する手法を説明する．エコー拡散法は単一エコー法 [118] を拡張したもので，1 つの大きなエコーの代わりに小さな PN 系列を並べることで秘匿性および検出性能の向上させた手法である．これらの手法は，人間の知覚特性である先行音効果 (precedence effect; 第一波面の法則) とエコー検知限 (echo-detection threshold) を利用している．先行音効果は，最初に到来した音が強調され，続いて到来する音はマスクされて知覚される特性である．エコー検知限は，到来した 2 つの音の時間差が短い場合にそれらが合成音として知覚される特性である．

現在までに、エコー拡散法を改善した手法もいくつか提案されている。Chou らは通常のエコー拡散法に加えて、カーネルの対称側にも PN 系列を加えた両側エコー拡散法 [119] を提案し、頑健性が向上することを確認している。さらに、周波数領域にて音質と検出性能の最適化する手法 [120, 121] や、奇数と偶数インデックスに分けて埋め込むことで、検出性能を向上させる手法 [122] など提案されている。

最終的には音質の評価を行うことになるが、その評価方法は主観評価 (subjective evaluation) と客観評価 (objective evaluation) の 2 つの大別できる。主観評価は実際に人を集めてアンケート等により音質を評価する手法である。代表的な手法は ITU (International Telecommunication Union) から勧告されている 3 刺激 2 重盲検法 (double-blind triple-stimulus with hidden reference) や MUSHRA (multi stimulus test with hidden reference and anchor) などである [123]。主観評価は直感的であるものの、多くの被験者と時間を必要とする欠点がある。

これを解決するために提案されたのが客観評価であり、音の物理量から音質の評価を行う。オリジナルの音源と加工済みの音源を、聴覚モデルを考慮した学習済みのニューラルネットワークに入力することで劣化量が得られる。代表的な手法として PEMO-Q [124] や ViSQOL [125] が提案されている。客観評価は、実験環境や被験者の属性によらない評価が可能であり、主観評価の手間がかからないため音響情報ハイディングの分野では広く利用されている。

第 3 章

鏡像法による送受信機間の擬似的な時刻同期手法

3.1 はじめに

本章では、送受信機間の時刻同期手法について説明する。音響測位技術の課題として、測位に必要なスピーカ数が多い点が挙げられる。この原因は、送信機であるスピーカと受信機であるマイクロフォンのマイクロ秒単位での時刻同期が困難であることに起因する。そのため、音響測位の分野では時刻同期を必要としない TDoA 方式による測位が一般的である。一方で、TDoA 方式は多くのスピーカ数を必要とするため、実環境での利用を想定する場合には時刻同期に基づく ToA 方式が望ましい。

本章では、受信するモバイルデバイス側で送信側の送信時刻の推定を行い、擬似的な時刻同期手法を提案する。提案手法では、屋内の設置スピーカから送信され、床や壁を介して反射される信号を、床や壁の鏡像位置にあるスピーカからの信号と捉える鏡像法 [13] および「鏡像スピーカ」の概念 [96] を利用する。提案する時刻同期手法では、特に床の鏡像スピーカに着目する。

本章の内容は、筆者の研究成果 [126] (© 2022 IEEE) を基にまとめたものである。

3.2 基本原理

屋内に設置されたスピーカ S_O の位置を $P_O = (x', y', z')$ とすると、床の鏡像スピーカ S_F の位置を $P_F = (x', y', -z')$ と定めることができる (図 3.1)。このとき、設置スピーカの位置 P_O は既知であり、推定したいモバイルデバイスの座標 $X = (x, y, z)$ は未知である。スピーカから信号が送信された時刻を t_e 、音速を c とすると、スピーカ S_O とモバイルデバイス間の距離は下記で表現できる。

$$\sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2} = c \cdot (t_O - t_e) \quad (3.1)$$

このとき、 t_O はスピーカ S_O から送信された信号の到来時刻である。同様に、鏡像スピーカ S_F とモバイルデバイス間の距離は下記で表現できる。

$$\sqrt{(x' - x)^2 + (y' - y)^2 + (z' + z)^2} = c \cdot (t_F - t_e) \quad (3.2)$$

このとき、 t_F はスピーカ S_F から送信された信号の到来時刻である。

2 台のスピーカは、 xy 平面において同一の座標を取るため、 t_e に対して式 (3.1) と式 (3.2) から下記の式 (3.3) を導出することができる。

$$\begin{aligned} t_e &= \frac{c^2(t_O^2 - t_F^2) + 4z'z}{2c^2(t_O - t_F)} \\ &= \frac{t_O + t_F}{2} + \frac{2z'z}{c^2(t_O - t_F)} \end{aligned} \quad (3.3)$$

ここで未知なのは床からの高さ z のみになるので、これを求めることで、 t_e を推定できることが分かる。

このとき到来時刻 t_O , t_F は村上らの提案手法 [96] に従い検出する。この検出手法では、まず相互相関によって得られた包絡線の平均値の 2 倍をしきい値として設定する。その後、しきい値を超えるピークを到来時刻の候補として選出する。本稿では到来時刻の候補を階層クラスタリングによって検出する。

3.3 アクティブセンシングによる床からの高さ推定

提案手法では、モバイルデバイスの高さ z を推定するために、BatMapper[81] で用いられるアクティブセンシングによる高さ推定を行う。アクティブセンシングでは、図 3.2 のように内蔵スピーカから音響信号を送信し、床を介した反射波をマイクロフォンで受信し処理することで床からの距離を推定する。

多くのモバイルデバイスには上部と下部にそれぞれ、スピーカとマイクロフォンが標準で搭載されている。このとき、受話用である上部スピーカは、モバイルデバイスの正面方向に対して指向性をもっているため、モバイルデバイスの画面を見ながら歩行する際、指向性は天井に対して強く、床に対して弱くなる。そのため、本稿では下部のスピーカとマイクロフォンを利用する。

3.3.1 送信信号の設計

送信信号には、帯域幅を確保しながら距離分解能が良好となるチャープ信号を用いる。以下の式は線形チャープ信号 $s(t)$ である。ここで f_0 は開始周波数、 f_1 は終了周波数であり、 T

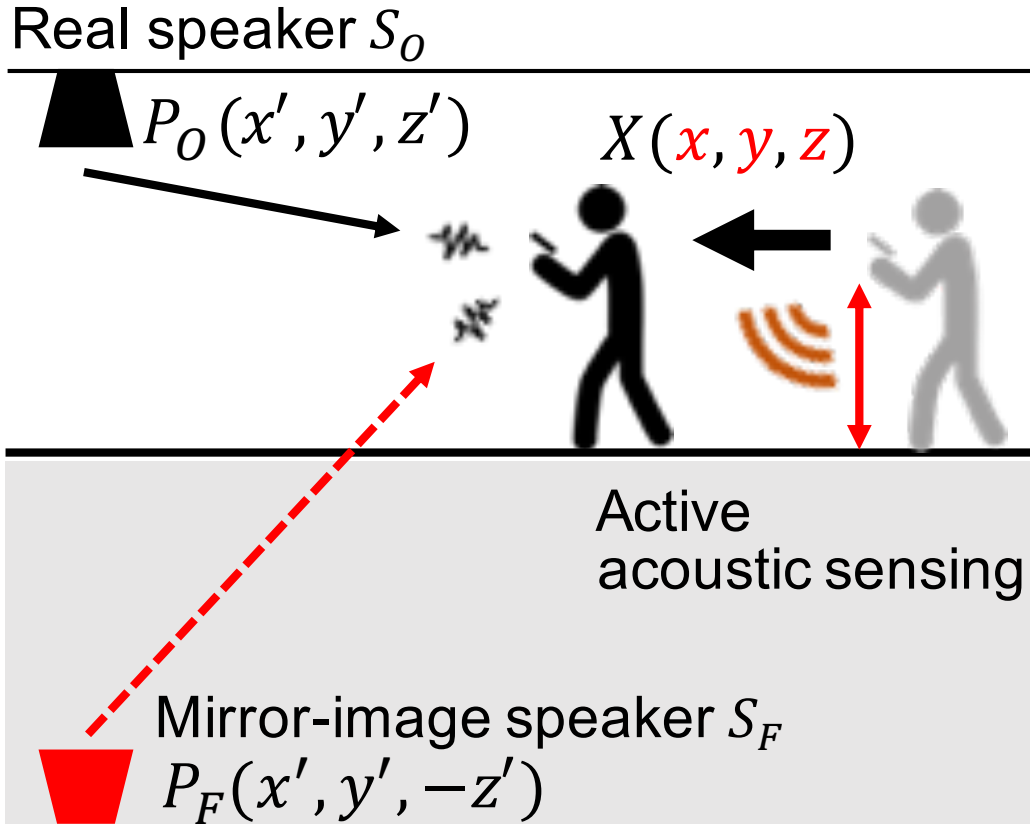


図 3.1: 時刻同期手法の概要 ([126], © 2022 IEEE)

は f_0 から f_1 までの遷移時間を表す.

$$s(t) = \sin 2\pi \left(f_0 t + \frac{k}{2} t^2 \right), \quad (3.4)$$

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

ユーザがモバイルデバイスを手に持って歩行する際の高さ z は, 0.7 m から 1.5 m 程度である. これを考慮して, 信号を送信している間に反射波が受信されることを避けるため, 信号長は $T = 3$ ms と設定した. これは, 0.5 m 以上離れた床からの高さであれば, 反射波が受信される前に送信が終了するためである. また, モバイルデバイスの内蔵スピーカからチャープ信号を送信する際に生じるクリッキングノイズ [85] の影響を避けるために, 送信信号に対してハン窓をかけることで振幅レベルが瞬間的に上昇することを防ぐ.

このとき, 送信時刻推定のターゲットである外部スピーカからも信号が送信されるが, その周波数は 12 k–22 kHz に設定する. アクティブセンシング側では環境ノイズや人の声の帯域を避けるため, 5 k–10 kHz の帯域を使用し, 1 秒間に 64 回の信号送信を行う.

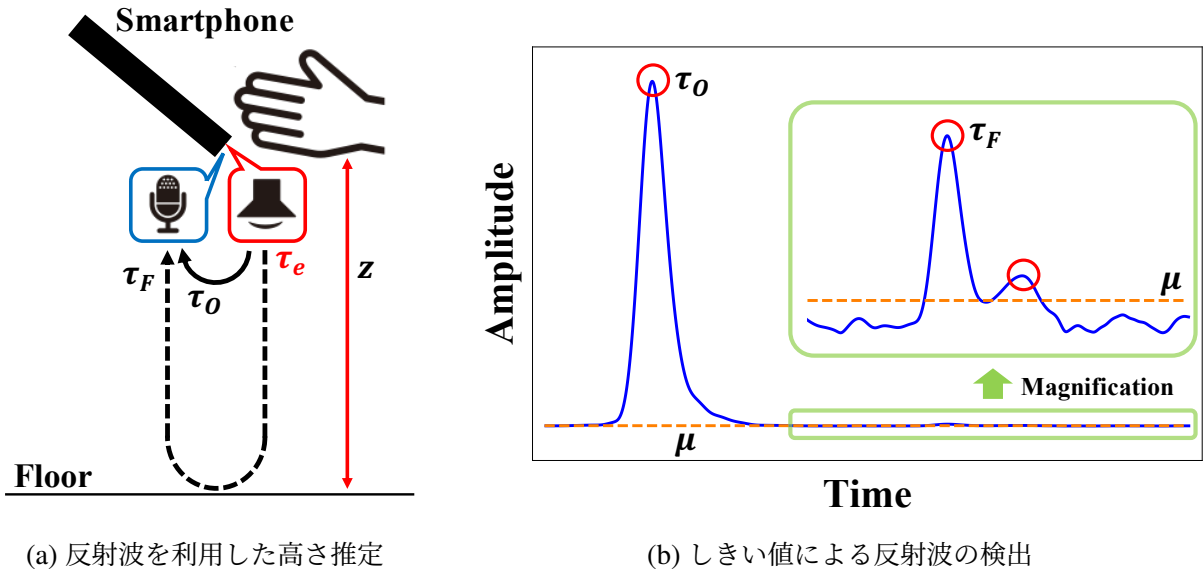


図 3.2: アクティブセンシングの概要 ([126], © 2022 IEEE)

3.3.2 高さ z の推定

アクティブセンシングにおいて送信された信号は、マイクロフォンに直接到来する直接波と、反射体を重ね合わせた形で受信される。受信された時系列データと送信信号との相互相関を算出することで、図 3.2b のように包絡線を算出することができる。このとき得られた包絡線の平均値の 2 倍をしきい値 μ として設定し、これを満たす最大値は反射波の到来時刻に相当する。

事前実験により、アクティブセンシングでは直接波と比較し、反射波の振幅が非常に小さいことを確認している。反射波がノイズに埋没ことを避けるため、算出される包絡線に対し 0.5 秒間、すなわち 32 回のインコヒーレント積分処理 [127] を行う。この平均化処理により、床以外の反射波の影響を軽減する効果も期待できる。図 3.3 はインコヒーレント積分の回数に応じた包絡線の変化を示しており、回数を重ねる毎に、床の反射波のピークが顕著になっていることが確認できる。

上記の処理で検出される、床を介した反射波の到来時刻を τ_F とすると、モバイルデバイスの高さ z は下記のように表現できる。

$$z = \frac{c \cdot (\tau_F - \tau_e)}{2} \quad (3.5)$$

τ_e は内蔵スピーカからの送信時刻だが、内蔵のスピーカとマイクロフォン間の内部システムでの時刻同期は難しく、送信時刻 τ_e を推定することは困難である。そのため図 3.2b にて直

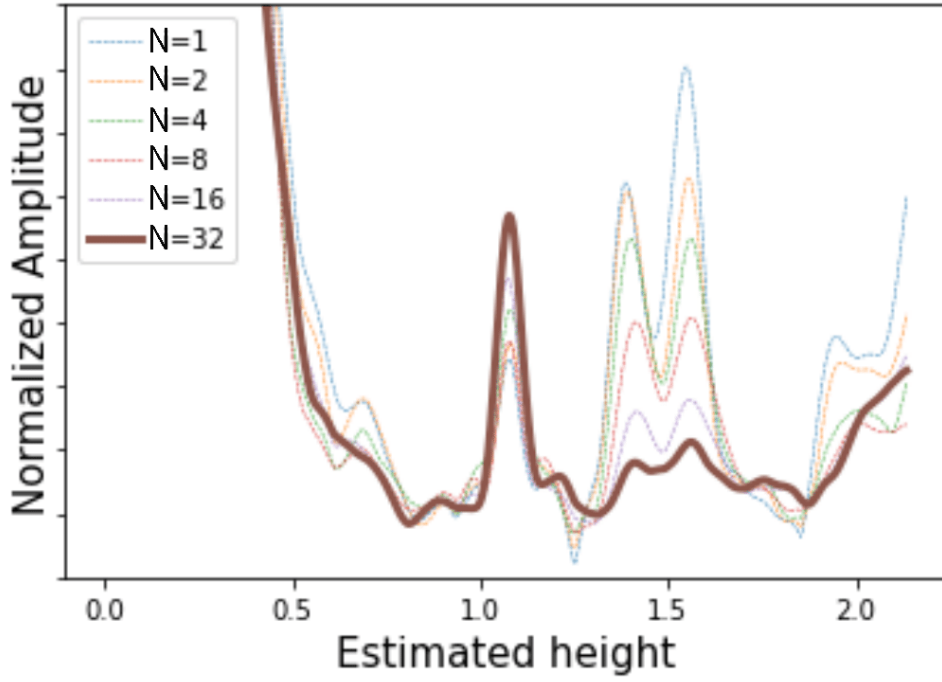


図 3.3: インコヒーレント積分の回数による比較 ([126], © 2022 IEEE)

接波の到来時刻 τ_O を送信時刻 τ_e として利用する.

$$z = \frac{c \cdot (\tau_F - \tau_O)}{2} \quad (3.6)$$

3.4 送信時刻推定

上述のアクティブセンシングにより, モバイルデバイスの高さ z の候補を算出することができる. また, 村上らの手法 [96] を活用することで, 設置スピーカ S_O の到来時刻 t_O に加え, 床の鏡像スピーカ S_F の到来時刻 t_F の候補を算出することができるが, z , t_F を一意に定めることは難しい. 次に, z , t_F を一意に定め, 本研究の目的である外部の設置スピーカからの送信時刻 t_e を推定する方法について説明する.

3.4.1 送信時刻の特徴

外部の設置スピーカからは一定の間隔 T_{fra} で信号が送信される. そのため推定された送信時刻の候補 $\tilde{t}_e$ も一定の間隔になることが予想できる. このとき送信間隔 T_{fra} を 1 フレームとすると, i 番目のフレームで観測した送信時刻の候補 $t_e^{(i)}$ は下記のように表現できる.

$$\tilde{t}_e^{(i)} = t_e^{(i)} - i \cdot T_{fra} \quad (3.7)$$

そのため、 z と t_O , t_F の正しい組合せで式 ((3.3)) により計算される送信時刻の候補 $\hat{t}_e$ は、ユーザの行動に関わらず一定となる。そこで本手法では、ユーザに歩行等のアクションをしてもらうことで、 z と t_O , t_F の誤った組合せで算出される送信時刻を除去し、正しい送信時刻 t_e を選択することを試みる。

3.4.2 ユーザのアクションに基づく選択

観測したフレーム数を N_{fra} とするとき、図 3.4 で示されるように、 N_{fra} フレーム間でのユーザの移動に対して、複数の送信時刻の候補の集合を獲得することができる。このとき、送信時刻は一定であるため、 N_{fra} フレーム間で推定された正しい送信時刻の集合は一点に集中する形で1つのクラスターを形成する。このクラスターを2つのステップで抽出する。

まず、大きな外れ値を除去する。提案手法では、偶然、正解値に近い送信時刻の候補が算出される場合がある。そのため、モバイルデバイスを見ながら歩行していても高さが不変であることを利用し、その高さを新たな指標として導入する。さらに、推定された送信時刻に音速 c を乗じ、送信時刻の候補 $\hat{t}_e^{(j)}$ を、距離の組合せ $p = (c \cdot \hat{t}_e^{(j)}, z^{(j)})$ に置き換える。このとき、 $\hat{t}_e^{(j)}$ は送信時刻の j 番目の候補であり、 $z^{(j)}$ は $\hat{t}_e^{(j)}$ に対応するスマートフォンの高さである。図 3.5 は、距離の組合せ p の集合を示している。 p と、異なる組合せ p' との距離 $\|p - p'\|$ が d_{out} 未満の組合せが N_{out} 未満であった場合に、外れ値とし除去する。実験に基づき、本稿では $d_{out} = 0.1$ m, $N_{out} = 0.2 \cdot N_{fra}$ と設定する。

次に、外れ値が除去された集合に対し、階層クラスタリングによって送信時刻候補のクラスター $\mathbf{C}^{(i)}$ ($i \in 1, 2, \dots$) を抽出する (図 3.6)。実験結果に基づき、下記の条件を満たすクラスター $\mathbf{C}^{(i)}$ を送信時刻 t_e のクラスターとして決定し定める。

- $|\mathbf{C}^{(i)}| \geq 0.8 \cdot N_{fra}$
- $c \cdot \sigma^{(i)} \leq 0.1$ m

このとき、 $|\mathbf{C}^{(i)}|$ はクラスター $\mathbf{C}^{(i)}$ の要素数であり、 $\sigma^{(i)}$ は送信時刻候補のクラスター $\mathbf{C}^{(i)}$ の標準偏差である。

3.4.3 ToA ベースの測距・測位

提案する時刻同期手法により推定される送信時刻 $\hat{t}_e$ を用いることによって、下記のように n 番目のスピーカ $S_{n,K}$ ($n \in \{1, 2, \dots\}, K \in \{O, F\}$) とモバイルデバイスとの距離を算出することができる。

$$\|X - P_{n,K}\| = c \cdot (t_{n,K} - \hat{t}_e) \quad (3.8)$$

このとき、 $P_{n,K}$ はスピーカ $S_{n,K}$ の3次元座標であり、 $t_{n,K}$ はスピーカ $S_{n,K}$ から送信された信号の到来時刻である。

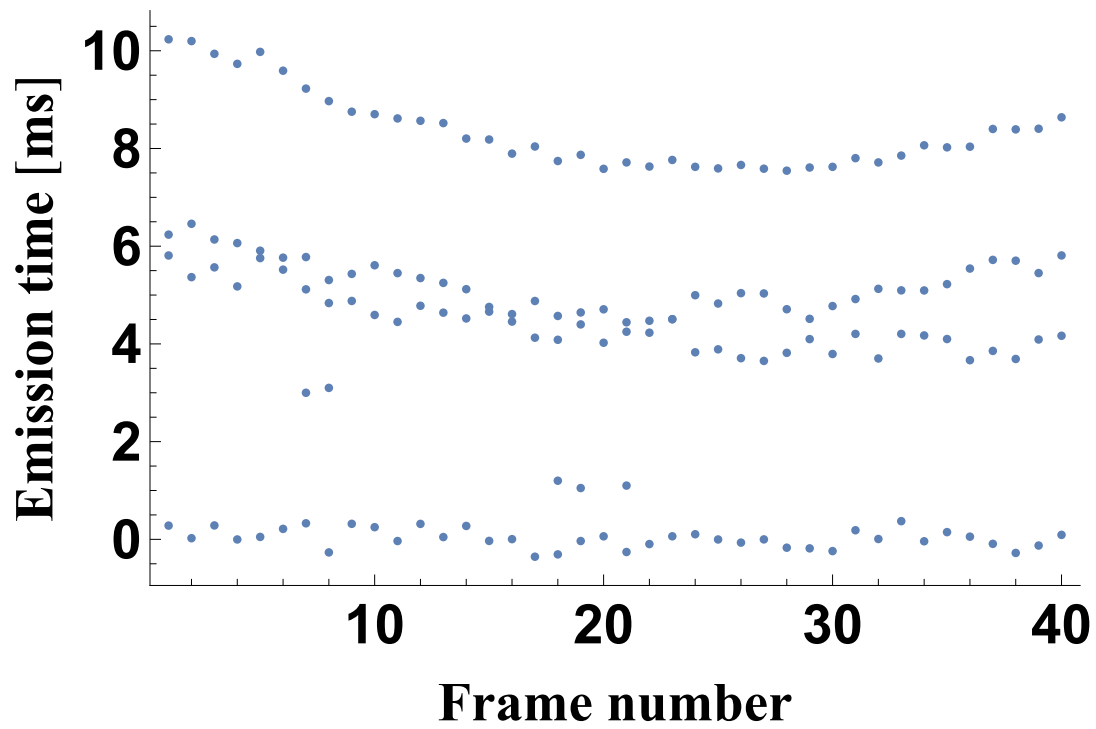


図 3.4: 各フレームで推定される送信時刻の候補

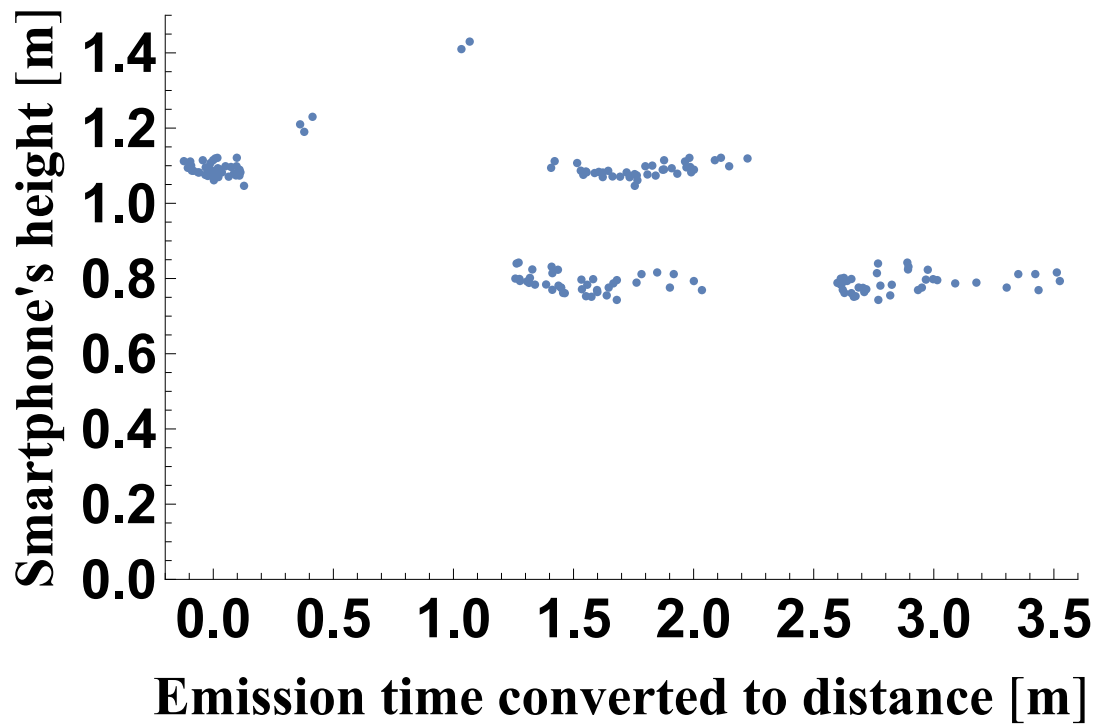


図 3.5: 推定された送信時刻に基づく距離の組合せ p の集合

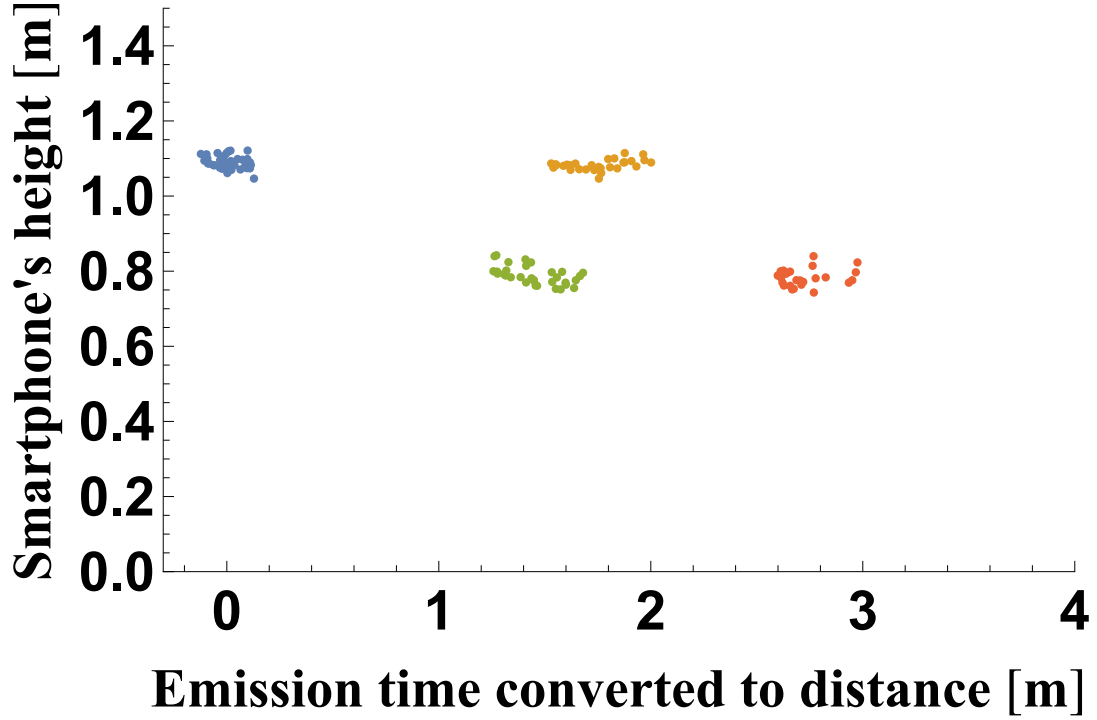


図 3.6: 推定された送信時刻候補のクラスター

ToA 方式の測位では 3 台のスピーカを必要とするが，反射波を活用する本手法では 1 台の設置スピーカから床の鏡像スピーカを含む 2 台のスピーカを仮想的に獲得することができる．そのため，2 台の設置スピーカで ToA ベースの測位を実現できる．このとき，4 つ ($|n| \times |K|$) の方程式が利用可能で，式 (3.9) に対し最小二乗法を適用することで，スマートフォンの 3 次元座標 $\mathbf{X}$ を算出する．

$$J = \sum_n \sum_K (d_{n,K})^2 \quad (3.9)$$

$$d_{n,K} = \|\mathbf{X} - \mathbf{P}_{n,K}\| - c \cdot (t_{n,K} - \hat{t}_e)$$

3.5 予備実験：スマートフォンの周波数特性

はじめに，代表的な Android スマートフォン 4 機種 of スピーカおよびマイクロフォンの周波数特性の調査を行った．一般的なスマートフォンには端末の上部および下部に，それぞれスピーカとマイクロフォンが搭載されている．この実験では各スピーカとマイクロフォンの特性を調査した．図 3.7 は 4k–23 kHz までの調査結果である．スピーカの特性を調べる際には性能が良好な標準マイクで録音し，マイクロフォンを調べる際にはスピーカとして Fostex FT200D を使用した．4 つの実験結果を通して，本章では特性が最も安定している Samsung Galaxy S10 plus を使用する．

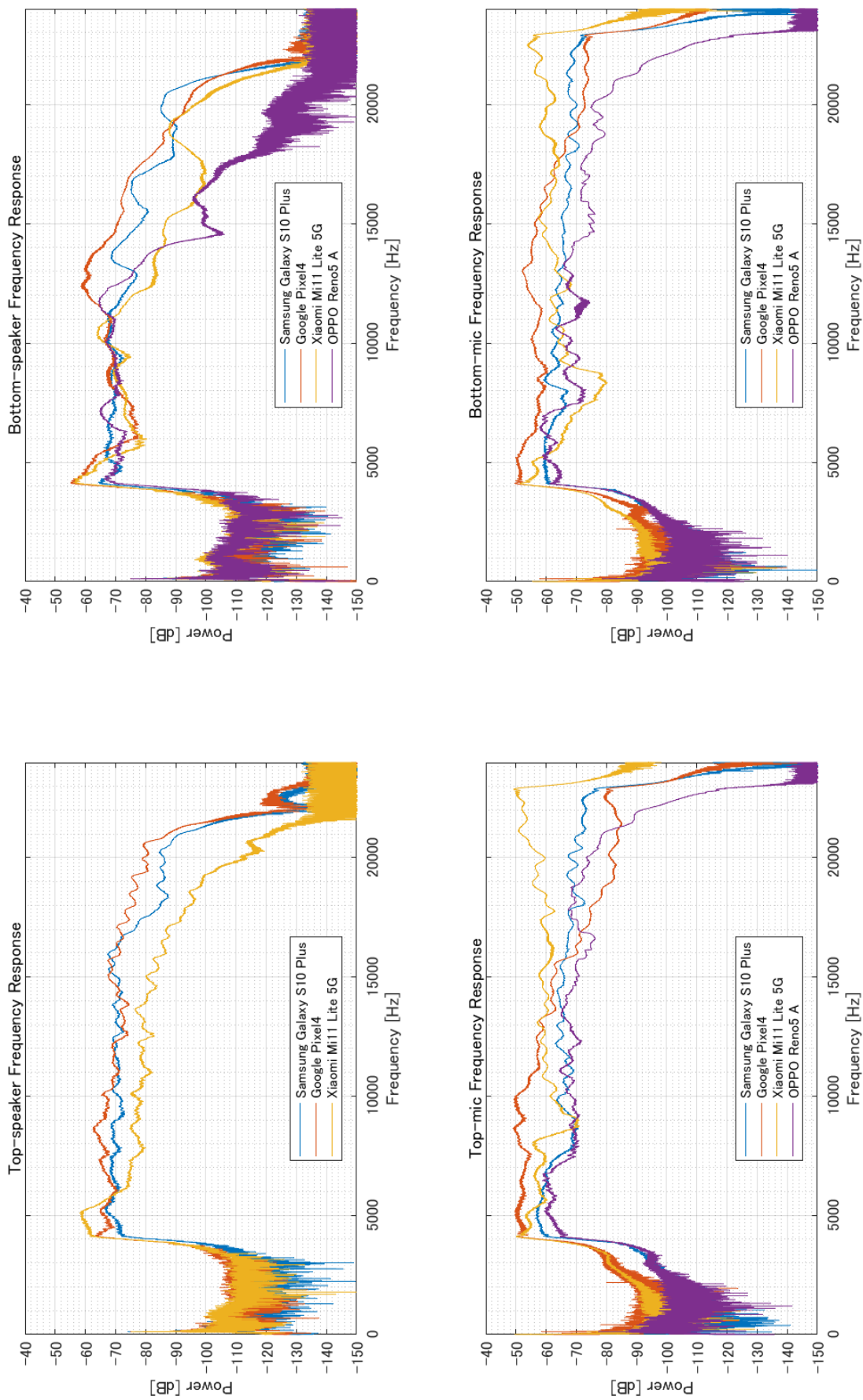


図 3.7: スマートフォンのスピーカとマイクロフォンの周波数特性：(左上) 上部のスピーカ、(右上) 下部のスピーカ、(左下) 上部のマイクロフォン、(右下) 下部のマイクロフォン

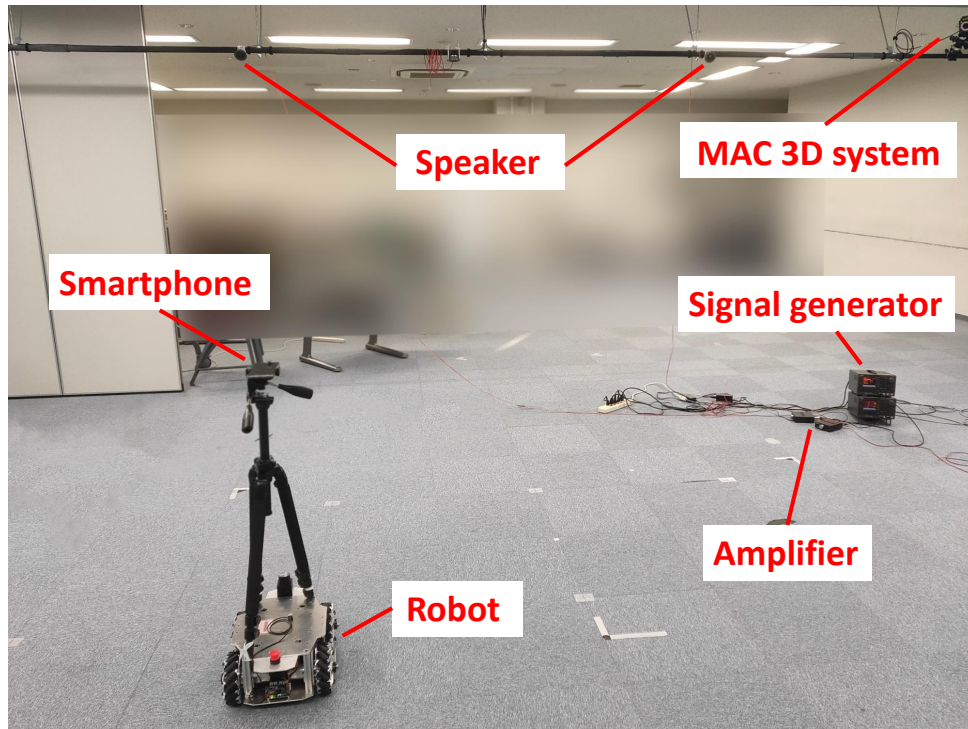


図 3.8: 実験環境 ([126], © 2022 IEEE)

3.6 時刻同期手法

3.6.1 実験設定

評価実験では、信号生成器とアンプ、スピーカ、スマートフォンを用いた。信号生成器は NF WF1948, アンプは Fostex AP20d を使用した (図 3.8)。スピーカは Fostex FT200D を使用し、測距実験での送信信号は、図 3.9 のように 1 台のスピーカから帯域幅が 12k–16 kHz, 信号長 10 ms のチャープ信号が 1 秒間に 8 回送信されるように設定した。このとき, Lazik らの手法 [85] と同様にクリッキングノイズを避けるため, フェードインおよびフェードアウトになるような信号を追加した。測位実験では 2 台のスピーカから帯域幅をそれぞれ 12k–16 kHz と 18k–22 kHz に設定し同時に送信した。また, スマートフォンは Samsung Galaxy S10 plus を使用し, マイクロフォンのサンプリングレートは 48 kHz に設定した。このとき, 録音された時系列データは信号処理時に 16 倍でアップサンプリングし, オフラインにて解析を行った。

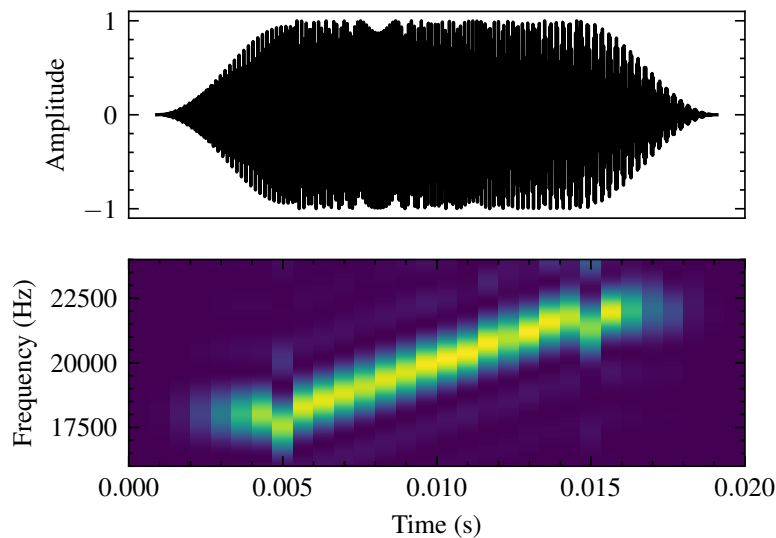


図 3.9: 送信信号の波形 (上) とその周波数 (下)
([126], © 2022 IEEE)



図 3.10: 参照スピーカとスマートフォン
の位置関係

3.6.2 時刻同期精度の評価

はじめに時刻同期精度に関する評価実験を行った．実験では時刻同期済みのスピーカ 2 台を用意し，同一信号を送信させた．そのうち 1 台は天井付近に設置し，もう 1 台の参照スピーカは図 3.10 のようにスマートフォンに密着させた．最初に参照スピーカとスマートフォンで時刻同期を行い，これを真値とした．その後，参照スピーカを十分遠い距離に移動させた．最後に提案手法を用いて天井付近の設置スピーカの送信時刻推定を行い，その推定結果と得られた真値との比較を行った．このとき床からの高さ z は既知とし，計測は 100 回行った．

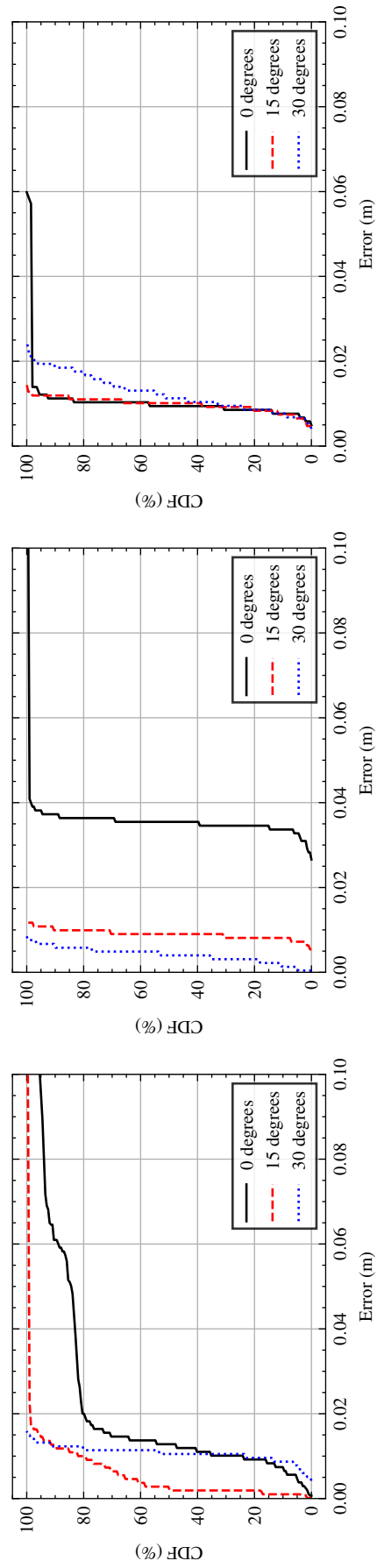
計測の結果，平均同期誤差が 0.1072 ms，音速を乗じて距離に変換すると 0.0369 m であることを確認した．

3.6.3 測距精度の評価

提案する時刻同期手法では，送信時刻 t_e を推定するためにアクティブセンシングを活用し，スマートフォンの高さ z を推定する．そのため，推定された高さ z の測距誤差が時刻同期の精度に直接影響する．そこでアクティブセンシングにおける z の測距精度を調査した．図 3.11 は，スマートフォンの高さを 1.0 m, 1.2 m, 1.4 m としたときの誤差に関する累積分布関数 (CDF) を示している．このとき，スマートフォンをピッチ方向に $0^\circ, 15^\circ, 30^\circ$ 傾けたときの精度も調査した．90th-percentile 誤差は，高さ 1.0 m のとき，順番に 6.10 cm,

1.27 cm, 1.23 cm (図 3.11a), 高さ 1.2 m のとき 3.72 cm, 0.99 cm, 0.68 cm (図 3.11b), 高さ 1.4 m のとき 1.12 cm, 1.19 cm, 1.84 cm(図 3.11c)であることを確認した。

また、提案する時刻同期手法で推定されるスピーカとスマートフォン間の同期誤差を調査した。図 3.12 は、スマートフォンの高さを 1.0 m に設定し、設置スピーカとスマートフォン間の距離を 2.0 m, 4.0 m, 6.0 m としたときの同期誤差に関する累積分布関数である。各距離における 90th-percentile での同期誤差は、それぞれ 0.10 ms, 0.14 ms, 0.16 ms であった。これに音速を乗じて距離に変換すると、それぞれ 3.62 cm, 4.92 cm, 5.68 cm になることを確認した。



(a) 高さ 1.0 m

(b) 高さ 1.2 m

(c) 高さ 1.4 m

図 3.11: アクティブセンシングによる高さ推定の精度 ([126], © 2022 IEEE)

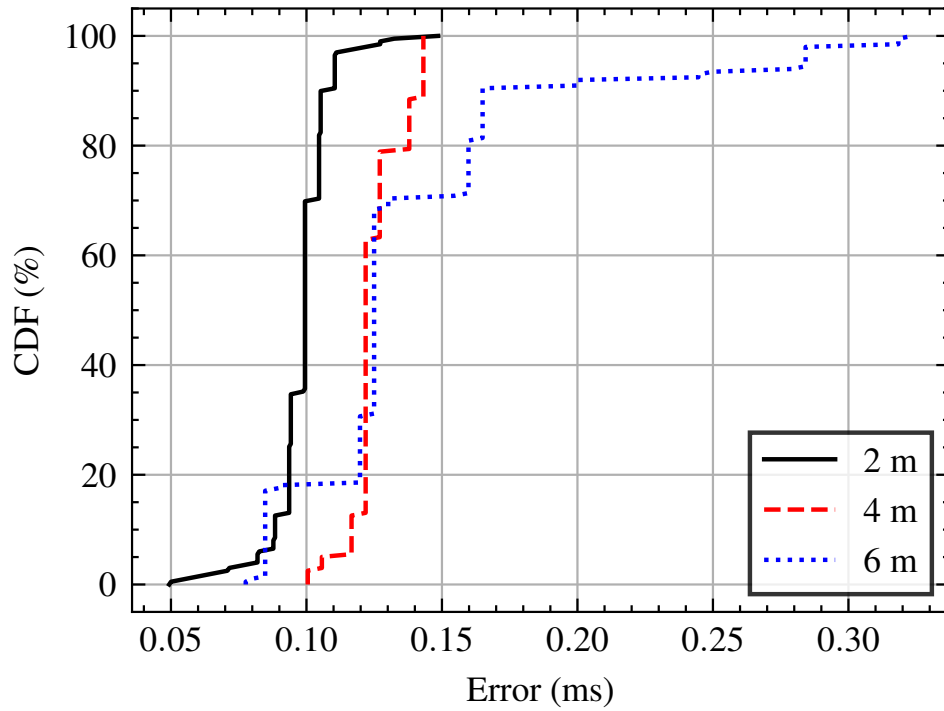
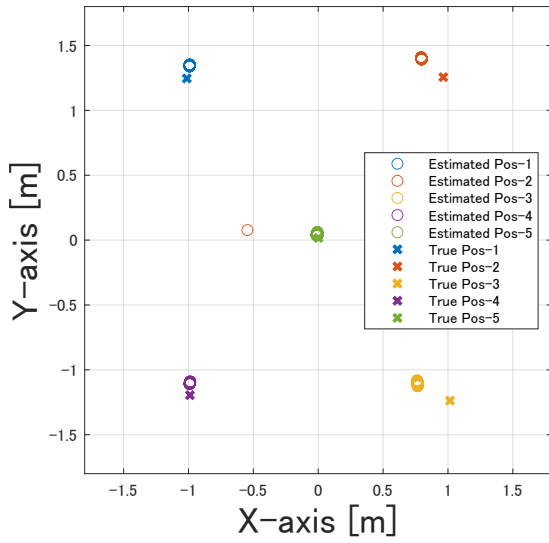


図 3.12: 距離による同期精度 ([126], © 2022 IEEE)

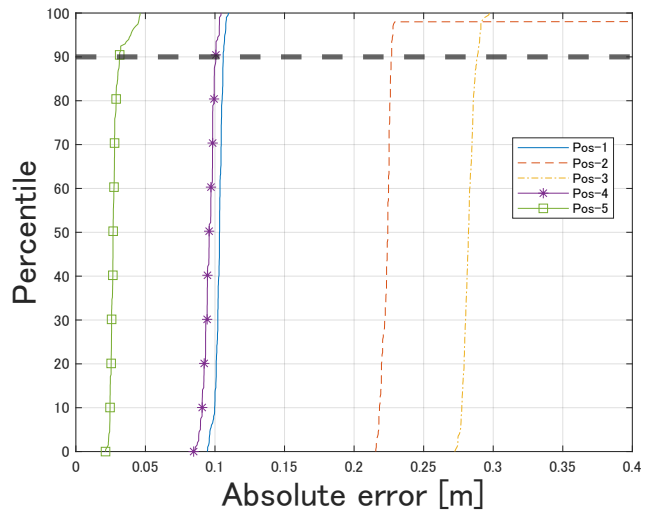
3.6.4 測位精度の評価

提案する時刻同期手法に基づく測位精度の評価実験を行った。固定時における測位精度を調査するため、三脚に固定されたスマートフォンを5カ所の地点で計測した。図 3.13a は、各地点における測位結果を示している。また、図 3.13b は、各地点における測位誤差の累積分布関数を示しており、90th-percentile での測位誤差は 28.9 cm であることを確認した。

次に、移動時における測位精度の評価実験を行った。ユーザに「M」を描くように移動してもらい、その測位誤差を計測した。スマートフォンの移動時では固定時とは異なり、アクティブセンシングで推定される高さ z の測距誤差が大きくなることが予想される。そこで、アクティブセンシングで生じる測位誤差の影響を明らかにするため、高さ z の真値を与えた場合の測位精度についても評価を行った。これは、ユーザではなくロボットによる移動で評価を行った。本論文では、正解値としてモーションキャプチャシステム (MCS) の MAC3D を利用した。図 3.14 は測位結果の軌跡、図 3.15 はその測位誤差の累積分布関数であり、90th-percentile での測位誤差が、高さ z を推定した場合で 38.9 cm、高さ z の真値を与えた場合で 28.8 cm であった。

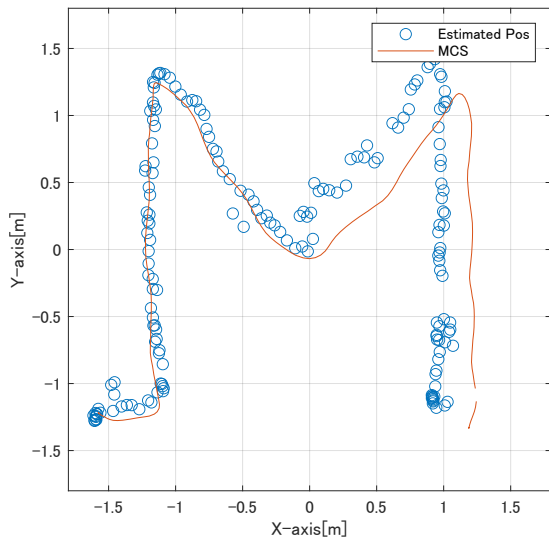


(a) 真値および推定座標

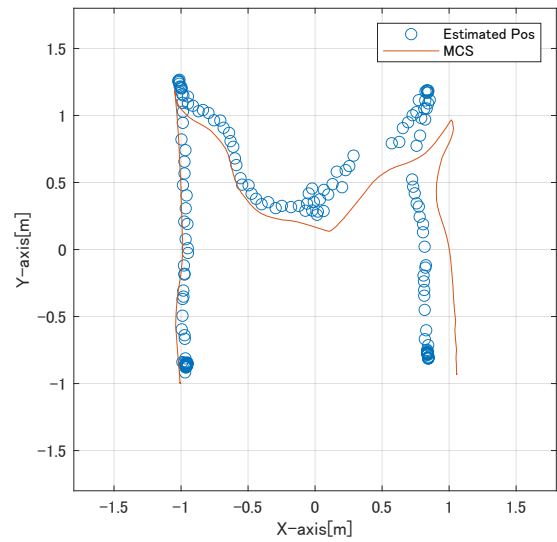


(b) 測位誤差の累積分布関数

図 3.13: 各計測地点での測位結果



(a) 推定した高さ z を用いた場合の軌跡



(b) 真値の高さ z を用いた場合の軌跡

図 3.14: 移動時の測位結果 ([126], © 2022 IEEE)

3.7 考察

3.7.1 測距および測位誤差の原因

図 3.15 より, ユーザが移動をしている場合の測位においても十分な精度で測位可能であることが確認できる. 一方で, アクティブ音響センシングで推定された高さ z を用いた時刻

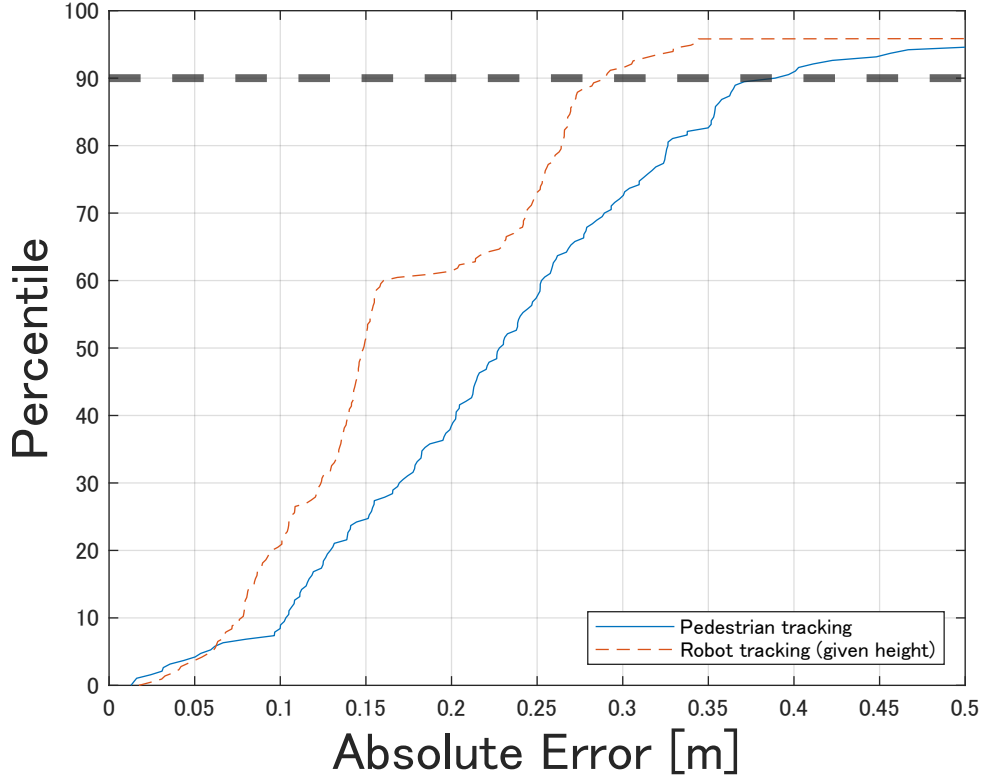


図 3.15: 測位誤差の累積分布関数 (CDF) ([126], © 2022 IEEE)

同期と、高さ z の真値を与えた時刻同期とで測位誤差に大きな差異があることも確認できる．そこで、ここではスマートフォンの高さの誤差に起因する測距誤差について議論する．スマートフォンの高さ誤差が Δz のとき、式 (3.3) によって推定される送信時刻で生じる測距誤差は次の通りである．

$$\begin{aligned}
 c \left(\frac{2z' \Delta z}{c^2(t_O - t_F)} \right) &= \frac{2z'}{d_O - d_F} \Delta z \\
 &= \frac{2z'}{\sqrt{d_{xy}^2 + (z' - z)^2} - \sqrt{d_{xy}^2 + (z' + z)^2}} \Delta z \\
 &= - \frac{\sqrt{d_{xy}^2 + (z' - z)^2} + \sqrt{d_{xy}^2 + (z' + z)^2}}{2z} \Delta z
 \end{aligned} \tag{3.10}$$

このとき、 d_{xy} は、 xy 平面上のスピーカとスマートフォン間の距離として定義され、この誤差は、パラメータ Δz 、 z' 、 z 、および d_{xy} の影響を受ける．スピーカの高さ z' は固定されており、スマートフォンの高さ z は大きく変化しないため、実験に基づき、 z' と z をそれぞれ 2.3 m と 1.0 m に設定する．図 3.16 は、高さの誤差 Δz に応じた、距離 d_{xy} で生じる測距誤差を示している．図 3.16 より、アクティブ音響センシングでの高さ推定精度が向上することで、より高精度な時刻同期による測距、測位が可能であることを確認できる．また、本

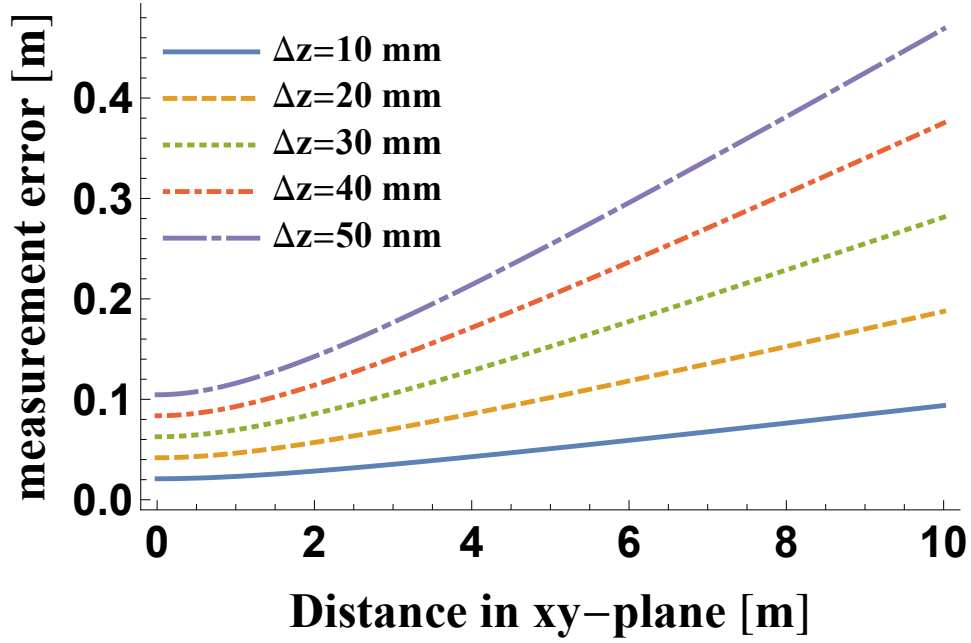


図 3.16: 高さの推定誤差 Δz で生じる距離 d_{xy} と測距誤差の関係 ([126], © 2022 IEEE)

研究では、各フレームで推定される送信時刻を各フレームにおける測距、測位に利用する。一方で、図 3.16 より、スピーカとスマートフォン間の距離が短いほど送信時刻の誤差が小さくなることを確認できる。そのため、スピーカとスマートフォン間の距離が短い地点で推定された送信時刻を全てのフレームに対して利用することで、測距、測位誤差の影響を軽減することが期待できる。

3.7.2 測位結果の曖昧性

提案手法では、2つの設置スピーカと2つの鏡像スピーカを含む4つのスピーカが xz 平面上に配置される。そのため、 xy 平面における2台の設置スピーカを結ぶ直線に対して対称な位置に、スマートフォンの2つの測位結果が算出される。本論文では、この問題を回避するために、実験環境において壁側に2つのスピーカを配置することで、フロアプランに基づき測位結果を一意に定めている。この問題は、今後の課題として取り組む必要があり、WiFi や BLE の RSS を活用し大まかな位置を推定することで、位置を一意に定める方法が考えられる。

3.8 まとめ

本章では床からの反射波に着目した擬似的な時刻同期手法についてまとめた。時刻同期精度は高さ推定に依存するが、正確な高さが与えられた場合は平均誤差が 3.69 cm になること

を確認した。この同期精度は目標である 10 cm 以下の誤差を達成するのに十分な精度である。一方で、高さ推定の誤差が測距誤差に影響を及ぼすことが課題だが、逐次的に推定を繰り返す、一度特定の既知の座標に設置する、などの方法で誤差を軽減することができる。

スピーカとマイクロフォン間でモジュールを用いない時刻同期技術は、電波の利用が困難な医療施設や原子力発電所などの特殊な環境での利用が期待される。また、この技術により、従来は 3 台必要とされていたスピーカが 2 台で測位を実現できるようになるため、設置コストの削減と測位範囲の拡大が可能となる。この利点は、音響測位の普及を促進する上で重要である。

第 4 章

FDM と TDM を組み合わせた測位領域の拡張手法

4.1 はじめに

本章では、測位領域の拡張手法について説明する。従来のパッシブ型の音響測位手法では FDM が主流だが、高精度測位が可能な領域が制限される課題がある。これは主に、モバイルデバイスに搭載された安価かつ低性能なマイクロフォンに起因する。モバイルデバイスに内蔵されたマイクロフォンはスピーカと常に正対状態であるとは限らず、受信時には到来角が大きくなる場合が多い。そのためモバイルデバイスの指向性によって受信信号が劣化し、帯域が足りず受信時刻の検出が困難となる。特に周波数が高くなるほど指向性による影響が大きくなる傾向がある。

提案手法では、TDM と FDM を組み合わせた送信方式を利用し、測位頻度と測位領域の課題に対処することを試みる。具体的には、FDM 方式に加えて送信毎に各スピーカに対する帯域の割り当てを変更することで指向性による周波数特性の影響を緩和する。提案手法の利点は、従来の FDM 方式と同等の性能および測位頻度で測位可能であり、また測位頻度を落とすことで高精度測位が可能な領域を拡張できる点である。

本章の内容は、筆者の研究成果 [128, 129] を基にまとめたものである。

4.2 送信方式

提案する送信方式では、図 4.1 に示すように、各スピーカからスケジューリングされた順番で毎回異なる帯域の信号を送信させる。このとき全てのスピーカは同時に送信するが、互いの帯域が干渉しないように設計する。例えば、3 台のスピーカを用いて 2 次元測位を行うことを想定する。このとき高域、中域、低域の 3 つの帯域の信号を用いる。これらの信号を

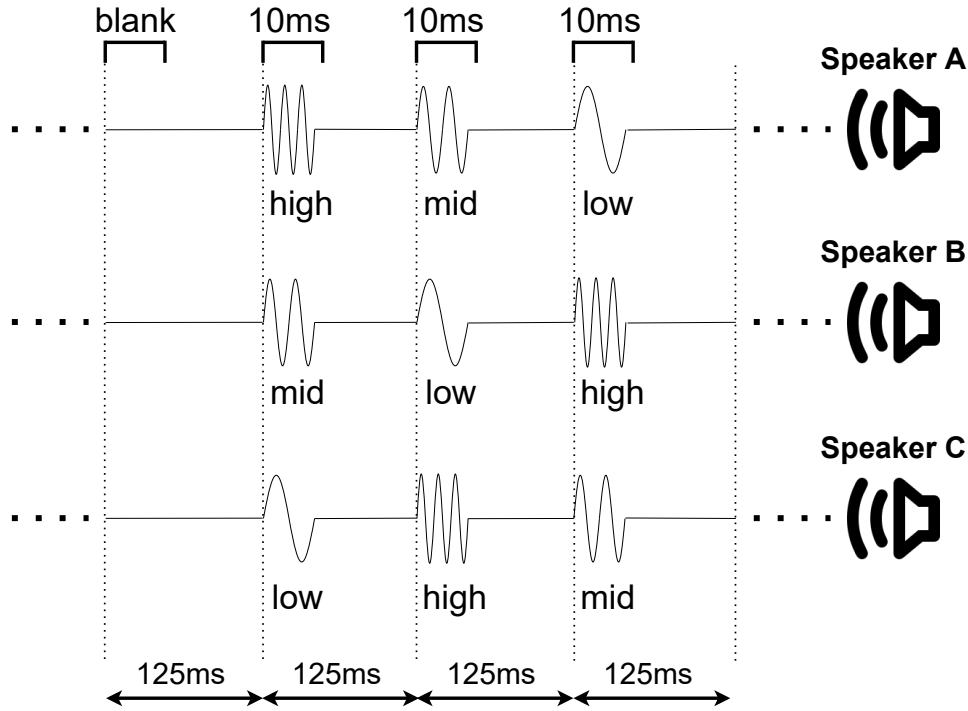


図 4.1: 提案送信方式の概要

4つのタイムスロットと呼ばれる時間単位に分割して送信を行い、これが1周する周期をフレームとする。最初のタイムスロットでは、一般に、パイロット信号と呼ばれる周期の開始を識別するための信号を送信する。本論文ではこのパイロット信号の代わりに無音時間を設けることで周期の開始を識別する。

この送信方式によりマイクロフォンの周波数特性による影響を軽減し、広範囲での測位が可能となる。指向性の影響で単一タイムスロットでの測位が困難な位置では複数タイムスロットを使い、受信しやすい帯域を使うことで測位不可となる要因を抑制することができる。

本提案方式の特徴は信号が十分受信できる位置において、FDM方式と同等の測位頻度を保つことができる点である。これは、TDM方式のみでは1回の測位に1フレーム必要であるのに対し、提案方式では単一タイムスロットでも全スピーカからの信号を受信できるためである。

提案する送信方式での送信信号はチャープ信号を用いる。チャープ信号は測距を目的として水中ソナーやレーダの分野で広く利用されている。使用する線形チャープ信号 $s^i(t)$ は周波数が線形に増加する特徴があり以下の式で表せる。

$$s^i(t) = \exp \left[j2\pi \left(f_0^i t + \frac{k}{2} t^2 \right) \right], \quad (4.1)$$

$$k = \frac{f_1^i - f_0^i}{T},$$

このとき i 番目のタイムスロットで送信する信号の f_0^i は開始周波数, f_1^i は終了周波数, T は f_0^i から f_1^i までのスweep時間である. $s^i(t)$ は複素信号であり, 実際に送信される信号は実部のみとなる.

4.3 受信信号の検出

MF (Matched Filter; マッチドフィルタ) は, レーダの分野で信号の受信時刻を検出するために広く利用されているパルス圧縮技術である. MF では受信信号と参照信号の畳み込み処理を行い, 最大値をとるインデックスを受信時刻とする. 畳み込み処理によるマッチドフィルタを以下に示す.

$$R^i(t) = \int_t^{t+T_s} r(t)s^i(\tau - t)d\tau \quad (4.2)$$

ここで T_s はタイムスロット長, $r(t)$ は受信信号, $s^i(t)$ は参照信号 (送信信号) である. $s^i(t)$ は複素信号であるため, 算出結果 $R^i(t)$ も同様に複素信号となり, その絶対値から包絡線を算出できる.

4.4 コヒーレント積分

提案手法では, コヒーレント積分 [127, 130] を応用し, 受信時刻の検出性能の向上を図る. 通常のコヒーレント積分では位相情報を保持している信号を加算処理し, SNR を向上させる. 加算された互いの信号の位相差が小さいと振幅は大きくなるが, 位相差が大きいと相殺され振幅は小さくなる. 特に同相で m 個の信号を加算した場合では SNR が m 倍に改善する効果がある. 本論文でも同様に MF 処理後に位相情報を保持した状態での加算処理を行い, MF で算出された受信時刻の位相差が 0 になる性質を利用する. 受信時刻付近では位相差が小さいため振幅が大きくなり検出しやすくなるが, 一方その他では位相差が大きいいため振幅は小さくなる.

提案手法では N を 1 フレームに含まれるタイムスロット数とすると, 1 フレームで $N - 1$ 個の信号を送信する. モバイルデバイスで受信した $N - 1$ 個の信号の MF 算出結果に対して加算処理を行い, 包絡線の最大値から受信時刻の検出を行う. コヒーレント積分を応用した受信時刻の検出手法を以下に示す.

$$t_{co} = \arg \max_t \left\| \sum_{i=1}^{N-1} R^i(t + T_s(i - 1)) \right\| \quad (0 \leq t < T_s) \quad (4.3)$$

ここで N はタイムスロットの総数, $R_i(t)$ は i 番目に受信した信号に対して MF で算出した複素信号である.

図 4.2 にシミュレーションにより得られる包絡線の結果を示す. ここでコヒーレント積分

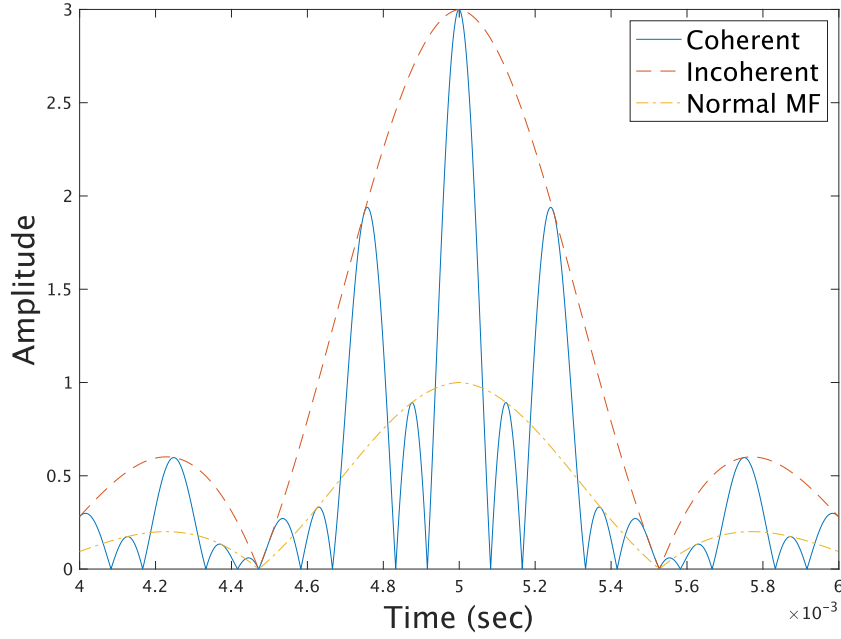


図 4.2: コヒーレント積分による包絡線のシミュレーション

は位相情報を含めた状態での加算処理を行うのに対し、インコヒーレント積分は包絡線の振幅のみでの加算を行う。このシミュレーション結果から、コヒーレント積分は他の処理と比較してメインローブが最も細く時間分解能が高いことから、精度の高い受信時刻検出ができる。シミュレーションでは大きいサイドローブが現れるが、メインローブより大きくなることは確認できなかったため、そのまま最大値を取るインデックスを選択することで受信時刻を推定する。

4.5 TDoA による測位

測位計算には TDoA を利用する。TDoA は送受信機間の時刻同期が困難である場合に利用される。一般に TDoA では M 次元の位置推定には少なくとも $M + 1$ 台のスピーカが必要となり、2 次元測位には少なくとも 3 台のスピーカが必要となる。設置されたスピーカの位置を既知として以下の連立方程式を立てられ、この方程式から x, y を求めることで位置推定ができる。

$$\begin{aligned}
 \sqrt{(x_1 - x)^2 + (y_1 - y)^2} - \sqrt{(x_2 - x)^2 + (y_2 - y)^2} &= c \cdot (t_1 - t_2), \\
 \sqrt{(x_1 - x)^2 + (y_1 - y)^2} - \sqrt{(x_3 - x)^2 + (y_3 - y)^2} &= c \cdot (t_1 - t_3),
 \end{aligned} \tag{4.4}$$

ここで x_j, y_j は j 番目の 2 次元平面でのスピーカの位置を表し、 x, y は推定するスマートフォンの位置である．また、 t_j は j 番目のスピーカからの受信時刻であり、式 (4.3) から各スピーカ毎に求める． c は音速であり、室温 H （摂氏）を用いて以下の式で計算を行う．

$$c = 331.3 + 0.606H \quad (4.5)$$

4.6 群遅延の補正

実環境での計測において、チャープ信号の周波数によって受信時刻に遅延が生じることを確認した．固定したデバイス間で送受信した時刻差から ToF を計測してみると、受信する周波数によって ToF が変化した．この遅延はチャープ信号の周波数を同じにして距離を変えても一定であった．この受信時刻の遅れは群遅延の影響だと考えられる [127]．群遅延は信号がフィルタを通過する際に生じる遅延時間と捉えられる．スマートフォンにはアンチエイリアシングフィルタ (AAF) と呼ばれる折り返しを防止するためのローパスフィルタが搭載されているため、これが遅延に起因していると考えられる．群遅延はフィルタの位相特性の微分で表現され、スマートフォンでの受信時では周波数が高くなるほど遅延時間および分散が大きくなる傾向がある．本論文では MF 処理後に、あらかじめ計測を行った値を用いて受信時刻の補正を行う．

4.7 評価実験

4.7.1 事前実験 1

事前実験としてモバイルデバイスの指向性による周波数特性を調査した．実験では図 4.3a に示すように距離を一定にしながらスマートフォンを 15 度ずつ傾け、そのとき得られた周波数特性を調査した．ここで 0 度はスピーカとスマートフォンが正対の状態を表す．図 4.3b は Sumsung Galaxy S10 plus の周波数特性であり、音響測位で使用する 7 kHz–23 kHz までは拡大した図である．この結果から周波数が高くなる、もしくは到来角が大きくなるほど周波数特性が劣化する傾向があることが分かる．

4.7.2 事前実験 2

続いて事前実験として、4 つの端末を用いて群遅延の調査を行った．実験では、スピーカとスマートフォンを 1.0 m 離して固定し、はじめにカットオフ周波数 (24 kHz) から十分離れた 10 kHz–12 kHz のチャープ信号を送信する．この信号をスマートフォンで受信した時刻を t_{ini} とする．その後、十分長い時間 T_{con} を空けてから、調査する周波数のチャープ信号を送信する．このとき受信した信号の受信時刻 t_{rec} を検出し、群遅延量 $t_{gd} = t_{rec} - t_{ini} - T_{con}$

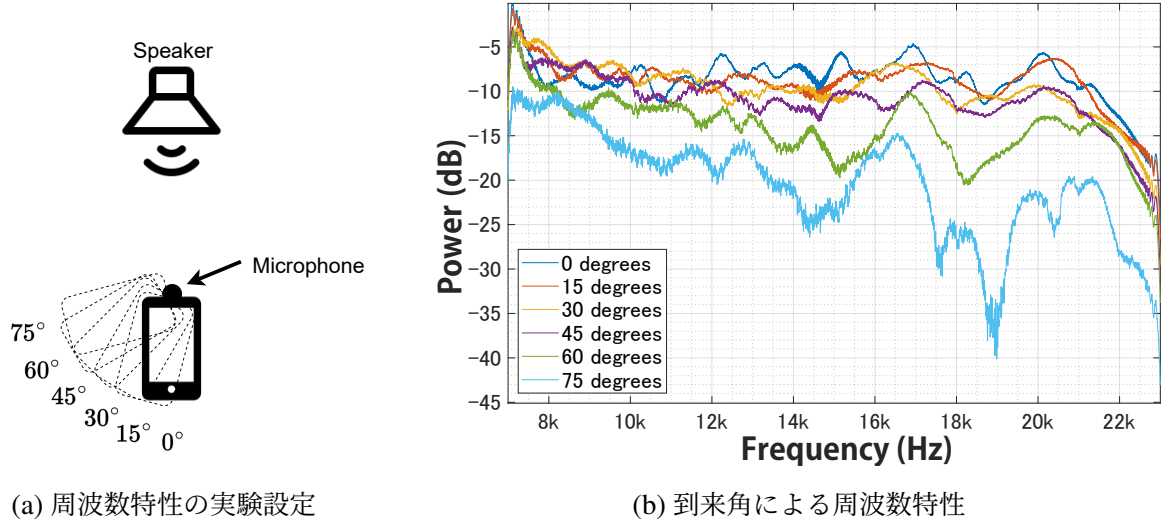


図 4.3: 実験設定と結果

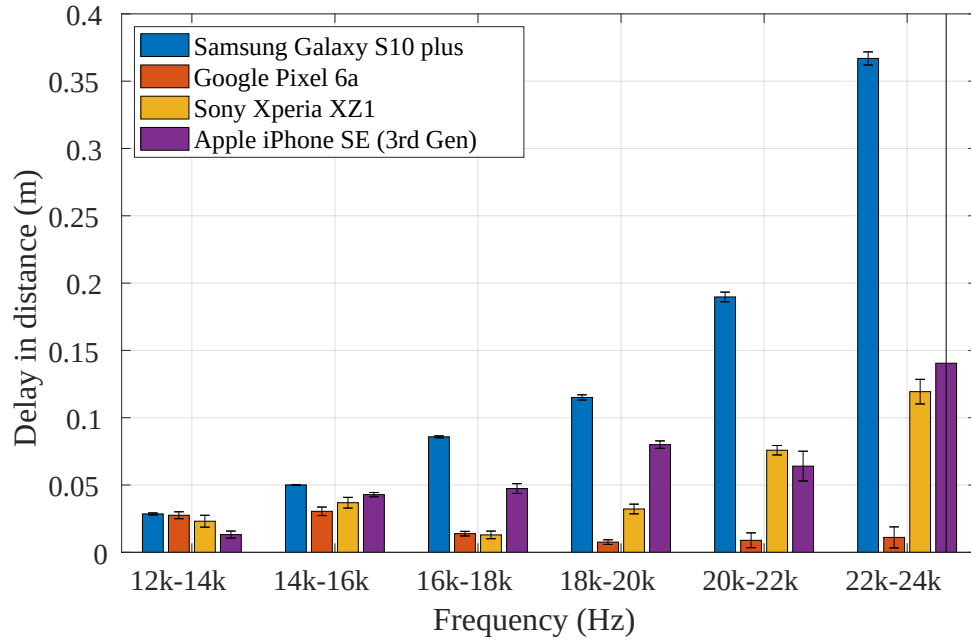


図 4.4: 機種ごとの群遅延の比較

を算出する．このとき t_{gd} は遅延が生じていなければ 0 になる．得られた計測結果を図 4.4 に示す．この図では各端末で 100 回計測して得られた平均値および標準偏差を示している．また、縦軸は遅延時間に音速を乗算して距離に変換している．この結果から端末によって群遅延量および分散が異なることが分かる．本論文では、群遅延の影響や補正の効果を検証するため分散が最小となった Galaxy S10 plus を用いた．群遅延は実際には指向性などの影響も受けるが、厳密な測定は困難であるため本論文では 100 回計測した平均値を補正值とする．

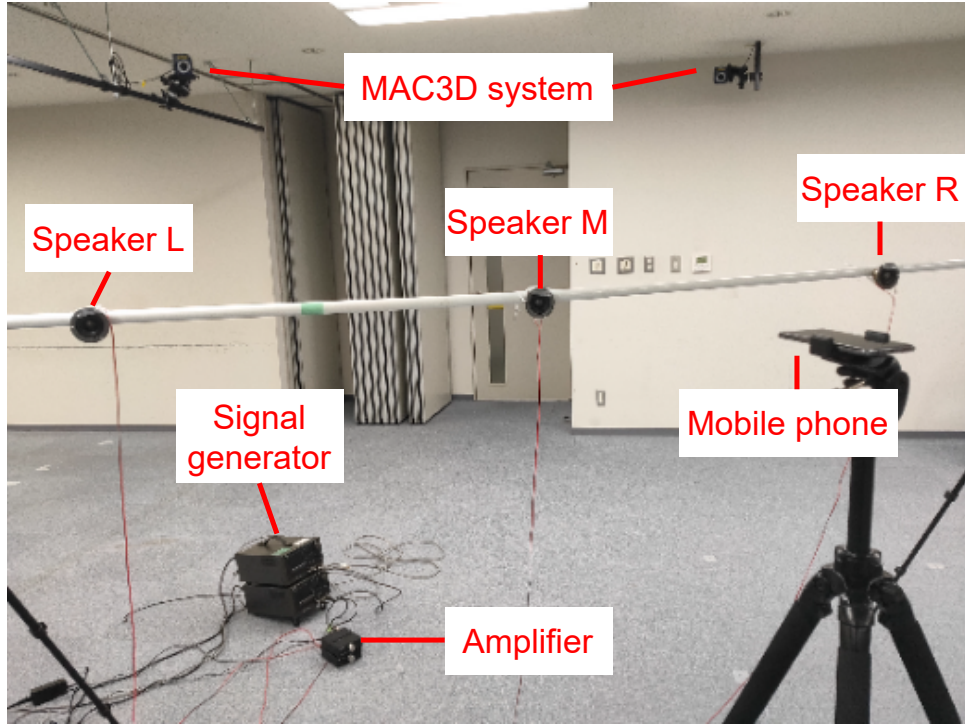


図 4.5: 実験環境

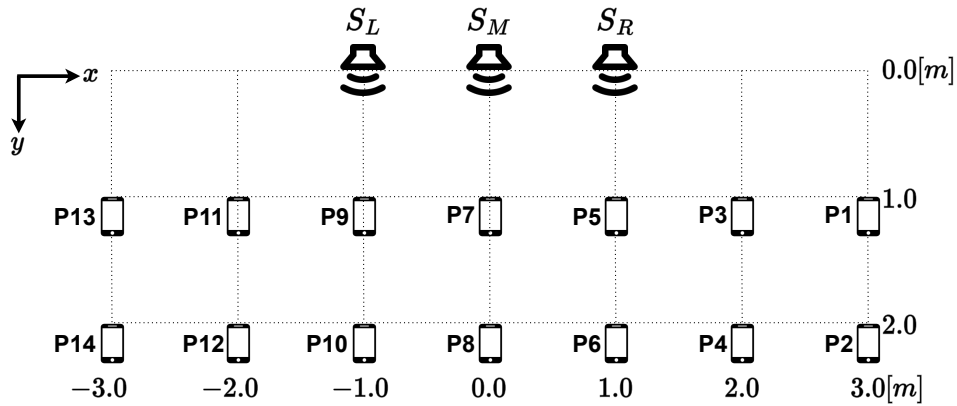


図 4.6: スピーカとスマートフォンの配置

4.7.3 測位実験の環境

提案手法の有効性を確認するために3台のスピーカを用いて TDoA による2次元測位実験を図 4.5 に示す環境で行った。スピーカは周波数特性が良好な Fostex 製の FT200D[131] を用いた。信号生成器として NF 製の WF1948, スピーカアンプとして Fostex 製の AP20d を用いた。計測配置は図 4.6 の14か所で三脚に固定して行った。このときスマートフォンの床からの高さは一定にし、 $-y$ 方向を向くように設置した。また、高精度モーションキャ

プチャシステムであるナックイメージテクノロジー製の MAC3D システムを用いて、スピーカと計測点の位置を測定し、これを真値として TDoA 測位の結果と比較を行った。実験を行ったときの室温は 20.9 度であった。スマートフォンのサンプリングレートは 48 kHz に固定し、スマートフォンの上部に内蔵されているマイクロフォンで録音を行い、その録音データはオフラインを解析した。

4.7.4 測位実験

信号処理におけるパラメータは帯域と信号長などのトレードオフを考慮することが重要である。本論文では全ての信号長は文献 [96] と同様に 10 ms に設定し、チャープ信号の帯域は 2 kHz になるように設定した。また信号の干渉を避けるためのガードバンドは 2 kHz に設定した。ここで 20 kHz–22 kHz, 16 kHz–18 kHz, 12 kHz–14 kHz のチャープ信号をそれぞれ、高域チャープ、中域チャープ、低域チャープとする。1 フレームは 500 ms に設定し、4 つの 125 ms のタイムスロットが含まれる。このとき最初のタイムスロットはパイロット信号として全てのスピーカから何も送信しない無音時間を設けることでフレーム周期を判別する。表 4.1 に各実験の条件を示す。この表におけるタイムスロット数は、測位を 1 回実行するのに使用する数を表している。

表 4.1: 実験条件

	実験 1: 提案手法	実験 2: 群遅延の補正	実験 3: コヒーレント積分	実験 4: タイムスロット数	実験 5: FDM	実験 6: TDM
送信方式	FDM+TDM	FDM+TDM	FDM+TDM	FDM	FDM	TDM
遅延補正の有無	有り	無し	有り	有り	有り	無し
加算処理	コヒーレント積分	コヒーレント積分	インコヒーレント積分	無し	コヒーレント積分	無し
タイムスロット数	3	3	3	1	3	3
周波数帯域	低域, 中域, 高域	低域, 中域, 高域	低域, 中域, 高域	低域, 中域, 高域	低域, 中域, 高域	低域

表 4.2: 測位実験の結果

実験	平均誤差 (cm)	平均 90th-pecentile(cm)	最小 / 最大 90th-percentile (cm)	標準偏差 (cm)	測位に 失敗した割合 (%)	誤差 1.0 m 以下 の割合 (%)
1	17.56	18.94	2.12(P9) / 63.25(P1)	0.93	0.0	100.0
2	25.58	27.96	2.13(P9) / 205.15(P1)	2.11	0.0	92.86
3	49.86	32.02	1.29(P6) / 121.34(P13)	25.61	4.21	92.92
4	65.67	98.37	5.36(P7) / 367.84(P14)	50.08	5.00	79.25
5	64.30	66.20	7.29(P5) / 216.44(P13)	1.89	0.43	78.77
6	28.45	39.37	4.20(P7) / 184.91(P14)	8.65	7.21	93.61

実験 1：提案手法

実験 1 では提案手法の評価を行う．図 4.1 のように，各スピーカは他のスピーカと干渉しない帯域のチャープ信号を送信する．右スピーカ S_R ，中央スピーカ S_M ，左スピーカ S_L からは以下の順番で送信するように設定した．

- S_R : 無音 → 高域 → 中域 → 低域
- S_M : 無音 → 中域 → 低域 → 高域
- S_L : 無音 → 低域 → 高域 → 中域

受信した信号に対して MF 処理後に群遅延の補正を行った．その後，3 つのタイムスロットを用いてコヒーレント積分による複素信号の状態での加算処理を行い，時間窓の最大となるピークのインデックスを検出した．最後に受信時刻を用いて TDoA による測位計算を行った．

実験 2：群遅延の補正

実験 2 では群遅延の補正の有無による測位精度の評価を行った．実験 1 で得られたデータに対して MF 処理後に群遅延の補正を行わずにコヒーレント積分を適用し，時間窓の最大となるピークを検出した．

実験 3：コヒーレント積分

実験 3 ではコヒーレント積分処理の測位精度の評価を行った．実験 1 で得られたデータに対して群遅延の補正後，包絡線で平均をとるインコヒーレント積分を適用したときの測位結果との比較を行った．コヒーレント積分は位相情報を含めた状態での加算処理を行うのに対し，インコヒーレント積分は包絡線の振幅のみでの加算を行う．インコヒーレント積分によって検出する受信時刻 t_{in} は以下の式で表せる．

$$t_{in} = \arg \max_t \sum_{i=1}^{N-1} \|R^i(t + T_s(i-1))\| \quad (0 \leq t < T_s) \quad (4.6)$$

実験 4：タイムスロット数

実験 4 では単一タイムスロットのみによる測位精度の評価を行った．実験 1 で得られた単一のタイムスロットのデータに対して群遅延の補正後に MF 処理によるピーク検出を行った．このときスピーカからは提案手法と同じ 4.1 の順序で送信を行い，フレーム当たり 3 回の測位を行った．この実験と提案手法である 3 つのタイムスロットを利用した手法との比較を行った．

実験 5：FDM

実験 5 では従来手法である FDM 方式による測位精度の評価を行った。実験では S_R, S_M, S_L にそれぞれ高域チャープ、中域チャープ、低域チャープを割り当て、各タイムスロットで同じ帯域のチャープ信号を送信させた。また提案手法と実験条件を合わせるため、群遅延の補正後に 3 つのタイムスロットを用いたコヒーレント積分を適用して受信時刻の検出を行った。

実験 6：TDM

実験 6 では従来手法である TDM 方式による測位精度の評価を行った。実験 1 で得られたデータに対して、信号の劣化が少ない低域チャープ信号を送信するタイムスロットのみスピーカ毎に MF 処理を行いピーク検出を行った。この実験では低域チャープのみを用いるため群遅延による受信時刻の補正值は全て同じであり、受信時刻差で測位を行う TDoA に影響しないことから補正は行っていない。

4.7.5 実験結果

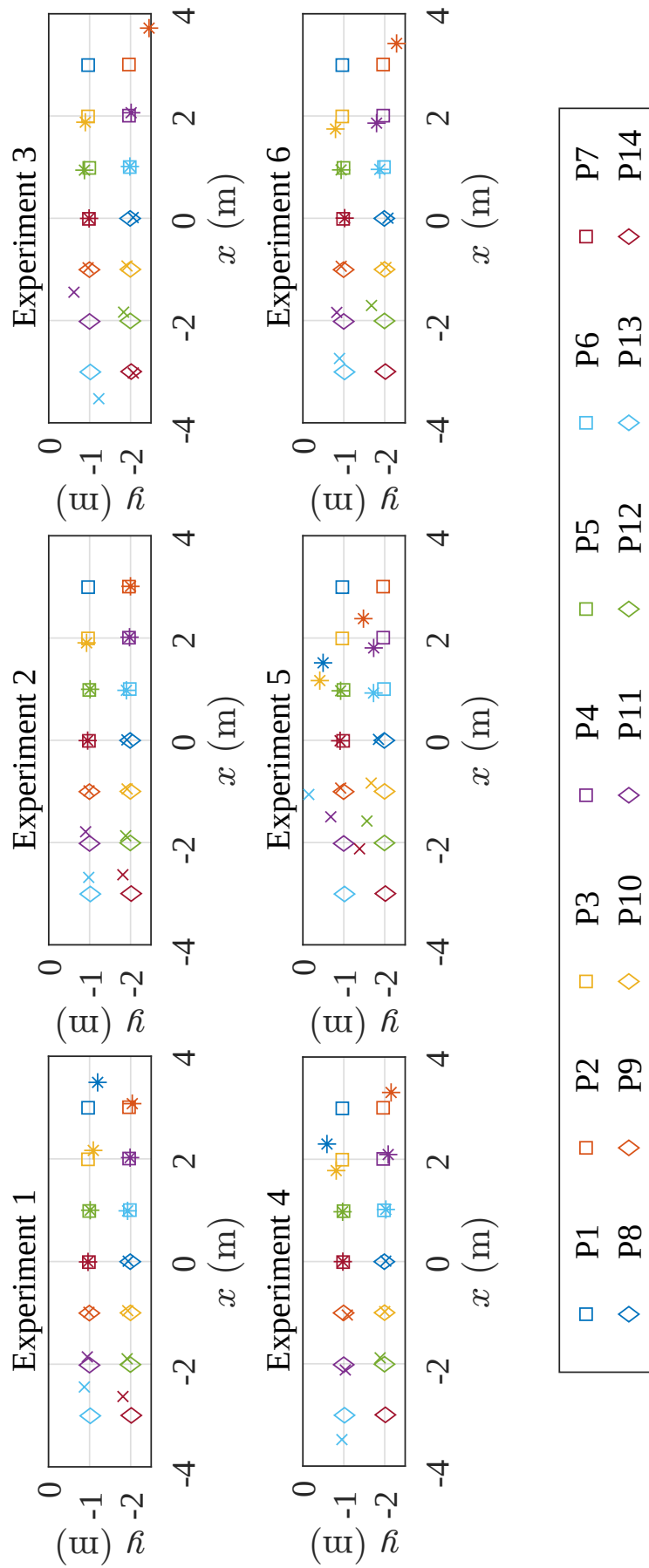


図 4.7: 各実験の測位結果 (□, ◇: 真の位置, ×, *: 平均推定位置)

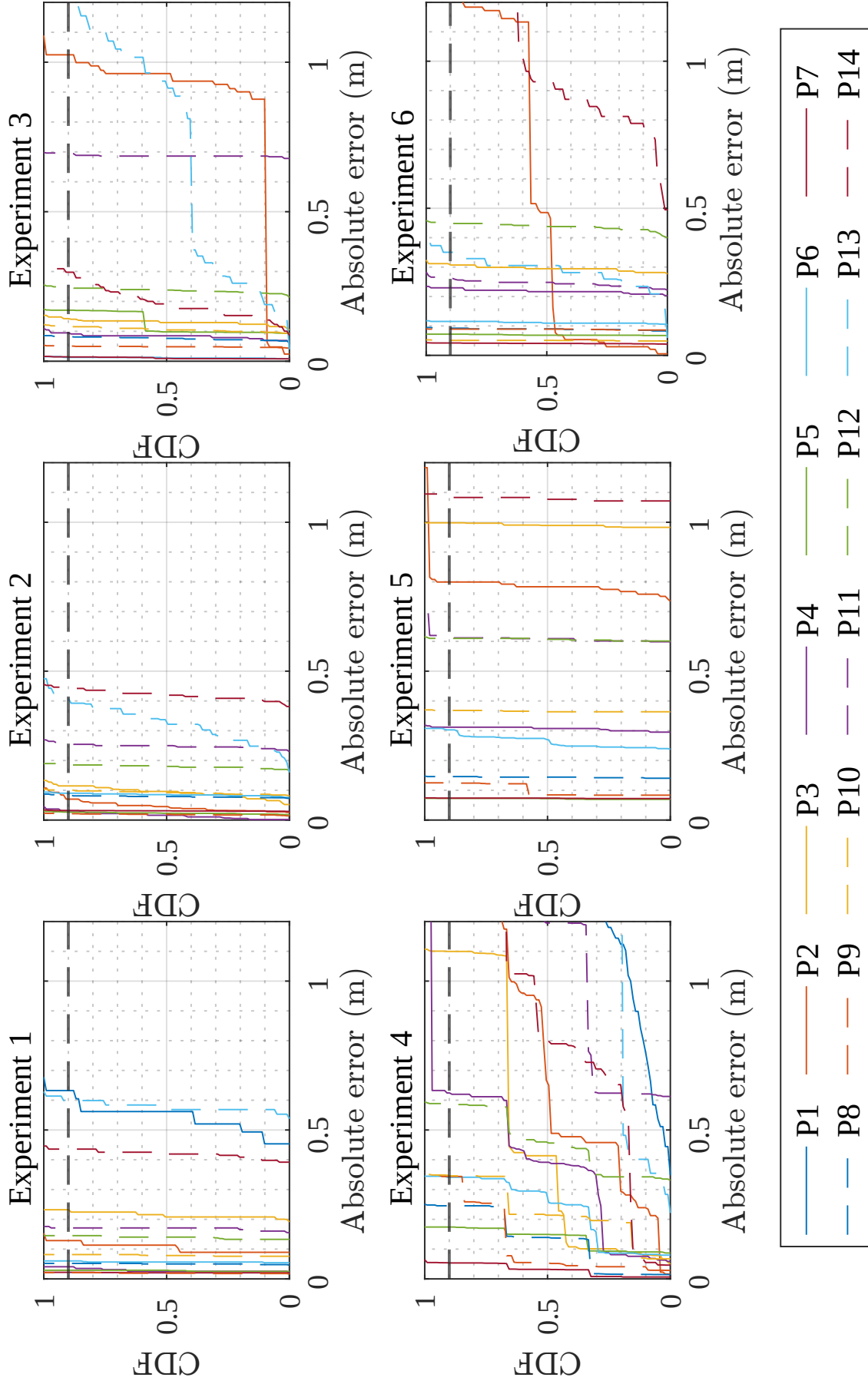


図 4.8: 各実験の誤差の累積分布関数 (CDF)

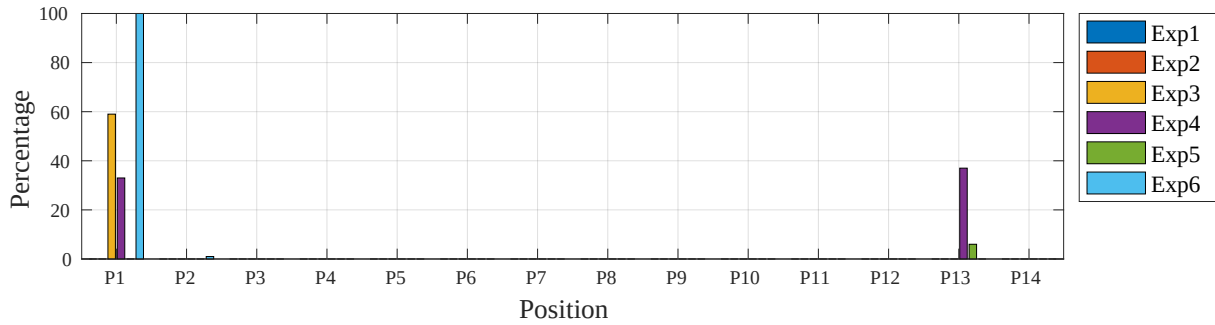


図 4.9: 各測位点での測位計算に失敗した割合

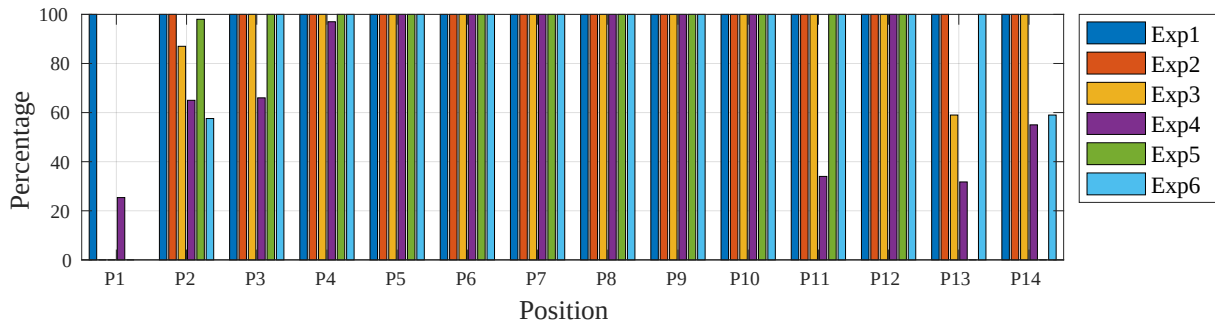


図 4.10: 各測位点での誤差が 1.0m 以下になった割合

実験は各計測点で 100 回測定を行い評価を行った．表 4.2 は各実験で得られた全体の結果をまとめたものである．この表では各実験の絶対誤差，90th-percentile 誤差，標準偏差，TDoA 処理時に解が得られなかった割合，最後に誤差が 1.0 m 以下となる割合を示している．また図 4.7 は測位結果を可視化したものであり，図 4.8 は誤差の累積分布関数を表す．この図では 1.20 m 以下の累積分布関数を表示している．

この結果から x 軸の端の測定位置 (P1, P2, P13, P14 など) では信号の到来角が大きくなるため，誤差と分散が大きくなる傾向があることが分かる．また x 軸の中央付近の測定位置 (P7, P8 など) では，到来角が小さくなるため誤差および分散は小さくなった．

続いて図 4.9 は，ピークの誤検出により測位計算に失敗した割合を示している．到来角が大きくなるような測定位置では，手法によって割合は高くなっているが提案手法では確認されなかった．

図 4.10 は各測位点での誤差が 1.0 m 以下に収まった割合を示している．この図より，中央付近では全ての実験で 1.0 m 以下に収まっていることが確認できるが，到来角が大きくなるような位置では提案手法のみ収まっている．

実験 2 の結果から，群遅延の補正を行う提案手法の方が分散が小さくなることを確認した．実験 3 の結果から，提案手法で使用するコヒーレント積分処理の方が測位誤差および分散が小さくなることを確認した．実験 4 の結果から，中央付近 (P5, P6, P7, P8, P9, P10)

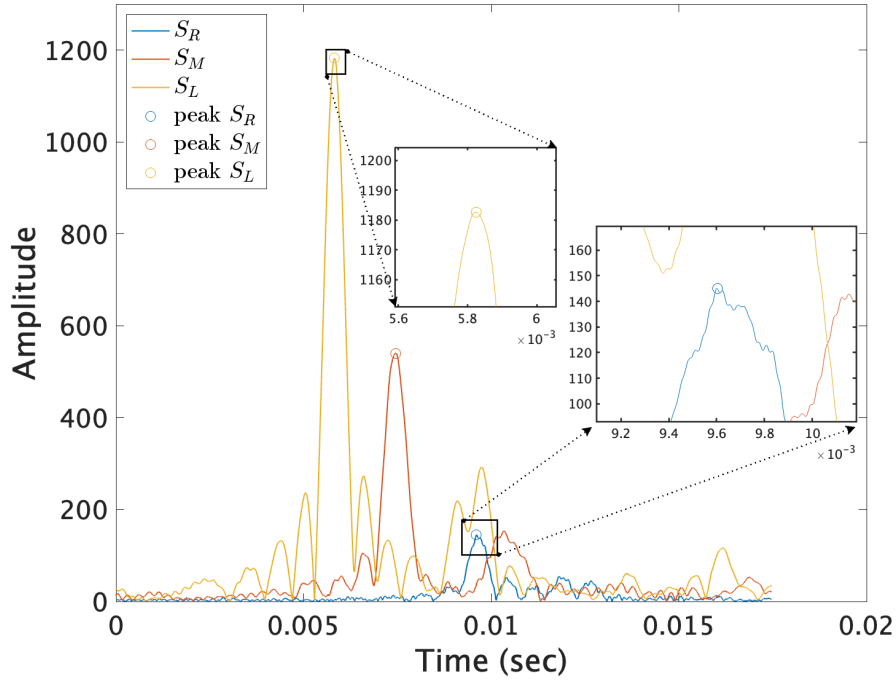


図 4.11: FDM 方式により P13 で得られた包絡線と拡大図

では単一タイムスロットのみでも安定した測位ができていることが分かる．一方，中央付近以外での誤差および分散は大きくなった．実験 5 の結果から，従来手法の FDM 方式では中央付近でのみ高精度測位が可能であり，到来角が大きくなる計測位置では誤差が大きくなることを確認した．実験 6 の結果から，信号の劣化が少ない低域チャープのみを用いた TDM 方式では，誤差が 1.0 m 以下に収まる割合が 93.61 % となり，広範囲で測位可能ではあるものの，提案手法との比較では性能が落ちることを確認した．

以上の結果から提案手法の有効性が確認できた．特に誤差が 1.0 m 以下となる割合が 100 % となり，到来角が大きい測定位置でも頑健な測位が可能になっていることが分かる．

4.8 考察

4.8.1 群遅延の補正

実験 2 の結果より群遅延の補正を行わない場合，測位時の分散が大きくなることを確認した．これは受信する帯域の信号によって受信時刻にズレが生じ，コヒーレント積分処理時に正常に加算されないためと考えられる．一方，図 4.7 の計測結果と図 4.8 の CDF の図より，大部分で測位精度は提案手法と同等であることが分かる．これは，事前実験 1 より群遅延はスピーカとマイクロフォン間の距離が長くなる方向で推定される．式 4.4 では受信時刻差に

よる測位計算を行うが、ここで群遅延量が相殺されるため測位への影響が小さくなったと考えられる。

4.8.2 コヒーレント積分

実験 3 の結果から、提案手法で使用するコヒーレント積分により測位誤差および分散が小さくなることを確認した。これは 4.2 のシミュレーションに示すように、振幅が最大となるメインローブのピークが細くなり、頑健なピーク検出および受信時刻の検出が可能になるためだと考えられる。

4.8.3 タイムスロット数

実験 4 の結果から、中央付近 (P5, P6, P7, P8, P9, P10) では単一タイムスロットのみでも安定した測位ができていることが分かる。一方、中央付近以外での誤差および分散は大きくなった。また、誤差の累積分布関数は 3 段の階段状になっているが、これは受信した信号の周波数によって測位誤差が変化しているためである。特に高域チャープが送信されたタイムスロットでの測位誤差が大きくなることを確認した。

単一タイムスロットでは広範囲での安定した測位は困難だが、中央付近の安定して受信可能な領域では、単一タイムスロットでも測位可能であることが確認できる。このような領域では測位頻度は FDM 方式と同等となる。提案手法では TDM と FDM の利点を活かし、測位誤差の小さい領域であれば単一タイムスロットで、困難な位置であれば 3 タイムスロットを使うことで高精度測位が可能となる。このとき必要タイムスロット数の判別は今後の課題としたい。

4.8.4 FDM による測位

実験 5 の結果より従来手法である FDM 方式では、広い領域における高精度測位が困難になることを確認した。一方、到来角が小さくなる中央付近では高精度測位が可能であり、分散も小さくなることを確認した。

図 4.11 は、FDM 方式による P13 で得られた包絡線である。このとき到来角が大きくなる S_R からの信号には、同図の拡大図に示すようなノイズが含まれていることが分かる。この結果、ピークが不安定となり結果的に間違った受信時刻を検出している。P1、P2、P3、P13、P14 でも同様の波形が観測された。これらの理由により、FDM ベースの手法では測位誤差が大きくなったことが考えられる。一方、中央付近や提案手法ではこのような凹凸のある不安定なピークは確認されず、滑らかで鋭いピークが検出できるため測位精度が向上したと考えられる。

4.8.5 TDM による測位

実験 6 の結果より従来手法である TDM 方式では、従来手法である FDM 方式と比較して誤差が小さくことを確認した。これは、TDM の実験では信号の劣化が少ない低域チャープを利用したためと考えられる。一方、TDM は 3 つのタイムスロットを用いて 1 回の測位を行うため測位頻度は少なくなる。また提案手法の実験結果との比較により、コヒーレント積分による平均化処理の有効性が確認できる。

4.9 限界と展望

周波数分割数

本論文では多重化を行うため、周波数分割数をスピーカ数と同じ $N = 3$ に設定して実験を行った。周波数分割数を増加させ、追加スピーカに帯域を割り当てることで測位可能なカバレージは拡大する。一方、スピーカ毎に割り当てられる帯域幅が狭くなるため、時間分解能が低下し、測位精度が悪化することも予想される。最適な周波数分割数の設定は、測位を行う環境、スピーカの特性、設置位置などを考慮する必要がある、今後の課題としたい。

障害物

提案手法ではスピーカとマイクロフォンが見通し内であることを想定している。そのため、間に障害物が存在する環境では音の性質上、対応することは難しい。また、到来角が 90 度以上になる場合、スマートフォンの筐体自体が障害物となるため、測位精度が悪化することが予想される。しかし、一般的なスマートフォンには 2 個以上のマイクロフォンが異なる位置に配置されているため、少なくとも 1 つのマイクロフォンは見通し内となる状況が多く、この問題は解決できると考える。これらのマイクロフォンをアレイとして捉えることで、さらなる測位精度の向上も見込める。

移動体

測位対象が移動体である場合では、TDM による測位と同様に測位誤差が大きくなる問題がある。提案手法では測位対象が静止している場合を想定しているため移動体を対象とした測位は難しい。これは、3 つのタイムスロットを用いたコヒーレント積分による加算処理では、受信時刻が各タイムスロットで異なるためピーク付近が正常に加算されず、正確な受信時刻の検出が難しいためである。この課題の解決案として、粒子フィルタやカルマンフィルタによる移動体のトラッキングによる補正法などが考えられる。

4.10 まとめ

本章では、測位領域の拡張を目的とした新しい送信方式を提案した。この方式では、FDMとTDMを組み合わせることで、従来問題となっていた指向性による受信信号の劣化を抑制し、測位精度を向上させた。

社会に音響測位を展開する際には、常に3台以上のスピーカから十分な信号強度を確保することは困難であると予想される。提案手法は、この課題を解決し、低コストで高精度な測位システムの実現に寄与するものであり、実用化に向けた有力なアプローチと考えられる。

第 5 章

エコー埋め込みを用いた BGM による屋内測位

5.1 はじめに

本章では、人間に知覚されにくいエコーを測位用信号として BGM に埋め込み、これを利用する新たな測位方式を提案する。屋内環境では、音源からの直接波と床や天井、家具などからの反射波が合成され伝わる。やまびのように直接波と反射波の到来時刻に大きな差がある場合には知覚されるが、その差が小さい場合には人間には知覚されにくいことが知られている [118]。本研究では、この性質を利用し、音響情報ハイディング技術であるエコー拡散法 [111] を用いて、BGM に微弱な PN (Pseudo Noise) 系列を埋め込み、これをスピーカの ID として利用する (図 5.1)。さらに、ケプストラム解析 [132] により、BGM と埋め込まれた PN 信号を分離する。これにより、複数のスピーカ間の到来時刻差 (Time Difference of Arrival, TDoA) を基にモバイル端末の位置を推定する。

音響情報ハイディング技術を屋内測位に応用した研究は少ないが、主に映画館での違法録音を検出することを目的とした中島らの一連の手法が提案されている [133, 134, 135, 136]。これらの手法では放映される音源に PN 系列を直接スペクトル拡散によって埋め込み、各スピーカから録音位置に到来する遅延時間を利用して、最尤推定による測位を行っている。評価実験の結果、誤差が 0.44 m であることが報告されている。

本章の内容は、筆者の研究成果 [137] を基にまとめたものである。

5.2 提案手法

本研究では、 N 台のスピーカから送信される音源を、モバイル端末で録音することで位置を推定する。その際、スピーカから送信される信号は、ジャズやクラシックなどの BGM

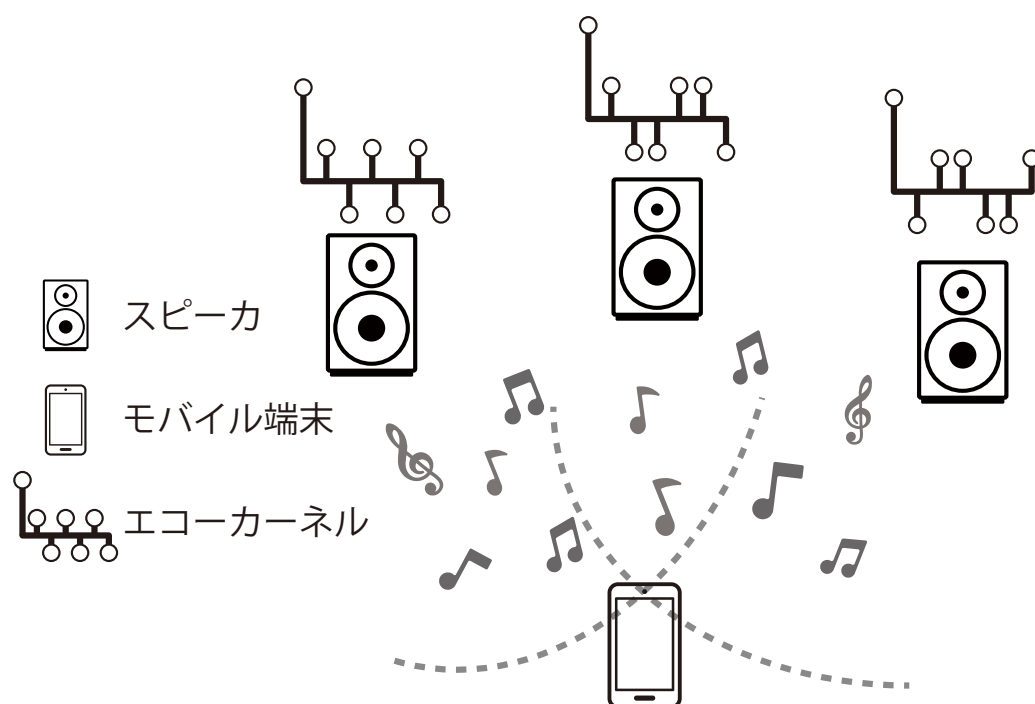


図 5.1: 提案システムの概要

であり，これにより人間に不快感を与えない測位を実現する． N 台のスピーカからは同じ BGM が送信されるが，遅延時刻を検出するため，エコー拡散法により互いに異なるカーネルを ID として埋め込む．本節では，はじめに各スピーカから送信する音源の加工方法について説明する．続いて，受信側で得られる信号のモデル化を行い，ケプストラム解析によって音源とカーネルを分離する方法を述べる．その後，分離されたカーネルから到来時刻差の推定方法，およびノイズ軽減手法を説明し，最後に TDoA アルゴリズムによる位置推定手法を述べる．

5.2.1 埋め込み方法

合計 N 台のスピーカがあり，スピーカ i に割り当てるカーネル $k_i(t)$ を以下のように表す． $\delta(t)$ はディラックのデルタ， $p_i(t)$ は系列長 L_{PN} の PN 系列で他と直交する．本研究では自己相関および相互相関性能が良好であり，スペクトル拡散技術でも広く利用されているゴールド系列を用いる．ゴールド系列の中でも，特に相関性能が良好になるプリファードペアとよばれるパラメータを使用する．ゴールド系列の利点は，同じペアで $2N + 1$ 個の互いに直交する系列を生成できるため，多くのスピーカに ID を割り当てられる点と，相関性能が他の PN 系列と比較しても良好である点である． α は PN 系列の振幅を表すパラメータで，大き

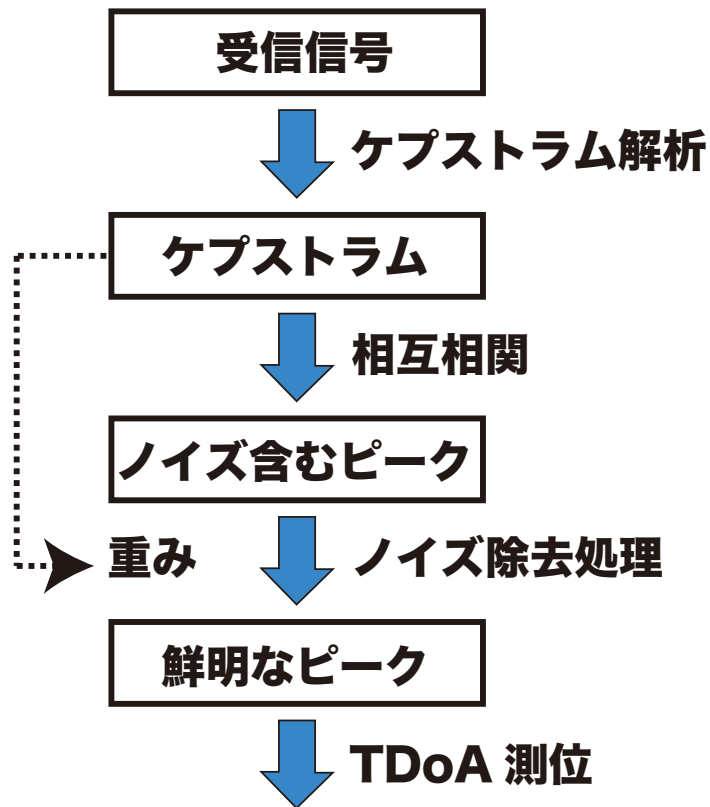


図 5.2: 検出手順の概要

くすると検出は容易になるが，音質が劣化することが予想される．

$$k_i(t) = \delta(t) + \alpha p_i(t) \quad (5.1)$$

カーネル $k_i(t)$ を音源 $s(t)$ との畳み込みを行い，それを送信信号 $w_i(t)$ とする．ここで $\otimes$ は畳み込みを表す記号である．

$$w_i(t) = s(t) \otimes k_i(t) \quad (5.2)$$

5.2.2 検出方法

検出の目的は，TDoA アルゴリズムによる測位を行うため，信号間のマイクロフォンに届くまでの到来時刻差の取得である．図 5.2 は，提案手法での検出までの手順を表した図である．受信するマイクロフォン側の視点に立つと，到来するカーネル $k_{rec}(t)$ は，図 5.3 のように各スピーカに割り当てたカーネルがそれぞれ空気中で遅延して合成されたものになる．この空気中で合成された合成カーネルは式 5.3 のように表せる．ここで τ_i はスピーカ i からの遅延時間である．また，最も早く到来した信号を基準とするため，その信号の遅延時間を 0 とする．

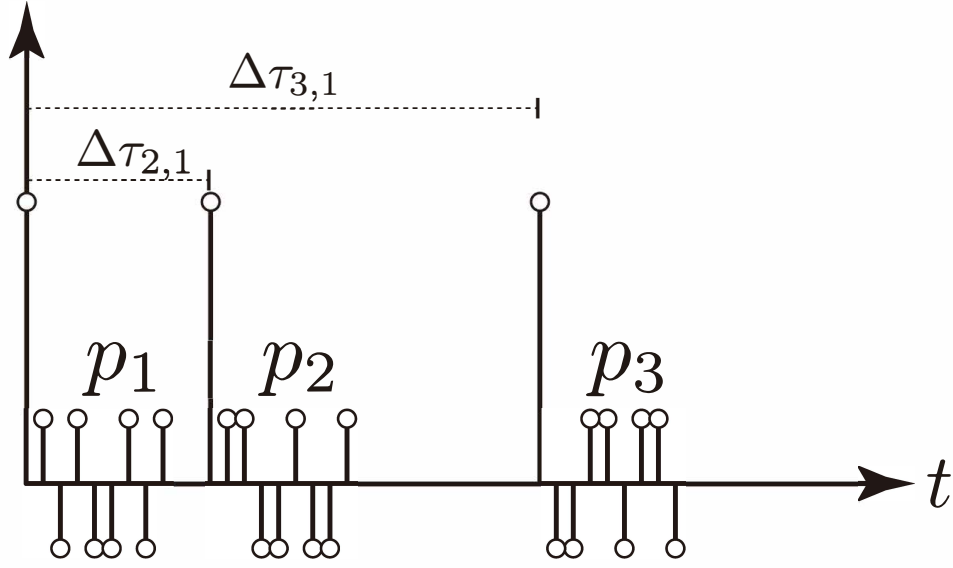


図 5.3: 到来する合成エコーカーネル $k_{rec}(t)$

$$k_{rec}(t) = \sum_{i=1}^N \left(\delta(t - \tau_i) + \alpha p_i(t - \tau_i) \right) \quad (5.3)$$

マイクrophonに到来する信号 $w_{rec}(t)$ は、伝送路のノイズを無視すると、音源 $s(t)$ と $k_{rec}(t)$ を畳み込むことで表せる。

$$\begin{aligned} w_{rec}(t) &= s(t) \otimes k_{rec}(t) \\ &= \sum_{i=1}^N s(t - \tau_i) + s(t) \otimes \left(\sum_{i=1}^N \alpha p_i(t - \tau_i) \right) \end{aligned} \quad (5.4)$$

$w_{rec}(t)$ の DFT(離散フーリエ変換) の結果は以下となる。ここで $\mathcal{F}[\cdot]$ は DFT を表し、 $j = \sqrt{-1}$, ω は角周波数である。式 (5.5) の第 2 項目は、畳み込みの演算が周波数領域では乗算になる性質を利用した。

$$\begin{aligned} \mathcal{F}[w_{rec}(t)] &= \sum_{i=1}^N S(\omega) e^{-j\omega\tau_i} + S(\omega) \sum_{i=1}^N \alpha P_i(\omega) e^{-j\omega\tau_i} \\ &= S(\omega) \sum_{i=1}^N \left(e^{-j\omega\tau_i} + \alpha P_i(\omega) e^{-j\omega\tau_i} \right) \end{aligned} \quad (5.5)$$

$\mathcal{F}[w_{rec}(t)]$ の対数を取ることによって、音源とカーネルを分離することができる。以降 $K_{rec}(\omega) = \mathcal{F}[k_{rec}(t)] = \sum_{i=1}^N \left(e^{-j\omega\tau_i} + \alpha P_i(\omega) e^{-j\omega\tau_i} \right)$ とする。

$$\begin{aligned} \log \mathcal{F}[w_{rec}(t)] &= \log (S(\omega) \cdot K_{rec}(\omega)) \\ &= \log S(\omega) + \log K_{rec}(\omega) \end{aligned} \quad (5.6)$$

$\log \mathcal{F}[w_{rec}(t)]$ の IDFT(逆離散フーリエ変換) を行った結果を $\tilde{w}(t)$ とする．この一連の処理はケプストラム変換と呼ばれ、 $\tilde{w}(t)$ の次元は周波数の逆数となるので時間に戻るが、「Frequency」のアナグラムとしてケフレンシー (Quefrency) とも呼ばれる．また、対数による非線形変換の影響を受けたものをそれぞれ $\tilde{\delta}(t)$, $\tilde{p}_i(t)$ とする． $\tilde{p}_i(t)$ は $\delta(t) \gg \alpha p_i(t)$ であることから、テイラー展開を行うことで $\tilde{p}_i(t) \approx p_i(t)$ と近似することができる [111]. $\mathcal{F}^{-1}[\cdot]$ は IDFT を表す．

$$\begin{aligned}
\tilde{w}(t) &= \mathcal{F}^{-1}[\log \mathcal{F}[w_{rec}(t)]] \\
&= \mathcal{F}^{-1}[\log S(\omega)] + \mathcal{F}^{-1}[\log K_{rec}(\omega)] \\
&= \tilde{s}(t) + \tilde{k}_{rec}(t) \\
&= \tilde{s}(t) + \sum_{i=1}^N \left(\tilde{\delta}(t - \tau_i) + \alpha \tilde{p}_i(t - \tau_i) \right)
\end{aligned} \tag{5.7}$$

続いて、遅延時間 τ_i を検出するために相互相関 $r_i(t)$ を算出する．相互相関はリファレンスを反転させて畳み込みを行う処理となる．ここで、式 (5.8) のように周波数領域での畳み込みにより、時間領域での計算と比較して計算量を削減することができる． $r_i(t)$ の中の適切なピークを検出することで、遅延時間 τ_i を得ることができる．* は共役を表す．

$$\begin{aligned}
r_i(t) &= \tilde{w}(t) \otimes p_i(-t) \\
&= \mathcal{F}^{-1}[\tilde{W}(\omega) \cdot P_i^*(\omega)]
\end{aligned} \tag{5.8}$$

ここで、遅延時間 τ_i を得るのに 2 つの情報がある．1 つ目の情報として、スピーカ i に割り当てた PN 系列との相互相関 $r_i(t)$ には、 τ_i の位置にピークが出現することを利用する．このピークは他の PN 系列とは直交することから、相互相関には現れないので、そのピークに対応するスピーカ ID を特定することができる．2 つ目は $\sum_{i=1}^N \tilde{\delta}(t - \tau_i)$ に対応する項である．ケプストラム $\tilde{w}(t)$ にも、 $\tilde{\delta}(t)$ の遅延時間に対応する位置にピークが出現する．そのため、この 2 つの情報を有効的に使い、より頑健なピーク検出を試みる．具体的には以下のように $\tilde{w}(t)$ を重みとして捉え、重み付けを行った $r_i^w(t)$ を算出する． $\|\cdot\|$ はノルムを表す．

$$r_i^w(t) = r_i(t) \cdot \|\tilde{w}(t)\| \tag{5.9}$$

$r_i^w(t)$ のノイズをさらに軽減するため、フレーム毎の平均化を行う．この手法はインコヒーレント積分とも呼ばれ、ノイズの平均値からのゆらぎを抑圧し、ピーク検出がしやすくなる効果がある．平均する回数 M を増加させることで、その効果は大きくなる．これにより、音源 $s(t)$ のケプストラム $\tilde{s}(t)$ は時間によって変化するので、これをノイズとして捉え、平均化によって音源そのものの影響を抑えることができる．

$$\overline{r_i^w}(t) = \frac{1}{M} \sum_{l=1}^M r_{i,l}^w(t) \tag{5.10}$$

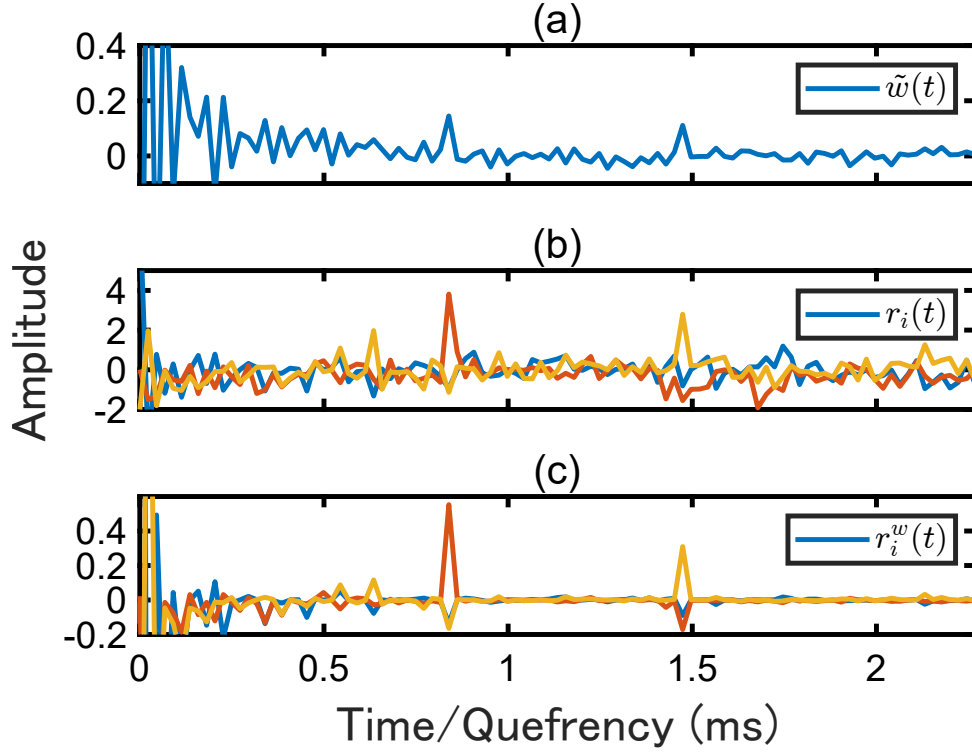


図 5.4: (a) 受信信号より得られるケプストラム, (b) 各スピーカに割り当てた PN との相関, (c)(b) の相関を (a) のケプストラムを用いてフィルタリングした結果

$\overline{r_i^w(t)}$ に存在するピークを検出することで到来時刻差 τ_i を推定する．本研究では，範囲内の最大値を取る位置を選ぶ．

5.2.3 測位アルゴリズム

測位には TDoA アルゴリズムを利用する． $\mathbf{p} = (x, y)^\top$, $\mathbf{p}_i^{sp} = (x_i^{sp}, y_i^{sp})^\top$ はそれぞれモバイル端末の座標，スピーカ i の座標を表す． c は音速である．モバイル端末の推定位置 $\hat{\mathbf{p}}$ は非線形最小二乗法によって求められる．基準となる $\mathbf{p}_1$ は最も早く到来したスピーカからの信号であり，そのときの $\tau_1 = 0$ を基準とする．この判定には直流成分，つまりインデックス 0 において最大値を取るものを選ぶ．

$$R_i = \|\mathbf{p} - \mathbf{p}_i^{sp}\| - \|\mathbf{p} - \mathbf{p}_1^{sp}\| - c\Delta\tau_{i,1}$$

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \sum_{i=2}^N R_i^2 \quad (5.11)$$

表 5.1: 音響処理に関するパラメータ

パラメータ	値
サンプリング周波数 (Hz)	44100
DFT サイズ (Sample)	32768(= 2^{15})
音源	ジャズ, クラシック, ポップ, アナウンス
α の大きさ	0.006, 0.01, 0.02
PN の長さ	1023, 2047
フレーム数	8
オーバーラップサイズ	DFT サイズの 1/8 (7/8 を重ねる)

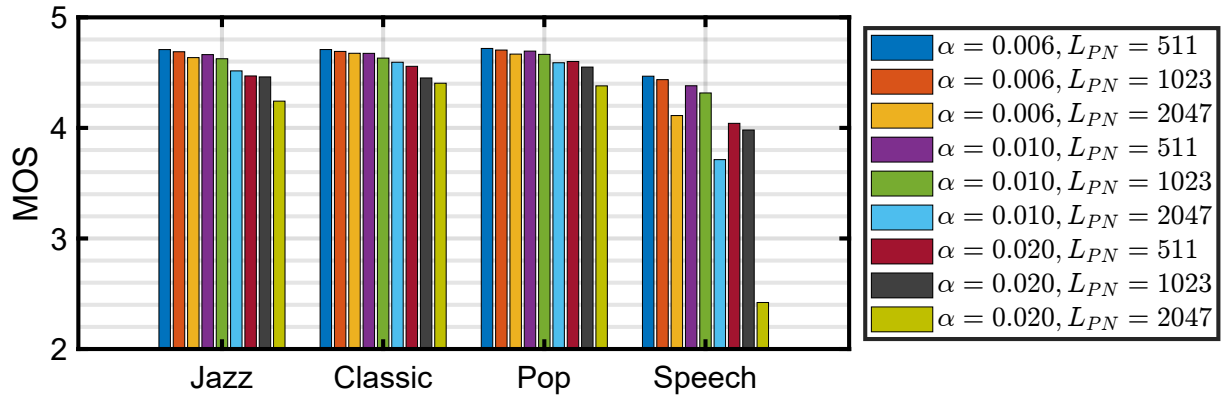


図 5.5: 各種パラメータによる音質

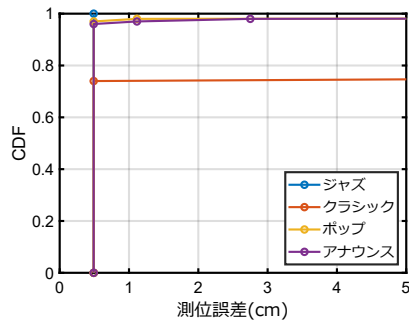


図 5.6: 音源のジャンル

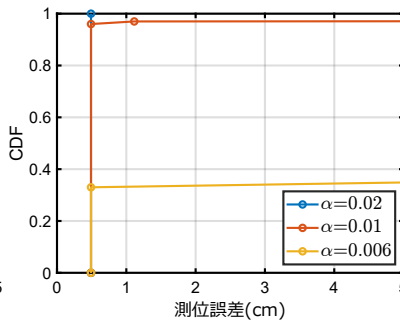


図 5.7: 埋め込み強度 α

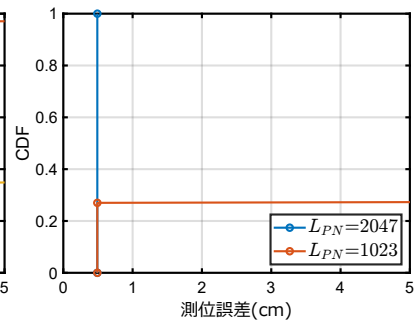


図 5.8: PN の長さ L_{PN}

5.3 評価実験

提案手法の音質に関するパラメータである埋め込み強度 α , および PN の信号長 L_{PN} と測位精度はトレードオフの関係にあることが予想できる. 埋め込み強度 α が大きいほどエ

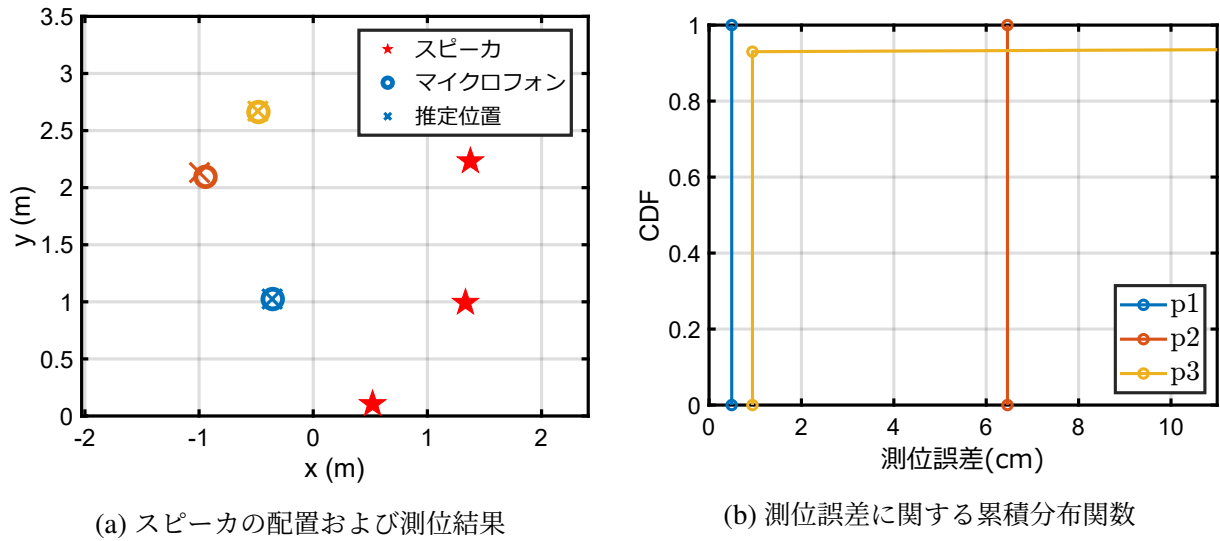


図 5.9: 測位実験の結果

表 5.2: MOS: Mean Opinion Scores

MOS	対応判断基準
5	劣化知覚不能
4	違いが分かるが気にならない
3	やや気になる
2	気になる
1	非常に気になる

コー成分の増加によって音質が劣化する一方で、PN 系列が強調されるため相関性能が向上し、結果的に測位精度も改善するはずである。同様に、PN 系列の長さ L_{PN} が長いほど、エコーの持続時間が長くなるため音質が劣化する一方、相関性能が向上するはずである。さらに、音源のジャンルでもその特徴によって音質と測位精度に影響を与えることも予想できる。本節では、各種パラメータを変化させ、測位精度および音質の調査を行う。本論文で設定するパラメータは表 5.1 の通りである。

5.3.1 設定と実験環境

送信側のハードウェア構成として、スピーカは Fostex FF105WK、スピーカボックスはバスレフ型の Fostex BK125WB2 を用いた。DAC は MOTU ultralite-mk5、アンプは Fostex AP20d を用いた。本研究では 3 つのスピーカから同時に音を送信することを想定しているが、Python のライブラリ sounddevice[138] を用いて時刻同期およびチャンネルの制御を行っ

た．受信側のマイクロフォンは Earthworks M30 を用いた．座標の真値は，誤差 1.0 mm 以下となるナックイメージテクノロジー製のモーションキャプチャシステムを用いて計測した．

測位実験では，スピーカ 3 台を用いて TDoA アルゴリズムによる 2 次元測位を行った．各スピーカに互いに直交する PN 系列を割り当て，カーネル埋め込み済みの音源を 3 台のスピーカから同時に送信するように設定した．送信および録音するときのサンプリング周波数は，多くの端末でサポートされている 44100 Hz，DFT サイズは高速計算の条件である 2 のべき乗の $32768=2^{15}$ に設定した．各計測は 100 回行い，16 bit PCM で録音したデータをオフラインにて解析を行った．計測時の気温は 28.0 度，騒音はレベル 44 dB であった．

5.3.2 音質

はじめに，埋め込み強度 α および PN の信号長 L_{PN} を変化させたときの音質の比較を行った．このとき α は埋め込み強度最大となる $1/\sqrt{L_{PN}}$ を超えないように設定した．音質は音質評価ソフト ViSQOL-v3 [125] を用いた．ViSQOL は主観と客観音響品質の相関が高いことから，音響情報ハイディングの分野で頻繁に利用されている．オリジナルの音源と埋め込み後の音源を入力すると，主観音質劣化量 MOS (Mean Opinion Scores) を数値化して出力する．劣化量の目安を 5.2 に示す．音源は [139] および [140] から合計 4 ジャンルを選んだ．

実験結果

ジャンル毎にパラメータを変えて音質評価を行った結果を図 5.5 に示す．この結果から，埋め込み強度 α が大きいほど，また PN 系列長 L_{PN} が長いほど MOS 値が劣化する傾向にあることがわかる．また，どのジャンルもパラメータによって同様の劣化具合になった．

5.3.3 音源のジャンルと測位精度

ここでは，音源のジャンルによる測位精度への影響について調査を行った．実験では $\alpha = 0.02$ ， $L_{PN} = 2047$ に固定し，ジャンルを変えて計測を行った．

実験結果

計測結果の CDF(累積分布関数)を図 5.6 に示す．この結果から，音源がジャズの場合で測位誤差が 0.49 cm，分散が 0 となり，高精度な測位が実現できていることが確認できる．ポップとアナウンスは同じような精度となり，クラシックの精度は若干低くなった．

5.3.4 埋め込み強度と測位精度

埋め込み強度による測位精度への影響について調査を行った。実験では $L_{PN} = 2047$ ，音源をジャズに固定し，埋め込み強度 α を変えて計測を行った。

実験結果

計測結果の CDF を図 5.7 に示す。この結果から $\alpha = 0.02$ と $\alpha = 0.01$ では同様の結果となり，測位精度に大きな差異は確認できなかった。一方で $\alpha = 0.006$ のとき，測位精度は大きく低下した。

5.3.5 PN 系列の長さで測位精度

PN 系列の長さによる測位精度への影響について調査を行った。実験では $\alpha = 0.02$ ，音源をジャズに固定し，PN 系列の長さ L_{PN} を変えて計測を行った。

実験結果

計測結果の CDF を図 5.8 に示す。この結果から PN 系列の長さによって測位精度が大きく変わることが確認できる。

5.3.6 計測位置と測位精度

ここでは計測する位置による測位精度への影響について調査を行った。実験では $\alpha = 0.02$ ， $L_{PN} = 2047$ ，音源をジャズに固定し，録音する位置を 3 箇所変えて計測を行った。

実験結果

スピーカとマイクロフォンの真値，および推定位置との測位誤差を図 5.9 に示す。この結果から，どの計測位置においても高精度な測位が実現できていることが確認できる。

5.4 考察

最適なパラメータ

実験結果 5.3.4 および 5.3.5 から，埋め込み強度 α が大きいほど，また PN 系列の長さ L_{PN} が長いほど高い測位精度に改善することが確認できる。一方で測位精度を重視したパラメータを用いると，結果 5.3.2 のように MOS 値が劣化するため，測位精度と音質はトレードオフの関係にあることが示される。

また，結果 5.3.5 から， $L_{PN} = 1023$ とすると測位精度が大きく低下することから， $L_{PN} = 2047$ は必須である．一方で結果 5.3.4 から，埋め込み強度 $\alpha = 0.01$ ， $\alpha = 0.02$ どちらでも測位精度は大きく変わらない．このときの MOS 値は 4.2 程度であり，これは埋め込み前の音源と比較して十分許容できる音質である．まとめると $\alpha = 0.01$ ， $L_{PN} = 2047$ が，測位精度と音質のバランスをとれた良好なパラメータであると結論付けられる．

音楽ジャンル

エコー拡散法は音源にカーネルを畳み込む処理を行うため，元の音源の振幅が小さいと，その分だけ埋め込まれるエネルギーも小さくなる．極端な例として，振幅が 0 の無音が続くような音源の場合，埋め込み後の振幅も 0 になる．このような特徴があることから，無音もしくは途切れる時間帯があるアナウンスを用いたときの測位精度は低下すると予想していたが，ポップと同等の精度になった．これは式 (5.10) により，移動しながら平均化処理を行っているため，無音になる時間帯の影響が緩和されていることが要因であると考えられる．また，音源としては周波数帯域の広いものが望ましいが，ドラム音を含むジャズやポップなどのジャンルは，瞬間的にインパルスのような広い帯域をとる時間帯があるので提案手法との相性が良いと考える．

測位の分散と誤差

実験 5.3.6 の CDF の結果から，どの計測位置でも外れ値を除いて分散が 0 になるような高精度測位になっていることが確認できる．音響を利用して測距を行う場合，1 サンプル点だけ受信時刻が移動すると推定距離も 0.78 cm 変化するが，この結果から距離分解能の高い推定ができていることがわかる．一般に，周波数帯域と距離分解能にはトレードオフの関係があるが，提案手法では広帯域かつ相関性能が良好な PN 系列を有効活用できていると言える．

また，計測箇所 p2 における系統誤差は p1, p3 と比較して大きくなった．この測位誤差の要因として，測位精度の劣化の程度を表す数値である DoP (Dilution of Precision) による影響が考えられる．DoP 値はスピーカの配置を工夫することによって改善できる．今回の計測環境では，スピーカ間の距離を比較的短くなるように設置したが，計測する環境を覆うように配置することで改善すると考える．

5.5 限界と展望

遠近問題

エコー拡散法による測位は，一種のスペクトル拡散による通信と捉えられるため，CDM 特有の遠近問題 (near-far problem) を考慮する必要がある．本研究の計測環境では発生しな

かったが、マイクロフォンの位置が1つのスピーカと近く、他と遠い場合には受信するエネルギーの差で遠いスピーカの信号を受け取ることが困難となる。この問題の解決には、近いスピーカの音量を小さく、遠いスピーカは大きくするなど、マイクロフォンの座標によって適応的な送信電力制御を行うことで解決できるが、スピーカ配置の最適化や単純にスピーカ数を増加させることでもある程度は回避できると考える。

ピーク検出

提案手法の主な誤差の要因は、SN 比の低下による誤ったピークの検出である。正しいピークが検出できる場合、数センチメートル以下の測位誤差になることが確認できたが、誤ったピークの検出時には的はずれな位置を推定してしまう。特に、0.2 ms 以下の短いケプレンシー領域には、音源によるピークが集中する傾向があるため、この範囲での正確なピーク検出は困難である。このような問題には、主にレーダの分野で利用される CFAR[141] などの適応的な閾値アルゴリズムを採用することで解決できると考える。さらに、得られたケプストラムの波形に対して、深層学習を用いたパターン認識やノイズリダクションを行うことで、さらなる測位精度の向上も見込める。

また本研究では、PN 系列として自己相関および相互相関性能が良好であるゴールド系列を採用したが、相関性能がさらに良好である Zadoff-Chu 系列 [142] や ZCZ 系列 [143] を採用することで、ピーク検出性能が改善するため、結果的に測位精度が向上する可能性がある。

見通し外と反射波

本研究で扱う音響には、電波のような透過力がないことから、障害物によって直接波が遮断されるような見通し外環境では測位ができない。一方で、反射波に対しては、PN 系列との相関値が直接波のものと比較して小さくノイズに埋没するため、他の音響測位手法と比較しても優位性があると考えられる。

他パラメータの調査

本研究の実験では調査しなかったが、DFT サイズ、平均化する回数などの最適なパラメータを探索する必要がある。DFT サイズを長くほど頑健な推定が可能になるが、測位更新頻度が低下してしまう。同様に平均化する回数も増加させるほどノイズを軽減できるが、測位更新頻度が低下するため、応用するアプリケーションを考慮してこれらのパラメータを決定する必要がある。

さらに本研究では3箇所のみでの計測に留まったが、計測箇所を増加させ、さらなる誤差の要因と傾向を究明する必要がある。

5.6 まとめ

本章では、エコー拡散法を応用した高精度測位手法を提案し、その有効性を示した。評価実験により、音質を損なうことなく数 cm での高精度測位が可能であることを示し、音質と測位精度のトレードオフ関係を明確化した上で、適切なパラメータの検討を行った。また、従来は課題であった汎用モバイルデバイスを対象に、可聴域内かつ人間に違和感を与えない測位信号を用いて目標である誤差 10 cm を十分に達成可能であることを示した。

第 6 章

結論と今後の展望

6.1 結論

本研究では、音響を用いた屋内測位技術において、3つの主要課題に取り組み、新しい手法を提案して有効性を検証した。

送受信機間の時刻同期手法では、床からの反射波を利用した擬似的な時刻同期手法を提案した。この手法により、専用のハードウェアなしでの時刻同期が可能になり、ToA方式での測位が可能になることを示した。この手法は、設備スピーカが少ない環境で TDoA方式が利用できない場合や、ネットワークを介さずにサーバとマイクロ秒単位の同期が必要な場合に有効である。特に2次元測位では2台のスピーカのみで測位が実現できるため、応用範囲は広い。

測位領域の拡張手法では、FDMとTDMを組み合わせた新しい測位方式を提案し、受信信号の周波数特性による測位精度への影響を抑制した。これにより、広範囲での高精度測位が可能であることを確認した。この手法も設備スピーカが少ない場合に有効であり、屋内測位を普及させる上で有用である。

エコー埋め込みによる測位手法では、可聴域を活用したエコー埋め込み技術を提案し、BGMに測位信号を埋め込むことで人に違和感を与えずに高精度測位が可能であることを示した。評価実験により、埋め込み強度や信号長さの調整により、音質と測位精度のバランスをとることが可能であることを示した。この手法は、数 cmレベルの高精度測位が求められる環境で有効であり、音響を活用した測位技術の新たな可能性を示すものである。

6.2 今後の展望

現実的な屋内測位技術の確立には電波、音波、可視光、PDRなど様々な技術を組み合わせる必要があると考えられている [1]。本論文では音響を用いた測位技術を中心に扱ったが、

実際に他の技術と組み合わせ、各技術の短所を補うような統合システムの開発が求められる。また、本論文での実験環境は理想的な環境での評価に留まっているため、実際の利用シーンに近い条件での評価と実環境特有の課題に対応する必要がある。

謝辞

本論文は、筆者が北海道大学 大学院情報科学院 情報理工学コース 知能情報学研究室に在籍中の研究成果をまとめたものになります。本論文の執筆にあたり、多くの方からご指導・ご支援を頂きました。

指導教官である、北海道大学 大学院情報科学院 情報理工学コース 数理科学分野 知能情報学研究室 杉本雅則教授には、不自由のない十二分な研究環境を提供していただくとともに、多大なるご指導・ご支援を賜りました。ここに深く感謝の意を表します。

同研究室 中村将成准教授には、研究内容だけでなく、日頃の学生生活や就職活動についても多くのご助言をいただきました。ここに感謝の意を表します。

同数理科学分野 田中章教授、今井英幸教授、松原崇教授には、フェローシップのメンターや学位論文審査をはじめとした様々な機会においてご指導をいただきました。ここに感謝の意を表します。

独立行政法人 大学改革支援・学位授与機構 橋爪宏達教授には、日頃の打ち合わせにて貴重なご指導・ご助言を賜りました。ここに感謝の意を表します。

はこだて未来大学 渡邊拓貴准教授には同研究室在籍中、研究室ゼミなどを通じて多くのご助言をいただきました。ここに感謝の意を表します。

東京大学 村上弘晃助教には、筆者が学部3年生のときから現在に至るまで、研究活動を進める上で多くのことを学ばせていただきました。ここに感謝の意を表します。

研究室秘書 村田恵美さん、新浜円さんには、快適な学生生活を送るにあたり、事務手続きなどさまざまな面で多大なるご支援をいただきました。ここに感謝の意を表します。

同研究室 博士後期課程 大西悠生君には、研究室の運営や後輩への指導にあたり、多くのご協力をいただきました。ここに感謝の意を表します。

本研究は、北海道大学博士人材 DX フェローシップ、並びに公益財団法人 立石科学技術振興財団 研究助成 (C) による支援を受けたものです。ここに感謝の意を表します。

最後になりますが、経済的・精神的に支援してくださった両親に感謝いたします。

参考文献

- [1] 国土交通省. 屋内測位環境構築の手引き. 国土交通省 不動産・建設経済局 情報活用推進課, 3 2022. 令和 4 年 3 月版.
- [2] Mai A Al-Ammar, Suheer Alhadhrami, Abdulmalik Al-Salman, Abdulrahman Alarifi, Hend S Al-Khalifa, Ahmad Alnafessah, and Mansour Alsaleh. Comparative survey of indoor positioning technologies, techniques, and algorithms. In *Proc. of IEEE International Conference on Cyberworlds*, 2014.
- [3] IndustryARC. Global indoor positioning and navigation market. <https://www.izindustryarc.com/Report/43/global-indoor-positioning-navigation-market.html>, 2024. Accessed: 2024-12-03.
- [4] 栗田雄一, 小林純也, 末永剛, 小笠原司. 動線情報に基づく個人識別と人間関係の可視化. 計測自動制御学会論文集, Vol. 46, No. 5, pp. 257–265, 2010.
- [5] Edward T Hall. The hidden dimension. *Garden City*, 1966.
- [6] Nobuo Kawaguchi, Kei Hiroi, Atsushi Shionozaki, Masamichi Asukai, Toshimune Nasu, Yu Hashimoto, Takeharu Nakamura, Tetsuya Gotou, and Shinsuke Ando. Wi-fi human behavior analysis and ble tag localization: A case study at an underground shopping mall. In *Proc. of EAI MobiQuitous*, 2016.
- [7] Weichao Liang, Zhiang Wu, Jie Cao, and Jun Gu. Understanding customer behavior in shopping mall from indoor tracking data. In *Proc. of IEEE CSCWD*, 2018.
- [8] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. The active badge location system. *ACM Trans. on Information Systems*, Vol. 10, No. 1, pp. 91–102, 1992.
- [9] Paramvir Bahl and Venkata N Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *Proc. of IEEE INFOCOM*, 2000.
- [10] Chao Cai, Rong Zheng, Jun Li, Linwei Zhu, Henglin Pu, and Menglan Hu. Asynchronous acoustic localization and tracking for mobile targets. *IEEE Internet of Things Journal*, Vol. 7, No. 2, pp. 830–845, 2019.
- [11] Tao Liu, Lei Wan, Zhenquan Qin, Chen Qian, and Xin Zhou. Passive acoustic localiza-

- tion based on cots mobile devices. In *Proc. of IEEE ICPADS*, 2018.
- [12] Yu-Ting Wang, Jun Li, Rong Zheng, and Dongmei Zhao. Arabis: An asynchronous acoustic indoor positioning system for mobile devices. In *Proc. of IEEE IPIN*, 2017.
 - [13] Jont B Allen and David A Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, Vol. 65, No. 4, pp. 943–950, 1979.
 - [14] AngelaD’ amico, Richard Pittenger. A brief history of active sonar. *Aquatic Mammals*, Vol. 35, No. 4, pp. 426–434, 2009.
 - [15] Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and James Collins. *Global positioning system: theory and practice*. Springer Science & Business Media, 2012.
 - [16] Junsung Cho, Jaegwan Yu, Sanghak Oh, Jungwoo Ryoo, JaeSeung Song, and Hyoungshick Kim. Wrong siren! a location spoofing attack on indoor positioning systems: The starbucks case study. *IEEE Communications Magazine*, Vol. 55, No. 3, pp. 132–137, 2017.
 - [17] Zsófia Papp, Garry Irvine, Roland Smith, Ferenc Mogyorósi, Péter Revisnyei, Istvan Törős, and Alija Pašić. Tdoa based indoor positioning over small cell 5g networks. In *Proc. of IEEE/IFIP NOMS*, 2022.
 - [18] Takayuki Akiyama, Masanori Sugimoto, and Hiromichi Hashizume. Time-of-arrival-based smartphone localization using visible light communication. In *Proc. of IEEE IPIN*, 2017.
 - [19] Chunyi Peng, Guobin Shen, Yongguang Zhang, Yanlin Li, and Kun Tan. Beepbeep: a high accuracy acoustic ranging system using cots mobile devices. In *Proc. of ACM SenSys*, 2007.
 - [20] Omar Hashem, Khaled A Harras, and Moustafa Youssef. Accurate indoor positioning using ieee 802.11 mc round trip time. *Pervasive and Mobile Computing*, Vol. 75, p. 101416, 2021.
 - [21] Guido De Angelis, Antonio Moschitta, and Paolo Carbone. Positioning techniques in indoor environments based on stochastic modeling of uwb round-trip-time measurements. *IEEE Trans. on Intelligent Transportation Systems*, Vol. 17, No. 8, pp. 2272–2281, 2016.
 - [22] Liyuan Song, Tingting Zhang, Xiaofei Yu, Chuan Qin, and Qinyu Zhang. Scheduling in cooperative uwb localization networks using round trip measurements. *IEEE Communications Letters*, Vol. 20, No. 7, pp. 1409–1412, 2016.
 - [23] 藤本弘. 信号処理の基礎と応用. コロナ社, 2020.
 - [24] Arthur Venon, Yohan Dupuis, Pascal Vasseur, and Pierre Merriaux. Millimeter wave fmcw radars for perception, recognition and localization in automotive applications: A

- survey. *IEEE Trans. on Intelligent Vehicles*, Vol. 7, No. 3, pp. 533–555, 2022.
- [25] Shuhui Yang, Zimu Yuan, and Wei Li. Error data analytics on rss range-based localization. *Big Data Mining and Analytics*, Vol. 3, No. 3, pp. 155–170, 2020.
 - [26] Sebastian Sadowski and Petros Spachos. Rssi-based indoor localization with the internet of things. *IEEE Access*, Vol. 6, pp. 30149–30161, 2018.
 - [27] Masanari Nakamura, Kento Fujimoto, Hiroaki Murakami, Hiromichi Hashizume, and Masanori Sugimoto. Indoor localization method for a microphone using a single speaker. In *Proc. of IEEE IPIN*, 2021.
 - [28] Yu Cao, Weiguang Shi, Lijun Sun, and Xiaomei Fu. Channel state information-based ranging for underwater acoustic sensor networks. *IEEE Trans. on Wireless Communications*, Vol. 20, No. 2, pp. 1293–1307, 2020.
 - [29] Suining He and S-H Gary Chan. Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons. *IEEE Communications Surveys & Tutorials*, Vol. 18, No. 1, pp. 466–490, 2015.
 - [30] Ramsey Faragher and Robert Harle. Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, Vol. 33, No. 11, pp. 2418–2428, 2015.
 - [31] Xinping Rao, Le Qin, Yugen Yi, Jin Liu, Gang Lei, and Yuanlong Cao. A novel adaptive device-free passive indoor fingerprinting localization under dynamic environment. *IEEE Trans. on Network and Service Management*, 2024.
 - [32] Zuoya Liu, Ruizhi Chen, Feng Ye, Guangyi Guo, Zheng Li, and Long Qian. Improved toa estimation method for acoustic ranging in a reverberant environment. *IEEE Sensors Journal*, Vol. 22, No. 6, pp. 4844–4852, 2020.
 - [33] Lei Zhang, Minlin Chen, Xinheng Wang, and Zhi Wang. Toa estimation of chirp signal in dense multipath environment for low-cost acoustic ranging. *IEEE Trans. on Instrumentation and Measurement*, Vol. 68, No. 2, pp. 355–367, 2018.
 - [34] Takayuki Akiyama, Masanori Sugimoto, and Hiromichi Hashizume. Syncsync: Time-of-arrival based localization method using light-synchronized acoustic waves for smart-phones. In *Proc. of IEEE IPIN*, 2015.
 - [35] Georg Oberholzer, Philipp Sommer, and Roger Wattenhofer. Spiderbat: Augmenting wireless sensor networks with distance and angle information. In *Proc. of ACM/IEEE IPSN*, 2011.
 - [36] 手嶋優風, 土屋隆生, 飛龍志津子. コウモリのエコーロケーション——音で“見る”術に学ぶ——. 電子情報通信学会 通信ソサイエティマガジン, Vol. 16, No. 1, pp. 6–12, 2022.

- [37] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, and John J Leonard. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans. on Robotics*, Vol. 32, No. 6, pp. 1309–1332, 2016.
- [38] Shoudong Huang and Gamini Dissanayake. Convergence and consistency analysis for extended kalman filter based slam. *IEEE Trans. on Robotics*, Vol. 23, No. 5, pp. 1036–1049, 2007.
- [39] Takafumi Taketomi, Hideaki Uchiyama, and Sei Ikeda. Visual slam algorithms: A survey from 2010 to 2016. *IPSJ Trans. on Computer Vision and Applications*, Vol. 9, pp. 1–11, 2017.
- [40] 丹羽健友, 岩田直泰, 山本俊六. 早期地震警報に向けた地震観測点サイト特性と震央位置推定誤差の関係. 日本地震工学会論文集, Vol. 21, No. 2, pp. 57–69, 2021.
- [41] 大賀寿郎, 山崎芳男, 金田豊. 音響システムとディジタル処理. コロナ社, 1995.
- [42] 浅野太. 音のアレイ信号処理-音源の定位・追跡と分離-. コロナ社, 2011.
- [43] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. on Antennas and Propagation*, Vol. 34, No. 3, pp. 276–280, 1986.
- [44] Bjorn Ottersten, Mats Viberg, and Thomas Kailath. Performance analysis of the total least squares esprit algorithm. *IEEE Trans. on Signal Processing*, Vol. 39, No. 5, pp. 1122–1135, 1991.
- [45] Tatsuya Iizuka, Takuya Sasatani, Toru Nakamura, Naoko Kosaka, Masaki Hisada, and Yoshihiro Kawahara. Millisign: mmwave-based passive signs for guiding uavs in poor visibility conditions. In *Proc. of ACM MobiCom*, 2023.
- [46] Charles Knapp and Glifford Carter. The generalized correlation method for estimation of time delay. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 24, No. 4, pp. 320–327, 1976.
- [47] Kouei Yamaoka, Robin Scheibler, Nobutaka Ono, and Yukoh Wakabayashi. Sub-sample time delay estimation via auxiliary-function-based iterative updates. In *Proc. of IEEE WASPAA*, 2019.
- [48] Richard O Duda. Modeling head related transfer functions. In *Proc. of IEEE ACSSC*, 1993.
- [49] Zhijian Yang, Yu-Lin Wei, Sheng Shen, and Romit Roy Choudhury. Ear-ar: indoor acoustic augmented reality on earphones. In *Proc. of ACM MobiCom*, 2020.
- [50] Joe C Chen, Kung Yao, and Ralph E Hudson. Source localization and beamforming. *IEEE Signal Processing Magazine*, Vol. 19, No. 2, pp. 30–39, 2002.
- [51] Seong-Hwang Kim, Seul-Yi Lee, Yali Zhang, Soo-Jin Park, and Junwei Gu. Carbon-

- based radar absorbing materials toward stealth technologies. *Advanced Science*, Vol. 10, No. 32, p. 2303104, 2023.
- [52] Taweesak Sanpechuda and La-or Kovavisaruch. A review of rfid localization: Applications and techniques. In *Proc. of IEEE ECTI-CON*, 2008.
 - [53] SL Ting, SK Kwok, Albert HC Tsang, and George TS Ho. The study on using passive rfid tags for indoor positioning. *Journal of Engineering Business Management*, Vol. 3, p. 8, 2011.
 - [54] Huixin Dong, Yirong Xie, Xianan Zhang, Wei Wang, Xinyu Zhang, and Jianhua He. Gpsmirror: Expanding accurate gps positioning to shadowed and indoor regions with backscatter. In *Proc. of ACM MobiCom*, 2023.
 - [55] Shahriar Nirjon, Jie Liu, Gerald DeJean, Bodhi Priyantha, Yuzhe Jin, and Ted Hart. Coin-gps: Indoor localization from direct gps receiving. In *Proc. of ACM MobiSys*, 2014.
 - [56] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 37, No. 6, pp. 1067–1080, 2007.
 - [57] Jie Xiong and Kyle Jamieson. Arraytrack: A fine-grained indoor location system. In *Proc. of USENIX NSDI*, 2013.
 - [58] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *Proc. of ACM SIGCOMM*, 2015.
 - [59] Kang Eun Jeon, James She, Perm Soonsawad, and Pai Chet Ng. Ble beacons for internet of things applications: Survey, challenges, and opportunities. *IEEE Internet of Things Journal*, Vol. 5, No. 2, pp. 811–828, 2018.
 - [60] Chao Xiao, Daiqin Yang, Zhenzhong Chen, and Guang Tan. 3-d ble indoor localization based on denoising autoencoder. *IEEE Access*, Vol. 5, pp. 12751–12760, 2017.
 - [61] Bardia Alavi and Kaveh Pahlavan. Studying the effect of bandwidth on performance of uwb positioning systems. In *Proc. of IEEE WCNC*, 2006.
 - [62] Rudolf Zetik, Jürgen Sachs, and R Thoma. Uwb localization-active and passive approach [ultra wideband radar]. In *Proc. of IEEE I2MTC*, Vol. 2, pp. 1005–1009, 2004.
 - [63] Davide Dardari, Raffaele d’Errico, Christophe Roblin, Alain Sibille, and Moe Z Win. Ultrawide bandwidth rfid: The next generation? *Proc. of the IEEE*, Vol. 98, No. 9, pp. 1570–1582, 2010.
 - [64] Minghui Zhao, Tyler Chang, Aditya Arun, Roshan Ayyalasomayajula, Chi Zhang, and Dinesh Bharadia. Uloc: Low-power, scalable and cm-accurate uwb-tag localization and tracking for indoor applications. In *Proc. of ACM IMWUT*, Vol. 5, No. 3, 2021.
 - [65] Pablo Corbalán, Gian Pietro Picco, and Sameera Palipana. Chorus: Uwb concurrent

- transmissions for gps-like passive localization of countless targets. In *Proc. of ACM IPSN*, 2019.
- [66] Ashutosh Dhekne, Mahanth Gowda, Yixuan Zhao, Haitham Hassanieh, and Romit Roy Choudhury. Liquid: A wireless liquid identifier. In *Proc. of ACM MobiSys*, 2018.
 - [67] Tianyue Zheng, Zhe Chen, Shujie Zhang, Chao Cai, and Jun Luo. More-fi: Motion-robust and fine-grained respiration monitoring via deep-learning uwb radar. In *Proc. of ACM SenSys*, 2021.
 - [68] Fusang Zhang, Zhaoxin Chang, Jie Xiong, Junqi Ma, Jiazhi Ni, Wenbo Zhang, Beihong Jin, and Daqing Zhang. Embracing consumer-level uwb-equipped devices for fine-grained wireless sensing. In *Proc. of ACM IMWUT*, Vol. 6, No. 4, pp. 1–27, 2023.
 - [69] Milad Afzalan and Farrokh Jazizadeh. Indoor positioning based on visible light communication: A performance-based survey of real-world prototypes. *ACM Computing Surveys (CSUR)*, Vol. 52, No. 2, pp. 1–36, 2019.
 - [70] Takuto Sato, Shota Shimada, Hiroaki Murakami, Hiroki Watanabe, Hiromichi Hashizume, and Masanori Sugimoto. Alisa: A visible-light positioning system using the ambient light sensor assembly in a smartphone. *IEEE Sensors Journal*, Vol. 22, No. 6, pp. 4989–5000, 2021.
 - [71] Wonho Kang and Youngnam Han. Smartpdr: Smartphone-based pedestrian dead reckoning for indoor localization. *IEEE Sensors journal*, Vol. 15, No. 5, pp. 2906–2916, 2014.
 - [72] Hang Yan, Qi Shan, and Yasutaka Furukawa. Ridi: Robust imu double integration. In *Proc. of ECCV*, 2018.
 - [73] Sachini Herath, David Caruso, Chen Liu, Yufan Chen, and Yasutaka Furukawa. Neural inertial localization. In *Proc. of IEEE/CVF CVPR*, 2022.
 - [74] Elahe Soltanaghaei, Akarsh Prabhakara, Artur Balanuta, Matthew Anderson, Jan M Rabaey, Swarun Kumar, and Anthony Rowe. Millimetro: mmwave retro-reflective tags for accurate, long range localization. In *Proc. of ACM MobiCom*, 2021.
 - [75] Andy Harter. The anatomy of a context-aware application. In *Proc. of ACM/IEEE MobiCom*, 1999.
 - [76] Fabian Höflinger, Rui Zhang, Joachim Hoppe, Amir Bannoura, Leonhard M Reindl, Johannes Wendeberg, Manuel Bühner, and Christian Schindelhauer. Acoustic self-calibrating system for indoor smartphone tracking (assist). In *Proc. of IEEE IPIN*, 2012.
 - [77] Joan Bordoy, Christian Schindelhauer, Fabian Höflinger, and Leonhard M Reindl. Exploiting acoustic echoes for smartphone localization and microphone self-calibration. *IEEE Trans. on Instrumentation and Measurement*, Vol. 69, No. 4, pp. 1484–1492, 2019.

- [78] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proc. of ACM MobiCom*, 2000.
- [79] Nissanka B Priyantha, Allen KL Miu, Hari Balakrishnan, and Seth Teller. The cricket compass for context-aware mobile applications. In *Proc. of ACM MobiCom*, 2001.
- [80] Yasuhiro Fukuju, Masateru Minami, Hiroyuki Morikawa, and Tomonori Aoyama. Dolphin: An autonomous indoor positioning system in ubiquitous computing environment. *WSTFES*, Vol. 3, p. 53, 2003.
- [81] Bing Zhou, Mohammed Elbadry, Ruipeng Gao, and Fan Ye. Batmapper: Acoustic sensing based indoor floor plan construction using smartphones. In *Proc. of ACM MobiSys*, 2017.
- [82] Swadhin Pradhan, Ghufraan Baig, Wenguang Mao, Lili Qiu, Guohai Chen, and Bo Yang. Smartphone-based acoustic indoor space mapping. In *Proc. of ACM IMWUT*, Vol. 2, No. 2, 2018.
- [83] Bing Zhou, Mohammed Elbadry, Ruipeng Gao, and Fan Ye. Battracker: High precision infrastructure-free mobile device tracking in indoor environments. In *Proc. of ACM SenSys*, 2017.
- [84] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cekan, Yingying Chen, Marco Gruteser, and Richard P Martin. Detecting driver phone use leveraging car speakers. In *Proc. of ACM MobiCom*, 2011.
- [85] Patrick Lazik and Anthony Rowe. Indoor pseudo-ranging of mobile devices using ultrasonic chirps. In *Proc. of ACM SenSys*, 2012.
- [86] Patrick Lazik, Niranjini Rajagopal, Oliver Shih, Bruno Sinopoli, and Anthony Rowe. Alps: A bluetooth and ultrasound platform for mapping and localization. In *Proc. of ACM SenSys*, 2015.
- [87] Patrick Lazik, Niranjini Rajagopal, Bruno Sinopoli, and Anthony Rowe. Ultrasonic time synchronization and ranging on smartphones. In *Proc. of IEEE RTAS*, 2015.
- [88] Ruizhi Chen, Zheng Li, Feng Ye, Guangyi Guo, Shihao Xu, Long Qian, Zuoya Liu, and Lixiong Huang. Precise indoor positioning based on acoustic ranging in smartphone. *IEEE Trans. on Instrumentation and Measurement*, Vol. 70, pp. 1–12, 2021.
- [89] Shuai Cao, Xiang Chen, Xu Zhang, and Xun Chen. Improving the tracking accuracy of tdma-based acoustic indoor positioning systems using a novel error correction method. *IEEE Sensors Journal*, Vol. 22, No. 6, pp. 5427–5436, 2021.
- [90] Kaikai Liu, Xinxin Liu, and Xiaolin Li. Guoguo: Enabling fine-grained smartphone localization via acoustic anchors. *IEEE Trans. on Mobile Computing*, Vol. 15, No. 5, pp. 1144–1156, 2015.

- [91] Shuai Cao, Xiang Chen, Xu Zhang, and Xun Chen. Effective audio signal arrival time detection algorithm for realization of robust acoustic indoor positioning. *IEEE Trans. on Instrumentation and Measurement*, Vol. 69, No. 10, pp. 7341–7352, 2020.
- [92] Xiang Chen, Yuheng Chen, Shuai Cao, Lei Zhang, Xu Zhang, and Xun Chen. Acoustic indoor localization system integrating tdma+ fdma transmission scheme and positioning correction technique. *MDPI Sensors Journal*, Vol. 19, No. 10, p. 2353, 2019.
- [93] Zheng Li, Ruizhi Chen, Guangyi Guo, Feng Ye, Long Qian, Shihao Xu, Lixiong Huang, and Liang Chen. Dual-step acoustic chirp signals detection using pervasive smartphones in multipath and nlos indoor environments. *IEEE Internet of Things Journal*, 2023.
- [94] Tao Liu, Xiaoji Niu, Jian Kuang, Shuai Cao, Lei Zhang, and Xiang Chen. Doppler shift mitigation in acoustic positioning based on pedestrian dead reckoning for smartphone. *IEEE Trans. on Instrumentation and Measurement*, Vol. 70, pp. 1–11, 2020.
- [95] Guangyi Guo, Ruizhi Chen, Ke Yan, Zheng Li, Long Qian, Shihao Xu, Xiaoguang Niu, and Liang Chen. Large-scale indoor localization solution for pervasive smartphones using corrected acoustic signals and data-driven pdr. *IEEE Internet of Things Journal*, Vol. 10, No. 17, pp. 15338–15349, 2023.
- [96] Hiroaki Murakami, Takumi Suzuki, Masanari Nakamura, Hiromichi Hashizume, and Masanori Sugimoto. Five degrees-of-freedom pose-estimation method for smartphones using a single acoustic anchor. *IEEE Sensors Journal*, Vol. 21, No. 6, pp. 8030–8044, 2020.
- [97] Hiroaki Murakami, Takuya Sasatani, Masanori Sugimoto, Issey Sukeda, Yukiya Mita, and Yoshihiro Kawahara. Syncecho: Echo-based single speaker time offset estimation for time-of-flight localization. In *Proc. of ACM SenSys*, 2024.
- [98] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. Fm-track: pushing the limits of contactless multi-target tracking using acoustic signals. In *Proc. of SenSys*, 2020.
- [99] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. Lasense: Pushing the limits of fine-grained activity sensing using acoustic signals. *Proc. ACM IMWUT*, Vol. 6, No. 1, 2022.
- [100] Lei Wang, Tao Gu, Wei Li, Haipeng Dai, Yong Zhang, Dongxiao Yu, Chenren Xu, and Daqing Zhang. Df-sense: Multi-user acoustic sensing for heartbeat monitoring with dualforming. In *Proc. of MobiSys*, 2023.
- [101] Juan He, Jie Xiong, Weihang Hu, Chao Feng, Enjie Yao, Xiaojing Wang, Chen Liu, and Xiaojiang Chen. Cw-acouslen: A configurable wideband acoustic metasurface. In *Proc. of MobiSys*, 2024.
- [102] Wenguang Mao, Jian He, and Lili Qiu. Cat: High-precision acoustic motion tracking. In

- Proc. of MobiCom*, 2016.
- [103] Wenchao Huang, Yan Xiong, Xiang-Yang Li, Hao Lin, Xufei Mao, Panlong Yang, Yunhao Liu, and Xingfu Wang. Swadloon: Direction finding and indoor localization using acoustic signal by shaking smartphones. *IEEE Trans. on Mobile Computing*, Vol. 14, No. 10, pp. 2145–2157, 2014.
 - [104] Yunting Zhang, Jiliang Wang, Weiyi Wang, Zhao Wang, and Yunhao Liu. Vernier: Accurate and fast acoustic motion tracking using mobile devices. In *Proc. of IEEE INFOCOM*, 2018.
 - [105] Haoran Wan, Lei Wang, Ting Zhao, Ke Sun, Shuyu Shi, Haipeng Dai, Guihai Chen, Haodong Liu, and Wei Wang. Vector: Velocity based temperature-field monitoring with distributed acoustic devices. *Proc. ACM IMWUT*, Vol. 6, No. 3, 2022.
 - [106] Weiguo Wang, Yuan He, Meng Jin, Yimiao Sun, and Xiuzhen Guo. Meta-speaker: Acoustic source projection by exploiting air nonlinearity. In *Proc. of ACM MobiCom*, 2023.
 - [107] Yijie Li, Juntao Zhou, Dian Ding, Yi-Chao Chen, Lili Qiu, Jiadi Yu, and Guangtao Xue. Mudis: An audio-independent, wide-angle, and leak-free multi-directional speaker. In *Proc. of MobiCom*, 2024.
 - [108] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury. Backdoor: Making microphones hear inaudible sounds. In *Proc. of ACM MobiSys*, 2017.
 - [109] 鎌倉友男. 非線形音響-基礎と応用-. コロナ社, 2014.
 - [110] 鶴木祐史, 西村竜一, 伊藤彰則, 西村明, 近藤和弘, 藺田光太郎. 音響情報ハイディング技術. コロナ社, 2018.
 - [111] Byeong-Seob Ko, Ryouichi Nishimura, and Yôiti Suzuki. Time-spread echo method for digital audio watermarking. *IEEE Trans. on Multimedia*, Vol. 7, No. 2, pp. 212–221, 2005.
 - [112] Nhut Minh Ngo and Masashi Unoki. Robust and reliable audio watermarking based on phase coding. In *Proc. of IEEE ICASSP*, 2015.
 - [113] Masashi Unoki and Ryota Miyauchi. Method of digital-audio watermarking based on cochlear delay characteristics. In *Multimedia Information Hiding Technologies and Methodologies for Controlling Data*, pp. 42–70. 2013.
 - [114] 村松巖, 荒川薫. オクターブ類似性に基づくオーディオ信号への電子透かし. 電子情報通信学会論文誌, Vol. 87, No. 6, pp. 787–796, 2004.
 - [115] Laurence Boney, Ahmed H Tewfik, and Khaled N Hamdy. Digital watermarks for audio signals. In *Proc. of IEEE ICMCS*, 1996.
 - [116] Hosei Matsuoka, Yusuke Nakashima, and Takeshi Yoshimura. Acoustic ofdm system and

- performance analysis. *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. 91, No. 7, pp. 1652–1658, 2008.
- [117] Tao Chen, Longfei Shangguan, Zhenjiang Li, and Kyle Jamieson. The design and implementation of a steganographic communication system over in-band acoustical channels. *ACM Trans. on Sensor Networks*, Vol. 19, No. 4, pp. 1–25, 2023.
 - [118] Daniel Gruhl, Anthony Lu, and Walter Bender. Echo hiding. In *Information Hiding: First International Workshop Cambridge*, pp. 295–315, 1996.
 - [119] Shuang-An Chou and Shih-Fu Hsieh. An echo-hiding watermarking technique based on bilateral symmetric time spread kernel. In *Proco. of IEEE ICASSP*, 2006.
 - [120] Guang Hua, Jonathan Goh, and Vrizlynn LL Thing. Time-spread echo-based audio watermarking with optimized imperceptibility and robustness. *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, Vol. 23, No. 2, pp. 227–239, 2015.
 - [121] Yong Xiang, Dezhong Peng, Iynkaran Natgunanathan, and Wanlei Zhou. Effective pseudonoise sequence and decoding function for imperceptibility and robustness enhancement in time-spread echo-based audio watermarking. *IEEE Trans. on Multimedia*, Vol. 13, No. 1, pp. 2–13, 2010.
 - [122] Yong Xiang, Iynkaran Natgunanathan, Dezhong Peng, Wanlei Zhou, and Shui Yu. A dual-channel time-spread echo method for audio watermarking. *IEEE Trans. on Information Forensics and Security*, Vol. 7, No. 2, pp. 383–392, 2011.
 - [123] International Telecommunication Union. Recommendation itu-r bs.1534-1: Method for the subjective assessment of intermediate quality level of coding systems. Technical report, ITU Radiocommunication Assembly, Geneva, Switzerland, 2003.
 - [124] Rainer Huber and Birger Kollmeier. Pemo-q—a new method for objective audio quality assessment using a model of auditory perception. *IEEE Trans. on audio, speech, and language processing*, Vol. 14, No. 6, pp. 1902–1911, 2006.
 - [125] Michael Chinen, Felicia SC Lim, Jan Skoglund, Nikita Gureev, Feargus O’Gorman, and Andrew Hines. Visqol v3: An open source production ready objective speech and audio metric. In *Proc. of IEEE QoMEX*, 2020.
 - [126] Takumi Suzaki, Hiroaki Murakami, Masanari Nakamura, Hiroki Watanabe, Hiromichi Hashizume, and Masanori Sugimoto. Pt-sync: Cots speaker-based pseudo time synchronization for acoustic indoor positioning. In *Proc. of IEEE IPIN*, 2022.
 - [127] Richard G Lyons. *Understanding digital signal processing*. Prentice Hall, Upper Saddle River, N.J., third edition, 2011.
 - [128] Takumi Suzaki, Masanari Nakamura, Hiroaki Murakami, Hiroki Watanabe, Hiromichi Hashizume, and Masanori Sugimoto. Expanding the positioning area for acoustic local-

- ization using cots mobile devices. In *Proc. of EAI MobiQuitous*, 2021.
- [129] 須崎太久弥, 中村将成, 村上弘晃, 渡邊拓貴, 橋爪宏達, 杉本雅則. 音響屋内測位におけるモバイルデバイスのための高精度測位領域の拡張手法. *情報処理学会論文誌*, Vol. 64, No. 11, pp. 1525–1536, 2023.
- [130] 伊藤信一. レーダシステムの基礎理論. コロナ社, 2015.
- [131] Fostex. Ft200d. <https://www.fostex.jp/products/ft200d/>. Accessed: 2024-12-09.
- [132] Donald G Childers, David P Skinner, and Robert C Kemerait. The cepstrum: A guide to processing. In *Proc. of IEEE*, 1977.
- [133] Yuta Nakashima, Ryuki Tachibana, and Noboru Babaguchi. Watermarked movie soundtrack finds the position of the camcorder in a theater. *IEEE Trans. on Multimedia*, Vol. 11, No. 3, pp. 443–454, 2009.
- [134] Yuta Nakashima, Ryosuke Kaneto, and Noboru Babaguchi. Indoor positioning system using digital audio watermarking. *IEICE Trans. on Information and Systems*, Vol. 94, No. 11, pp. 2201–2211, 2011.
- [135] 兼頭亮介, 中島悠太, 馬場口登. 音響電子透かしの検出強度を用いた位置推定. *電子情報通信学会総合大会講演論文集*, Vol. 2009, No. 2, 03 2009.
- [136] 兼頭亮介, 中島悠太, 馬場口登. 音響電子透かしを用いた屋内での録音位置推定. *電子情報通信学会総合大会講演論文集*, Vol. 2010, No. 2, 03 2010.
- [137] 須崎太久弥, 中村将成, 村上弘晃, 橋爪宏達, 杉本雅則. エコー埋め込みを用いた bgm による屋内測位. *情報処理学会 UBI 研究会*, 2024.
- [138] Python. sounddevice. <https://python-sounddevice.readthedocs.io/>.
- [139] European Broadcasting Union. Sound quality assessment material recordings for subjective tests. EBU Technology & Innovation, 2008.
- [140] Michaël Defferrard, Kirell Benzi, Pierre Vandergheynst, and Xavier Bresson. Fma: A dataset for music analysis. *arXiv preprint arXiv:1612.01840*, 2016.
- [141] Hermann Rohling. Radar cfar thresholding in clutter and multiple target situations. *IEEE Trans. on Aerospace and Electronic Systems*, No. 4, pp. 608–621, 1983.
- [142] David Chu. Polyphase codes with good periodic correlation properties (corresp.). *IEEE Trans. on Information Theory*, Vol. 18, No. 4, pp. 531–532, 1972.
- [143] Zhengchun Zhou, Dan Zhang, Tor Helleseth, and Jinming Wen. A construction of multiple optimal zcz sequence sets with good cross correlation. *IEEE Trans. on Information Theory*, Vol. 64, No. 2, pp. 1340–1346, 2017.