

特別研究報告

題目

マイクロホンアレイの無線ネットワーク化による野外環境下での
音源位置推定システムの実装と評価

指導教員

村田 正幸 教授

報告者

細川 侑嗣

平成 30 年 2 月 14 日

大阪大学 基礎工学部 情報科学科

マイクロホンアレイの無線ネットワーク化による野外環境下での
音源位置推定システムの実装と評価

細川 侑嗣

内容梗概

近年、生物の優れた機能や行動などに注目したバイオメティクスに関する研究が多くなされている。情報通信分野においても、群知能と呼ばれる、個体の自律的動作から生じる協調的な社会的行動をネットワーク制御に応用した研究が数多くみられる。我々は、群知能における生物同士のコミュニケーション、特にニホンアマガエルにみられる音声を用いたコミュニケーションに着目しており、ネットワーク制御への応用を目指している。音声コミュニケーションは多くの動物において見られるものであり、索敵などの群れの生存に関わる重要な戦略が潜んでいると考えられる。音声コミュニケーションを解析するためには、どの個体がいつ、どこで、どのようなシグナルを送受しているのかを計測することが重要になる。しかしながら、生物の多くは野外環境に生息し、発見が難しい種も多くいる。生物が発する音声を用いて、音源である生物の位置推定を行うことで、より容易に音声コミュニケーションを解析することが可能となる。近年は、音の到来方向を計測するためのマイクロホンアレイが比較的安価に入手できるようになっており、本報告では、実環境での携行や設置の容易さを考慮して、無線ネットワークにより相互接続したマイクロホンアレイを用いた音源位置推定システムを実装し、性能評価を行う。音源位置推定手法として、観測領域を格子状に分割し、各格子の中心位置に対して幾何学的に定義したコストが最小となるような格子を探索し、その中心座標を推定位置とする手法を用いる。構築した位置推定システムを用いた実験によって、見通しの良い屋外環境において、15 m×5 m の観測領域内の音源の位置を、平均 84 cm の誤差で推定できることを示した。またこの際、長さ 30 s の録音データを用いた音声到来方向の推定に要した時間は 4.45 s であり、推定した音声到来方向情報を用いた位置推定に要した時間は 0.099 s であった。

主な用語

音源位置推定、無線ネットワーク、音声到来方向、ニホンアマガエル

目次

1	はじめに	5
2	観測対象と音源位置推定システムの性能要件	7
3	関連研究	9
3.1	音声到来方向 (DOA) の推定手法	9
3.2	DOA を用いた位置推定手法	10
3.2.1	最尤推定法を使った位置推定手法	11
3.2.2	Grid-based method	12
3.3	実装する位置推定手法	12
3.3.1	位置推定範囲の拡張	13
3.3.2	単一音源の場合の位置推定	15
3.3.3	複数音源の場合の位置推定	16
4	音源位置推定システムの実装	18
4.1	使用機器とシステム構成	18
4.2	無線ノードの実装	21
4.2.1	音声データの取得機能の実装	21
4.2.2	MUSIC による DOA 推定機能の実装	21
4.3	位置推定サーバの実装	26
5	実装した音源位置推定システムの性能評価	29
5.1	屋外環境下での DOA 推定実験	29
5.1.1	実験環境	29
5.1.2	実験結果	29
5.2	屋外環境における位置推定実験	32
5.2.1	実験環境	32
5.2.2	実験結果	34
6	おわりに	38
	謝辞	39
	参考文献	40

目 次

1	ニホンアマガエル	7
2	DOA を用いた位置推定の例	10
3	Grid-based method の概要	13
4	マイクロホンに囲まれる領域の外側を対象とした位置推定	14
5	推定誤りの例	14
6	探索する格子数の削減	16
7	使用した無線ノード	19
8	位置推定システムの構成	20
9	位置推定システムの動作シーケンス	20
10	誤差のヒートマップ ($N_x = 100, N_y = 100$)	27
11	誤差のヒートマップ ($N_x = 200, N_y = 200$)	28
12	DOA 推定結果のヒストグラム	30
13	DOA 推定実験における音源とマイクロホンの位置	30
14	音源-マイクロホン間距離と DOA 推定精度	31
15	屋外実験の様子	32
16	音源位置推定システム	33
17	位置推定実験における音源とマイクロホンアレイの位置	34

表 目 次

1	使用機器のスペック	18
2	MUSIC のパラメータ	22
3	$N_x = 100, N_y = 100$ の際の位置推定精度	24
4	$N_x = 200, N_y = 200$ の際の位置推定精度	24
5	$N_x = 300, N_y = 300$ の際の位置推定精度	25
6	$N_x = 400, N_y = 400$ の際の位置推定精度	25
7	$N_x = 500, N_y = 500$ の際の位置推定精度	25
8	位置推定手法のパラメータ	26
9	位置推定実験結果 (単位はすべて m)	35
10	位置推定実験における各マイクロホンアレイの DOA	35

1 はじめに

生物の振る舞いに着想を得ることで新たな技術を創り出す、バイオミメティクスに注目が集まっている。このような学際的な研究進歩の背景には、様々な実験技術の開発とコンピュータ性能の向上によって、生物システムの数学的モデル化に関する数多くの研究が可能になったことがある。我々は情報通信技術の分野において、生物の自律分散的な振る舞いに着目し、ロバストで適応性のある制御技術へと応用してきた [1]。個体の自律的動作から生じる協調的な社会的行動は群知能と呼ばれ、ネットワーク制御に群知能の知見を応用した研究も多くなされている [2]。我々は生物の中でも、ニホンアマガエル（以降カエルとも表記）の持つ興味深い生態に着目している。カエルの種の多くでは、オスがメスを引き寄せるために広告音と呼ばれる求愛のための鳴き声を発する。広告音の役割は、求愛の他にオス同士の縄張りの主張であるとも考えられている [3]。ある種のカエルでは、広告音を同相同期的に発する、あるいは逆相同期的に発することが知られている。同相同期的に発声するとメスを自身に引き寄せる効果は薄れるものの、外敵に自身を発見されるリスクは低くなる。逆相同期的に発声すると、メスを自身に引き寄せる効果が期待できるが、一方で外敵に発見されるリスクが高まる。同相同期と逆相同期の違いのように、カエルの種類が異なると個体同士のコミュニケーションに差異が現れるのは、彼らの存在する外的環境の違いが影響していると予想され、群れの生存率を高めるための重要な戦略がその中に潜んでいる可能性がある。ニホンアマガエルは逆相同期的に広告音を発するカエルの一種であり、我々は、ニホンアマガエルの音声コミュニケーションを解析し、情報ネットワークへの応用を目指している。この取り組みにおいては、いくつかの研究成果も得られている [4]。

音声コミュニケーションは多くの動物に見られる普遍的な行動の一つであり、群れに社会性を生み出す重要な要素である。音声コミュニケーションを分析する際には、どの個体がいっ、どこで、どのようなシグナルで互いにコミュニケーションを取っているかを観測することが重要になる。しかしながら、生物の多くは野外環境に生息し、その発見が難しい種も数多く存在する。音を利用した位置推定は、生物が聴覚を獲得して以来用いてきた技術であり、現代ではマイクロホンによる録音が可能となったことで、音声信号を利用した位置推定を工学的に扱うことが可能となった。音源位置推定に関する研究はこれまでに様々になされており、複数のマイクロホンに音源からの音が到着した時刻の差 (Time Difference of Arrival; TDOA) を利用した手法、音源からの音の到来方向 (Direction of Arrival; DOA) を利用した手法に大別される [5]。

TDOA を用いた位置推定手法は、各マイクロホンの座標及び TDOA の値から、各マイクロホンを焦点とした双曲線を推定位置として求めるものである。二次元平面上においては、少なくとも三つのマイクロホンを同一直線状に並ばないように設置することで、音源の推定

位置が一点に定まる。信号の TDOA を求める方法には、GCC-PHAT [6] のように二つの信号の相互相関を求めるものがあるが、複数の音源から音声と同時に到着した場合には区別が困難であるという問題がある。また、TDOA を計算する際には、マイクロホン間の正確な時刻同期が必要となる。

DOA を用いた位置推定手法は、マイクロホンの位置を始点として音声到来した方向へ向かう半直線を推定位置として求めるものである。二次元平面上においては、少なくとも三つのマイクロホンを同一直線状に並ばないように設置することで、音源の推定位置が一点に定まる。信号の DOA を求める方法は Multiple Signal Classification (MUSIC) [6] のように信号部分空間を利用するものが、ノイズへの耐性が強いことから広く使われている。DOA を計算するためには複数の時刻同期したマイクロホンが必要であり、多くの場合、複数のマイクロホンをハードウェア的に接続することで時刻同期を実現する。このような機器はマイクロホンアレイと呼ばれ、マイクロホンアレイを用いることで高精度な DOA を得ることが可能である。また、MUSIC やその派生の DOA 推定手法では、マイクロホンアレイ内のマイクロホンの数を M とするとき、 $M - 1$ 個の音源から同時に発声が届いた場合にも、それぞれの DOA を推定することが可能である。

本研究の目的は、屋外環境において複数音源の位置を高精度に推定することが可能な音源位置推定システムを構築することである。屋外環境では様々なノイズや複数の音源からの音声が観測される可能性があるため、DOA を用いた位置推定手法を実装に用いる。具体的な目標精度や推定可能な音源数については、位置推定の対象とするニホンアマガエルの習性を考慮して設定を行う（詳細は 2 章に記載する）。

本報告では文献 [7] の位置推定手法を実機に実装し、評価を行う。DOA の推定方法については MUSIC を用いることとし、構築したシステムについて屋外環境下での実験による性能評価を行う。文献 [7] において提案した位置推定手法に対して、推定した DOA を入力として与えて音源の推定位置を得る。そのため、提案システムの位置推定精度は、MUSIC と位置推定手法の推定精度の両方に依存する。そこで、はじめにシミュレーションを用いて、DOA の推定精度に関する要件を明らかにする。次に、得られた要求精度を満たすように、実機に MUSIC を実装する。所望の精度が得られない場合には、位置推定手法において、利用するマイクロホンアレイの位置を変更する、数を増やすなどの緩和を行う。最後に、位置推定システムを構築し、その推定精度や実行時間などの性能を明らかにする。

本研究報告の構成は以下のとおりである。2 章に本研究で観測対象とするニホンアマガエルの特徴と、その特徴を踏まえたシステムの要件について説明する。3 章では、DOA の推定手法と DOA を用いた位置推定手法について関連研究を述べ、本報告で実装する手法についても説明を行う。4 章で具体的な実装方法について説明する。5 章で実験結果を示し、6 章で結論を述べる。



図 1: ニホンアマガエル

2 観測対象と音源位置推定システムの性能要件

ニホンアマガエルの体長は 2.0 ~ 4.5 cm 程度であり、水田や森林に生息する（図 1）。ニホンアマガエルは春が繁殖期であり、繁殖期では夕方から 0 時頃までの間、広告音と呼ばれる、オスがメスに自分の存在を知らせる鳴き声を発する。ニホンアマガエルは、1 匹のニホンアマガエルが鳴きだすと、それに合わせて鳴き始める（コーラス）。その際、2 ~ 4 匹程度の少数のニホンアマガエルがそれぞれが同時に鳴かないように逆相同期して鳴くことが知られている [8]。またこのとき、互いに 1 m 以上離れて鳴くことが多い。ニホンアマガエルは水中や水上ではあまり鳴かず、また、鳴いている間は移動をせず、一度合唱を始めると、5 分以上コーラスを続けることも多い。ニホンアマガエルの一回の広告音は 0.1 s から 0.2 s ほどの長さであり、その基本周波数は約 2,000 Hz である。

想定する環境 ニホンアマガエルが生息し、活発に鳴いている春～夏季の田圃を想定する。田圃の一辺は長くても 20 m 程度であり、田圃の周りには多くの場合、背の高い木はないが草が生い茂っている。田圃には水が張られており、畔はぬかるんでいる上に平坦な部分は少ない。本稿では単一の田圃を対象とし、棚田等、複数の田圃に渡っての推定は行わない、すなわち、鉛直方向の推定を行わないこととし、機器の設置できる場所は同一平面上であることとする。

システム要件 提案システムはカエルの観測だけでなく、音声コミュニケーションを行う様々な生物の観測に応用可能であるが、ここではニホンアマガエルの特徴およびその生息環境での利用を考慮してシステムの設計を行う。前述のとおり、推定システムは田圃において利用する。この際には、地面のぬかるみや様々な障害物、天候、地形などの要因があるため、想定通りに機器を配置することが必ずしも保証できるものではない。多数の機器を敷設する場合には、その準備やメンテナンスコストも問題となる。以上から、少数の機器により構成されるシステムであること、機器の設置位置にある程度の任意性があること、機器の調整を容易に行えることがシステムに求められる。

本報告においては、音源から発生する音声を複数のマイクロホンで取得し、無線ネットワークを用いることで計算機に必要な情報を収集することで位置の推定を行う。二次元平面状での位置推定は、同一直線状に並ばない3台の機器それぞれからDOAを収集すれば実現可能である。また、無線ネットワークを用いることで、機器同士が互いに無線通信範囲内に存在すれば互いに通信を行えるため、敷設のコストが軽減される。

性能についての要件としては、前述のニホンアマガエルの特徴を考慮して、以下の数値を目標と設定する。

- 推定精度：誤差 50 cm 以下
- 推定時間：5 分以下
- 推定可能領域： $20 \times 20 \text{ m}^2$

目標値について、推定誤差は1 mの間隔をあけて鳴いている二匹のカエルの識別が可能になるように設定した。推定時間については、カエルが継続して鳴き続ける時間を計測し、ヒストグラムを作成したところ、90 %以上を含んでいた300 sを今回は目標とした。観測領域については、単一の田圃を観測対象と考えており、 $20 \times 20 \text{ m}^2$ の領域での推定が可能であれば十分であると考えられる。

3 関連研究

3.1 音声到来方向 (DOA) の推定手法

音声到来方向 (DOA) の推定方法としては、MUltiple SIgnal Classification (MUSIC) [9] やその派生である Root-MUSIC [10]、ESPRIT [11] などが広く知られおり、文献 [12] で MUSIC を用いた鳥の DOA 推定も行われている。ここでは、本研究で用いる MUSIC について詳細を述べる。

MUSIC は周波数推定の方法として提案された手法であり、応用の一つとして信号の到来方向推定が着目されている。MUSIC では、観測信号の固有ベクトルによって張られる空間 (信号部分空間) に、利用するマイクロホンアレイにおける音声伝達特性を表すステアリングベクトル $\mathbf{A}(\omega, \theta)$ を変換する。ここで ω は周波数、 θ は角度を表しており、ステアリングベクトルが伝達関数形式であることを前提としている。観測音源以外の音源がなく、ノイズのない環境においては、観測信号の信号相関行列の固有値は、音源の数だけ正の値を取り、それ以外は 0 となる (ただし、いずれの信号も互いに無相関であることを前提としている)。音声到来方向を θ^* とするとき、0 固有値に対応する信号部分空間が $\mathbf{A}(\omega, \theta^*)$ と直交することを MUSIC では利用する。

マイクロホンアレイによって観測された時間領域の音声信号 $\mathbf{x}(\tau)$ のフーリエ変換を $\mathbf{X}(\omega)$ と表すこととする。それぞれの要素数は、アレイに含まれるマイクロホンの数 (M) である。また、 $\mathbf{S}(\omega)$ を音源の信号を表す複素数ベクトルとする (\mathbf{S} の要素数は音源の数 (L))。マイクロホンアレイが観測する信号 \mathbf{X} は、 \mathbf{S} が空間を伝播して届いた際には元の \mathbf{S} から振幅と位相が変化する。 θ の方向から到来する音源信号に関して、観測信号 \mathbf{X} はステアリングベクトルを用いて以下のように表される。

$$\mathbf{X}(\omega) = \mathbf{A}(\omega, \theta)\mathbf{S}(\omega) \quad (1)$$

このとき、観測信号の空間的性質を示す統計量である信号相関行列 $\mathbf{R}(\omega)$ は、以下の式で定義される。

$$\mathbf{R}(\omega) = \mathbf{X}(\omega)\mathbf{X}^H(\omega) \quad (2)$$

ただし、 \mathbf{X}^H は \mathbf{X} のエルミート転置を表す。 $\mathbf{R}(\omega)$ を固有値分解すると、

$$\mathbf{R}(\omega) = \mathbf{E}(\omega)\mathbf{\Lambda}(\omega)\mathbf{E}^{-1}(\omega) \quad (3)$$

と表される。ただし、 $\lambda_i(\omega)$, $\mathbf{e}_i(\omega)$ ($1 \leq i \leq M$) をそれぞれ $\mathbf{R}(\omega)$ の固有値、固有ベクトルとすると、 $\mathbf{E}(\omega) = \{\mathbf{e}_1(\omega), \mathbf{e}_2(\omega), \dots, \mathbf{e}_M(\omega)\}$ であり、 $\mathbf{\Lambda}(\omega)$ は対角成分 (i, i) を $\lambda_i(\omega)$

とする対角行列である。 $\lambda(\omega)$ は各音のエネルギーを表しており、音源の数を L とするとき、 $\lambda_i(\omega)$, $\mathbf{e}_i(\omega)$ は、 $1 \leq i \leq L$ の場合は音源 i に対応する固有値および固有ベクトルになる。

以上から、ステアリングベクトル $\mathbf{A}(\omega, \theta)$ の方向 θ と音源からの音声到来方向が一致すると、 $L+1 \leq i \leq M$ の場合、固有値は 0 となるため、式 (4) が得られる。

$$|\mathbf{e}_i(\omega)\mathbf{A}(\omega, \theta)|^2 = 0 \quad (4)$$

式 (4) を $L+1 \leq i \leq M$ の範囲で和を取り、その逆数に対して正規化項を乗じることで、式 (5) の空間スペクトル (MUSIC スペクトル) が得られる。

$$P(\omega, \theta) = \frac{|\mathbf{A}^H(\omega, \theta)\mathbf{A}(\omega, \theta)|}{\sum_{i=L+1}^M |\mathbf{A}^H(\omega, \theta)\mathbf{e}_i|^2} \quad (5)$$

P は前述のとおり、ステアリングベクトル $\mathbf{A}(\omega, \theta)$ の方向 θ と音源からの音声到来方向が一致すると、部分空間の直交性により分母が 0 となり (実際には、ノイズの影響で完全に 0 にはならない)、 P は無限大 (ピーク) になる。観測音声信号に対して MUSIC スペクトルを計算し、 P がピークを持つような θ を探索することで、音声の到来方向、すなわち DOA を推定することができる。

3.2 DOA を用いた位置推定手法

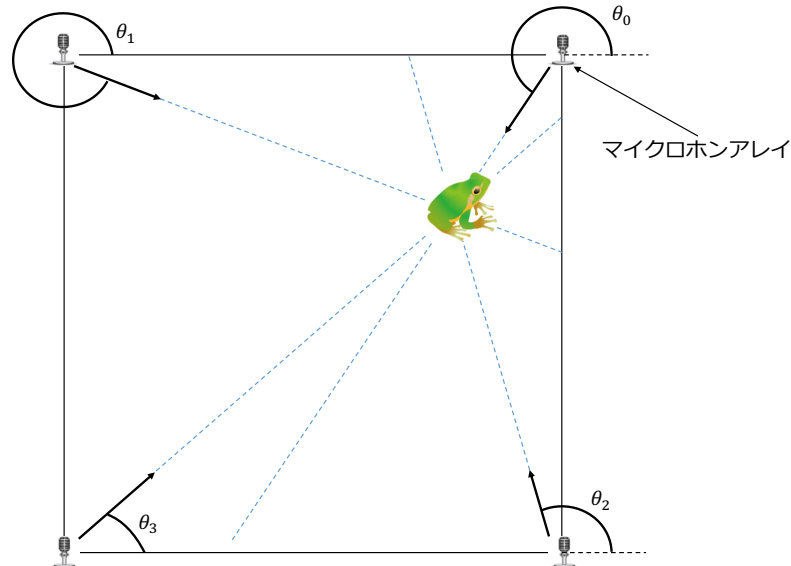


図 2: DOA を用いた位置推定の例

DOA を用いた音源の位置推定手法には様々なものがあるが、その基本的な考え方は、図 2 に示すようにマイクロホンアレイを始点として DOA の方向に伸びる半直線の交点を音源の位置とすることである。この方法は単純なため計算量が小さいが、複数の音源が存在する場合には、無関係な交点が発生するため一意に定まらない。また、DOA に推定誤差がある際には、DOA が一点上で交わらない状況が生じる。そこで、多くの場合は DOA を用いた幾何学的な推定により音源の位置を推定する。

以降の節では、DOA を用いた位置推定手法について関連研究を述べる。

3.2.1 最尤推定法を使った位置推定手法

文献 [13] では、最尤推定 (Maximum Likelihood; ML) 法を用いた位置推定手法を提案している。DOA の推定結果に誤差が含まれると、図 2 のようにはならず、DOA の交点が複数現れる場合がある。そこで、DOA の誤差が正規分布に従うと仮定し、ML 法によって位置推定を行う。具体的な推定方法は、式 (6) を最小化するような音源の位置 x を求めるものである。

$$J_{ML}^{(DOA)}(x) = \sum_{m=1}^M \frac{1}{\sigma_m^2} (\hat{\theta}_m - \theta_m(x))^2 \quad (6)$$

ここで、 M はマイクロホンアレイの数であり、 σ_m^2 はマイクロホンアレイ m における DOA の誤差の分散である。また $\hat{\theta}_m = \theta_m + n_m$ (ただし、 $n_m \sim N(0, \sigma_m^2)$) である。文献 [13] では DOA の誤差が音源とマイクロホン間の距離に依存する場合としない場合で、以下の 2 種類の手法を提案している。

Position Maximum Likelihood(P-ML) 音源とマイクロホンアレイの距離が増加するときに、DOA の誤差が増加するモデルを仮定した際に、最尤推定で音源位置を決定する。具体的には、式 (7) のコスト C_{P-ML} を最小化する音源位置 \hat{P} を推定位置とする。ここで、 $\angle(m-P)$ はマイクロホン m とある座標 P がなす角度を表す。

$$C_{P-ML} = \sum_{m=1}^M \left(\frac{1}{\sigma_m^2} |\hat{\theta}_m - \angle(m-P)|^2 + \ln \sigma_m^2 \right) \quad (7)$$

Position Nonlinear Least Squares(P-NLS) DOA の誤差が音源とマイクロホンアレイの距離に依存しないモデルを仮定した際に、最尤推定で音源位置を決定する。具体的には、式 (8) のコスト C_{P-NLS} を最小化する \hat{P} を推定位置とする。

$$C_{P-NLS} = \sum_{m=1}^M |\hat{\theta}_m - \angle(m-P)|^2 \quad (8)$$

3.2.2 Grid-based method

グリッドベースの位置推定アルゴリズム (Grid-based method [14]) では、観測領域を $N \times N$ 個の格子に分割し、各格子の中心座標と、計測した DOA を元に、各格子に対して幾何学的に定義されるコストを計算する (図 3)。Grid-based method は、このコストが最も小さい格子の中心座標を位置推定結果とするものであり、以下の手順 1.~5. により位置推定を行う。

1. 観測対象範囲を含んだ正方形領域を位置推定対象範囲とし、その四隅にマイクロホンアレイを設置する。
2. 位置推定対象範囲を $N \times N$ 個の格子に分割し、各格子の中心座標を計算する。
3. 各マイクロホンアレイから各格子の中心へ向かう角度 Ψ を計算する。 Ψ は $(M \times N)$ の行列で、各要素 $\psi_{m,n}$ は m 番目のマイクロホンアレイから n 番目の格子へ向かう角度を表す。
4. コスト関数 $\sum_{m=1}^M \left[A(\hat{\theta}_m, \psi_{m,n}) \right]^2$ を最小化する格子 n を求める。ここで、 $A(\hat{\theta}_m, \psi_{m,n})$ はマイクロホンアレイ m が推定した DOA ($\hat{\theta}_m$) と、マイクロホンアレイ m から n 番目の格子の中心に向かう角度 ($\psi_{m,n}$) との角距離を表す。角距離 A は文献 [14] で定義されており、 $A(\theta_1, \theta_2) = 2 \sin^{-1}((\exp(j\theta_1) - \exp(j\theta_2))/2)$ で与えられる (j は虚数単位)。
5. 格子の 1 辺の長さが閾値よりも長ければ、4. で求めた格子を位置推定対象範囲として 2. に戻る。そうでなければ、得られた格子の中心座標を推定結果とし、位置推定を終了する。

コスト関数は前述の ML 法のものと同様だが、位置推定対象範囲を離散化することで計算量を削減することに着眼している。文献 [14] の著者らは、手順 2. から 4. を再帰的に行うことで計算コストを抑えつつも、高精度の位置推定が可能であると主張している。Grid-based method の拡張として、同じ著者による文献 [15] や文献 [16] などがある。

3.3 実装する位置推定手法

本稿では文献 [7] の位置推定手法を実装する。この手法は前述の Grid-based method における位置推定対象範囲を、マイクロホンアレイで囲まれた外側の領域に拡張した手法である (図 4)。この手法について、以降に詳細を述べる。

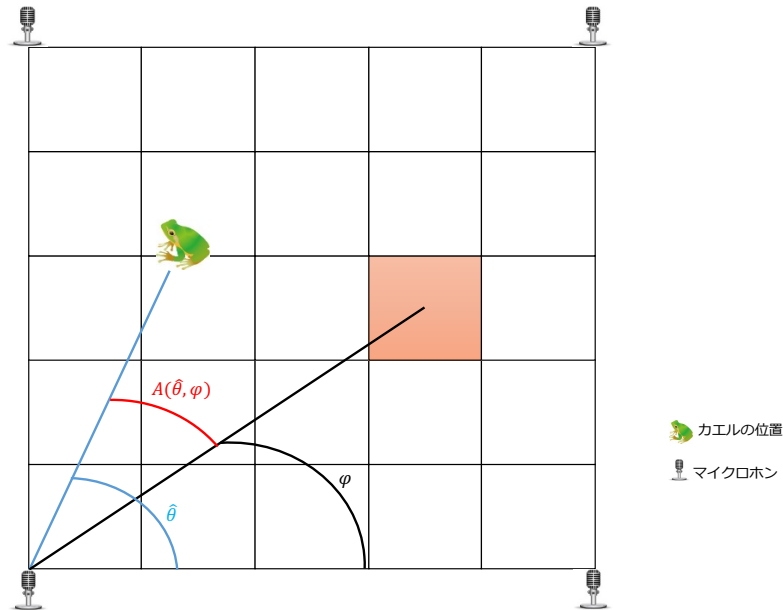


図 3: Grid-based method の概要

3.3.1 位置推定範囲の拡張

既存の位置推定手法の問題点として、実環境における携行・設置の困難さがある。既存手法の多くは、観測領域を囲むようにマイクロホンアレイを設置し、それらのマイクロホンアレイの正確な座標が利用可能であることを前提としている。しかしながら、実環境においては、各マイクロホンアレイの正確な位置の取得（あるいは任意の位置への設置）が困難である場合や、マイクロホンアレイが観測領域内の音源の DOA を推定するために十分な SN 比（信号対雑音比）を必ずしも獲得できない場合が考えられる。特に後者について、二次元平面上で位置推定が成功するためには、少なくとも二つのマイクロホンアレイの、DOA 推定可能範囲が重複する必要がある。広い領域をカバーするためにマイクロホンアレイ間の距離を大きく取ると、この重複面積は小さくなり、位置推定可能な領域が小さくなってしまう。上述の問題に対して、マイクロホンアレイ間の距離を近づけ、その上でマイクロホンアレイにより囲まれる領域の外側に存在する音源位置を推定することで解決を図る。各マイクロホンアレイの相対位置をあらかじめ物理的に固定しておくことで、各マイクロホンアレイの座標情報を正確に取得することが可能となり、さらに、設置・携行も非常に容易となる。

一方で、マイクロホンアレイで囲まれた範囲外の位置推定を行う場合には、文献 [14] のように再帰的に格子を探索する方法が利用できない。これは、音源と音源が含まれる格子の中心との距離が大きくなると、図 5 に示すように、推定誤りが生じることが理由である。図 5 では、音源が存在する格子（横線模様）とは異なる格子（縦線模様）の方が、音源か

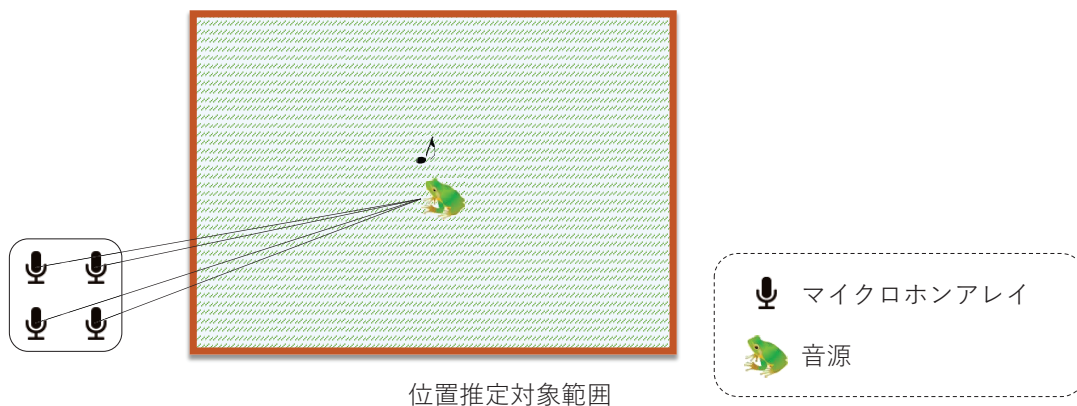


図 4: マイクホンに囲まれる領域の外側を対象とした位置推定

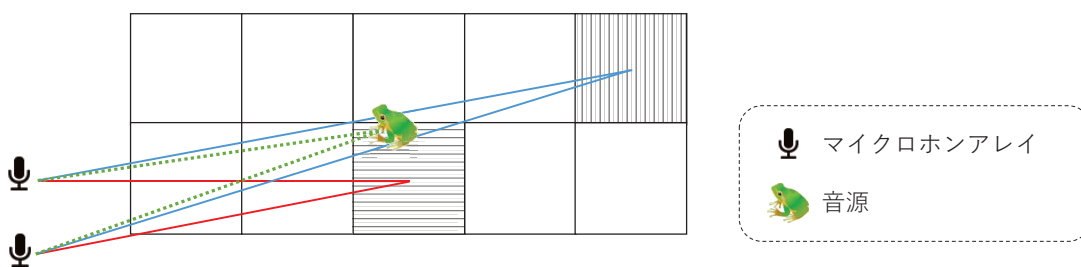


図 5: 推定誤りの例

らの DOA に近くなっている。最初に位置推定対象範囲をより小さな格子に分割することでこれを回避することが可能であるが、探索する格子の数が増えるために計算量の増加が問題となる。そこで、マイクロホンアレイで囲まれた範囲外を対象として位置推定を行う場合には、初めに、マイクロホンアレイの重心を始点として音源の存在する方向に向かう半直線を推定することとする。この半直線と交わる格子のみを対象に探索を行うことで、計算量の削減が可能である。

3.3.2 単一音源の場合の位置推定

まず、音源が単一の場合について説明を行う。位置推定は以下の三つのステップで行う。

Step 1. 観測対象範囲を含んだ長方形領域を位置推定対象範囲とし、位置推定対象範囲を $N_x \times N_y$ 個の長方形格子に分割する

Step 2. 各マイクロホンアレイから送られてきた DOA をもとに、マイクロホンアレイの重心を始点とした半直線を求め、この半直線と交差する格子を求める。

Step 3. Step 2. で得られた格子の中から、コストが最小となるものを求め、その中心座標を位置推定結果とする

マイクロホンアレイの位置に関しては、位置推定対象範囲内の任意の位置にある音源を対象に DOA の推定を行うために十分な SN 比が得られるのであれば、任意の位置に設置して良い。Step 1. では、あらかじめ観測領域を $N_x \times N_y$ 個の格子に分割する。この格子は必ずしも正方形である必要はない。格子の集合を P とし、各格子の中心の座標をあらかじめ計算しておく。この際の格子数は、所望の推定精度に応じて設定する。Step 2. において、探索する格子を決定するために、全マイクロホンアレイの重心から音源に向かう角度を推定する。マイクロホンアレイの重心から、この推定角方向に引いた半直線と交差する格子のみを、Step 3. におけるコスト計算の対象とする。探索する格子数の削減について、概要を図 6 に示す。

全マイクロホンアレイの重心を始点としたときに音源が存在する方向を推定するために、まず、式 (9) に示すコスト関数 C_{angle} を定義する。 C_{angle} は、各マイクロホンアレイから得た DOA と、角度 θ_i との角距離 A の二乗和であり、 C_{angle} が最小となる θ_i^* を推定結果とする。

$$C_{angle}(i) = \sum_{m=1}^M \left[A(\hat{\theta}_m, \theta_i) \right]^2 \quad (9)$$

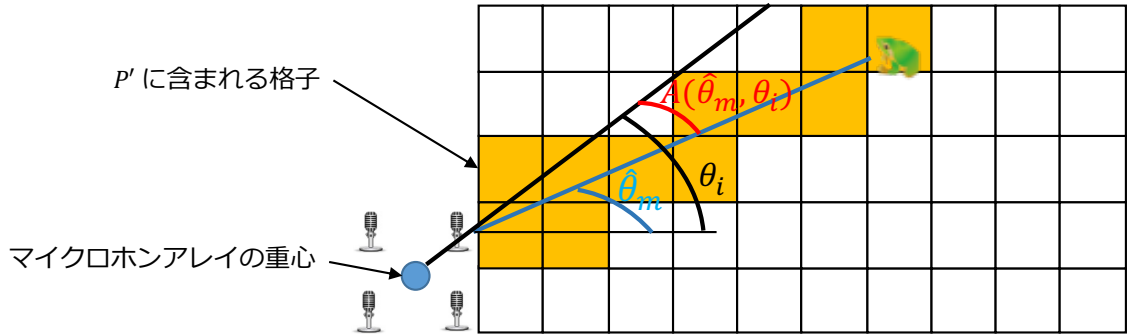


図 6: 探索する格子数の削減

$$\theta_i^* = \arg \min_{0 \leq \theta_i < 2\pi} (C_{angle}(\theta_i)) \quad (10)$$

ただし、 $\hat{\theta}_m$ は、マイクロホンアレイ m から得た DOA の推定値である。 $0 \leq \theta_i < 2\pi$ であり、 θ_i の粒度については実装に依存する。

推定の対象である格子の集合を P' とするとき、 P' の j 番目の要素に対するコスト $C_{cell}(j)$ を計算し、コストが最小となるような j^* を求め、その中心座標を推定結果とする。

$$C_{cell}(j) = \sum_{m=1}^M \left[A(\hat{\theta}_m, \theta_{m,j}) \right]^2 \quad (11)$$

$$j^* = \arg \min_{j \in P'} (C_{cell}(j)) \quad (12)$$

ここで、 $\theta_{m,j}$ は、マイクロホンアレイ m から P' の j 番目の格子の中心に向かう角度である。

3.3.3 複数音源の場合の位置推定

音源が複数個存在する場合には、マイクの数と音源の数の積だけ DOA が得られる。さらに、音源が存在する格子の組み合わせ全てについて、得られた DOA を元にコストを計算する必要がある。この際の格子の組み合わせ数は非常に大きくなり、計算量が問題となる。そこで、本手法では以下の四つのステップからなるヒューリスティックを用いる。音源の数を S 、マイクロホンアレイの数を M とする。ここで音源数は既知としているが、実装においては、DOA の推定の際に得られる値を用いる。

Step 1. 各マイクロホンアレイで得られた DOA から一つずつ取り出すときのすべての組み合わせ K を生成する

Step 2. 全ての $k \in K$ を用いて、単一音源の場合の位置推定を行う

Step 3. Step 2. において最も小さい C_{cell} を与える k^* を求め、その際の結果を位置推定結果とする

Step 4. k^* を求めた後に、 k^* に含まれる DOA を持つ組を K から除き、Step 2. に戻る

以上の処理を K が空集合になるまで続けることで、複数音源の位置推定を実施する。

表 1: 使用機器のスペック

項目	Raspberry Pi 3 Model B	Panasonic Let's note CF-SX1
クロック周波数	1.2 GHz, 4 コア	1.9 GHz, 2 コア
RAM	1 GB	8 GB
OS	Raspbian GNU/Linux 9.1 (stretch)	Windows 7

4 音源位置推定システムの実装

この章では、提案手法の実装方法について述べる。本稿において実装したシステムは、無線機器に接続したマイクロホンアレイと位置推定サーバによって構成される。以降でシステムを構成する機器と実装方法について詳細を述べる。

4.1 使用機器とシステム構成

マイクロホンアレイとしてたまご型マイクアレイ (TAMAGO-03; 株式会社システムインフロンティア) を使用する。TAMAGO-03 は高さ 12 cm、8 チャンネルのマイクロホンアレイであり、アナログ音声信号をサンプリング周波数 16 kHz で 24 bit の振幅情報として、デジタル変換する。他の機器とのインタフェースとして USB 端子を備えている。

マイクロホンアレイと位置推定サーバを無線接続するために、無線ノードとして Raspberry Pi 3 Model B (Raspberry Pi Foundation 社) を TAMAGO-03 に接続して利用する (図 7)。Raspberry Pi 3 Model B には無線 LAN アダプタ (IEEE 802.11b/g/n (2.4GHz) 対応) が標準搭載されており、実装するシステムにおいては、全端末と位置推定サーバがアドホックネットワークを構成する。

位置推定サーバとして、屋外での利用を考慮し、ノートパソコン (Panasonic Let's note CF-SX1) を用いる。これらの機器のスペック情報について、表 1 に示す。

これらの機器を用いたシステム構成を図 8 に示す。図に示すように、4 台の無線ノードと 1 台の位置推定サーバを用いる。本音源位置推定システムによる音源の位置推定は以下の順序で実施される。

1. 各マイクロホンアレイは T 秒間録音を行い wav ファイルに出力する
2. 各マイクロホンアレイは wav ファイルから Δ 秒ごとに音源の DOA を推定する
3. マイクロホンアレイは推定した DOA を位置推定サーバに送信する
4. 位置推定サーバは受信した DOA を元に位置推定を行い結果を出力する



図 7: 使用した無線ノード

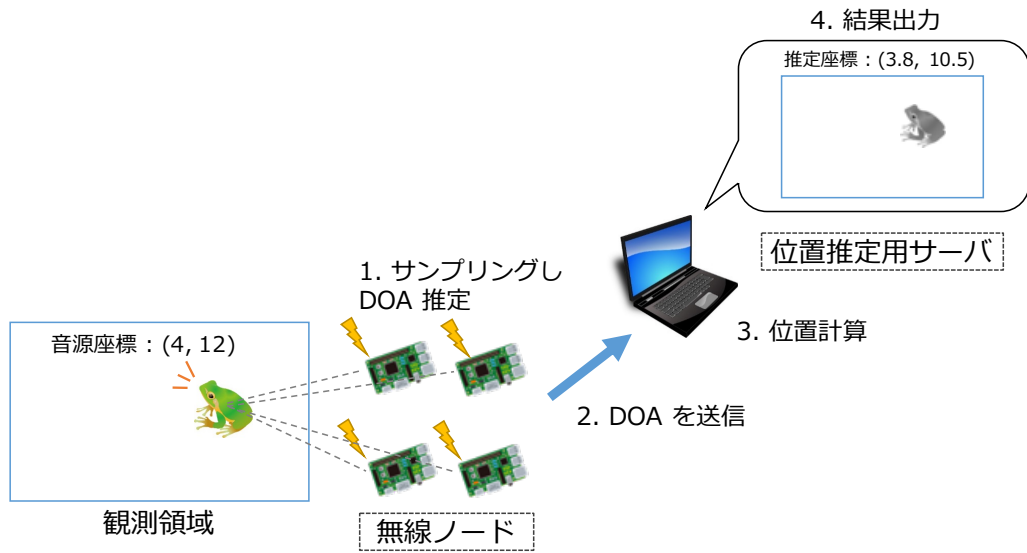


図 8: 位置推定システムの構成

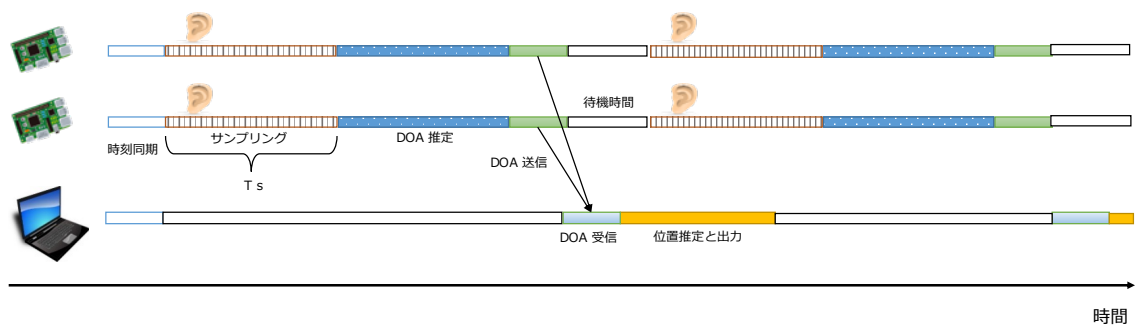


図 9: 位置推定システムの動作シーケンス

無線ノードと位置推定サーバの動作シーケンスは図 9 に示す通りである。マイクロホンアレイと位置推定サーバはあらかじめ ntp (network time protocol) により時刻同期しており、位置推定サーバには、全無線ノードから同時刻に DOA の推定値が届く。全ての無線ノードから DOA の推定値が届くと、位置推定サーバは位置推定を行い、結果を出力する。

以降の節では無線ノードの実装及び位置推定サーバの実装について詳細を述べる。

4.2 無線ノードの実装

無線ノードに必要な機能は、音声データの取得、音声データからの DOA 推定、推定結果の送信である。以降の節でこれらの実装について詳細を述べる。

4.2.1 音声データの取得機能の実装

無線ノードは位置推定システムの起動時に、位置推定サーバと ntp を用いた時刻同期を行う。これは、何らかの理由で推定した DOA の送信に失敗した際に、オフラインで推定結果の整合性を確認するためである。時刻同期後、python2.7 で実装したプログラムで T 秒の間、16 kHz での音声サンプリングを開始し、wav ファイルを作成する。 T は録音する音声長であり、任意に設定可能である。

4.2.2 MUSIC による DOA 推定機能の実装

wav ファイルを入力とし、c++で実装した MUSIC によって DOA の推定を行う。MUSIC による DOA 推定については 3.1 節に述べたとおりである。録音した音声には無音の区間が含まれることもあるため、 T 秒の音声を Δ 秒ごとに区切り、各区間で音圧値の二乗和が閾値よりも高い場合にのみ DOA の推定を行う。推定結果は、逐次テキストファイルに書き込む。 T 秒間の音声全ての DOA 推定が終わると、直ちに位置推定サーバに DOA を記録したテキストファイルを送信する。MUSIC では音声を読み出し、信号相関行列を求め、その固有値と固有ベクトルを導出する。この固有値はマイク数 M 、音源数 L に対して、 $M - L$ 個だけが非常に小さな値を取り、それらの固有値に対応する固有ベクトルが成す空間は雑音成分による部分空間である。音声到来方向と一致するステアリングベクトルの要素はこの空間と直交するため、DOA の推定に用いることができる。提案システムにおいて MUSIC を実装する際に設定するパラメータについて、表 2 にまとめる。

これらの内、SAMPLING_FREQUENCY、NUM_FREQ、DEGREE_SIZE については、利用するステアリングベクトルによって定まる。ステアリングベクトルはマイクロホンアレイの信号伝達特性を表すものであり、取得の方法は大きく二つある。一つは各マイクロホ

表 2: MUSIC のパラメータ

Parameters	Value	Description
BLOCKTIME	0.5 s	Δ の長さ
SAMPLING_FREQUENCY	16,000 Hz	サンプリング周波数
FRAME_LENGTH	512	BLOCKTIME 中のフレームの数
WINDOW_TYPE	HAMMING	窓関数の指定 (ハミング窓)
NUM_FREQ	512	離散フーリエ変換における周波数の分割数
FREQUENCY_RANGE_L	1,500 Hz	MUSIC スペクトルを計算する周波数の下限
FREQUENCY_RANGE_U	3,000 Hz	MUSIC スペクトルを計算する周波数の上限
NUM_SOURCE	1	音源の数 S
NUM_MIC	8	マイクロホンアレイの持つマイクロホンの数 M
DEGREE_SIZE	5°	DOA 推定における角度の粒度

ンの位置関係から解析的に求める方法であり、もう一つは実測により求める方法である。今回使用する TAMAGO-03 においては、卵型の筐体の周囲にマイクロホンがあるため、音の回り込みの特性を解析的に求めることは困難である。そのため実測により求めたステアリングベクトルの利用が有効であると考えられる。ただし、ステアリングベクトルの正確な計測は専用の機器が必要である。本報告においては、ロボット聴覚オープンソースソフトウェアである HARK (Honda Research Institute Japan Audition for Robots with Kyoto University) [17] が公開しているステアリングベクトルファイルを利用した。

BLOCKTIME は 0.5 s であり、SAMPLING_FREQUENCY が 16 kHz であるため、この間のサンプル数は 8,000 となる。FRAME_LENGTH で指定した値でこのサンプルを分割し (剰余は切り捨てる) 複数のフレームを作成する。このフレームごとに窓関数をかけ合わせた後に離散フーリエ変換によって周波数成分に変換する。窓関数には一般的に利用されるハミング窓 [18] を利用している。離散フーリエ変換では、SAMPLING_FREQUENCY を NUM_FREQ で除算した値である 31.25 Hz ごとの周波数成分を求める。これ以降は、各周波数成分ごとに個別に計算を行う。この際、FREQUENCY_RANGE_L および FREQUENCY_RANGE_U で指定した周波数の下限と上限の範囲内のみを対象とする。各フレームの信号相関行列を求め、その総和をこの 0.5 s 間の信号相関行列とする。

信号相関行列を求めた後は、その固有値および固有ベクトルを求める。固有値の大きいものから NUM_SOURCE 個だけ取り除き、残った固有値に対応する固有ベクトルと、ステアリングベクトルを元に MUSIC スペクトルを求める。このようにして得られた MUSIC スペクトルは各周波数成分ごとに求めたものであるため、最終的な DOA の推定には、これらの

重み付け和を用いる。重みとしては各周波数成分における最大固有値の平方根を用いる [17]。ここで、推定できる DOA の精度はステアリングベクトルに依存する。HARK の公開しているステアリングベクトルは 5° (=DEGREE_SIZE) ごとに計測した値であるため、DOA は 5° ごとに出力される。

ステアリングベクトルの補間 本実装においては、MUSIC による DOA の推定精度を向上するために、文献 [19] の補間法でステアリングベクトルを補間する。この補間法は周波数領域での線形補間と、時間領域での線形補間を組み合わせた方法 (Frequency- and Time-Domain Linear Interpolation (FTDLI)) であり、位相特性と振幅特性の両方について精度良く補間可能である。FTDLI は任意の粒度に補間可能であるが、DOA の推定に要する計算量は DEGREE_SIZE に反比例するため、推定精度と計算量がトレードオフとなる。そこで、どの程度の粒度で補間するかを事前に定める必要がある。以降では、補間する粒度を定めるための計算機シミュレーションの結果を示す。

本報告では位置推定が可能な領域の広さとして、 $20\text{ m} \times 20\text{ m}$ を目標とした。ここでは、その範囲の中心にマイクロホンアレイを設置したと仮定し、対称性を考えて、 $20\text{ m} \times 20\text{ m}$ の領域を四分割した $10\text{ m} \times 10\text{ m}$ の領域の一つについて位置推定精度を評価する。四つのマイクロホンアレイを一辺 1 m の正方形の各頂点上に設置し、それぞれの座標を $(-1, -1)$ 、 $(-1, 0)$ 、 $(0, 0)$ 、 $(0, -1)$ とする。また、観測領域の各頂点の座標を $(0, 0)$ 、 $(0, 10)$ 、 $(10, 10)$ 、 $(10, 0)$ とする。音源については、この観測領域を 100×100 の格子に分割し、それぞれの格子の中心に設置する。

位置推定手法の精度は、観測領域を分割する格子数、DOA 推定精度、マイクロホンアレイの位置と数によって決まる。マイクロホンの位置と数は前述のように設定しているため、ここでは、観測領域を分割する格子数、DOA 推定精度と位置推定精度の関係を明らかにする。前述の通り、DOA の推定結果は、ステアリングベクトルの DEGREE_SIZE に丸められる。ここでは、各音源の位置とマイクロホンアレイの位置から求まる真の DOA を用いる場合、DEGREE_SIZE を整数倍した値の中で、真の DOA から最も近い値を用いる場合について調べる。DEGREE_SIZE の値を 0° 、 1° 、 2° 、 3° 、 4° 、 5° と設定し、 N_x 、 N_y の値をともに 100 、 200 、 300 、 400 、 500 と設定する。ただし、DEGREE_SIZE が 0° とは、DOA が真の値であることを表す。

位置推定精度について、最小誤差、最大誤差、平均誤差は表 3~7 に示す結果となった。DEGREE_SIZE が小さいほど誤差は小さくなるのが分かる。最小誤差、平均誤差については、いずれの場合でも目標精度である 50 cm 以下となっているが、最大誤差については DOA として真の値が得られた場合のみ 50 cm 以下を達成できていることが分かる。ただし、今回のシミュレーションでは全てのマイクロホンアレイで正しい DOA が取得できるこ

表 3: $N_x = 100, N_y = 100$ の際の位置推定精度

DEGREE_SIZE	0°	1°	2°	3°	4°	5°
平均誤差 (cm)	14.5	17	23.6	30.5	39.8	48.5
最大誤差 (cm)	55.2	99.2	148.5	199.1	223.3	257
最小誤差 (cm)	0	0	7.1	7.1	7.1	0

表 4: $N_x = 200, N_y = 200$ の際の位置推定精度

DEGREE_SIZE	0°	1°	2°	3°	4°	5°
平均誤差 (cm)	7.3	12.1	20.5	29.3	38.7	47.3
最大誤差 (cm)	27.6	77.9	138.1	191.7	208.7	244.9
最小誤差 (cm)	0	3.5	3.5	3.5	3.5	3.5

とを仮定して、その上で値の丸め処理を行っているが、実際には DOA 推定自体に誤差が含まれることも考えられる。

DEGREE_SIZE=1° とした際の、次章の実験に用いる機器において、マイクロホンアレイで録音した 30 s の音声を用いて MUSIC による DOA 推定を行った結果、4.45 s となった。DEGREE_SIZE が大きくなるにつれ、計算時間は短くなると考えられるため、DEGREE_SIZE=1° の場合でも計算時間は目標時間である 5 分を下回っているため、DEGREE_SIZE は 1° と設定することとする。

誤差に関しては、その空間的な分布がどのようになっているのかも重要である。そこで、誤差の空間分布を調べるためにヒートマップを作成した。前述の通り、シミュレーションにおいては、10 m×10 m の観測領域を 100×100 の格子に分割し、それぞれの格子の中心に音源を置いている。各格子に置いた音源に対する位置推定を行い、その誤差に応じた色で格子内を塗りつぶし、誤差の分布をヒートマップで表現した。

$N_x = 100, N_y = 100$ および、 $N_x = 200, N_y = 200$ の場合について、図 10、11 で示した結果が得られた。図 10 から分かるように、マイクロホンアレイに対して水平あるいは垂直の方向に近づくほど誤差が大きくなる、また、マイクロホンアレイから離れるほどで大きな誤差が生じやすいことが分かる。前者については、位置推定手法において、初めに推定する半直線が必ずしも音源のある格子を通過しないことが原因となっている。後者については、マイクロホンアレイからの距離が大きい地点では、各マイクロホンアレイ DOA の推定結果の差が小さくなるため、DOA 推定の際の粒度の影響がより大きく現れ、結果的に誤差が大きくなっている。これらの点についてはマイクロホンアレイの数を追加する、あるいはマイクロホンアレイ間の距離を大きくすることで、誤差を抑えることが可能であることが分かつ

表 5: $N_x = 300, N_y = 300$ の際の位置推定精度

DEGREE.SIZE	0°	1°	2°	3°	4°	5°
平均誤差 (cm)	4.8	11.0	20.1	29.1	38.5	47.2
最大誤差 (cm)	21.7	80.2	139.4	212.8	203.9	251.9
最小誤差 (cm)	0	2.4	2.4	2.4	2.4	2.4

表 6: $N_x = 400, N_y = 400$ の際の位置推定精度

DEGREE.SIZE	0°	1°	2°	3°	4°	5°
平均誤差 (cm)	3.6	10.6	19.9	29.0	38.5	47.2
最大誤差 (cm)	21.3	76.1	137.9	205.7	210.5	256.3
最小誤差 (cm)	0	1.8	1.8	1.8	1.8	1.8

表 7: $N_x = 500, N_y = 500$ の際の位置推定精度

DEGREE.SIZE	0°	1°	2°	3°	4°	5°
平均誤差 (cm)	2.9	10.4	19.8	29.0	38.6	47.3
最大誤差 (cm)	21.0	67.9	134.4	201.4	207.2	251.1
最小誤差 (cm)	0	1.4	1.4	1.4	1.4	1.4

表 8: 位置推定手法のパラメータ

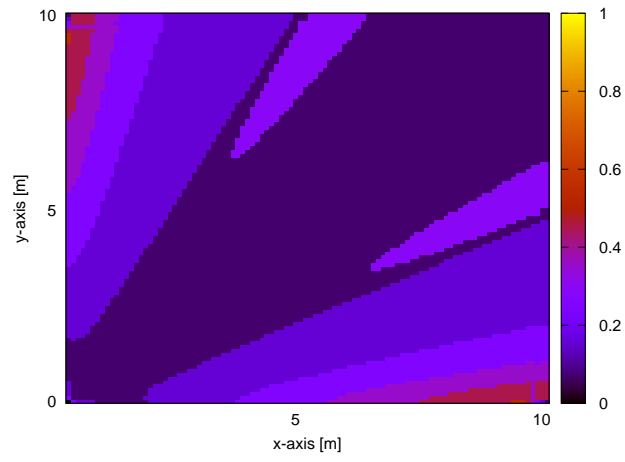
Parameters	Value	Description
N_x	100	x 軸の格子数
N_y	100	y 軸の格子数
MIC_NUM	4	マイクロホンアレイの数
RES_ANGLE	0.05°	θ_i の粒度

ている [20]。一例として、DEGREE_SIZE を 1° 、 N_x 、 N_y を 500 と設定した際に、マイク間距離を 2 m に変更したところ、平均誤差は 6.6 cm、最大誤差は 47.0 cm、最小誤差は 0 cm となり、目標精度を達成可能である結果が得られた。ただし屋外環境においては、マイクロホンアレイ間の距離が大きくなるほど、マイクロホンアレイの相対位置の正確な計測が困難となる点は留意すべきである。

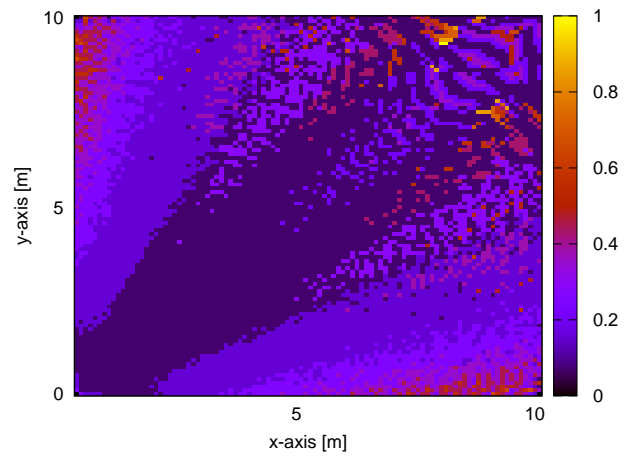
4.3 位置推定サーバの実装

無線ノードから位置推定サーバに届いた推定結果は、 T 秒の音声に関して Δ 秒ごとに求めた DOA 推定値である。この際、DOA は真値に対して大きな誤差を含む場合がある。そこで、各マイクロホンアレイから届いた T 秒の DOA 推定結果について、最頻値を求め、その値を位置推定の入力として与えることとする。最頻値を用いるのは、DOA の誤差がノイズや他の音源によるものである場合には、推定結果が真値に依存しないためである。最後に位置推定結果をテキストファイルに出力する。

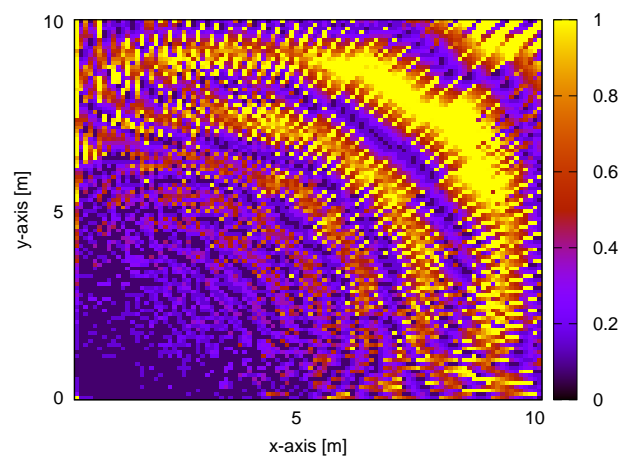
位置推定手法は 3.3 節で述べたように、複数のマイクロホンアレイに届いた DOA から、音源の存在する可能性のある半直線を導出し、その半直線上と交差する格子を探索することで、音源の位置を推定する。位置推定手法実装におけるパラメータ設定については表 8 の通りである。



(a) DEGREE_SIZE=0°

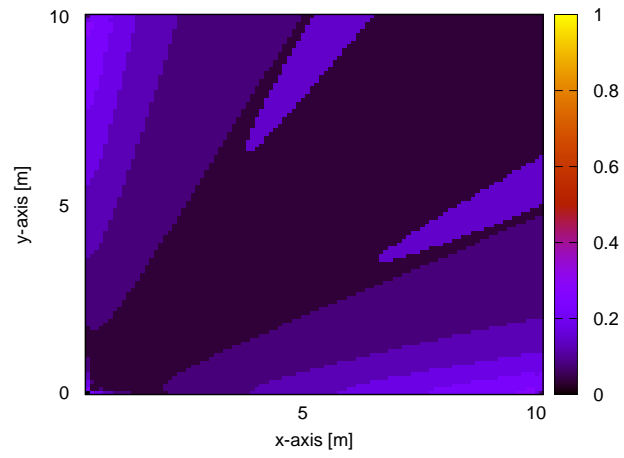


(b) DEGREE_SIZE=1°

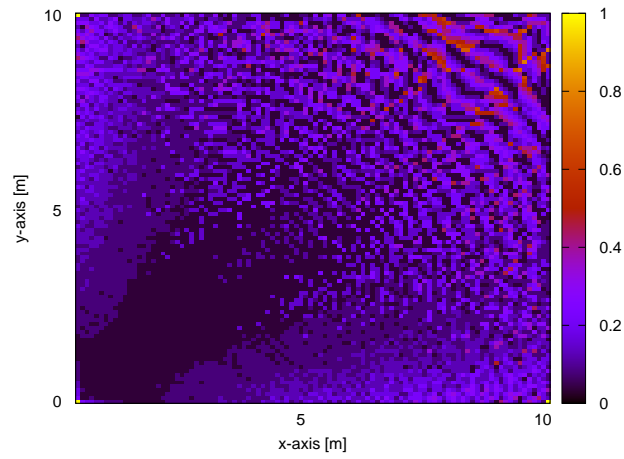


(c) DEGREE_SIZE=5°

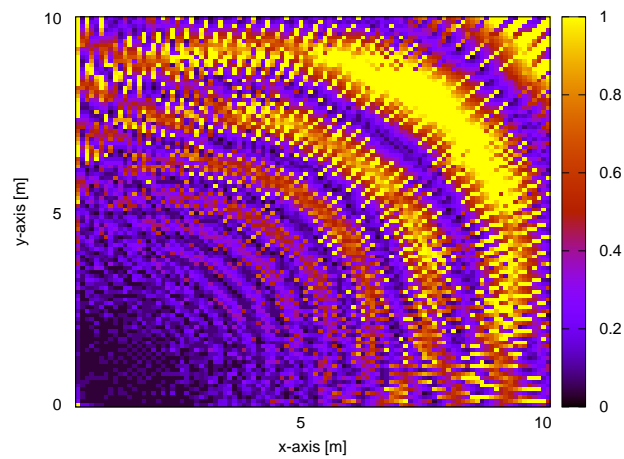
図 10: 誤差のヒートマップ ($N_x = 100, N_y = 100$)



(a) DEGREE_SIZE=0°



(b) DEGREE_SIZE=1°



(c) DEGREE_SIZE=5°

図 11: 誤差のヒートマップ ($N_x = 200, N_y = 200$)

5 実装した音源位置推定システムの性能評価

4章で実装したシステムを用いて位置推定実験を行い、その精度および計算時間を評価する。3.3節で述べたように、位置推定精度は、DOAの推定精度に依存するため、まず屋外環境におけるDOAの推定精度を評価する実験を行った。その後、構築したシステムを用いた屋外環境での位置推定実験を行い、システムの位置推定精度を評価した。

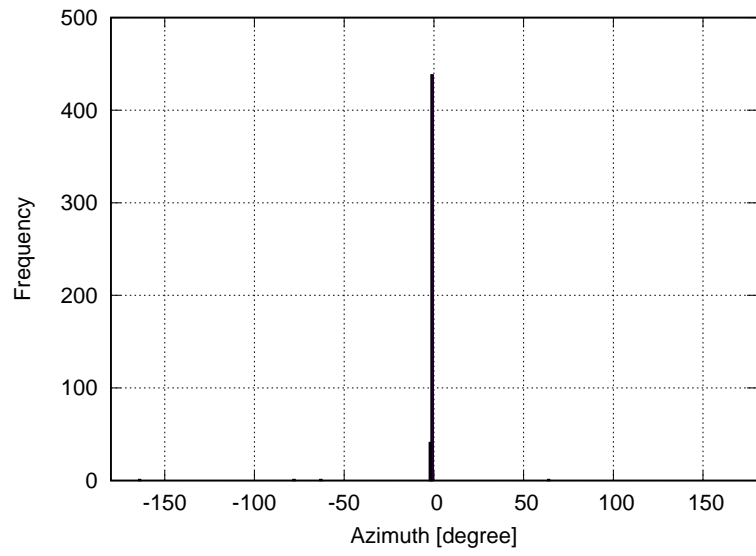
5.1 屋外環境下でのDOA推定実験

5.1.1 実験環境

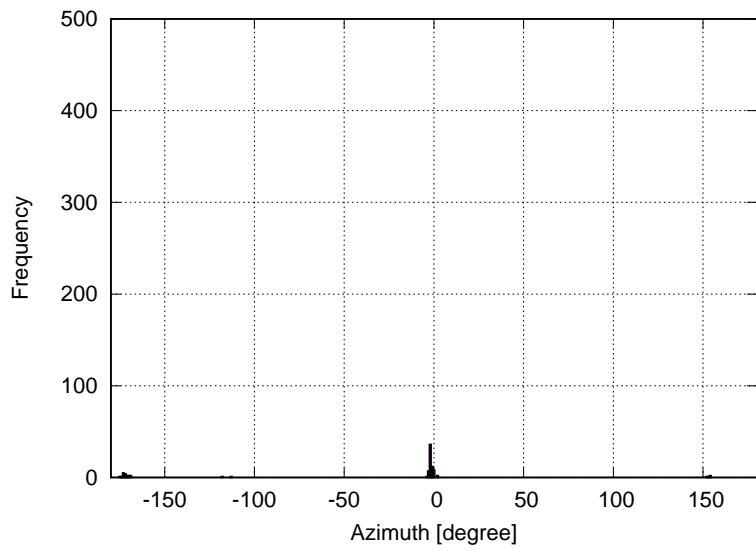
音源-マイクロホン間の距離とDOA推定の結果との関係を明らかにするため、周囲に障害物のないグラウンドにおいて、図13マイクロホンアレイを一つ設置し、音源の位置をマイクロホンアレイから1m離すごとにDOAの推定を行った。距離測定のため、最大距離を25mとした。音声としてニホンアマガエルの広告音を録音したものをスピーカーで再生した。各音源位置に関して、この音声をマイクロホンアレイにより5分間録音し、0.5sごとにDOAの推定を行う（合計600回）。ここで、音源音圧を計測した所、最大値が80dBであった。実際に屋外で夜間のカエルの鳴き声を計測した際には最大音圧は110dBを超えることもあり、その音圧よりは小さい値であるが、スピーカーの性能上この値が限界であった。マイクロホンアレイから音源への角度は 0° となるようにスピーカーを設置し、推定結果のヒストグラムを得る。ここで、得られたヒストグラムにおいて、最頻値が真の角度から誤差 1° 以内となるような音源-マイクロホンアレイ間の最大距離を、DOA推定可能距離とする。

5.1.2 実験結果

本実験の結果では、DOA推定可能距離は実験における最長距離の25mとなった。図12は、音源とマイクとの距離が1m, 20mのそれぞれの場合におけるDOA推定結果をヒストグラムとして示したものである。頻度の総和が600とならないのは、DOA推定を行う際には音圧値の二乗和を用いて静音区間のフィルタリングを行っているためである。図12(b)では -170° 付近に小さなピークが現れており、ノイズが特定の角度から到達する場合には真の角度と全く異なる角度に推定結果が現れることが分かる。そのため、DOAの推定結果を単に平均・分散で評価することは不適切であると考えられる。そこで、DOAの推定精度を評価するための指標として、推定結果の中で、真の角度に一致した結果の割合、および絶対誤差が 1° 以下となる割合を評価する（絶対誤差の 1° は、人による設置の誤差と考えて設定した）。



(a) 音源-マイク間距離 1 m



(b) 音源-マイク間距離 20 m

図 12: DOA 推定結果のヒストグラム

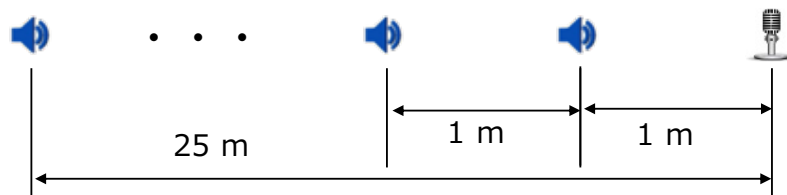


図 13: DOA 推定実験における音源とマイクロホンの位置

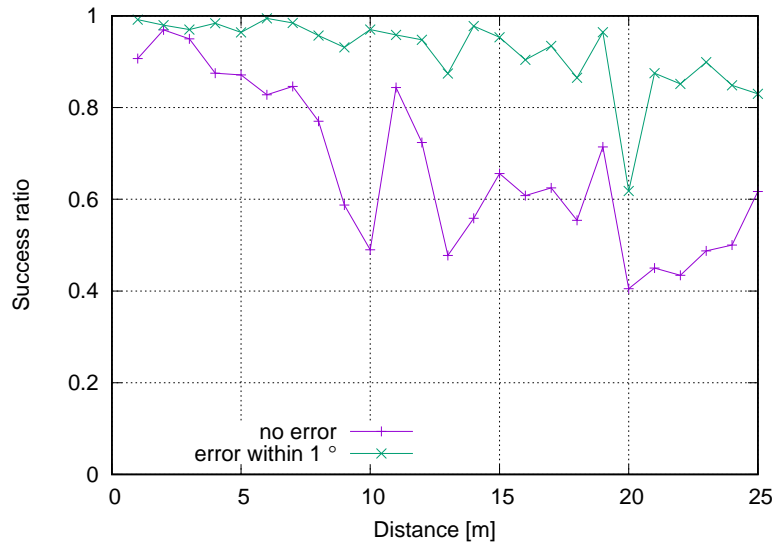


図 14: 音源-マイクロホン間距離と DOA 推定精度

評価結果を図 14 に示す。横軸を音源-マイクロホンアレイ間の距離としており、縦軸はその角度に推定された割合を表している。推定精度が低くなる点は、設置誤差が原因と考えられ、こういった落ち込みはサンプル数を増やしたとしても避けるのは難しいと考えられる。前述のとおり、音源とマイクロホンとの距離が 1~25 m の範囲では DOA 推定の結果の最頻値が音源とマイクロホンとの真の角度 $\pm 1^\circ$ の範囲に含まれた。またその割合は図 14 に示すように音源とマイクロホンが遠いほど低い傾向がみられるものの、25 m 離れた際にも 8 割以上となっている。



図 15: 屋外実験の様子

5.2 屋外環境における位置推定実験

5.2.1 実験環境

実装した音源位置推定システムの性能を評価するために、屋外環境における位置推定実験を行った。DOA の推定実験と同様に、障害物のないグラウンドで実験を行った。実験の様子を図 15 に示す。屋外環境における位置推定精度を評価するためには、音源の正確な座標が必要となる。しかしながら、グラウンドでは正確な座標を取得することが困難である。そこで、まず座標 $(0,0)$ と定めた点から、メジャーを用いた測定により、座標 $(5,0)$ の地点を決定し、これらを基準点とする。その後、基準点からある程度距離を離しつつランダムに音源を設置し、各音源位置において、それぞれの基準点からの距離をレーザー測距計で測定し、三点測量で音源の座標を求めた。用いたレーザー測距計はライカ社のライカディスト D210 であり、誤差は 1 mm 程度である。実験は昼間に行ったため、基準点から離れるほど、基準点にレーザー照準を定めることが困難となった。そのため、音源とマイクロホンアレイ間の最大距離は 10 数 m 程度となっている。

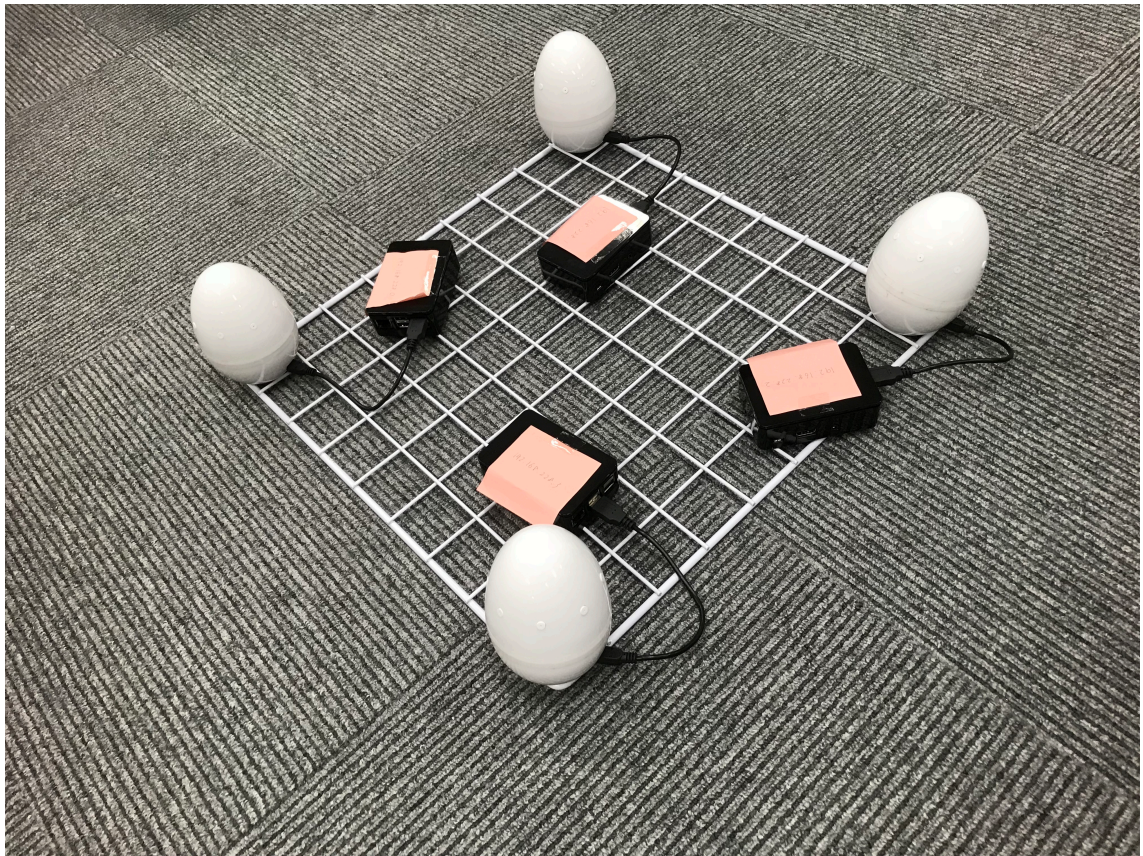


図 16: 音源位置推定システム

実験においては、マイクロホンアレイの相対位置を正確に計測する必要がある。そこで、格子状の金網を用意しマイクロホンアレイの位置を固定した音源位置推定システムを構築した(図 16)。この金網の四隅にマイクロホンアレイを固定した際には、マイクロホンアレイ間の距離は 40 cm となる。実験時には金網を四枚用いることで、マイクロホンアレイ間の距離を 85 cm として位置推定を行った。四台のマイクロホンアレイの座標はそれぞれ (0.85, 0), (0, 0), (0.85, -0.85), (0, -0.85) である(順に node 0、node 1、node 2、node 3 とする)。これらを図示すると、図 17 のようになる。

また、実験においてはマイクロホンアレイの相対方向も事前に計測する必要がある。本実験では、はじめに、キャリブレーションを行うためのオフセットを導出する。座標 (5, 0) に音源を設置し、各マイクロホンアレイで DOA の推定を行う。同時にマイクロホンアレイの座標と音源の座標を元に DOA の真値を求めておき、推定結果との差分をオフセットとして以降の位置推定に用いることとする。

位置推定手法におけるパラメータとしては、表 8 と同様である。音源の設定は DOA 推定実験と同様であり、無線ノードは各音源位置について 30 s の間録音を行う。

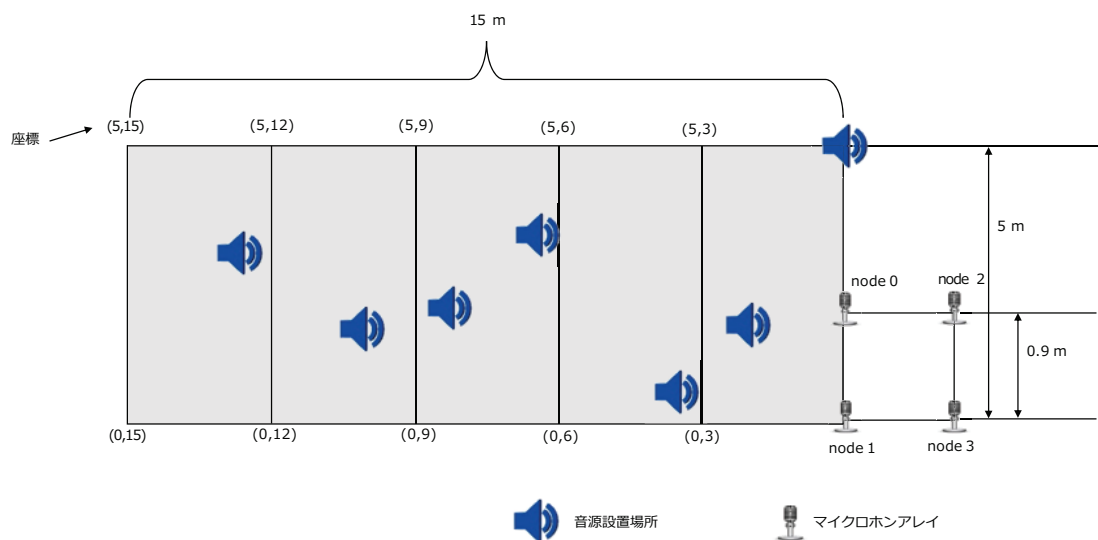


図 17: 位置推定実験における音源とマイクロホンアレイの位置

5.2.2 実験結果

位置推定結果を表 9 に示す。音源位置については座標 (5, 0) を含めた合計 7 箇所で行った。このとき、平均誤差は 0.84 m、最大誤差は 1.87 m、最小誤差は 0.079 m となった。また、30 s の音声の DOA 推定には 4.45 s、位置推定には 0.099 s 要した。位置推定にお

表 9: 位置推定実験結果 (単位はすべて m)

音源座標	マイクロホンアレイの重心 と音源との距離	屋外実験		シミュレーション	
		推定結果	誤差	推定結果	誤差
(5,0)	4.59	(4.975, 0.075)	0.08	(4.975,0.075)	0.08
(2.26, 2.47)	3.43	(1.975, 2.175)	0.41	(2.275,2.475)	0.02
(0.88, 3.13)	3.59	(0.975, 3.225)	0.13	(0.875,3.075)	0.06
(3.33, 6.79)	7.77	(3.225, 6.075)	0.72	(3.325,6.825)	0.04
(1.23, 8.32)	8.79	(1.875, 9.825)	1.63	(1.225,8.325)	0.004
(2.83, 10.24)	10.93	(3.625, 11.925)	1.87	(2.875,10.425)	0.20
(2.89, 12.92)	13.57	(3.725, 13.575)	1.06	(2.925,13.125)	0.21

表 10: 位置推定実験における各マイクロホンアレイの DOA

音源座標	推定結果 (キャリブレーション後)	真値	誤差
(5,0)	($-90^\circ, -90^\circ, -78^\circ, -80^\circ$)	($-90^\circ, -90^\circ, -78^\circ, -80^\circ$)	($0^\circ, 0^\circ, 0^\circ, 0^\circ$)
(2.26, 2.47)	($-29^\circ, -44^\circ, -21^\circ, -29^\circ$)	($-30^\circ, -42^\circ, -23^\circ, -34^\circ$)	($-1^\circ, 2^\circ, -2^\circ, -5^\circ$)
(0.88, 3.13)	($-3^\circ, -19^\circ, -2^\circ, -10^\circ$)	($-1^\circ, -16^\circ, -1^\circ, -13^\circ$)	($2^\circ, 3^\circ, 1^\circ, -3^\circ$)
(3.33, 6.79)	($-21^\circ, -30^\circ, -20^\circ, -22^\circ$)	($-20^\circ, -26^\circ, -18^\circ, -24^\circ$)	($1^\circ, 4^\circ, 2^\circ, -2^\circ$)
(1.23, 8.32)	($-6^\circ, -14^\circ, -6^\circ, -6^\circ$)	($-3^\circ, -8^\circ, -2^\circ, -8^\circ$)	($3^\circ, 6^\circ, 4^\circ, -2^\circ$)
(2.83, 10.24)	($-13^\circ, -19^\circ, -13^\circ, -13^\circ$)	($-11^\circ, -15^\circ, -10^\circ, -14^\circ$)	($2^\circ, 4^\circ, 3^\circ, -1^\circ$)
(2.89, 12.92)	($-12^\circ, -18^\circ, -12^\circ, 11^\circ$)	($-9^\circ, -13^\circ, -8^\circ, -12^\circ$)	($3^\circ, 5^\circ, 4^\circ, -1^\circ$)

る誤差の大きさを評価するために、表には、すべてのマイクロホンアレイで得られた DOA に誤差が含まれない場合のシミュレーション結果も記載している。この結果と比較すると、DOA に誤差がない場合には、誤差は最大でも 0.21 m であるため、DOA 推定における誤差が位置推定誤差に大きく起因していると考えられる。

ここで、各マイクロホンアレイが推定した DOA を表 10 に示す。キャリブレーションを行なった後の (node 0, node 1, node 2, node 3) の推定した DOA を表に記載している。前述の通り、音源が (5,0) の場合の DOA の真値とのオフセットを利用してキャリブレーションを行っており、キャリブレーション後の DOA の誤差は平均で 2.75°、最大では 6° となっている。この DOA の推定誤差によって位置推定に誤差が生じている。5.1 節で示したように、音源とマイクロホンアレイ間との距離が 14 m のとき、誤差 1° 以内での DOA が推定可能である。しかしながら、実験を行った場所は前節と同様であるものの、DOA の推定誤差が増加している。

MUSIC はノイズには強いことが知られているが、一方で反響によって生じる相関波が届くことによって精度が落ちることが知られている。前節の実験と同じ場所ではあるが、音源の具体的な設置場所は異なっているため、この点が誤差の原因に関連していると考えられ、今後検証を進める。本研究ではカエルの生息する野外環境として田圃の周辺を想定しており、比較的見とおしの良い開けた場所であることが多い。しかしながら、近隣に建造物が存在する場合や、棚田となっている場合には反響が生じやすくなると考えられ、DOA 推定において反響への対応は十分に考慮しておく必要がある。近年は反響に対してロバストな DOA 推定手法も提案されており [21]、より実環境に対して有効な手法を用いる必要があると考えている。

本位置推定実験においては、目標とする位置推定精度を、すべての音源位置に対して実現することができなかった。しかしながら、位置推定実験に利用している機器は比較的安価であり、無線ノードの数を増加することは容易であり、無線ノード数の増加による精度向上は現実的な解決方法である。この際、音源数が少数であれば、無線ノード数が増加することによる、位置推定に要する時間の増加はほとんど生じない。なぜなら本研究で構築したシステムでは、MUSIC の実行は無線ノードが行っているため DOA 推定に要する時間の増加は生じず、位置推定サーバには DOA の推定結果のみを送信しているため、通信に要するオーバーヘッドもほとんど増加しない。また位置推定手法に要する計算時間は DOA 推定よりも非常に小さく、マイク数の増加に対しても計算量の増加は少ない。ただし、音源数が大きくなると計算量は指数関数的に増加するため、実行時間によってはより計算量の少ない方法を検討する必要がある。現在は推定に要する目標時間を 5 分としているが、生物の観測や生態の調査といった応用を考えると、位置推定はリアルタイムに行えることが望ましい。本実験では、単一音源に関しては、位置推定に用いた録音音声の長さよりも短い時間で位置推定を

行えることを示した。現状は一定時間の録音の後に DOA の推定を逐次的に行っているが、この点を並列化することでよりリアルタイム性を向上できると考えている。今後は、単一音源に対する位置推定精度を向上しつつ、リアルタイム性を備えた音源位置推定システムの構築を行うことを目指す。

6 おわりに

生物同士の音声コミュニケーションは社会性を生み出す重要な要素である。本報告では、野外環境下における生物の音声コミュニケーション解析のための、音源位置推定システムの実装を行った。実装したシステムは、無線センサーノードとマイクロホンアレイを接続することで、観測情報を用意に収集することが可能であり、また、敷設や携行が容易である利点を持つ。位置推定手法には音声到来方向を用いる手法を用いており、無線センサーノードが音声到来方向を推定し、位置推定用のサーバが無線通信によって届けられた音声到来方向情報を用いて音源位置を推定する。

実装したシステムを用いた実験によって、周囲に障害物の存在しない環境では、最小 8 cm、平均 84 cm の誤差で音源位置の推定が可能であることを示した。位置推定には 30 s の録音音声を用いており、音声到来方向の推定には 4.45 s、推定した音声到来方向を用いた音源位置推定には 0.099 s 要した。本報告では観測対象の生物としてニホンアマガエルを想定しており、その生態から、推定誤差 50 cm 以下、推定時間 5 分以下を目標と設定した。推定時間の点では目標を満たしているものの、推定精度の点では目標を満たしていない結果も得られた。音源座標を元に計算して得られる、正確な DOA が得られた場合には、実装した位置推定手法を用いた場合に目標精度を満足することが確認できたため、DOA の推定誤差を抑えることで位置推定の精度を向上することが可能である。

今後の課題として、位置推定精度の改善が第一であり、そのために、DOA 推定誤差の改善を行う。位置推定精度が十分改善された後に、実際の屋外環境での生物の位置推定実験を行う。また、ニホンアマガエルは鳴いている間には移動しないため、推定時間の目標を 5 分としたが、音声コミュニケーションを行う他の生物には、例えば鳥のように音を発しながら移動するものも存在する。実装したシステムの応用性の向上のために、位置推定のリアルタイム化にも取り組む。また本研究では単一音源の場合のみを扱っているが、実際の屋外環境では複数のカエルが鳴いていることが考えられる。実装した音源位置推定システムでは複数音源にも対応可能であると考えているが、音源を複数設置した屋外での実験などによる評価を行い、複数音源の場合でも位置推定が可能であることを示していく。

謝辞

本報告を終えるにあたり、日頃よりの確なご指導、ご教授を頂きました大阪大学大学院情報科学研究科の村田正幸教授に心より感謝申し上げます。ならびに本報告の作成に丁寧なご指導をいただきました大阪大学大学院経済学研究科の小南大智助教に深く感謝いたします。カエルの生態や実験機器などに様々なご助言をいただき、また、屋外での調査にご協力をいただいた、筑波大学システム情報工学研究科の合原一究助教に厚く御礼申し上げます。平素から適切なお助言やご指導をいただきました大阪大学大学院情報科学研究科の荒川伸一准教授、大下裕一助教、久世尚美特任助教に心より厚く御礼申し上げます。最後に、日頃より様々な点でお世話になった、村上雅哉氏、志垣沙衣子氏、屋外での実験に協力してくれた松田拓己氏をはじめとする村田研究室の皆様、友人、家族のご厚意に感謝の意を表して謝辞といたします。

参考文献

- [1] A. Mutazono, M. Sugano, and M. Murata, “Energy Efficient Self-Organizing Control for Wireless Sensor Networks Inspired by Calling Behavior of Frogs,” *Computer Communications*, vol. 35, pp. 661–669, Mar. 2012.
- [2] Z. Zhang, K. Long, J. Wang, and F. Dressler, “On Swarm Intelligence Inspired Self-Organized Networking: Its Bionic Mechanisms, Designing Principles and Optimization Approaches,” *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 513–537, 1st Quarter 2014.
- [3] K. Wells, *The Ecology and Behavior of Amphibians*. University of Chicago Press, Feb. 2010.
- [4] 合原 一究, 小南 大智, 平野 康晴, 村田 正幸, “カエルの合唱法則に基づく無線センサネットワーク制御手法の提案,” 電子情報通信学会技術研究報告 (IN2016-155) , vol. 116, pp. 347–352, Mar. 2017.
- [5] M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee, “A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks,” *Wireless Communications and Mobile Computing*, vol. 2017, Aug. 2017.
- [6] C. Knapp and G. Carter, “The Generalized Correlation Method for Estimation of Time Delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, pp. 320–327, Aug. 1976.
- [7] 細川 侑嗣, 小南 大智, 合原 一究, 村田 正幸, “無線ネットワークにより相互接続したマイクロホンアレイを用いた音源位置推定手法の実装,” 電子情報通信学会技術研究報告 (ASN2017-87) , vol. 117, pp. 23–28, Jan. 2018.
- [8] I. Aihara, T. Mizumoto, T. Otsuka, H. Awano, K. Nagira, H. G. Okuno, and K. Aihara, “Spatio-Temporal Dynamics in Collective Frog Choruses Examined by Mathematical Modeling and Field Observations,” *Scientific reports*, vol. 4, Jan. 2014.
- [9] R. Schmidt, “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276–280, Mar. 1986.

- [10] B. D. Rao and K. S. Hari, “Performance analysis of root-MUSIC,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 1939–1949, Dec. 1989.
- [11] R. Roy and T. Kailath, “ESPRIT-estimation of signal parameters via rotational invariance techniques,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 984–995, July 1989.
- [12] R. Hedley, Y. Huang, and K. Yao, “Direction-of-arrival estimation of animal vocalizations for monitoring animal behavior and improving estimates of abundance,” *Avian Conservation and Ecology*, vol. 12, June 2017.
- [13] L. M. Kaplan, Q. Le, and N. Molnar, “Maximum likelihood methods for bearings-only target localization,” in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 3001–3004, May 2001.
- [14] A. Griffin, A. Alexandridis, D. Pavlidi, and A. Mouchtaris, “Real-Time Localization of Multiple Audio Sources in a Wireless Acoustic Sensor Network,” in *Proceedings of European Signal Processing Conference (EUSIPCO)*, pp. 306–3100, Sept. 2014.
- [15] A. Alexandridis and A. Mouchtaris, “Multiple sound source location estimation and counting in a wireless acoustic sensor network,” in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 1–5, Oct. 2015.
- [16] A. Alexandridis and A. Mouchtaris, “Multiple sound source location estimation in wireless acoustic sensor networks using doa estimates: The data-association problem,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 2, pp. 342–356, 2018.
- [17] “HARK SaaS.” <https://api.hark.jp/web/login>. available at Feb. 2018.
- [18] R. Hamming and C. Yuen, “Digital filters,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 10, pp. 689–689, Oct. 1980.
- [19] K. Nakamura, K. Nakadai, and H. G. Okuno, “A Real-Time Super-Resolution Robot Audition System that Improves the Robustness of Simultaneous Speech Recognition,” *Advanced Robotics*, vol. 27, pp. 933–945, May 2013.
- [20] Y. Hirano, “Collecting and analyzing chorus of japanese tree frogs and its application to energy-efficient control of wireless sensor networks,” Master’s thesis, Osaka University, 2017.

- [21] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, and H. Li, “A learning-based approach to direction of arrival estimation in noisy and reverberant environments,” in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2814–2818, Apr. 2015.