

少数マイクロフォンを用いる環境雑音を考慮した  
音源の特徴抽出及び位置推定に関する研究

2017年 3月

川岸 卓司

少数マイクロフォンを用いる環境雑音を考慮した  
音源の特徴抽出及び位置推定に関する研究

川岸 卓司

システム情報工学研究科  
筑波大学

2017年 3月

## 概要

近年、コンピュータの処理能力が飛躍的に伸びたことにより従来から用いられていた音響信号の高速かつ複雑な解析が可能となり、超音波などを用いた計測や携帯端末での音の認識など様々な環境で利用されている。一方、音声認識と比べ橋梁の打音検査や漏水検査、身体検査時の聴診から機械の異常音など人が聴くことによってその状態を判断し、音の認識や検出をしており、コンピュータでの全自動的な判断には至っていない。本論文は、雑音を有する様々な環境下で収録された見本となる音を元に少数のマイクロフォンを用いて特定の音の検出及びその音源の位置推定を行うものである。

音の検出のためには雑音の影響の少ない見本となる音を大量に用意する必要があり、用意が難しい音や見本となる音に雑音を含む場合への有効的な手法は提案されていない。そこで本研究では単一のマイクロフォンを用い、畜産業のひとつである養豚豚舎内で豚が呼吸器感染症になった際に生じるくしゃみ音の検出方法について検討を行った。豚舎内は、常に排気ファンやヒータ、冷房などの音に加えてくしゃみ音とは異なる豚の鳴き声や柵などに当たる活動騒音、餌作業や清掃作業などに伴う作業音など多くの雑音が発生している環境であり、対象となるくしゃみ音は検出のための見本の音となるデータは存在しておらず雑音環境から収集する必要がある。また、このくしゃみ音は常時発生していることがなく収集が難しい上に、従来の主要な音声認識法である大量学習によるモデル化ができない。そこで、このような雑音環境で発生回数が少ない音に対する音源の特徴抽出法及びその特徴を用いる検出法について検討を行った。また特徴抽出法の汎用性を評価するため異なる環境で収録された音に対し実験を行い提案法の評価をした。

特定の音を検出する手法として、くしゃみ音に含まれる 30 kHz 以上にもなる高周波成分の特徴や周波数毎の継続時間から 5 つのテンプレートでくしゃみ音を検出し、誤検出の少なさを示す識別率は 63%、見逃しにくさを示す感度は 85% となり先行研究における感染症の陰陽性の識別に有効であることが示された。一方、他の豚房で本手法の有効性を検証したところ識別率及び感度は共に 10% 程度となり、汎用性に問題があることがわかった。そこでそれぞれの豚房で収録できた少数のくしゃみ音を時間と周波数軸上の音圧レベルで示すスペクトログラムから計算される相乗平均で特徴を抽出し、見本となる音のデータとして用いる手法を提案し、定常雑音の低減や周波数の時間方向の変動を特徴とする手法を提案した。本手法によって、2 つの異なる豚舎においても識別率 54% 以上、感度 93% 以上となり、音による豚呼吸器感染症の発見法として提案手法の有効性が示された。

次に検出した音の発生位置を推定する手法について検討した。従来、マイクロフォンを用い音の発生位置を求めるためには大量のマイクロフォンを同時に収録する必要があった。本研究では上記の検出対象であるくしゃみ音に限らず、様々な環境で音の発生位置を推定する

ため、3つの音源を対象に少数マイクロフォンでの位置推定の検討を行った。豚のくしゃみ音に対しマイクロフォンとカメラを組み合わせた位置推定法や、カメラや電磁波の届かない土中のドリル先端位置を模した音源を検出対象とし、土中の音源が地表面での反射音する特徴を用い位置推定精度を向上させる方法について検討した。また屋外で飛行するドローンの飛行音から位置推定を行う手法について検討を行った。

音の発生位置の特定のため複数のマイクロフォンによって音の到達時間差が生じる原理を用い、4つまたは8つの少数のマイクロフォンによって音源の位置を推定する手法について検討をした。豚のくしゃみ音に対してカメラとマイクロフォンを組み合わせ角度推定から位置を推定する手法と従来法によるマイクロフォンのみでの位置推定についてその推定精度を検証し、カメラとマイクロフォンによる位置推定では50%、マイクロフォンのみでは20%程度で発生位置を特定した。また、土中にマイクロフォンを差し込むジオフォンアレイを作成し、地表面で反射する音波を用いる手法によって少数のマイクロフォンでも位置推定精度を向上させる方法について検討し、実際の土中において提案手法の評価を行った。その結果、地表面に置く従来方法に対し、提案手法は3.4倍の精度向上が期待できることがわかった。また、音の到来時間差の発見が難しい連続音に対して立体構成マイクロフォンと音の類似度によって重みづけされた最適化計算による手法によって空中のドローンに対する音源位置推定法を検討した。実際のドローンへ本手法を適応したところ、従来正しい位置を推定できなかった屋外で他の雑音がある中で約6m離れたドローンに対し0.5m程度で位置推定できることがわかった。また類似度を用いない手法に比べ3.2倍の位置推定精度向上が可能であることが示された。

本研究によって従来は多くの見本となるデータや統計的な解析が必要であった特定の音を検出する手法に対し、雑音環境下かつ発生回数が少ない特定の音に対する特徴の抽出及びその検出法を提案することができた。本手法は雑音の排除が難しい環境での収録や、単一のマイクロフォンによって音源との距離が数mの比較的遠い環境でも有効であることから様々な環境での特定の音の検出法として応用が期待される。さらに、その発生位置を少数のマイクロフォンでも高精度に特定が可能であることから、特定の音の検出とその音源位置の情報によって今後の音を使ったセンシング技術に大きく貢献できると考えられる。

# 目次

第1章	序論	1
1.1	研究背景	1
1.1.1	音の研究と適応例	1
1.1.2	音の認識と問題点	2
1.1.3	感染症早期発見のための豚くしゃみ音検出	4
1.2	研究目的及び論文構成	8
第2章	音源の特徴抽出	11
2.1	はじめに	11
2.2	音源の特徴抽出の現状	12
2.3	周波数特徴と音源継続時間に着目する検出法	14
2.3.1	見本データの収集と解析	15
2.3.2	検出実験	17
2.3.3	実験結果と考察	19
2.4	まとめ	21
第3章	異なる雑音環境下における特定音源の検出	23
3.1	はじめに	23
3.2	少数の見本データを用いる音源特徴の抽出と検出法	26
3.2.1	高域強調フィルタと定常雑音の低減法	26
3.2.2	スペクトログラム相乗平均による特徴抽出と検出法	30
3.2.3	異なる雑音環境下における検出実験と結果	33
3.2.4	まとめ	36
3.3	特定音の検出精度向上の手法と実験	37
3.3.1	DP マッチングと動的特徴量を用いる検出法の原理	37
3.3.2	提案手法による検出実験と結果	45
3.3.3	まとめ	48
3.4	まとめ	49

第4章	少数マイクロフォンによる音源位置推定	51
4.1	はじめに	51
4.2	原理	52
4.2.1	相互相関法による到来時間差の計算方法	52
4.2.2	到来時間差からの方向・位置の計算原理	53
4.3	カメラとマイクロフォンを用いる豚舎内の位置推定	55
4.3.1	カメラとマイクロフォンによる位置推定の計算手法	55
4.3.2	豚舎での実験条件	55
4.3.3	実験結果と考察	55
4.3.4	まとめ	57
4.4	反射音を用いる土中の位置推定	58
4.4.1	少数センサによる土中位置推定の背景	58
4.4.2	反射音を用いる位置の計算手法	59
4.4.3	土中における実験条件	64
4.4.4	結果と考察	66
4.4.5	まとめ	67
4.5	立体構成マイクロフォンによる位置推定	68
4.5.1	マイクロフォンを用いるドローン位置推定の背景	68
4.5.2	最適化手法を用いる連続音の位置の計算手法	69
4.5.3	ドローンに対する実験条件	73
4.5.4	結果と考察	74
4.5.5	まとめ	76
4.6	まとめ	76
第5章	結論	77
	謝辞	79
	参考文献	81
	付録A	91
	付録B	95

# 目次

Fig. 1.1	音の認識の基本原理	2
Fig. 1.2	音の特定度と時間的参照範囲	3
Fig. 1.3	センサ距離と雑音による問題点	3
Fig. 1.4	豚の咳・くしゃみ音検出システムの概要	6
Fig. 1.5	システム費用とシステム有無による損益比較	7
Fig. 1.6	システム有無による対施設工事費の減価償却比較	7
Fig. 1.7	論文構成	9
Fig. 2.1	2215 RPMにおけるスクリュ音, (a) 収録信号, (b) 周波数及び(c) ケプストラム	13
Fig. 2.2	識別率と感度の定義	15
Fig. 2.3	畜産草地研究所における実験機材の構成	16
Fig. 2.4	豚舎内発生音のスペクトログラムと周波数及び時間特徴	17
Fig. 2.5	周波数特徴と時間特徴を用いる検出法のフローチャート	18
Fig. 2.6	テンプレート数と識別率・感度の関係	20
Fig. 3.1	全農飼料中央研究所における実験機材の構成	23
Fig. 3.2	異なる場所におけるくしゃみ音のスペクトログラム	25
Fig. 3.3	冒頭 25 msec におけるくしゃみ音の振幅スペクトルの比較	26
Fig. 3.4	高域強調フィルタの特性	27
Fig. 3.5	高域強調フィルタ前後のくしゃみ音のスペクトログラム	28
Fig. 3.6	高域強調フィルタ前後のスペクトルサブトラクション法を適応したくしゃみ音のスペクトログラム	28
Fig. 3.7	音声信号と SNR 0 dB でのスペクトログラム	31
Fig. 3.8	SNR 0 dB 及び SNR 3 dB でのスペクトログラム相乗平均による結果	32
Fig. 3.9	くしゃみ音の見本データの作成に用いるフローチャート	32
Fig. 3.10	14 個のくしゃみ音によるスペクトログラム相乗平均による特徴抽出結果	33
Fig. 3.11	検出法のフローチャート	34
Fig. 3.12	誤識別の原因となった豚が金属柵にぶつかった音のスペクトログラム	36

Fig. 3.13	音の静的特徴と動的特徴	38
Fig. 3.14	くしゃみ音と金属柵音の動的特徴	39
Fig. 3.15	DP マッチングでの見本データと受信信号の類似度	40
Fig. 3.16	DP マッチングの入場ペナルティと移動コスト	40
Fig. 3.17	DP マッチングの最短経路の計算結果	41
Fig. 3.18	音声とサイン波の DP マッチングの結果	43
Fig. 3.19	音声と雑音付加音声の DP マッチングの結果	43
Fig. 3.20	音声と時間変化のある雑音付加音声の DP マッチングの結果	44
Fig. 3.21	検出法のフローチャート	46
Fig. 4.1	到来時間差と音源方向	53
Fig. 4.2	到来時間差と音源位置推定	54
Fig. 4.3	4つのマイクロフォンとカメラによる位置推定	54
Fig. 4.4	カメラにおける豚舎内における方向及び位置推定結果の例	56
Fig. 4.5	誘導式水平ドリル工法とロケータによる位置推定の概要	58
Fig. 4.6	地表面反射による音波伝搬経路の例	59
Fig. 4.7	反射波を含む受信信号の例	60
Fig. 4.8	反射波を含む相互相関波形	61
Fig. 4.9	2つのジオフォンアレイによる音波伝搬経路	62
Fig. 4.10	音速の測定と実験概要	65
Fig. 4.11	提案手法を用いるドローン位置推定の応用例	68
Fig. 4.12	実験で使用した AR. Dorone 2.0	69
Fig. 4.13	ドローン音の無響室及び屋外の周波数スペクトログラム	69
Fig. 4.14	環境音とドローン音比較の実験概要	70
Fig. 4.15	ドローン音の到来時間差の計算結果	71
Fig. 4.16	環境音とドローン音の騒音レベルの比較	71
Fig. 4.17	到来時間差の原理とマイクロフォンアレイ計測	71
Fig. 4.18	ドローンの位置推定のための実験構成図	73

# 表目次

Table 2.1	実験条件 .....	16
Table 2.2	くしゃみ音 5 つを用いる検出結果 .....	20
Table 3.1	実験条件 .....	24
Table 3.2	畜産草地研究所のデータに対する実験結果 .....	35
Table 3.3	全農飼料中央研究所のデータに対する実験結果 .....	35
Table 3.4	類似度の計算結果 .....	44
Table 3.5	畜産草地研究所のデータに対する実験結果 .....	47
Table 3.6	全農飼料中央研究所のデータに対する実験結果 .....	47
Table 4.1	豚舎内の方向及び位置推定結果 .....	56
Table 4.2	土中の位置推定結果の比較 .....	66
Table 4.3	2 組のマイクロフォン毎の計測結果.....	74
Table 4.4	すべてのマイクロフォンの組み合わせによる計算結果.....	75

# 第1章 序 論

## 1.1 研究背景

### 1.1.1 音の研究と適応例

本研究は、少数のマイクロフォンによって様々な環境下で特定の音源特徴を抽出し、その音源特徴を用い対象とする音の検出やその位置の推定を行うものである。近年、コンピュータの処理能力が飛躍的に伸びたことにより、従来用いられていた音響信号をより高速かつ様々な解析が可能となり、計測や音の認識などを様々な環境で用いる研究がなされている<sup>[1]</sup>。音波は電磁波や光に比べ伝搬速度が遅いことから、到達時間差などを計測する場合でも高時間分解能を有する計測機器を必要とせずに測定することが可能である。また、光や電磁波では土中や水中の減衰が大きく利用が難しい媒質でも良好な透過性を持つ音響信号が用いられている<sup>[2-4]</sup>。例えば、水中では音響信号による海洋ソナーや通信、位置探査などが研究されており<sup>[5,6]</sup>、土中では埋設物のソナーやトンネル工事の自己位置探査、位置探査などが研究されている<sup>[7,8]</sup>。空気中では人の声の分析<sup>[9-14]</sup>、異音検知やその音源定位<sup>[15-21]</sup>、音波伝搬を利用する温度計測など<sup>[22,23]</sup>多岐に渡る。この際、送受信器を利用し、送信信号と受信信号を比較する事により、その伝搬特性、温度や対象物の位置を探査するアクティブな手法や、受信器のみを利用するパッシブな手法がある。通信や探査などのアクティブな手法では、発生させる音響信号を計測者が任意に操作できる一方、送信のタイミングや信号の種類を設定し、環境に合わせチューニングをする必要があり、送受信器を用意するため機材が大型化する傾向がある。受信機のみを用いるパッシブな手法は、音源で発生する音やセンサと音源の間に含まれる環境情報などを得ることができる。一方で対象とする音源の数、種類や音圧、反射音や他の雑音のレベルなどを考慮しセンサと対象の距離や配置を検討するとともに、目的とする音の特徴を適切に用いなければならない。本研究ではこの受信機のみを用いるパッシブな手法の一つであるマイクロフォンを用い、様々な環境下において目的とする音の特徴抽出及び検出方法と発生位置の推定方法について述べる。本章ではマイクロフォンで音を収集し、その音の特徴から目的の音を検出して出力する最も身近で発展している音声認識による機械への入力や未だ人の手で行っている音を聞いて状態を判断する例を基に、先に挙げたマイクロフォンのみを用いる手法の問題点や、本研究が対象とした実験環境や音源について述べる。

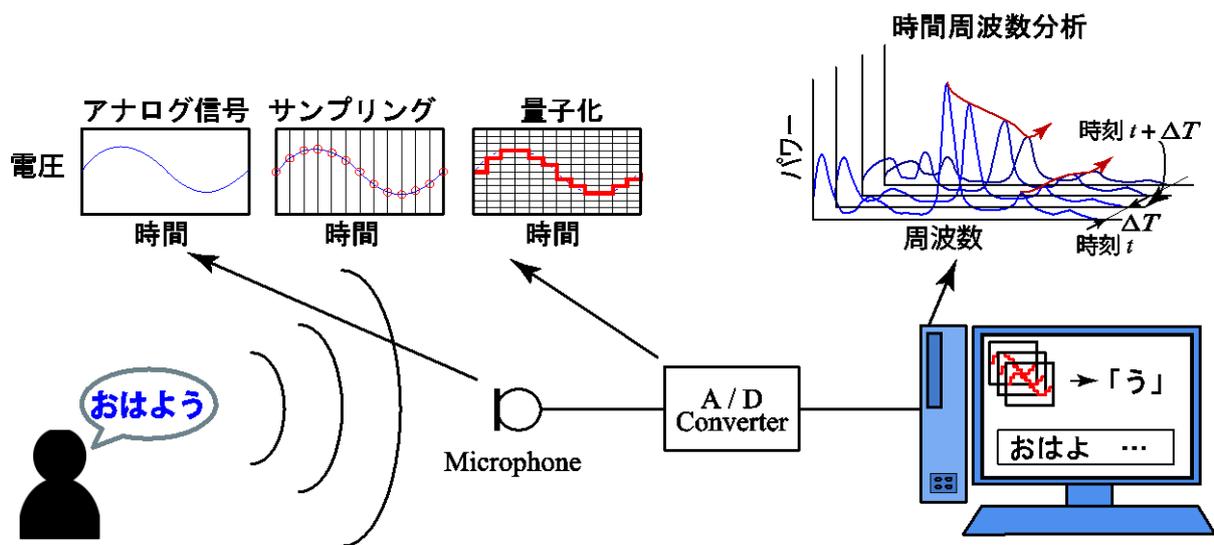


Fig. 1.1 音の認識の基本原理

### 1.1.2 音の認識と問題点

人が声を用いて機械へ入力する場合には音声認識と言語認識の2つの処理が必要となる。**Figure 1.1**のように図部屋で「おはよう」と話した際の処理を例にすると、人から発せられた言葉は、空間中を伝わりその音圧がマイクロフォンによって電気信号に変換される。この電気信号はアナログディジタルコンバータ(AD-Converter)によってディジタル信号に変換されコンピュータへ入力される。この際、アナログ信号をディジタル信号へ変換するため単位時間あたりどれくらいの計測点を記録するかを定める縦軸に当たるサンプリングと、電圧値をどれくらいの細かさで記録するかを定める横軸の細かさに当たる量子化を設定する必要がある。さらに、この信号を短時間毎に分割し短時間フーリエ変換によって周波数領域に変換し、単位時間あたりの特徴的な周波数やその時間変動を計算し、コンピュータに事前に用意された見本データにもっとも近い特徴を持った「う」を選択することとなる。この短時間に分割した信号をシフトさせていくことで「お」「は」「よ」「う」と認識される。以上が音声認識の原理であるが、実際には人の言語や年齢や声、なまり、性別などに対応する必要がある膨大な見本となるデータからこれらを統計的に処理し類推することで実現している。このため携帯端末などでは情報を分割してサーバ上で上記の処理を行っている。2つ目の処理である言語認識は、つぎはぎの文字列から「おはよう」が挨拶で最初に用いる言葉と認識することで、長い発言の際の異音同義語に対応することができ、人の音声はマイクロフォンを通して機械へ入力することができる<sup>[24,25]</sup>。このように実用可能となった音声認識の背景には膨大な見本となるデータをデータベースで整理し、統計的なモデルを確立できたことが大きい<sup>[26]</sup>。つまり、音声認識の手法を用いる場合には統計処理が可能な程度の見本となるデータが必要である。

その一方で人が音を聞くことで機械や身体など様々な状態を予測できることから、これをコンピュータに実施させる方法が検討されている。例えば、身体検査での医者の聴診や、打音検査、漏水検査、回転機械の異常なども音声認識がこれだけ進んでいるにもかかわらず、

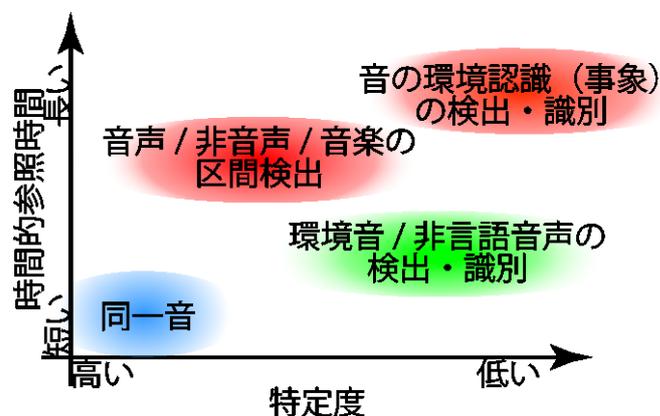


Fig. 1.2 音の特定度と時間的参照範囲



Fig. 1.3 センサ距離と雑音による問題点

未だに人が音を聞くことによって検査をしている。これは、状態を予測するための学習データを集めるのが難しいことに加え適切な音源特徴，検出法が確立されていないことに基因している。また人も正常な状態の学習データを持っている場合に判別可能であるが素人が身体を聴診しても異常がわからないのと同様にこの情報が少ない場合には検査できない。これらをコンピュータに実施させるには大量の学習データに加え，音声認識のように対象の特徴を分析し類似度を比較できる手法を検討する必要がある<sup>[15-21]</sup>。

**Figure 1.2** はどのような音かを機械によって特定する際の特定度の高低について 2014 年の論文<sup>[27]</sup>を基に著者が作成したものである。横軸に音を入力したときにその音が何であるかを特定する特定度，縦軸にどれくらいの時間を分析のために参照するかを示したものである。最も特定度が高く時間的参照時間が短いのは同一音であり，ブザーやクラクションなどの機械音が例に挙げられる。これらは周波数上の特徴など比較的簡単に特定できるとされる。最も特定が難しく参照時間が長いのが音による環境認識とされており，その音が交差点であるのか会議室での議事なのか，鉄道のアナウンスであるのかなどを特定することとされる。

## 1.1 研究背景

本研究では時間的参照範囲が短いと特定が難しいとされる環境音や非言語音声の検出・識別分野として豚のくしゃみ音やドローンのプロペラ飛行音などの音源を研究対象とした。

音源の特徴から分析や特定をする際には、音源と計測するセンサの距離にも留意する必要がある。人が聞き判断する場合であれば近づいたり離れたり、自分が聞きやすく判断しやすい距離に調整可能であるが、コンピュータや決まった位置にセンサがある場合は見本となるデータの収録時の距離と同じ距離でないと検出できないことが多く、**Fig. 1.3**のように距離の変動が検出に大きく影響することがわかっている<sup>[28]</sup>。これは距離によって周波数の減衰がある上に、途中の雑音がセンサに入ることが原因となっている。そこでこれらの問題に対処するため、距離の増減に対応する手法やセンサを複数設置することが検討されている。距離の増減をモデル化して対応する手法は、無響室や会議室などの一定の環境で検出率を向上させることがわかっているが、それぞれの環境モデルを構築する必要がありモデルのために雑音がない環境で事前計測し音の反響などをフィルターとして用いる必要があり様々な環境への応用が難しい<sup>[29]</sup>。マイクロフォンを複数構成し、同時に用いるものをマイクロフォンアレイと呼び特定の方向や位置の音源を強調し、他の雑音を低減することができるためこれらの問題を解決する手法として注目されている。一方でマイクロフォンの数だけ同期させ信号処理が必要のため、数を増やすほど音源強調や雑音低減が可能な一方で高性能な演算機が必要となりリアルタイム性が求められる処理に向かない特性がある。このように、対象とする音源の検出やその発生位置を少数のマイクロフォンで実現できれば、様々な環境で音による監視などに応用可能であると考えられる。

### 1.1.3 感染症早期発見のための豚くしゃみ音検出

現在日本国内で飼育される豚は約 980 万頭であり、その頭数は年々増加している。しかし、養豚農家の戸数は減少の一途をたどり、農家一戸当たりの飼育頭数は約 1,800 頭となり負担が増加しており<sup>[30]</sup>、一人当たりの豚の飼育数は約 450 頭になり飼育管理の効率化が求められている<sup>[31]</sup>。さらに、近年では、集中管理による飼育密度の上昇に伴い、各種の感染症発生時の被害リスクが懸念されている。特に豚呼吸器感染症は伝搬力が強く繁殖障害を引き起こすため、養豚においては生産を阻害する大きな要因であることが報告されている<sup>[32,33]</sup>。このため、現在は感染症を予防・発見する手法として各種抗体検査や RFID (Radio Frequency Identification) 等の無線通信による非接触認識技術を利用した管理の研究が実施されている<sup>[34]</sup>。しかしながら各種抗体検査では、一頭一頭の血液等の採取や分析には時間や費用がかかり、作業員の少ない環境では頻繁に行うことが困難である。RFID での手法は、カードやタグ状の媒体に電波によって情報の読み書きが行うことができ、豚の状態などを個別で管理することが可能で一方で、送受信距離が約 300 mm 程度であり、送受信場所の確保や莫大な設備投資が必要である。さらに、RFID を家畜それぞれに埋め込むため、出荷サイクルの短い豚は管理維持費用の点から、すべての豚舎での導入は難しいと考えられる。

そこで感染症時に発するくしゃみ音や咳の発生回数を測定することにより、呼吸器感染症の発症を発見する方法が報告されている<sup>[35]</sup>。これによると、くしゃみに関して肥育前期の豚

は呼吸器感染症の陽性・陰性の間で約 8 回/10 min (100 頭当たり) の違いがあることがわかっている。この手法は作業員が発生回数を測定するだけで陰陽性の識別を行うため、時間や費用がかからない一方、数人で飼育を行っている現在の養豚農家では手間がかかり継続的な実施が困難である。

そこでくしゃみ音を自動で検出し、どの場所で発生したものかを推定しある豚房の豚が呼吸器感染症の陽性反応がでた場合に速やかに養豚農家の作業員に知らせるシステムを構築することで、少人数でも効率的に生産管理が行えることとなる。この様に豚舎の発生音に着目した研究では、豚の咳音に着目しリアルタイムに豚舎内で発生する咳音を検出する手法や咳音が豚舎内のどの付近で発生しているかを判別するシステムを提案されている<sup>[36]</sup>。しかしながら、その測定精度は豚房 1 つ程度であり、個体の特定には至っていない。豚房での飼育密度は離乳期で約 0.28 頭/m<sup>2</sup>、肥育前期で約 0.55 頭/m<sup>2</sup>、飼育後期で約 0.75 頭/m<sup>2</sup>が適正であるとされており、一般的な豚房で約 12~20 頭が飼育されている<sup>[37]</sup>。このため、くしゃみ音が発生した豚房が特定できた場合においても、豚房の豚すべてを再度調べる必要がある。そこで本研究ではマイクロフォンを用いたくしゃみ音の検出を行い陰陽性の識別とその発生位置の特定を行うことで、音による感染症の早期発見手法を検討する。

ここで **Fig. 1.4** のような音によるくしゃみ音検出の装置がもたらす効果について試算する。2015 年時点で日本国内で出荷された豚 1 頭当たりの資本利子・地代全額算入生産費 (以下生産費という。) は 33,828 円となり前年比マイナス 2.6%程度で推移している。このうち、65%に当たる 22,177 円が豚の食事の費用となる飼料費、次いで 12%の 4,062 円が労働費、4%が獣医や医薬品代であり 2,125 円と推計されている<sup>[38]</sup>。一方感染症などで肥育中に死亡してしまった場合に生産費は 0 円となり、それまでにかかった飼料費、労働費などがマイナスとして計上される。さらに、死亡に至らない場合も多く、母豚の繁殖障害や肥育豚の食欲減衰による影響などが確認されており、その被害額は日本国内において 280 億/円と試算されている<sup>[39]</sup>。このような影響を世界中で与えている感染症に豚繁殖・呼吸障害症候群(porcine reproductive and respiratory syndrome: PRRS)が挙げられている。これはくしゃみや咳による空気伝搬によって被害拡大し多大な影響を与えること国際獣医疫事務局の指定疾病及び国内での届出伝染病に指定され、今現在でも経済被害を与え続けている<sup>[40]</sup>。

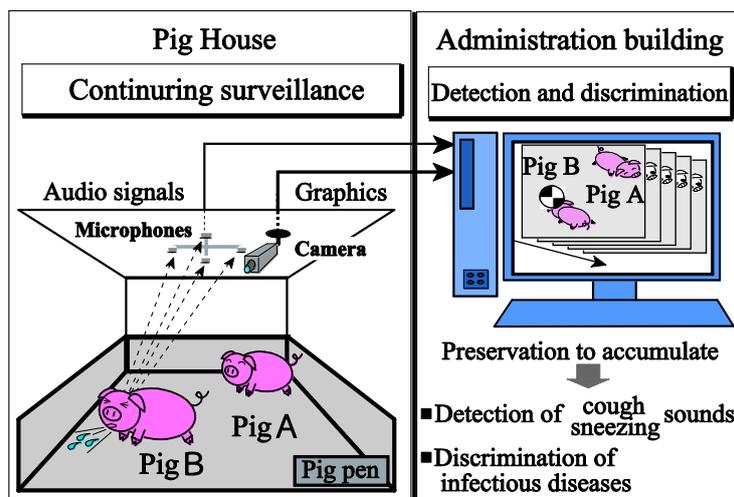


Fig. 1.4 豚の咳・くしゃみ音検出システムの概要

そこである豚舎において音響的手法によって感染症が早期発見できた場合の費用対効果について検討する．算出条件は政府統計<sup>[41]</sup>及び養豚経営に関する記事<sup>[42]</sup>，論文<sup>[38,40,43,44]</sup>から以下のように設定した．一戸 1,500 匹/年を出荷する養豚農家であり，出荷対象となる肥育豚を基本に計算する，なお母豚は 100 頭とし統計上の日本の平均的な養豚農家とした．肥育豚舎は 1 基あたり約 500 匹で 3 基保有している．このような肥育豚舎の設備設立には 1 基あたり平均 2760 万円の初期投資が必要となる．ここでマイクロフォンを各豚房(20 匹)設置したとすると，1 基あたり 25 個のマイクロフォンが必要となる．現状の実験に使用している計測システムでは 4 個のマイクロフォンから A/D コンバータ，10 万円程度の PC1 台でリアルタイム処理が可能であり，マイクロフォン 4 個あたり 20 万程度となることから，豚舎 1 基あたり 140-200 万程度の初期投資（計算では 200 万）とマイクロフォンの交換などを含め年間 20 万円程度のメンテナンス費用と試算した．なお，損害の推計については PRRS 発生時の事故率<sup>[44]</sup>を基に，感染症が発病した場合としなかった場合を含め出荷までに育成されずに死亡した割合及び体重減少した場合の損失算定値を用いた．

Figure 1.5 にシステム費用とシステムの有無による事故率から損益を比較した．シミュレーションでは PRRS 発生時に起きる体重減少が肥育豚の 50%，繁殖障害による肥育豚の死亡率が 25%として 15 年間のうち 5 年目，10 年目に発生したと仮定した．システム無しの場合には論文による損害<sup>[44]</sup>をそのまま適応し，システム有の場合には，豚房 1 つ以内で発見し対処（20 匹のみ:全数の 1.3%）したとして計算した結果を示す．まずシステム無しの赤線をみるとシステム費用がかからない一方で，被害が起きた際には損益が発生する．一度でも死亡事案が発生した場合にシステム費用を上回る損益となるが，体重減少のみの場合では被害額は 15 年運用したシステム費に満たないことがわかった．一方で緑線に■マークを見ると，

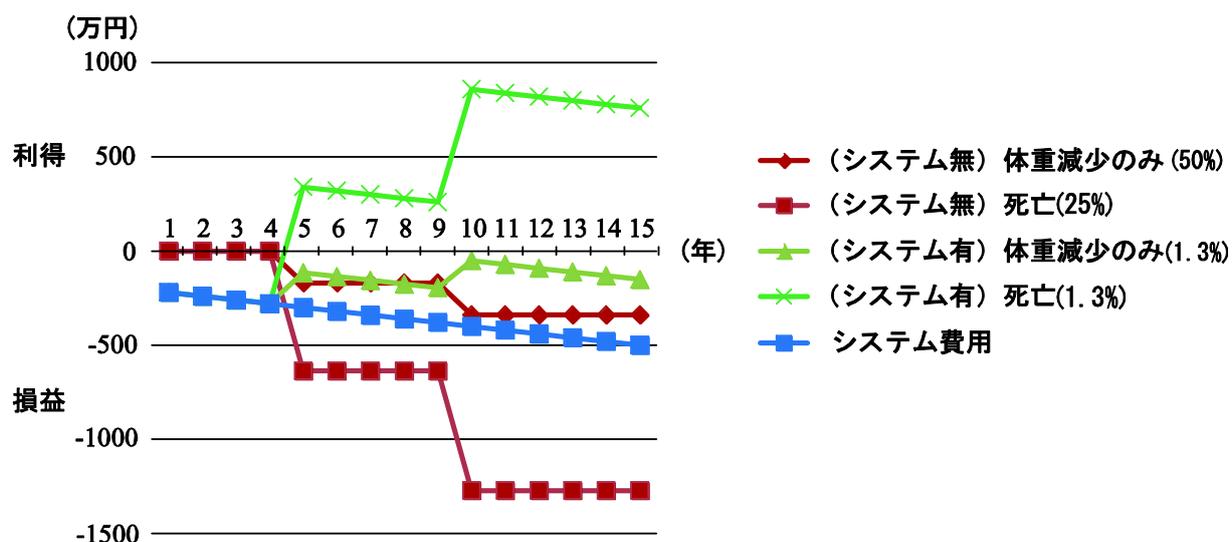


Fig. 1.5 システム費用とシステム有無による損益比較

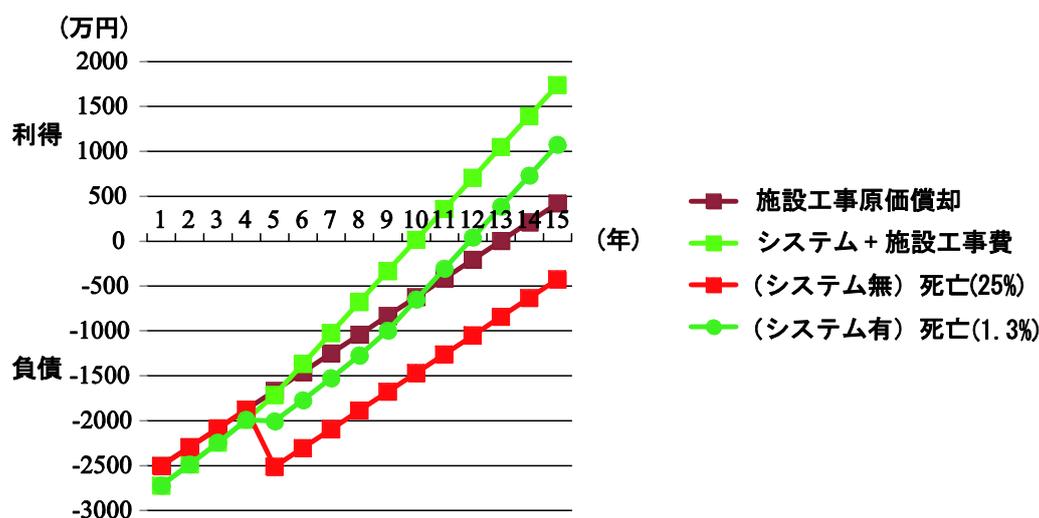


Fig. 1.6 システム有無による対施設工事費の減価償却比較

被害を1個豚房程度に抑えられたため、実際には25%の損害であったところ1.3%の処分で済んだことから約300万程度の利得が発生することがわかった。もし2回同規模の死亡事故が発生した場合にはシステム費用を盛り込んだとして約750万程度の利得を得ることができる。また体重減少のみでシステムの効果を見た緑線の▲マーカをみると1回程度の体重減少での損益はシステム有無で損益が同じことがわかった。結果的に、体重減少のみのPRRS事故率での損益は同程度だが、死亡率が含まれると15年で1度でも発生した場合にシステムによる利得を得ることができることがわかった。

次に養豚農家が肥育豚舎を工事した際の一匹あたりの減価償却額(1,391円)から、死亡事故が1回発生した際のシステム有無による負債状況の変化についてFig. 1.6示す。まず施

設工事原価償却を見ると、15年間異常がなかった場合に13年程度で原価償却できることがわかる。一方で緑線と■のマーカで示すシステム有の場合にはシステム費用に施設工事費を含んだものだが、システムによって獣医師を頻繁に呼ぶ必要がなくなり獣医師費・医薬品費(2,125円)が浮き、段階的に北海道程度の費用(1,215円)になった<sup>[4]</sup>と仮定すると、事故の有無にかかわらず原価償却が通常よりも早く済むことがわかった。これらの分析は、価格変動や物価変動などを考慮しておらず、音による効果も仮定に留まるものの導入コストや現在の損益に対する有効的な手法が少ない中で有益であることが考えられる。

## 1.2 研究目的及び論文構成

本研究の目的は、雑音を有する様々な環境下で少数のマイクロフォンを用いて特定の音源の検出及びその発生位置推定を行うことである。まず特定の音源を検出するため、前節の問題に焦点を当てる。音の検出のためには雑音の影響のない見本となる音を大量用意する必要がある。しかし、大量に用意できないような音源や雑音環境下でしか対象の音源を収集できない環境では有効な手法は提案されていない。そこで本研究では単一のマイクロフォンを用いて畜産業のひとつである養豚豚舎内で豚が呼吸器感染症になった際に生じるくしゃみ音の検出方法について検討を行った。豚舎内は、常に排気ファンやヒータ、冷房などの音に加え豚の鳴き声や柵などに当たる活動騒音、餌作業や清掃作業などに伴う作業音など多くの雑音が発生している環境であり、対象とするくしゃみ音は音声認識で用いられる見本データとなる音声コーパスがなく雑音環境から収集する必要がある。またこのくしゃみ音は常時発生していることがなく収集が難しく、従来の主要な認識方法である大量学習によるモデル化ができない。そこで、このような雑音環境で発生回数が少ない音に対する音源の特徴抽出法について検討をする。また特徴抽出法の汎用性を評価するため異なる環境での適応実験を行い提案法の評価をした。

次に様々な環境で音の発生位置を推定するため、以下の3つの音源を対象に少数マイクロフォンでの位置推定を行った。1つは豚舎内で発生する豚のくしゃみ音をマイクロフォンとカメラを用いる推定法を検討した。2つ目に土中のドリル先端位置を模した音源を用い、土中の音源が地表面での反射音する特徴を用い位置推定精度を向上させる方法について検討した。3つ目に屋外で飛行するドローンの飛行音から位置推定を行う手法について検討を行った。

本論文の構成を **Fig. 1.7** に示す。1章では音に関する研究やその背景について述べるとともに、本研究が対象とする豚のくしゃみ音源の検出と豚呼吸器感染症による被害抑制に関する費用の検討を行った。

2章においては、音を検出する研究や現状について先行研究などを説明し、豚舎での発生音の分析及び音の基本的な特徴である周波数特徴と音の継続時間を用いる検出法を実際の豚舎で収録したデータに対し適応を行い提案手法の検出精度の評価を行った。

3章では収録環境が異なり，発生する雑音が異なる環境においてもくしゃみ音の検出可能な手法を検討するため，豚舎において収録した少数の豚くしゃみ音データを基に音源特徴を抽出する手法について検討した．またその特徴を用いる検出法について検討を行い2つの豚舎で収録したデータに対し異なる豚舎においても同一の手法で検出精度の評価を行った．さらに精度向上のための手法についてくしゃみ音に含まれる音の特徴と類似度検出の手法について述べる．なお，2,3章では音を収集するための最小個数である単一のマイクロフォンのみでの検出法を検討した．

4章では複数のマイクロフォンによって音の到達時間差が生じる現象を用い，位置推定のために必要な最小個数である4つのマイクロフォン及び計8つ程度の少数のマイクロフォンによって音源の位置を推定する手法について，豚のくしゃみ音に対してカメラとマイクロフォンを組み合わせ角度推定から位置を推定する手法と従来法によるマイクロフォンのみでの位置推定についてその推定精度を検証した．また，土中においてはカメラなどが用いることができず，電磁波も透過が難しいことから土中にマイクロフォンを差し込むジオフォンアレイを作成し，地表面反射によって少数でも位置推定精度を向上させる方法について検討をした．最後に立体構成マイクロフォンを作成し，音の到来時間差の計測が難しい小型飛行ロボットを用いその位置の推定法について検討を行った．

5章は本研究の結論となっている．

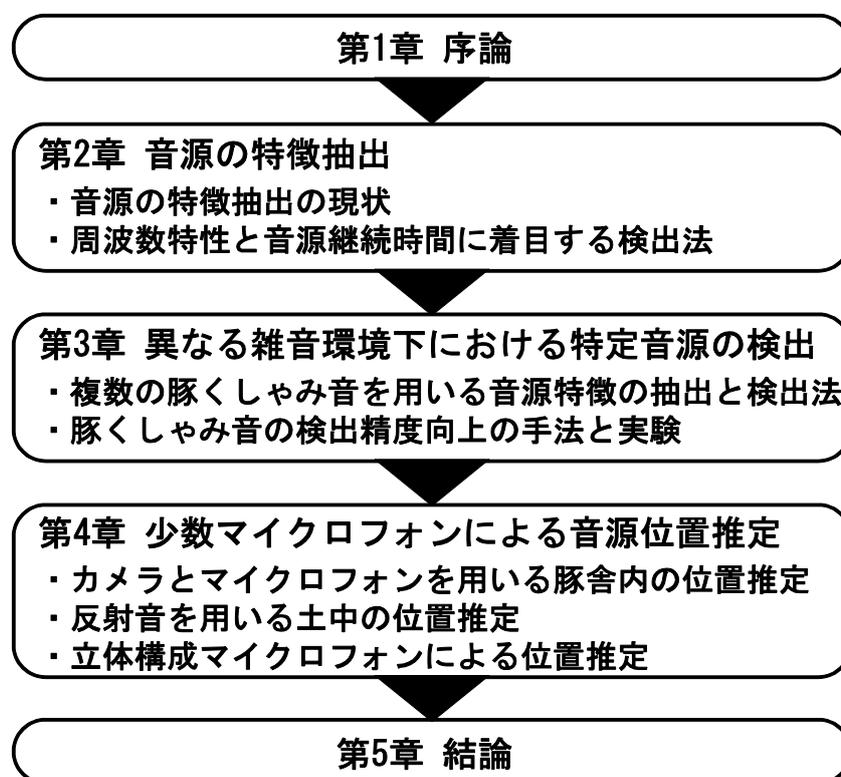


Fig. 1.7 論文構成

## 第2章 音源の特徴抽出

### 2.1 はじめに

本章では、本研究における検出の対象とする音の定義とその検出法に関する現状及び本研究で取り上げる検出対象音と従来法による分析結果について示す。まず検出対象の一つである異常音の定義について示す。異常音は一般的に通常の状態ではない音とされている。例えば、医者が受診時に聴診器で身体の音を聞いているのは通常の状態と異常な状態で音が異なることがわかっているため、これを聞くことで身体状態を解明する指標の一つとしている。橋梁検査のための打音検査や、漏水検出の聴音検査もこのような音を聞くことで状態の把握や漏水場所の位置などを検出する手法として一般的となっている。また、一般の人でも自転車や自動車、使っているパソコンなどでいつもと違う音がした際には異常音として認識し、異常の有無や故障診断を行っている。

このように、音で何かを判断する場合は通常の状態の音を知っているからこそ、身体的な異常や故障時の音などを人が聞き分けることができている。人が聞き分ける際には主に音の周波数特徴と時間特徴、音の大きさを用いて脳が、聞いた音の特徴と過去の通常の状態の音と比較し、この情報と差異があるとき異常がある音と認識される。つまり、人が異常音を検出しようとする場合には、過去に学んだ通常の状態が重要であり、過去の情報なしには異常音を発見することはできない。一方で通常の状態の音においてもその頻度や程度によっては異常音と定義できるものがある。例えば、くしゃみ音や咳音などは日に1回であれば特段異常はないと考えられるが、短時間で大量に発生した場合は異常音とも考えられる。他にも道路における単位時間あたりのクラクション数、人の泣き声や叫び声なども該当する場合がある。このように、異常音を検出することを考えると、前述の見本を必要とする異常音は「今までにない未知の音」であり後述の異常音は「既知の特定の音」である。異常音を機械的に発見する場合には、検出の対象がどちらの分類なのかを見極め、検出法を検討する必要がある[45-48]。

本研究では検出対象の音源を「既知の特定の音」として豚が疾病時に伴うくしゃみ音やスクリーモータの軸ぶれの音や飛行中のドローン、土中でのドリルビット音などとして検出法や位置推定法について検討した。本章では、特定の音を検出する関連研究と問題点、周波数特徴と時間特徴を用いる豚くしゃみ音の検出法及び結果について述べる。

## 2.2 音源の特徴抽出の現状

機械によって音声<sup>[9-11]</sup>や異常音<sup>[15-17]</sup>や、特定の音<sup>[18,20,21]</sup>などを検出する場合についても人間と同様に周波数特徴と時間特徴を主に用いている。しかし、対象の音が Fig. 1.2 のように同一音でない場合にはこの2つの特徴を適切に抽出することが難しい場合がある。例えば音声のように「あ」という音でも性別、年齢などによって周波数特徴や時間特徴が僅かに異なることがわかっているが、人が「あ」という音を聞いたときに性別や年齢にかかわらず「あ」という音であると特定できる。これは周波数特徴や時間特徴が異なる場合でも「あ」という音に含まれる共通の特徴を脳が学習していることに基因している<sup>[26]</sup>。そこで人が聴いた特性として音声の特徴を抽出する研究が実施されており、代表的な手法としてメル周波数ケプストラム係数 (MFCC: Mel-frequency Cepstral Coefficients) が挙げられる<sup>[49]</sup>。この手法は人が声を発する時は音源の特性（「あ」という周波数情報）と声道特性（喉や鼻、口などを含めた性別、年齢、個人で異なる特性）の畳み込みであると考えで実現している手法である。音源から声道を通り音が発せられマイクロフォンで収録される経路を考慮すると、収録信号は音源の周波数と、声道特性の積で示すことができる。よって両辺に対数をとることで、和の形にすることができる。この信号をさらにフーリエ変換をかけたものをケプストラム分析とよび音源信号の主要な周波数と声道を分離することができる。これに人の聴こえ方を模したメル周波数フィルタによる重み付きをすることによって実現され、音声認識の多くに実装されている<sup>[50]</sup>。一方で、この手法は雑音や反響の影響に弱く走行中の車内のような環境ではセンサと音源の距離が近い場合でも著しく認識率が悪化することがわかっている<sup>[51]</sup>。ここで Fig. 2.1 にモータで動く船のスクリュが軸ぶれした際の様子について示す<sup>[52]</sup>。(a)が通常のスクリュで(a')が軸ぶれのスクリュの時間波形、(b),(b')が周波数、(c),(c')がそれぞれのケプストラムの結果となる。ケプストラムは音声認識分野の他に、回転機などではピッチの微細変化などを周波数領域よりも正確に検出できることがわかっている<sup>[21,53]</sup>。本実験は水中でマイクと対象物の距離を一定で計測したが、2つの状態で音の大きさが異なる場合はセンサ間の距離によって変化するため一概に判定はできない。周波数領域では 17 kHz から 20 kHz に特徴があり、検出の特徴として使用できる可能性がある。ケプストラムの結果によって、軸ぶれ特有のうなり (7.3 Hz 及び 6.1 Hz) は確認でき特徴量として使用できる可能性があるが、ケプストラムは周波数領域信号の微細構造を強調しており、雑音環境では著しく検出精度が劣化することから非接触による水中での分析手法としては問題がある。

特定音としては同一音や環境音や非言語音声などの音響イベントなどが挙げられる<sup>[27]</sup>。同一音ではブザーやクラクションなど、銃声や爆発音など社会的な安全対策のため研究されている<sup>[54]</sup>。検出方法としては先に挙げた周波数特徴や継続時間特徴などをデータベースに集め、同一の音を検出して識別するものとなっている。音響イベントの検出では、データベースに入れる音響イベントをあらかじめ定義する必要である。例えば、ドアの開閉音として収録信号を音響イベントのドア開閉音として登録する。これと同一の特徴を持つ音が検出された際に、音響イベントのドア開閉音が起こったと検出するものである。

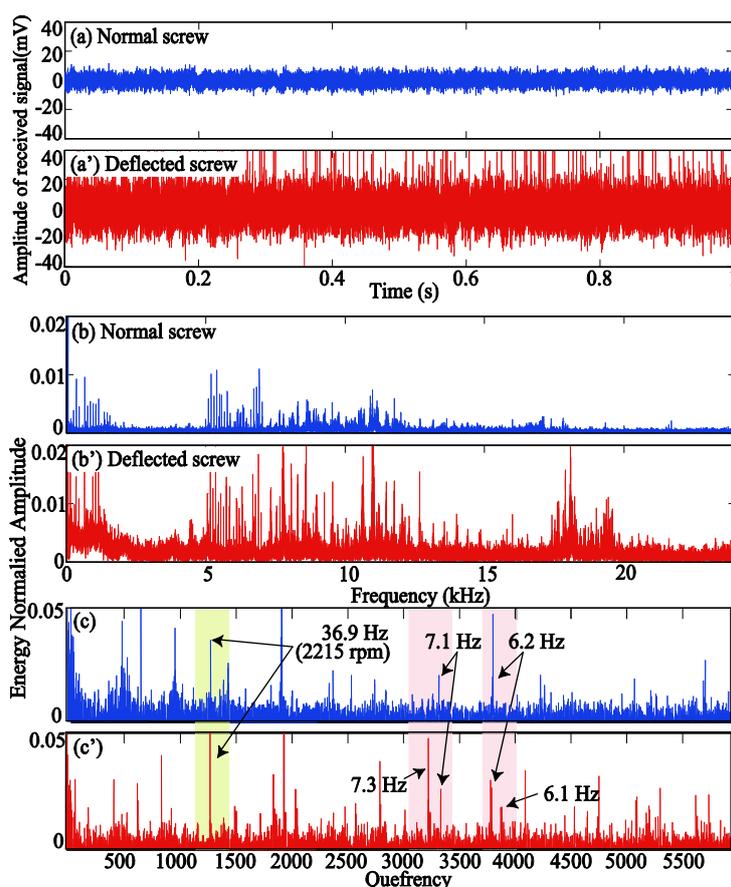


Fig. 2.1 2215 RPM におけるスクリュ音, (a) 収録信号, (b) 周波数及び(c) ケプストラム

一方でドアには多くの種類があり、様々なドア開閉音の検出のためには多くのドア開閉音を用意し、特徴を統計的にまとめる必要がある。検出方法としては、音声認識と同じくメル周波数ケプストラム係数等を特徴量とし、統計的なモデルである隠れマルコフモデルを用いることで検出率が高い結果となっている<sup>[55]</sup>。また特徴の抽出方法や検出の識別方法として、メル周波数ケプストラム係数ではなく非負行列因子分解<sup>[56,57]</sup>や独立成分分析<sup>[58,59]</sup>を用いる特徴の抽出や音のモデル化の手法として GMM(Gaussian Mixture Model)<sup>[60]</sup>なども検討されており、これら特徴と SVM(Support Vector Machine)<sup>[61]</sup>や NN(Neural Network)<sup>[62]</sup>などの機械学習の組み合わせによって音の検出を行っている研究が多く存在する。ただし、これらの手法で用いられる見本となる様々な音は各国研究機関で提供されるデータ<sup>[63,64]</sup>に致し研究されており、国内においては音声資源コンソーシアムが提供する数百種類、数万サンプルもの音響データベース<sup>[65]</sup>などの基準となるデータがあり、これに対し観測信号に雑音を付加し性能評価している。一方で、見本となるデータについては雑音が存在すると正しく特徴の抽出ができないためこれらの研究の多くはデータとして提供される雑音の少ない見本となる音の特徴を抽出した後に、雑音が付加された音などから検出を行っている。雑音を含んだ音を見本データとして検出を試み会議室の音から音響イベントなどを検出する CLEAR (Classification of Events, Activities and Relationships)などの競争型ワークショップなども行われているが、最も認識率が良いチームでも 30%前後の値となっている<sup>[66]</sup>。

## 2.3 周波数特徴と音源継続時間に着目する検出法

本節では音源の対象として豚舎内における豚のくしゃみ音を、最も基本的な音源の特徴抽出方法である周波数特徴と音源継続時間に基づき検出方法を検討する。豚のくしゃみ音は前節での音声コーパスなどが存在せず、豚舎内で自ら収録する必要がある。このため雑音の少ない環境での見本となる音の収録が難しく、検出対象である音と共に常に雑音が含まれる状態となる。このため、豚舎内で収録した音を発生した音源の種類と共に分類をする。その上で検出対象とするくしゃみ音独自の周波数特性や音源継続時間などをそれ以外の雑音と比較し、くしゃみ音のみを正確に検出できる手法の検討を行う。また、本手法での性能評価を行うため豚舎で収録したデータについて検出を行い、**Fig. 2.2**のように識別率と感度の基準を定めた。Aは豚舎内で収録した実際のくしゃみ回数と雑音の発生数である。くしゃみ音の発生回数については収録時間中常に2人以上で計測を行い目と耳でその場でくしゃみ音が起きたものを記録し、録音データ及び録画データから発生を確認したものであり、雑音については常時稼働している換気ファンより大きな音が発生した回数となる。この音には、作業員の清掃や餌やりの音も含まれている。また豚が叫んだり、騒いだり等で発生する様々な音についてこの雑音に含まれている。Bが提案手法での検出結果となる。このためaに当たる数字は、実際に発生したくしゃみ音を提案手法でくしゃみ音として検出できた数となり、bは提案手法ではくしゃみ音と検出したが実際には雑音だった回数となる。ここで提案法によってくしゃみ音と判別した数当たりにもどれだけ正しいくしゃみ音が含まれていたかを  $a / (a + b)$  の識別率という値を用いる。この値が高いほど雑音をくしゃみ音として誤ってしまう誤検出の可能性が少なくなる。またcの値は実際にはくしゃみ音が発生していたにも関わらず検出法では雑音として分類して見逃してしまった数となり、実際に発生したくしゃみ音の回数あたりの検出法でのくしゃみ音発見回数を  $a / (a + c)$  の感度とする。感度が高いほど提案手法でのくしゃみ音の見逃しが少ないと言える。このような評価手法は、天気予報<sup>[67]</sup>や臨床疫学<sup>[68]</sup>のスクリーニング検査の効果指標などで使用されており、くしゃみ音のような対象の音が雑音の発生回数より圧倒的に少ない場合などに有用とされる。例えば天気予報の精度において、晴れと雨のみのような判定で年間1日しか雨の降らない場所があった際に予報としてはすべて晴れで出しても精度が高くなってしまふ。そこで、晴れを正しく晴れ、雨を雨と判定できた精度を調べるため上記のような評価手法が提案された。本研究においてもくしゃみ音の検出を行う性能評価をこの2つの値を用いることとする。

		B		Total
		Sneezing	Noise	
A	Sneezing	a	c	a+c
	Noise	b	d	b+d
Total		a+b	c+d	

A : Correct value

B : Result of value of the discrimination

Discriminative rate :  $a / (a+b)$

Sensitivity :  $a / (a+c)$

Fig. 2.2 識別率と感度の定義

### 2.3.1 見本データの収集と解析

実際の豚舎の音を国立研究開発法人 農業・食品産業技術総合研究機構 畜産草地研究部門（収録当時は、独立行政法人 農業・食品産業技術総合研究機構 畜産草地研究所）において2012年6月27日から29日において録音したうち、人が2名以上でくしゃみ音の発生及びその位置を確認した。音声の収録にはレコーダ（TASCAM, DR - 680）、アンプ（SONY, PCM-D50）、コンデンサマイクロフォン Panasonic, WM62PC）を使用した。実験に用いたマイクロフォンは、高周波対応のマイクロフォン（Brüel & Kjær, Type 4939-A-011）と比較実験の結果、20-40 kHzの周波数帯域において最大-20 dB程度の低下で収録できることを確認している（詳細は付録Aに記す）。また、位置参照用にWebカメラ（SANWA, CMS-010BK）の画像をPC（HP, 6535s）で記録した。Webカメラの解像度は640×480画素であり、角度1 deg.に対する画素数は12であった。対象とした豚房には生後43日の豚11頭が飼養されており、舎内温度は平均27.5℃であった。実験構成をFig. 2.3、実験条件をTable 2.1に示す。

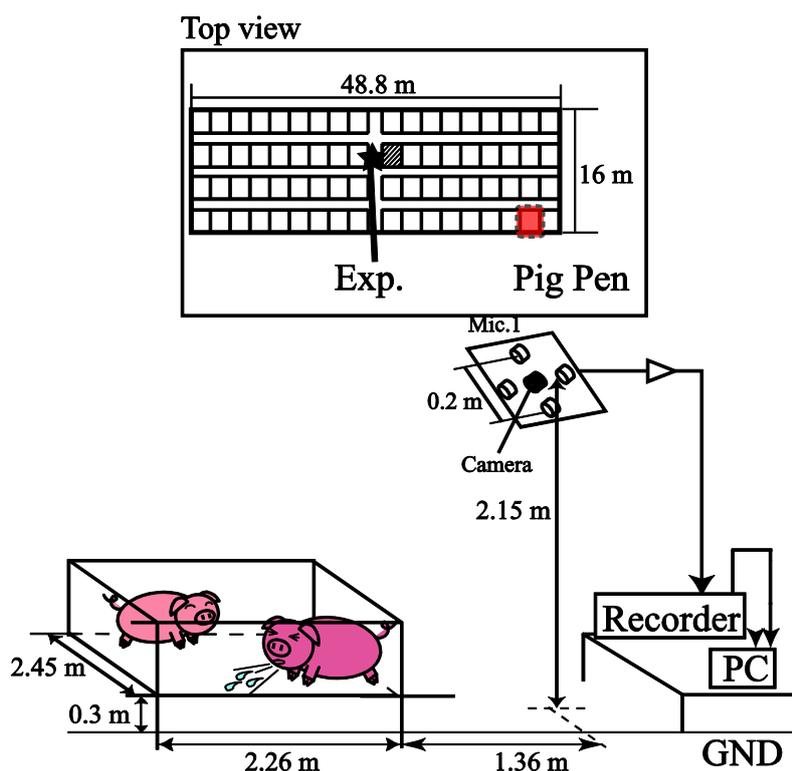


Fig. 2.3 畜産草地研究所における実験機材の構成

Table 2.1 実験条件

Total time for data	11,178	s
Sampling frequency	96	kHz
Temperature	27.5	°C
Sound velocity	348	m/s
Distance of Mic.( <i>d</i> )	0.2	m
Window length	0.5	s

収録結果、豚舎内で発生する豚の音声にくしゃみ(Sneezing), 咳(Cough) の他に代表的な鳴き声 Grant, Bark, Squeal があり<sup>[69]</sup>, 他の音として金属の柵がぶつかる音や、換気音が検知された. **Figure 2.4** にそれぞれのスペクトルグラムを示す. FFT 点数は 480 点とし長さ 5 ms の Hanning 窓を使用した. 横軸を時間, 縦軸を周波数とし, 音圧を輝度で示す. Fig. 2.4 (a) のくしゃみ音は, 0 - 40 kHz までの周波数を含み時間長は 0.1 s 程度である. Fig. 2.4 (b) の咳音は, 0 - 20 kHz までの周波数を含み, 0.1 s 前後の区間である. これに対して Fig. 2.4(c) の Grant は, 0 - 30 kHz の周波数が支配的である. また, くしゃみ音と比較して長い時間長を持つことがわかる. Fig. 2.4 (d) の Bark も同様に 0 - 30 kHz の周波数を含むが時間長は 0.15 - 0.5 s に分布している. Fig. 2.4 (e) の Squeal は, 全体的に音圧が大き

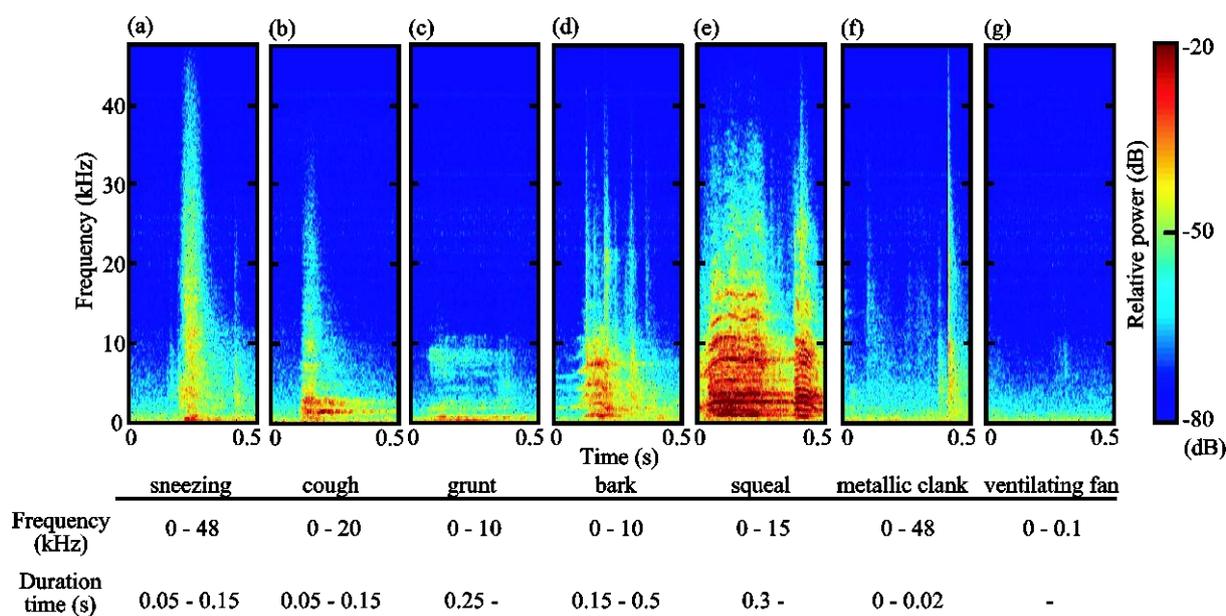


Fig. 2.4 豚舎内発生音のスペクトログラムと周波数及び時間特徴

く 0 - 10 kHz の周波数が主であり、時間長が他と比較し長い。Fig. 2.4 (f) は、金属音はくしゃみ音と近似しており 0 - 40 kHz の周波数を含むが時間長が約 0.01 s と非常に短い特徴がある。Fig. 2.4 (g) は、換気扇は 100 Hz 以下の低周波数であった。他にも作業員の声や、清掃作業に伴う音など様々な音が確認された。本収録で得た様々な音に比べ、くしゃみ音の特徴として明かになったのは 30 kHz の高周波を含み、音の継続時間が 0.05-0.15 程度であることがわかった。これらの特徴から豚舎の雑音内から目的音であるくしゃみ音を検出する手法について示す。

### 2.3.2 検出実験

前節の収録の同様の場所、豚房において約 3 時間収録したデータに対し検出法を評価する。なお、評価対象のデータ中の音は見本とするデータと同一場所の別日に計測したくしゃみ音 5 つを使用した。検出のフローチャートを Fig 2.5 に示す。

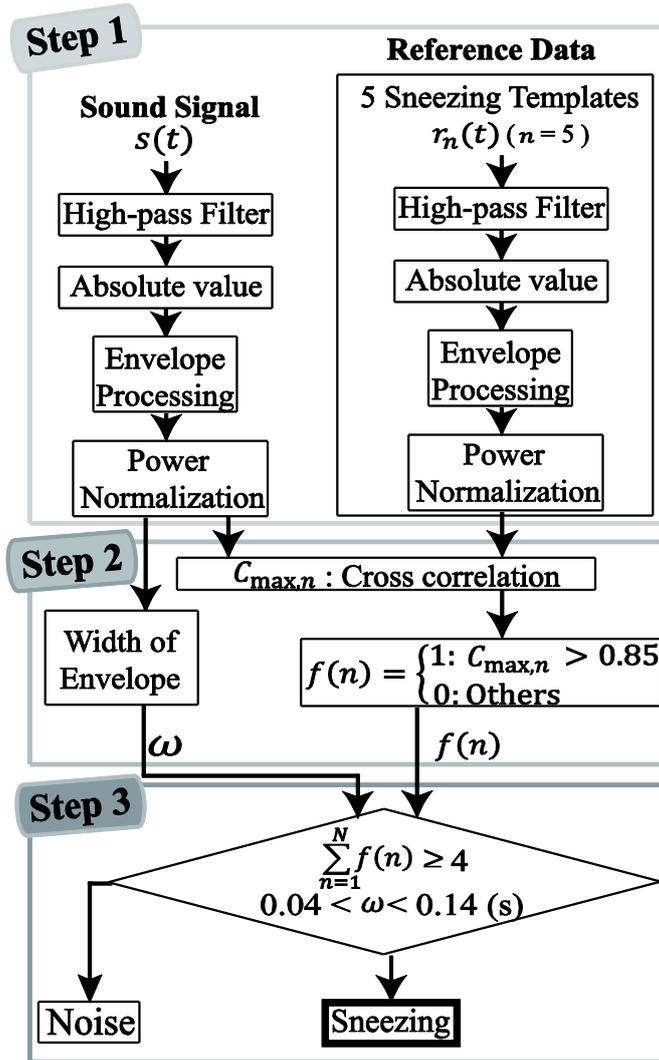


Fig. 2.5 周波数特徴と時間特徴を用いる検出法のフローチャート

**Step1)**前節の特徴からくしゃみ音は雑音となる他の豚舎発生音より 30 kHz 以上の周波数成分を多く含んでいることから，受信信号およびテンプレートに対し 4 次の Butterworth 型ハイパスフィルタを通過させる．さらに信号の絶対値をとった後に切り出し区間の信号の総エネルギーが等しくなるよう正規化し，その包絡線を受信信号  $s(t)$  とし， $N$ 個のテンプレートを  $r_n(t)$  ( $n=1, 2, \dots, N$ )とする．本章では五つのくしゃみ音 ( $N=5$ ) をテンプレートとして使用した．

**Step2)**信号の類似度を判定するため，正規化相互相関関数  $R_{cc-n}$  を以下に示す．

$$R_{cc-n}(\tau) = \frac{\sum_{t=0}^T s(t)r_n(t+\tau)}{\sqrt{\sum_{t=0}^T s^2(t)\sum_{t=0}^T r_n^2(t+\tau)}}, \quad (2.1)$$

と定義し、 $s(t)$  と  $r_n(t)$  の相関計算を行い  $R_{cc-n}$  を求め各テンプレートとの類似度を判定する。この時、受信信号とテンプレートの包絡線が類似しているとき  $R_{cc-n}$  が大きくなる。  $N=5$  の場合は五つのくしゃみ音テンプレートとそれぞれと受信信号の類似度を比較している。この時、正規化のため  $R_{cc-n}$  の最大値は 1 となり、本章では  $R_{cc-n} > 0.85$  となる時収録信号をくしゃみ音の候補として抽出するものとした。なお、0.85 のパラメータ値は実験により感度と識別率が最大となる値とした。また、豚舎内の雑音のうち、ハイパスフィルタを通過するものに金属音がある。Fig. 2.3(f) の金属音は、くしゃみ音に近い特徴を持ち、高い相関度を示すことがあるが、くしゃみ音に比べ発音時間が極端に短い。そこで、信号の包絡曲線の最大値となる付近の包絡線幅を  $w$  と定義することで、信号の持続時間がくしゃみ音と一致するものを抽出する。

**Step3)** 提案手法では  $R_{cc-n} > 0.85$  となるテンプレートが 4 つ以上あり、 $\omega$  が前節のくしゃみ音の時間領域における特徴を元に定義した  $0.04 \leq \omega \leq 0.14$  (s) の両方の判定を満たしたデータをくしゃみ音として検出した。

### 2.3.3 実験結果と考察

**Table 2.2** に結果について示す。収録した 61,938 区間中、くしゃみ音が含まれる区間は 14、雑音区間が 61,924 であった。従来手法、提案手法ともに実際に発生したくしゃみ音を検知する検出感度は 85.7% であった。両手法ともに検知できなかった 2 区間のくしゃみは他の豚房で発生したくしゃみ音（収録時に発音位置を確認済）であり、信号対雑音比が小さく高周波成分が含まれておらず包絡線に変化が現れないものであった。識別率は 63.2% であり、くしゃみ音として識別した 19 区間中で 12 区間が実際のくしゃみ音であり、7 区間のみ雑音をくしゃみ音として誤識別していた。よって、提案手法による識別率は 50% 以上であり、くしゃみ音の自動判別による陰陽性の判別に有効であることを確認した。一方でテンプレート 5 つを用い検出法を検討する際は、通常それぞれ異なる特徴を持ったくしゃみ音 5 つを指定し、それに高い相関（類似度）を持った値を採用する手法が一般的である。本手法を検討する際、各テンプレートについて実験した識別率と感度について **Fig 2.6** に示す。横軸が各テンプレートの数であり、5 つ使用したもの以外については各テンプレートでの結果の平均値及びその偏差を示している。まず各テンプレートを 1 つずつ使用した結果を見ると、感度は 80% 近くとなり、1 つの見本データのみでも見逃しをする可能性が低いことがわかる。一方で、1 つのみを使用した場合は雑音をくしゃみ音としてしまう識別率が悪化し、誤検出の数が多いことがわかる。そこでテンプレートの数を増やしていったところ、 $N=4$  が感度、識別率共に高くなることがわかった。5 つのテンプレートをすべて採用した場合には、誤検出の割合は低下するが、見逃しを示す感度も低下してしまうことがわかった。

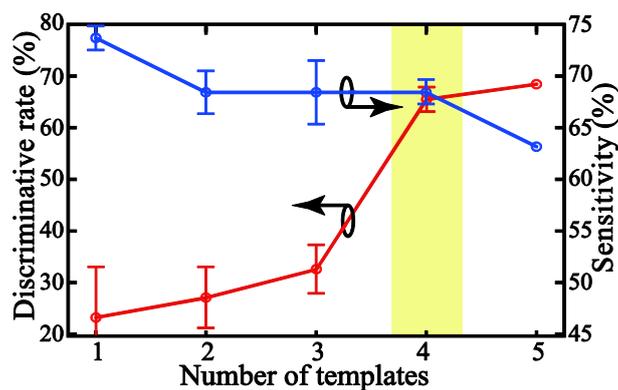


Fig. 2.6 テンプレート数と識別率・感度の関係

Table 2.2 くしゃみ音5つを用いる検出結果

		B		Total
		Sneezing	Noise	
A	Sneezing	(a) 12	(c) 2	14
	Noise	(b) 7	(d) 61,917	61,924
Total		19	61,919	

A : Correct value

B : Result of value of the discrimination

Discriminative rate (  $a / (a+b)$  ) : 63 %

Sensitivity (  $a / (a+c)$  ) : 85 %

## 2.4 まとめ

本章では音源特徴の抽出として、同一音や特定の音、音声などの特徴抽出の現状について、大量のデータと雑音の少ない見本となるデータがあれば統計的な手法を用いた特徴抽出法が実用化されている例やその値を機械学習することによる検出法などについて述べた。一方で、見本とするデータ自体に雑音が入っていた場合に適切な特徴の抽出法や検出法などがなく、豚のくしゃみ音のような特定の音に対しては、収録から検出法までの検討を行う必要があり豚舎内の音の分類によって高周波成分まで含む特徴と広帯域に渡り音源継続時間に特徴があることを抽出することができた。この特徴を用い検出法を検討し、実際の豚舎で収録したデータを基に実験を行った。その結果、5つの見本データを多数決のようなアルゴリズムで類似度を選択する手法によって提案手法の性能評価の一つであるシステムの見逃し辛さを測る感度は85%であり、誤検出を測る識別率は63%の結果となった。これは、先行研究による感染症の陰陽性識別のための基準を超えるものであり、感染症の発見に有効であることが示された。

問題点として本手法は、収録環境に合わせパラメータを調整しており他の環境で収録された音に対しても提案手法を適応して汎用性を確認する必要がある。

## 第3章 異なる雑音環境下における特定音源の検出

### 3.1 はじめに

前章の手法を異なる豚舎で収録したデータに対し適応するため、全農飼料畜産中央研究所の養豚研究室の協力<sup>[70]</sup>を得て豚舎内の音を収録した。2014年11月18日から19日において録音し、事前に訓練を受けた計測員が2名以上でくしゃみ音の発生及びその位置を確認した。音声の収録にはNational Instruments社のLabviewをインストールしたPC、アンプ(kitsrus, KIT-98)、コンデンサマイクロフォン(Panasonic, WM61A)を使用した。また、位置参照用に広角のWebカメラ(BUFFALO, BSW20KM11BK)をPCに記録した。対象とした部屋には生後40日前後の豚約100頭が飼養されており、収録された音には対象豚房以外の豚による音も多く含まれる。舎内温度は平均26.0℃でジェットヒータによって常に暖房で一定温度に保たれていた。実験構成をFig. 3.1、実験条件をTable 3.1に示す。対象となる豚はFig. 3.1の斜線部分である。

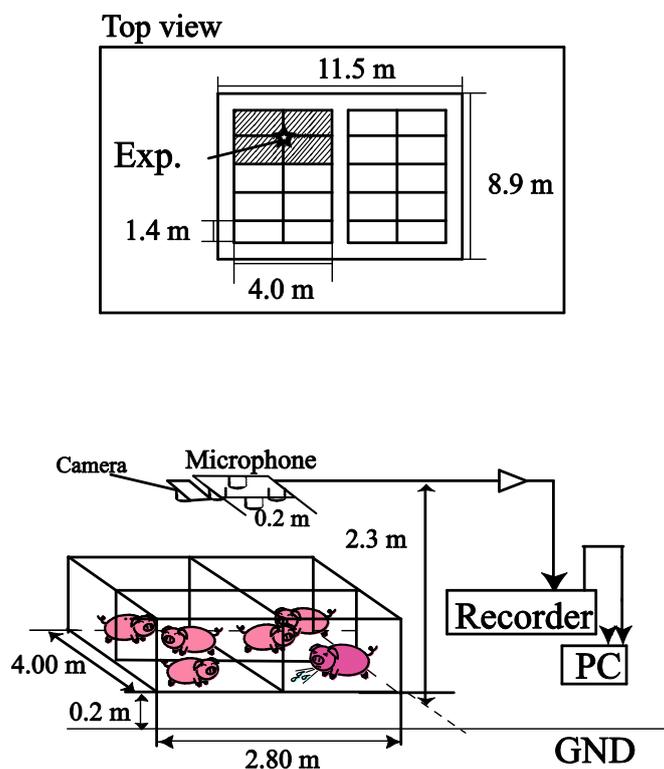


Fig. 3.1 全農飼料中央研究所における実験機材の構成

Table 3.1 実験条件

Total time for data	1,800 s
Sampling frequency	100 kHz
Temperature	26.0 °C
Sound velocity	347 m/s
Distance of Mic.( <i>d</i> )	0.2 m
Window length	0.5 s

全農飼料中央研究所の豚舎で収録したデータを2章の検出法へ適応したところ、識別率・感度ともに10%程度となった。誤検出の指標である識別率が悪化した原因として、豚舎によって雑音が変わることから、前章の検出のための特徴であった高周波帯域まで含む特徴と同じ雑音が多く発生する環境であったことが挙げられる。また、見逃しの指標である感度が悪化した原因として特徴抽出のひとつである音の継続時間が密閉された空間では反響し残響時間が異なることから<sup>[71]</sup>、開放空間であった前章の収録データの見本データと特徴が異なり、Fig. 3.1ではくしゃみ音を見逃していたと考えられる。二つ目の原因としてくしゃみ音の特徴であった高周波帯域を含む現象がくしゃみ音の発生位置が遠方となると距離によって減衰するため<sup>[72]</sup>、検出ができず見逃しが発生した。2つの豚房で発生したくしゃみ音のスペクトログラムをFig. 3.2に示す。畜産草地研究所での収録音には30 kHz付近に機械の雑音が確認できるが、全農での実験データには同一の雑音がなく、115 dB程度の倍音成分のような雑音が複数確認でき雑音環境が異なることがわかる。またくしゃみ音の特徴であった高周波帯域までの成分が25 kHz付近から減衰し確認できないことがわかった。

以上より、他の豚房で収録したデータに対して同じ見本データがそのまま使用できないことがわかった。そこで、本章では環境によって異なる雑音に対応する検出方法について雑音を低減する手法を検討するとともに、雑音環境下で収録された見本データから特徴を抽出する方法を検討する。

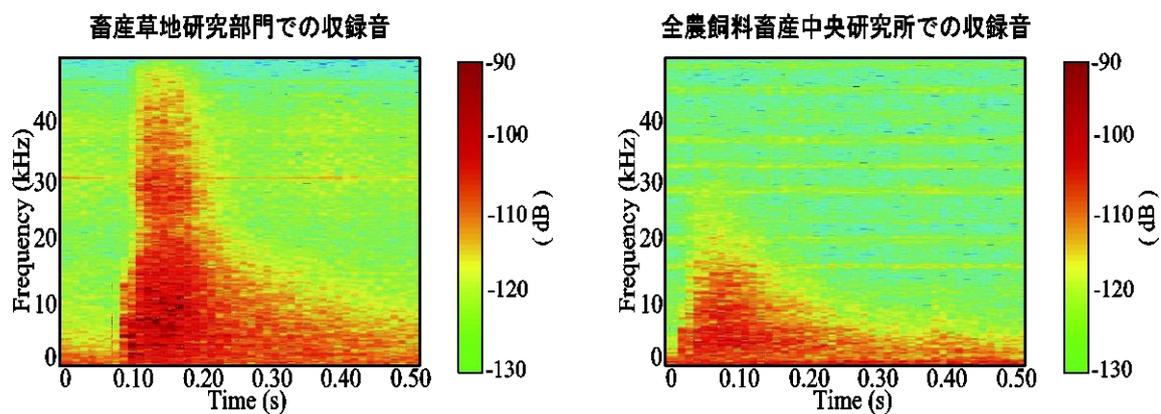


Fig. 3.2 異なる場所におけるくしゃみ音のスペクトログラム

## 3.2 少数の見本データを用いる音源特徴の抽出と検出法

### 3.2.1 高域強調フィルタと定常雑音の低減法

本節では、別々の豚舎で収録したくしゃみ音の見本データからくしゃみ音の特徴を抽出する手法及び異なる雑音環境下においても対応するため、雑音低減手法について述べる<sup>[73,74]</sup>。音の認識では一般的に 25 msec 程度に音を切り出しその区間内の主要な周波数やその包絡線の形状を特徴として用いている<sup>[26]</sup>。そこでくしゃみ音が開始する時点から 25 msec を切り出した振幅スペクトルを **Fig. 3.3** に示す。くしゃみ音を比較すると、主要な周波数や共通の倍音成分も確認できず、全農でのくしゃみ音に高周波が含まれていないことがわかる。よって従来までの検出法による主要な周波数や包絡線などの特徴からの検出が難しく、別の特徴を抽出する必要がある。そこで、**Fig. 3.2** のスペクトログラムの形状に着目して抽出法を検討する。主要な周波数や含まれる帯域も異なっているが、この 2 つのスペクトログラムは三角形のような形状をもっており、他のくしゃみ音も共通にこの形状を確認することができた。そこで、雑音を低減した後にスペクトログラムの形状から見本データを作成する検出法を検討する。

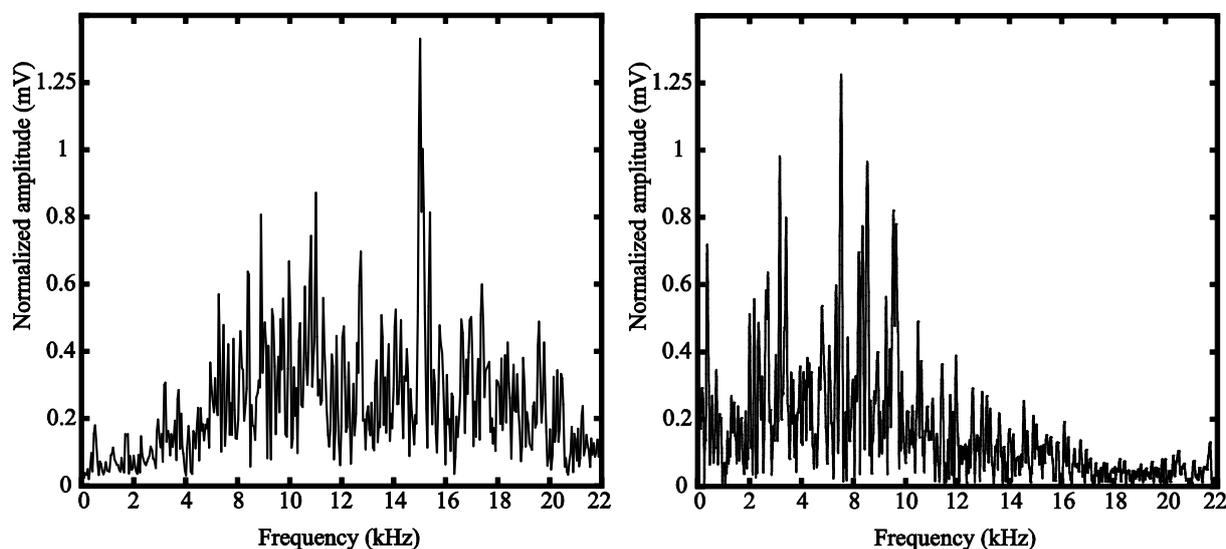


Fig. 3.3 冒頭 25 msec におけるくしゃみ音の振幅スペクトルの比較

まず雑音環境で収録されたくしゃみ音から、検出の際の見本となるデータを作成する前処理として、くしゃみ音に含まれる高域を強調する高域強調フィルタによる特徴の強調と定常的な雑音の低減処理を実施する。音声認識分野においては人の声に含まれる高域成分が認識に有効であるとされており、以下の伝達関数をもつ1次元のデジタルフィルタを通すことで高域強調を行っている<sup>[26]</sup>。

$$H(z) = 1 - az^{-1}, \quad (3.1)$$

$a$  は音声認識の分野では一般的に 0.95 が最適とされており、サンプリング周波数 8 kHz の条件で 6 dB/oct 程度の高域強調が可能である。一方で、くしゃみ音ではサンプリング周波数がこの条件の 10 倍以上にもなり、信号の有無にかかわらず高周波が強調されることになる。そこで、35 kHz 付近で 10 dB ほど強調できるフィルタのため本研究における高域強調フィルタは  $a = 0.80$  とした。高域強調フィルタの特性を Fig. 3.4 に、フィルタ処理前後のくしゃみ音のスペクトログラムを Fig. 3.5 に示す。

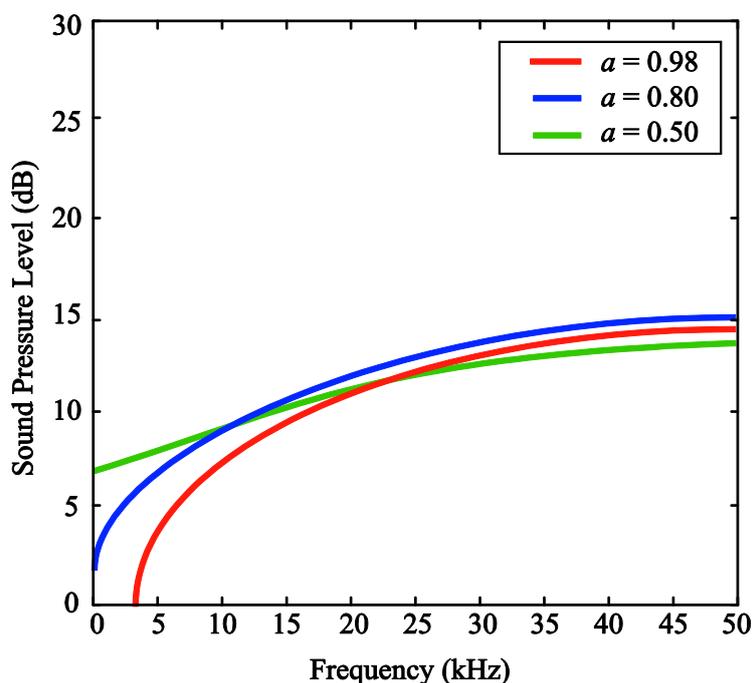


Fig. 3.4 高域強調フィルタの特性

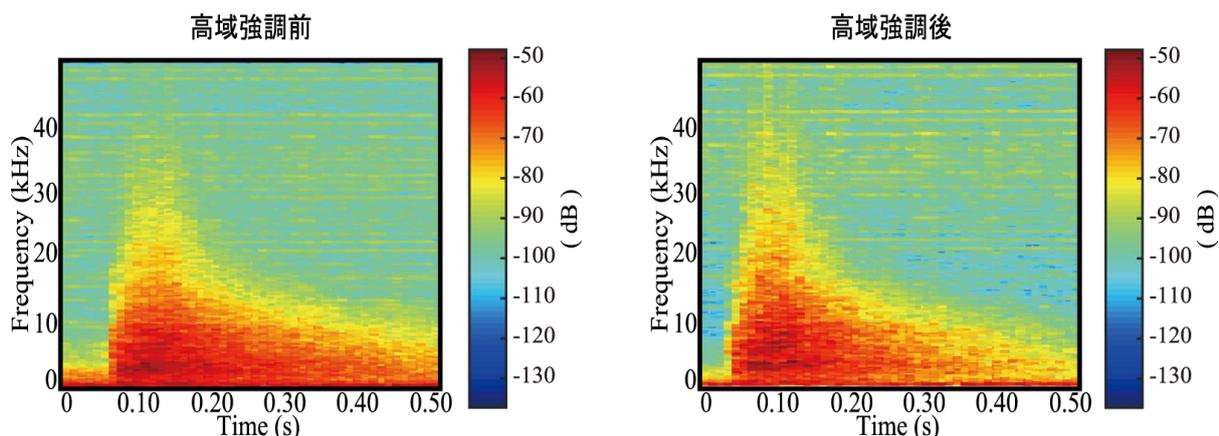


Fig. 3.5 高域強調フィルタ前後のくしゃみ音のスペクトログラム

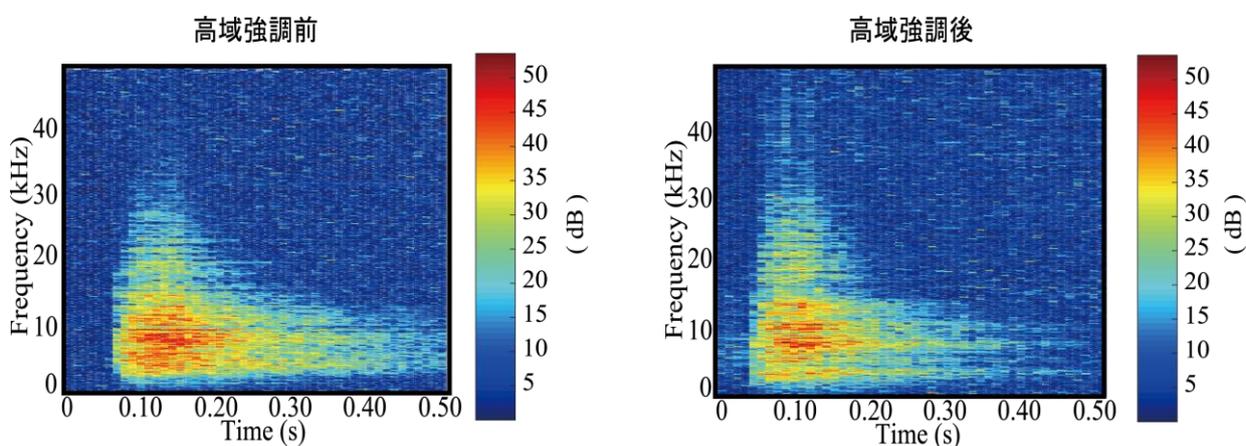


Fig. 3.6 高域強調フィルタ前後のスペクトルサブトラクション法を適応したくしゃみ音のスペクトログラム

次に定常的な雑音の低減方法について検討する. Figure 3.5 のどちらの図にも 2 kHz 程度の帯域にくしゃみ音の開始前から定常的な雑音を含んでいることがわかる. また, 高域強調前後にも機械雑音の一部とみられる高周波に渡る雑音が -80 dB 程度確認することができる. そこで, このような定常的な雑音を低減するためスペクトルサブトラクション法<sup>[75]</sup>をスペクトログラムへの応用を検討した. スペクトルサブトラクション法とは, 雑音低減手法の一つであり, 観測された音を周波数領域で雑音除去を行う手法である. マイクロフォンで収録された音が Fig. 3.5 のように定常的な雑音と目的の音が合わさった信号だとすると, 時刻  $t$  における収録信号  $s(t)$  は雑音  $n(t)$  と目的の音  $o(t)$  から,

$$s(t) = n(t) + o(t), \quad (3.2)$$

と示すことができる。スペクトログラムは短時間フーリエ変換によって細かいフレーム毎に区切られた信号を周波数領域に変換したものであるため、時刻に対応する  $k$  フレーム目の収録信号のスペクトル  $S(k, f)$  は、

$$S(k, f) = N(k, f) + O(k, f), \quad (3.3)$$

ここで  $N(k, f)$ ,  $O(k, f)$  はそれぞれ雑音と目的の音のスペクトルとなる。今 Fig. 3.5 のように収録信号の最初の数フレームにのみ定常的な雑音があると仮定すると、最初から  $K$  個目までのフレームによって定常的な雑音を推定すると、推定された雑音信号を  $\overline{N}(k, f)$  として、

$$|\overline{N}(k, f)| = \frac{1}{K} \sum_{k=1}^K |S(k, f)|, \quad (3.4)$$

と示すことができる。式(3.3)の雑音のスペクトルが推定された雑音と一致するとき、収録信号全体から推定された雑音信号を減算することによって、以下の式によって定常的な雑音を低減し目的の音を得ることができる。

$$O(k, f) = S(k, f) - |\overline{N}(k, f)|, \quad (3.5)$$

実際の収録信号では推定された雑音と、もともとの雑音が一致することはないため式(3.5)のように完全に目的の音のみにはならない。また目的の音と雑音信号が同一のスペクトル帯域をもった場合、目的の音のスペクトルも減算してしまうため注意が必要である。音声信号の場合は MAP 推定法<sup>[76]</sup>と呼ばれる雑音と音声の振幅・位相が統計的に独立であると仮定した雑音推定法などが存在するが、雑音として他の鳴き声や統計的な独立性を担保できない豚舎内の音についてはフレーム毎の単純な減算処理のみを適応し検出法を検討する。

**Figure 3.6** に高域強調フィルタ前後でのスペクトルサブトラクション法を適応したくしゃみ音について示す。本手法によって Fig. 3.5 の処理前の収録信号にあった定常的な雑音を低減するとともに、くしゃみ音の特徴である高周波まで含む特徴を確認することができた。Figure 3.6 は全農飼料中央研究所のあるひとつのくしゃみ音であり、他のくしゃみ音に本手法を適応し、くしゃみ音のスペクトログラムを確認した。その結果、観測されたすべてのくしゃみ音がまったく同じ形状ではないものの、Fig. 3.6 のような三角形に類似したスペクトログラムの形状を確認することができた。一方で、本手法の適応後でもくしゃみ音には音声認識の特徴で用いられる短時間あたりの主要な周波数を確認することはできなかった。次節では、これらのくしゃみ音から共通の特徴を抽出する手法について検討するとともに、収録音からくしゃみ音を検出する手法について提案する。

## 3.2.2 スペクトログラム相乗平均による特徴抽出と検出法

本項では、雑音環境で収録した見本となるデータに対し、雑音の影響を低減した音の特徴をスペクトログラムから抽出する方法及びその特徴を用いた検出法の検討を行う。スペクトログラムから音の特定を行う研究は、2005年程度から研究が始まった音声認識や環境認識の分野の方法の中では比較的新しい手法であるといえる<sup>[27]</sup>。これは、スペクトログラムを作成するために短時間フーリエ変換によってフレーム毎に処理をする従来の手法に加え、そのスペクトログラムから画像認識のような特徴抽出<sup>[19]</sup>や類似度検索が必要<sup>[77]</sup>であり計算負荷が高くなることから、近年のコンピュータの高速化に伴い研究が進められていると考えられる。一方で、スペクトログラムによって雑音環境で収録された信号から少ないサンプル数のみで見本となる音の特徴の抽出法について検討したものはなく、本研究ではこの音の特徴抽出の方法について検討した。

まず前項の雑音低減手法と高域強調によってくしゃみ音のスペクトログラムが類似していることをもとに、特徴の抽出方法について述べる<sup>[74]</sup>。様々な環境で収録されたくしゃみ音に含まれる雑音のうち定常的な雑音については、前項の手法で低減されており残さの雑音については収録されたデータ間において無相関であると仮定する。一方でくしゃみ音のスペクトログラムの形状が似ていることから、このスペクトログラムの画像の相加平均や相乗平均を求める<sup>[75]</sup>ことで、くしゃみ音の形状に含まれる特徴を抽出することができるのではないかと考えた。今  $x_1$  から  $x_n$  までの  $n$  個のデータがあった場合の相加平均  $\bar{x}$  と相乗平均  $x_G$  は

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (3.6)$$

$$x_G = \sqrt{x_1 \times \dots \times x_n} = \sqrt[n]{\prod_{i=1}^n x_i}, \quad (3.7)$$

となる。このデータをスペクトログラムに拡張するため、あるスペクトログラムの画像を行列  $\mathbf{X}$  とすると、

$$\mathbf{X} = [x_{a,b}] = \begin{bmatrix} x_{1,1} & \cdots & x_{1,b} \\ \vdots & \ddots & \vdots \\ x_{a,1} & \cdots & x_{a,b} \end{bmatrix}, \quad (3.8)$$

ここで  $[x_{a,b}]$  の  $a$  及び  $b$  は、スペクトログラムの周波数軸と時間軸に対応するもので、共に正の値とする。実際の処理では、定常的な雑音成分の減算処理の際に負になることがあるため、負の値については1と置換して計算を行うこととした。よって式(3.6)と式(3.7)はそれぞれ、

$$\bar{\mathbf{X}} = \frac{\mathbf{X}_1 + \dots + \mathbf{X}_n}{n} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i, \quad (3.9)$$

$$\mathbf{X}_G = \sqrt{\mathbf{X}_1 \times \dots \times \mathbf{X}_n} = \sqrt[n]{\prod_{i=1}^n \mathbf{X}_i}, \quad (3.10)$$

することができる. ここで  $n$  は用いる見本となるくしゃみ音の個数 (以下 テンプレート数と呼ぶ) となる. また相加平均は, 一つのデータが著しく小さいまたは大きい際に相乗平均と比べ大きな誤差となりえる ( $\because \frac{a+b}{2} \geq \sqrt{a \cdot b}$ ) ことから, 本研究では式(3.10)の相乗平均による手法を採用した. 本手法によってスペクトログラムから特徴の抽出が可能であるかシミュレーションによって検討する.

シミュレーションには MathWorks 社の数値計算ソフト MATLAB 2016b を用い付属の音声(女性が matlab と発話している音声)と白色雑音を付加させ SNR(Sound Noise Ratio)を, 雑音と音声の信号パワーの比が 0 dB(SNR = 0 dB)と雑音を少し下げた 3 dB(SNR = 3 dB)で比較をした. なお, 用いるテンプレート数を 10 個とし雑音は繰り返し回数の度にランダムで付加した. 元の音声信号と SNR 0 dB でのスペクトログラムを Fig. 3.7 に, SNR 0 dB, SNR 3 dB での相乗平均による結果を Fig. 3.8 示す.

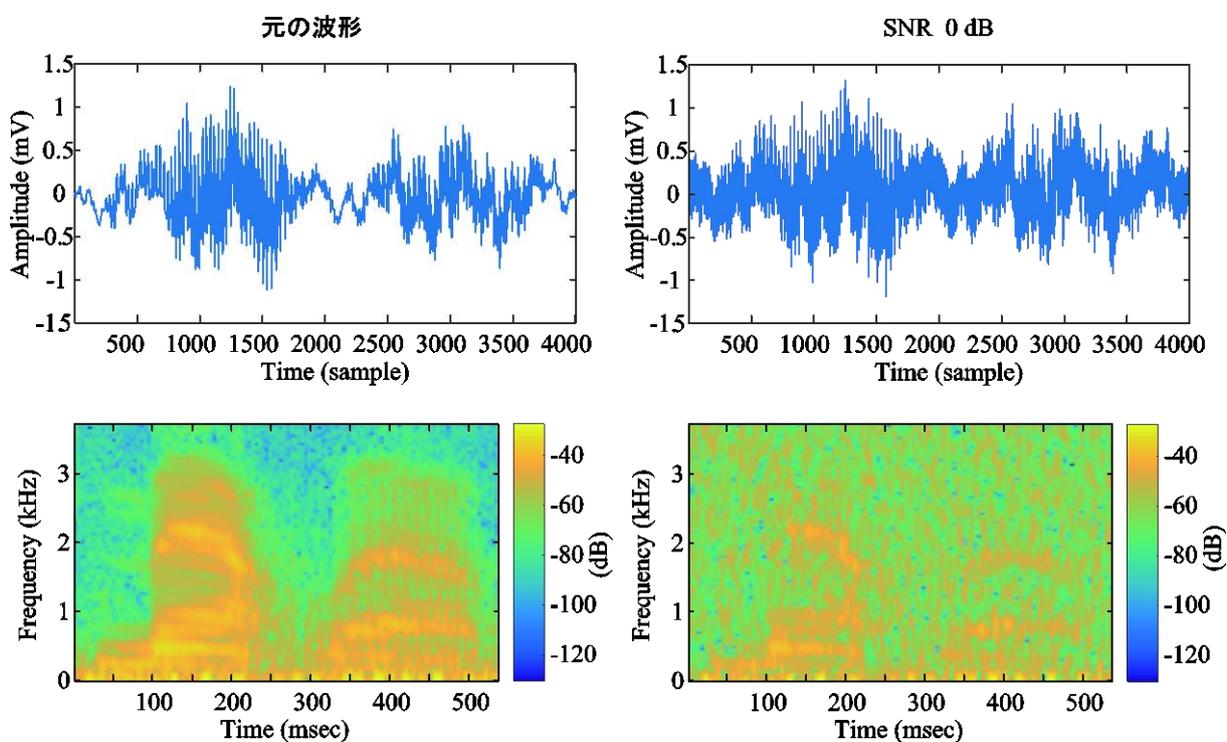


Fig. 3.7 音声信号と SNR 0 dB でのスペクトログラム

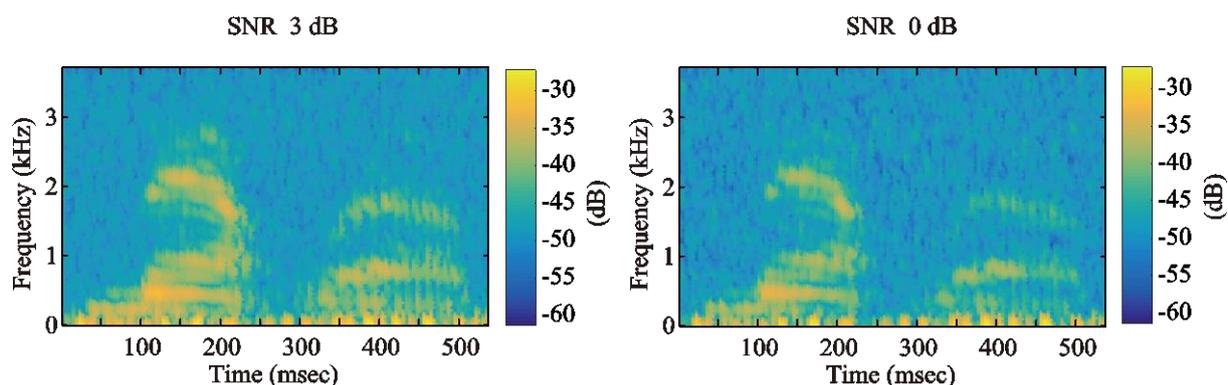


Fig. 3.8 SNR 0 dB 及び SNR 3 dB でのスペクトログラム相乗平均による結果

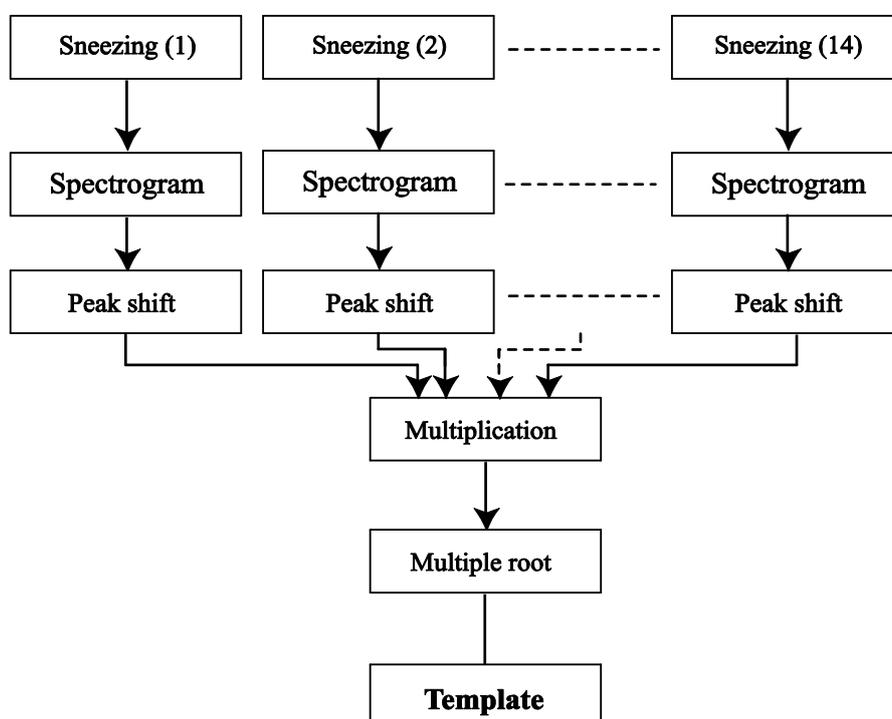


Fig. 3.9 くしゃみ音の見本データの作成に用いるフローチャート

この手法によって雑音に埋もれた音についてもスペクトログラムの形状を復元することが可能であることがわかった。

本手法をそれぞれの豚舎で確認されたくしゃみ音 14 個に対し適応のフローチャートを Fig. 3.9 に、その結果を Fig. 3.10 に示す。くしゃみ音をそろえるため、最大値及び音の開始時間について目視で確認し、くしゃみ音を合わせた。検出法ではこのくしゃみ音の相乗平均の結果を見本データとして扱うこととする。

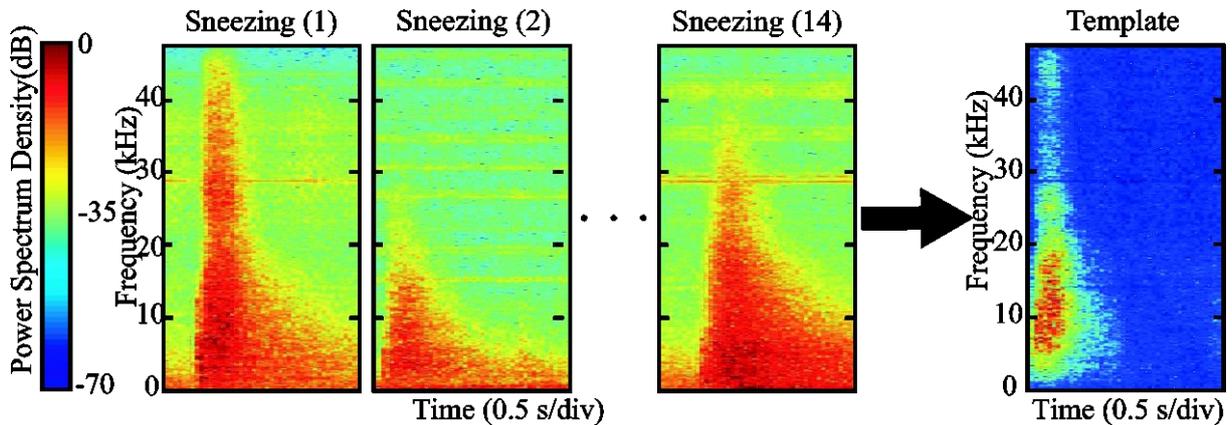


Fig. 3.10 14個のくしゃみ音によるスペクトログラム相乗平均による特徴抽出結果

この見本データと似ている信号を検出するため、スペクトログラム上の類似度を以下の計算<sup>[79]</sup>によって求める。検出対象となる収録音を切り出したスペクトログラムを式(3.8)の定義を適用し  $y_{a,b}$ 、見本となるくしゃみ音を  $x_{a,b}$  とすると、

$$C_{mac} = \frac{\sum \sum \{x_{a,b} \cdot y_{a,b}\}}{\sqrt{\sum \sum (x_{a,b})^2} \sqrt{\sum \sum (y_{a,b})^2}}, \quad (3.11)$$

として見本となるくしゃみ音と収録音での類似度を求めることができる。

### 3.2.3 異なる雑音環境下における検出実験と結果

本項では提案手法の有効性を評価するため、2つの豚房で収録した音源を用いて実験を行った。くしゃみ音の発生数をそろえるため、畜産草地研究所（豚房内 11 匹）でのデータ長を 30 min、全農飼料畜産中央研究所（豚房内 約 100 匹）のデータ長を 180 min 使用した。検出法のフローチャートを Fig. 3.11 に示す。まず、豚舎内で収録された音は 5 kHz 以上のハイパスフィルターによって人の声や豚の鳴き声を含む低周波の影響を減じる。これは、Fig.3.10の合成されたテンプレートにも 5 kHz 以下の情報が少ないことを考慮して決定した。次に収録された音を、周波数スペクトログラムに変換し、時間と周波数の情報を二次元の画像化を行った。この豚房の音の画像と、合成したくしゃみ音の画像で式(3.11)に基づく類似度計算を行いその値を  $C_{max}$  とした。画像で示される音の形状が似ていれば値が大きくなり、異なっていれば小さくなることで類似度を判定する。本実験では、正規化を行い 0-1 の範囲で類似度を判定できるようにし、類似度 0.9 以上であれば、識別機に入った信号がくしゃみ音であると判定できるとした。この値を高くすればするほど、雑音をくしゃみ音として誤識

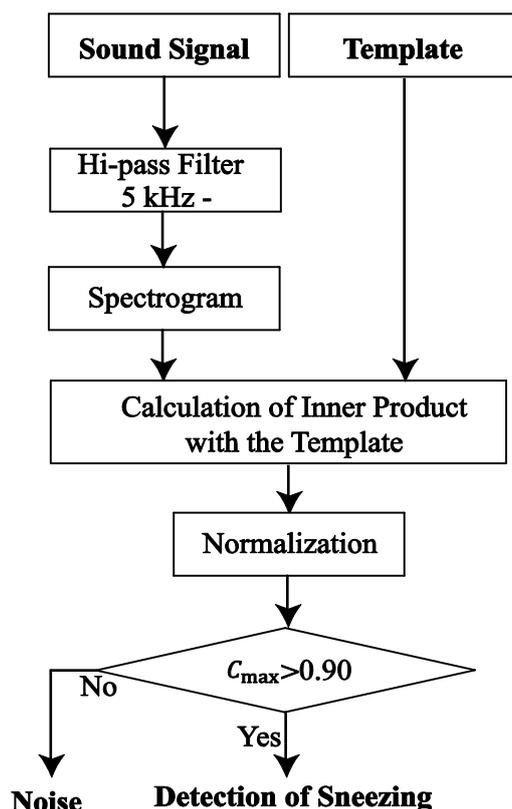


Fig. 3.11 検出法のフローチャート

別する確率は低下するが、小さい音や背景で別の音がしている際のくしゃみ音を検出できなくなる。逆に、低くすればするほど見逃すくしゃみ音が少なくなる一方で、誤識別の可能性があがるトレードオフの関係がある。本実験では、見逃しを少なくする設定を行った。また検出の切り出し時間を 0.5 s とし、切り出した時間にまたがる信号に対応するため 125 ms ずつ切り出す時間をシフトさせオーバーラップを 70 %とした。

**Table 3.2** および **Table 3.3** に実験結果について示す。まず **Table 3.2** では収録した区間中、くしゃみ音が含まれる区間は 14、雑音区間が 61924 であった。この区間は、収録の際に実際に人間がその場で聴音し、くしゃみ音の発生を確認し記録したものをくしゃみ音が含まれる区間と定義した。提案手法では、識別率が 13%、感度は 86%であった。識別率が低い原因として、パラメータの調整がくしゃみ音の見逃しを少なくする感度を上げる方向へ合わせたため、その分誤検出の回数が増加した。誤検出の原因として、くしゃみ音のスペクトログラムの形状に類似した信号が挙げられる。最も誤検出の原因となった音は、豚が金属の柵に当たった際に生じる金属音であった。**Fig. 3.12** にそのスペクトログラムを示す。高周波の特徴は異なるものの、2つのスペクトログラムは類似しており、本手法で閾値を変化させた場合にも区別することができなくなった。また、くしゃみ音 14 回を計測するため 3 時間程度収録したため、多くの雑音が発生しておりその分誤識別する回数 (83 回) が増えた。一方で **Table 3.3** では 30 分ほどで 15 回のくしゃみ音が発生しているため、発生した雑音も少なく 569 回

であり識別率は 53%、感度が 93%であり感染症の陰性・陽性の判定に必要な 50%以上の識別率および感度を満たしていることがわかった。

Table 3.2 畜産草地研究所のデータに対する実験結果

**Time : 3 hour**

		B		Total
		Sneezing	Noise	
A	Sneezing	(a) 12	(c) 2	14
	Noise	(b) 83	(d) 61841	61924
Total		95	61843	

A : Correct value

B : Result of value of the discrimination

Discriminative rate (  $a / ( a+b )$  ) : 13 %

Sensitivity (  $a / ( a+c )$  ) : 86 %

Table 3.3 全農飼料中央研究所のデータに対する実験結果

**Time : 30 minute**

		B		Total
		Sneezing	Noise	
A	Sneezing	(a) 14	(c) 1	15
	Noise	(b) 13	(d) 556	569
Total		27	557	

A : Correct value

B : Result of value of the discrimination

Discriminative rate (  $a / ( a+b )$  ) : 53 %

Sensitivity (  $a / ( a+c )$  ) : 93 %

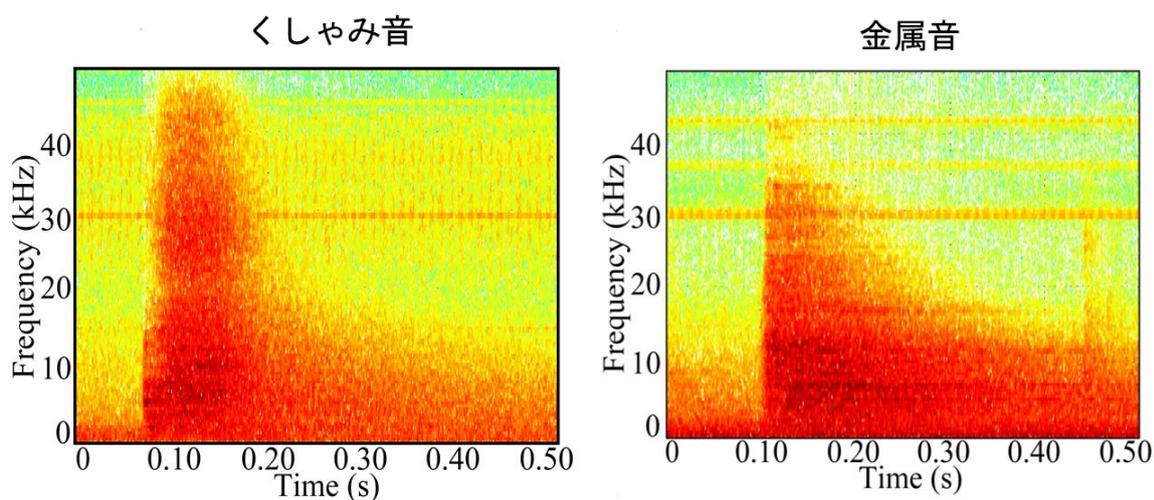


Fig. 3.12 誤識別の原因となった豚が金属柵にぶつかった音のスペクトログラム

### 3.2.4 まとめ

本節では、複数の豚くしゃみ音を用いる音源特徴の抽出と検出法として定常的な雑音の低減方法及びくしゃみ音の特徴である高周波の強調を行う手法を用い、複数の豚くしゃみ音からスペクトログラムの相乗平均によって特徴を抽出する手法を提案した。また検出法として特徴を抽出したスペクトログラムの画像を見本となるテンプレートとして用い、内積計算に基づく類似度の計算方法を提案した。異なる環境下である畜産草地研究所及び全農飼料中央研究所で収録した2つのデータに対し提案手法を適応した結果、全農飼料中央研究所では識別率は53%、感度が93%であり感染症の陰性・陽性の判定に必要な50%以上の識別率および感度を満たしていることがわかった。

一方で、畜産草地研究所では識別率が13%、感度は86%であり、この値は先行研究における感染症の陰性・陽性の判定に必要な識別率を満たしておらず、識別率を向上させる必要があることがわかった。識別率が低い原因となった音には、豚が金属の柵に当たった際に生じる金属音があり、検出法に用いるくしゃみ音と同じ形状のスペクトログラムであり、識別率の向上には、この金属音とくしゃみ音を区別する必要があることがわかった。また2つの収録データで見逃してしまったくしゃみ音の原因としては、ひとつは信号が小さく収録音で聞き取ることが難しくスペクトログラムに表れないものであった。これは、全農飼料中央研究所のくしゃみ音で1つ確認でき、発生が遠方であったことが原因であった。一方で距離が近い畜産草地研究所で収録したくしゃみ音では、信号も大きく高周波が含まれていたものの、シフト幅125msの75%オーバーラップのパラメータで計算した際、見本となるテンプレートと検出対象のくしゃみ音が僅かにずれたため、類似度の値が低く算出されることが原因であった。検出の際にオーバーラップさせるシフト幅を細かくすることで僅かなずれには対応できるが、計算量が増大しリアルタイム性が失われるため類似度の算出法についても検討する必要がある。

### 3.3 特定音の検出精度向上の手法と実験

#### 3.3.1 DP マッチングと動的特徴量を用いる検出法の原理

くしゃみ音の検出精度を上げるため、誤検出の原因となった豚が柵に当たる金属音とくしゃみ音の識別方法について検討すると共に、見逃しの原因でもあった見本となるテンプレートとのずれに対応する手法について検討する。本節では僅かなずれである時間変化に対応するため文字列の類似度や音声認識に用いられる DP マッチング(Dynamic Programming)[80]を類似度の計算に応用することを検討する。また、豚くしゃみ音は声帯から発する音声と考えらるが人の声のような特徴的なフォルマントは確認できない。これは、声帯振動の声と異なり鼻や口などを抜ける有声破裂音と同様の構造をしていると考えられる。ここで、音声認識分野において有声破裂音の特徴を表現できるとされる動的特徴量<sup>[81]</sup>を検出の特徴のとして用いることを検討する。

本項では、まず生体の発するくしゃみ音と豚が柵に当たった際に生じる金属音を区別するため、動的特徴量のスペクトログラムへの導入を説明する。その後、類似度の計算方法である DP マッチングについて述べる。

動的特徴量とは、1986年に提案された音声認識のための特徴抽出法の一つである<sup>[82]</sup>。

**Figure 3.13**のように、ある時刻  $t$  における周波数とパワーを時間  $T$  づつ区切ったものを静的特徴と呼び、 $f_1$  と  $f_2$  における矢印が示すように各々の時間軸方向の周波数変動を動的特徴と呼ぶ。例えば、 $f_1$  の周波数を見ると、時間軸方向に減少していく傾向がある。時刻  $t$  と  $t + T$  における頂点を結ぶ直線を描くとき、直線の式を周波数に対応するパワーを  $y$  として目的変数とし、時間軸に対応する説明変数を  $x$  とすると  $y = ax + b$  と示すことができる。このとき  $f_1$  の  $a$  は負の値を示す。一方で  $f_2$  は一定であるため、 $a$  は概ね 0 付近であることが予想できる。この傾きを示す  $a$  を動的特徴量またはデルタ特徴量と呼ばれ、第 2 章の音声認識に研究ではメルケプストラム係数などに適応し、メルケプストラム係数と共に特徴量として用いることで高い認識率を示している。周波数を時間軸方向の変動として捉えることで、音声認識のための特徴として用いることができる。この動的特徴は子音や有声破裂音などをよく表現できるとされており<sup>[81]</sup>、主要な周波数が見られないくしゃみ音に有効であると考えられる。

一方で Fig. 3.13 のように間隔  $T$  で 4 つの周波数とパワーを各々のフレームとして考えると、何フレーム毎にこの傾きを定めるかを検討する必要がある。複数フレームで動的特徴量である傾きを求める場合には、単回帰分析による手法が一般的である。最小二乗法による傾きの求め方について以下に示す。今、ある周波数  $f$  の  $i$  フレームにおける動的特徴量である傾き  $a_f$  を求めるとする。1 から  $i$  フレームの周波数に対応するパワーを  $y_1, y_2, \dots, y_i$  として目的変数とし、スペクトログラムの時間軸に対応する説明変数を  $x_1, x_2, \dots, x_i$  とすると、各々の点と描く直線との距離  $D_f$  を最小とするため、

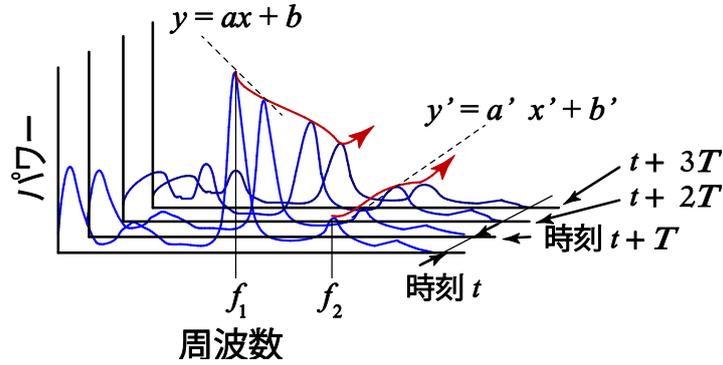


Fig. 3.13 音の静的特徴と動的特徴

$$D_f = \sum_{n=1}^i \{ y_n - (a_f x_n + b_f)^2 \}, \quad (3.12)$$

式(3.12)に、 $a_f$ と $b_f$ の偏微分条件を $\frac{\partial D_f}{\partial a_f} = 0, \frac{\partial D_f}{\partial b_f} = 0$ を用いると、それぞれは

$$b_f = \frac{\sum_{n=1}^i y_n - a_f \sum_{n=1}^i x_n}{i}, a_f = \frac{\sum_{n=1}^i x_n y_n - \frac{1}{i} \sum_{n=1}^i x_n \sum_{n=1}^i y_n}{\sum_{n=1}^i x_n^2 - \frac{1}{i} \sum_{n=1}^i x_n \sum_{n=1}^i x_n}, \quad (3.13)$$

と求めることができる。動的特徴量をスペクトログラムにおいて適応するためには、出力されたスペクトログラム画像の各々の周波数における時間軸方向の成分を取り出し、指定した各フレーム数で式(3.13)を適応することで周波数毎の動的特徴量を算出することができる。

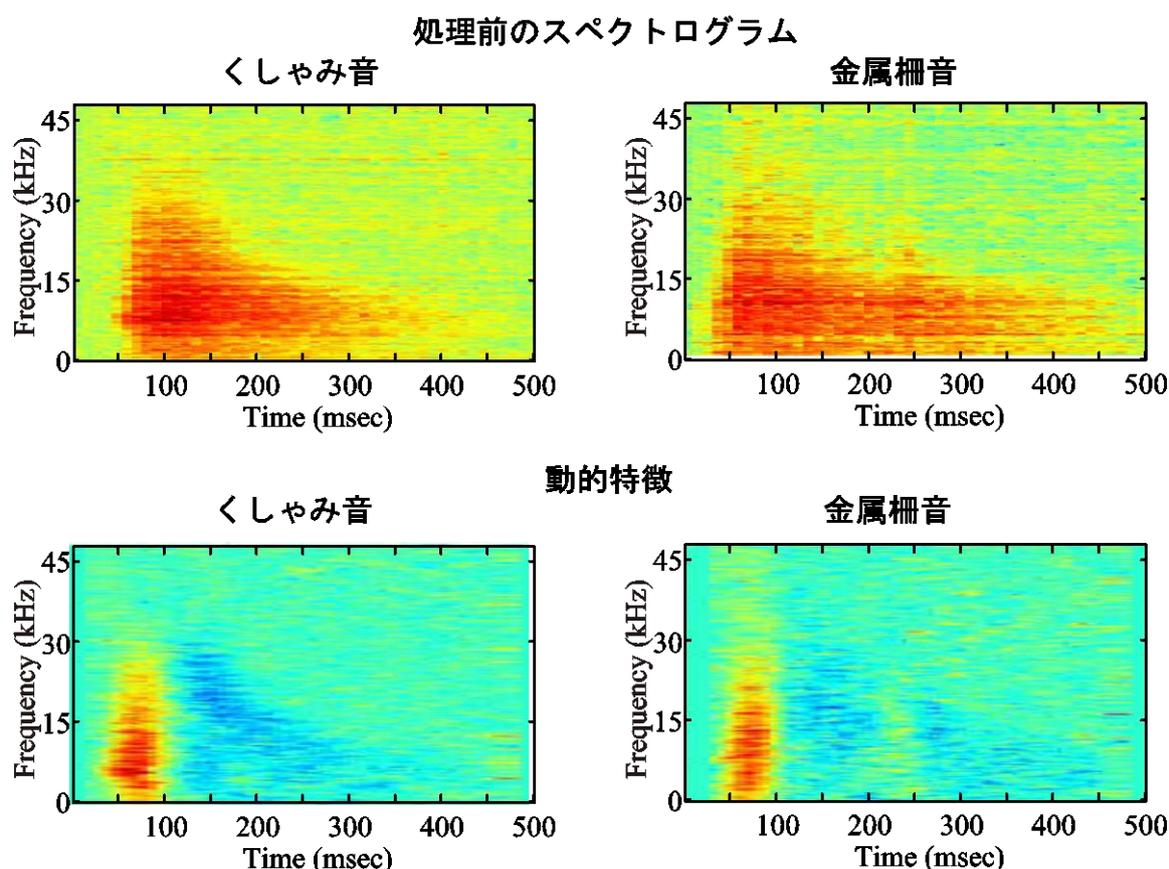


Fig. 3.14 くしゃみ音と金属柵音の動的特徴

ここで **Figure 3.14** に定常雑音を除去後のくしゃみ音と豚が柵に当たった音（以下 金属柵音と呼ぶ）へ動的特徴を適応した結果について示す．なお，誤検出の原因の音のうちスペクトログラムでの類似度が最も高くしゃみ音と金属柵音を示した．動的特徴の正の値を赤，負の値を青によって示す．**Figure 3.14** を見ると，処理前のスペクトログラムは大変似ていることがわかる．一方で動的特徴をみると 100 ms 付近の傾きが正となる区間はほぼ同じような特徴であるものの，250 ms 付近ではくしゃみ音が負の値で徐々に減衰する一方で，金属音は減衰が収まり一定の時間が存在する．この動的特徴によって金属柵音などの音では音の減衰率が異なることから，この特徴を用いることでくしゃみ音との識別が可能であることがわかった．

次に類似度の計算方法として DP マッチングについて述べる．この手法は時間伸縮された音や信号，文字列などのパターンマッチングの手法として用いられてきた．原理としては 1953 年に数学的最適化の手法として提案され<sup>[83]</sup>，近年の統計的な音のモデル化や混合ガウスモデルによる手法以前では見本データとの類似性を測る手法として多く用いられてきた．ここでスペクトログラムへの適応方法について例と定式化の後に示す．

まず簡単のため文字列での例として、**Fig. 3.15**に見本データと収録信号の類似度を示す。見本データとしてくしゃみ音である「ああくしゅん」、収録信号で「いぶしゅんい」のような信号が入力されたとする。この際「あ」や「い」は雑音だとする。この2種類を時間軸に並べ直交させると図のように示すことができる。それぞれの信号で共通な信号は「しゅ」と「ん」である。DP マッチングでは、青い矢印で示す経路の最短距離を求めることで類似度を計算することができる。ここでそれぞれの信号を升目状で示し、不一致の場合は-3のペナルティを与え、一致した場合には0を与えるとするとの**Fig. 3.16**の入場ペナルティのように示すことができる。上記の例の場合、一致する信号である「しゅ」と「ん」については0となり他の信号には-3が不一致の入場ペナルティとして与えられる。雑音や見本となる信号が一致した場合には図で示すように**Fig.3.15**のような斜めの経路で一致した値となる0が代入されていることがわかる。次に、時間軸の開始地点から矢印の終端である時間軸の終了地点までの移動経路について考える。

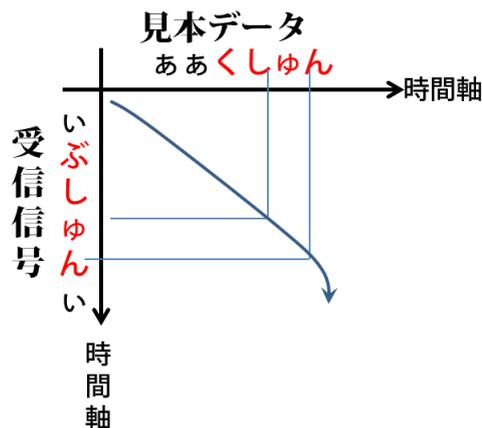


Fig. 3.15 DP マッチングでの見本データと受信信号の類似度

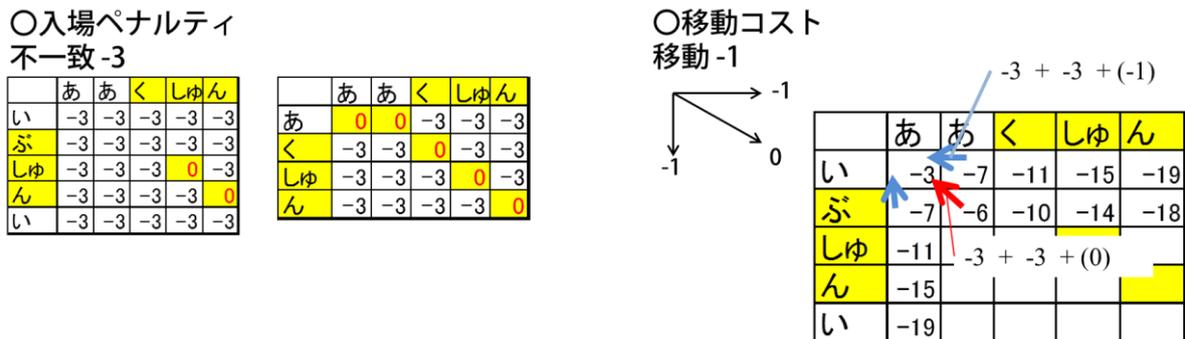


Fig. 3.16 DP マッチングの入場ペナルティと移動コスト

	あ	あ	く	しゅ	ん
い	-3	-7	-11	-15	-19
ぶ	-7	-6	-10	-14	-18
しゅ	-11	-10	-9	-10	-11
ん	-15	-14	-13	-12	-10
い	-19	-18	-17	-16	-13

Fig. 3.17 DP マッチングの最短経路の計算結果

DP マッチングでの移動経路は Fig. 3.16 の移動コストとして考えることができる。これは、斜め方向の時間軸に沿った昇目の移動であれば移動にかかるコストは0であり、反対に見本データの時間軸方向の移動や、収録信号方向の移動についてはペナルティと同様に-1のコストを加算するものとする。Figure 3.16 を例に具体的に計算をすると、時間軸の開始地点の「あ」と「い」に対応するブロックでは、不一致のため最初にペナルティである-3が与えられ、収録信号の時間軸方向の「あ」と「ぶ」に対応する値を求める際には、下方方向の移動コストが加算され次式のように-7と計算できる、

$$-7 = (-3)[\text{不一致}] + (-3)[\text{不一致}] + (-1)[\text{移動コスト}], \quad (3.14)$$

本式の2つの不一致は移動前の入場ペナルティと移動先のブロックでの入場ペナルティとなる。さらに、収録時間方向に進んだため、移動コストが付加され-7の結果が代入される。同様に、見本データの時間軸方向にも不一致と入場ペナルティから-7が代入されることとなる。次に斜め方向を計算すると、

$$-6 = (-3)[\text{不一致}] + (-3)[\text{不一致}] + (0)[\text{移動コスト}], \quad (3.15)$$

となり、この3つの経路での最短経路は赤の矢印方向である斜め方向となる。この式(3.14)及び式(3.15)を繰り返し計算することによって、ブロック全体の数値を計算することができ、赤い矢印で示す Fig. 3.17 のような最短経路を示すことができる。このとき一番右下となる時間軸方向の終点地点の-13という値が小さいほど類似度が高いとして算出ができる。

今式(3.8)のように収録信号  $\mathbf{X}$ 、見本データを同様に  $\mathbf{Y}$  とし、それぞれの信号の終端時刻を  $I, J$  とし、時刻毎のスペクトログラムの周波数の値をベクトルとして、時系列のベクトルで表現すると、

$$\mathbf{X} = \mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_I, \mathbf{Y} = \mathbf{y}_1 \mathbf{y}_2 \cdots \mathbf{y}_J, \quad (3.16)$$

と示すことができる. ここで最終的に求めたい二つの信号の類似度である距離を  $D_c(\mathbf{X}, \mathbf{Y})$  とし、式(3.16)の時系列毎の周波数に対応するベクトル間の距離を、

$$d(i, j) = \|\mathbf{x}_i - \mathbf{y}_j\|, \quad (3.17)$$

とすると、以下の式を計算することで最短経路を計算することができる.

$$g(1,1) = d(1,1), \quad (3.18)$$

$$g(i, j) = d(i, j) + \min \begin{bmatrix} g(i-1, j) \\ g(i-1, j-1) \\ g(i, j-1) \end{bmatrix}, \quad (3.19)$$

$$D_c(\mathbf{X}, \mathbf{Y}) = g(I, J), \quad (3.20)$$

式(3.18)は、開始地点の初期条件であり、式(3.19)はブロックを移動する3方向を示している. 式(3.19)を計算していき終端時刻まで計算した際の値が収録信号間の類似度として出力される結果となる. このようなアルゴリズムを計算することで、収録信号と見本データとなる信号の類似度を時間軸が多少ずれた状態でも出力できる.

本手法の有効性を検証するため、3.2.2節での音声を用いシミュレーションを行った. プログラムは D. Eillis 氏の発表論文[84]及びコロンビア大学の研究室サイト[85]を基に作成を行った. 評価対象は類似度を比較するため、見本データを付属の音声とし収録データとしてサイン波を用いた結果を **Fig. 3.18** に示す. また、同様の見本データで付属の音声に白色雑音を付加した音声及び、時間軸のずれに対応できているかを検討するため、収録信号の時間軸を変えた信号について提案手法を適応した. それぞれの結果を **Fig. 3.19** 及び **Fig. 3.20** に示す. なお、DP マッチングによる最適経路の計算結果を赤線で示した. この距離が類似度の指標となる. なお DP マッチングの結果である明度は入場ペナルティと移動コストが加味された結果が出力されており、この明度が高い場所を結んだものが最短経路として計算されている.

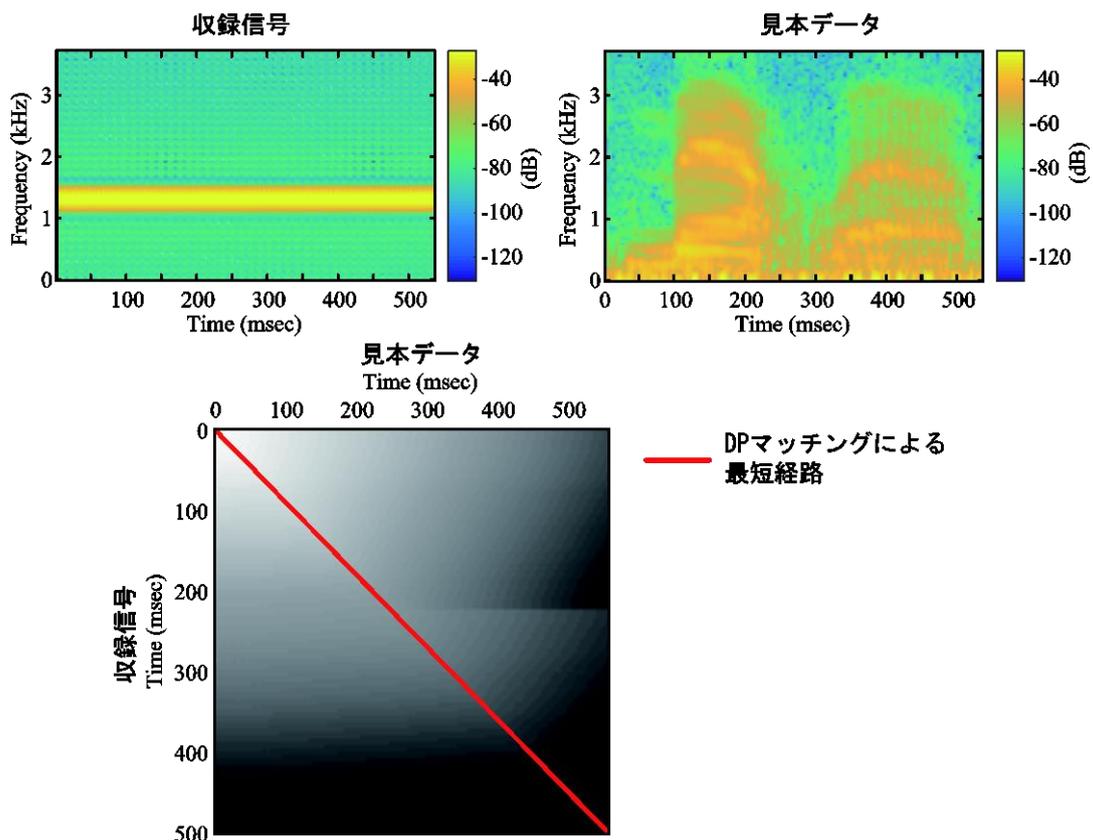


Fig. 3.18 音声とサイン波の DP マッチングの結果

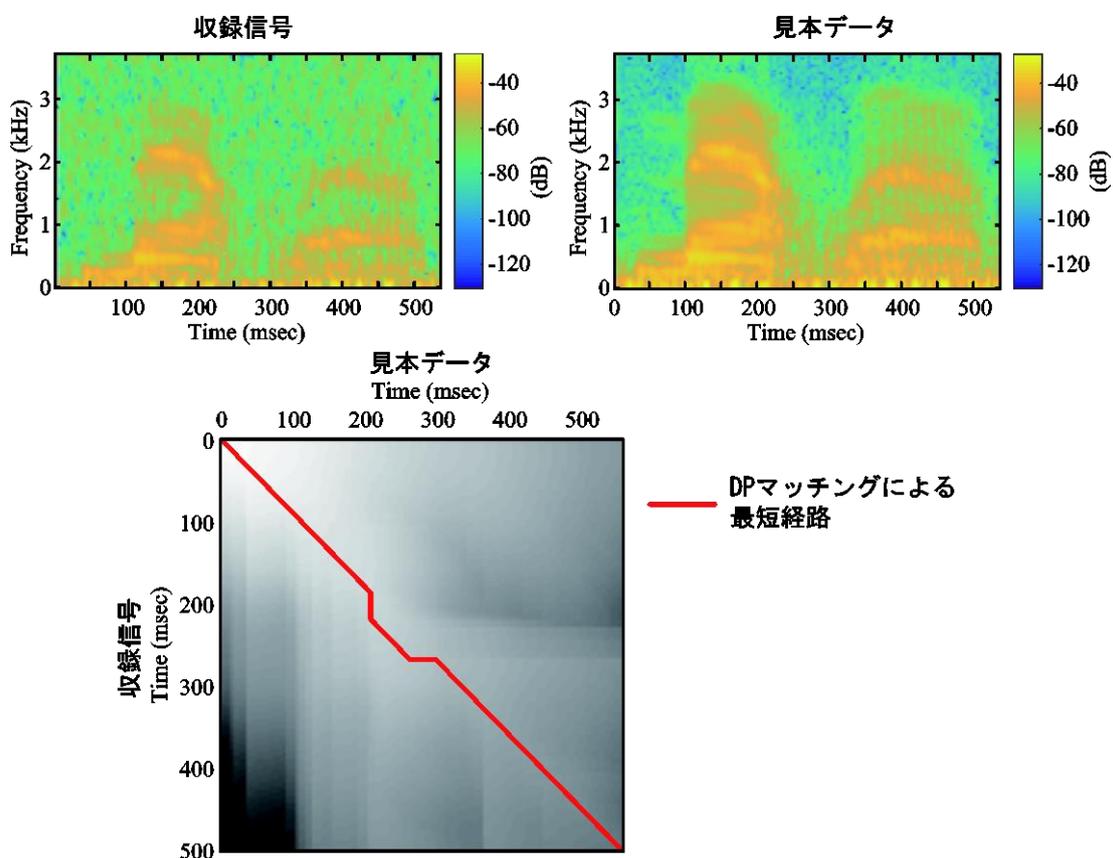


Fig. 3.19 音声と雑音付加音声の DP マッチングの結果

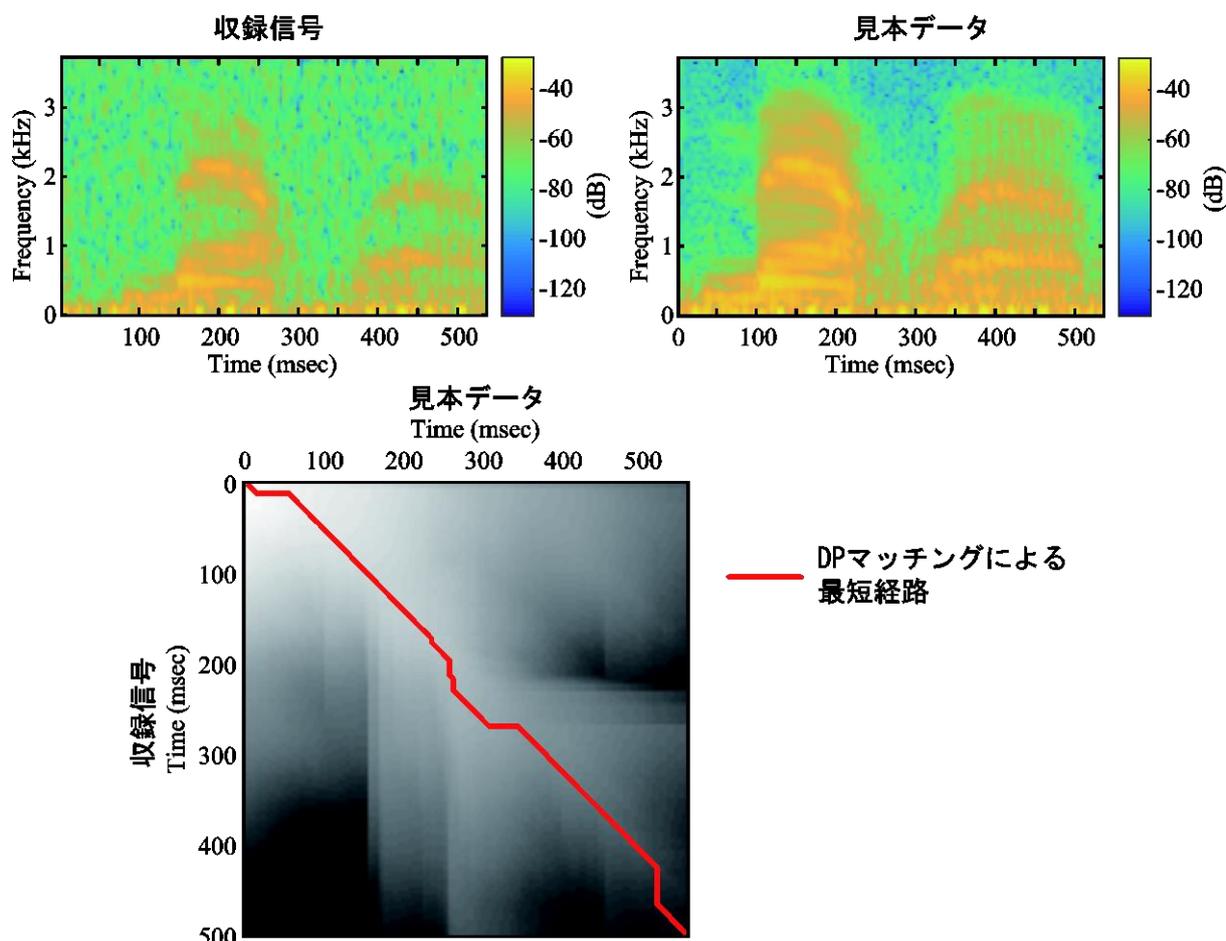


Fig. 3.20 音声と時間変化のある雑音付加音声の DP マッチングの結果

Table 3.4 類似度の計算結果

白色雑音	3.19
サイン波	3.87
雑音付加音声	1.34
時間変化のある雑音付加音声	1.39
同一信号	$4.1 \times 10^{-13}$

なお、時間変化のある雑音付加音声には 50 ms のずれを行い、見本データと時間の差異を発生させた。それぞれの類似度の計算結果と白色雑音のみ、同一音での実験結果を、Table 3.4 に示す。まったく別の信号である白色雑音やサイン波のみでは値は 3 以上となり、見本となる音声信号と収録信号におけるパターン間における距離が長いことがわかる。反対に雑音が付加された音声とそれをずらした音声についても 1.3 前後の結果となり、別の信号と比べパターン間の距離が短いことがわかった。また見本となる音声信号と同一信号で比較した結果

$4.1 \times 10^{-13}$  と 0 に近い値となり、類似度の指標としてスペクトログラムに対して使用可能であることがわかった。また Figure 3.20 の最短経路を示す赤線を見ると、開始地点と終端付近に時間変化に対応する経路の変化が見れとれる。この変化によって 0.05 程度の増加で類似度を計算できていることがわかる。なお、サイン波の赤線よりも最短経路は他の 2 つの経路が長く見えるが、実際には白黒の濃淡に示す入場ペナルティの大きさが加味されており、経路とその入場ペナルティの合計から類似度が計算されるため、見かけ上は短く見えるが濃淡を見ると、白のペナルティが少ない最短経路を選んでいることがわかる。

### 3.3.2 提案手法による検出実験と結果

本項では提案手法の有効性を評価するため、前節と同じ 2 つの豚房で収録した音源を用いて実験を行った。くしゃみ音の発生数をそろえるため、畜産草地研究所 (NILGS) (豚房内 11 匹) でのデータ長を 30 min, 全農飼料畜産中央研究所 (CRIFL) (豚房内 約 100 匹) のデータ長を 180 min 使用した。検出法のフローチャートを Fig. 3.21 に示す。前項との違いとして、特徴抽出法に動的特徴量を導入しスペクトログラムの類似度比較と動的特徴量の比較と 2 段階の類似度判定となったことに加え、それぞれの類似度の算出には前節の DP マッチングによる手法によって算出されるものとした。閾値については、それぞれの最大値で正規化した値を用い 2 つの検出結果を比較しつつ感度及び識別率が共に最大となるように設定した。なお、提案手法の汎用性を検証するため、前節と同じくそれぞれの収録データに対するパラメータは同一の条件で実施した。

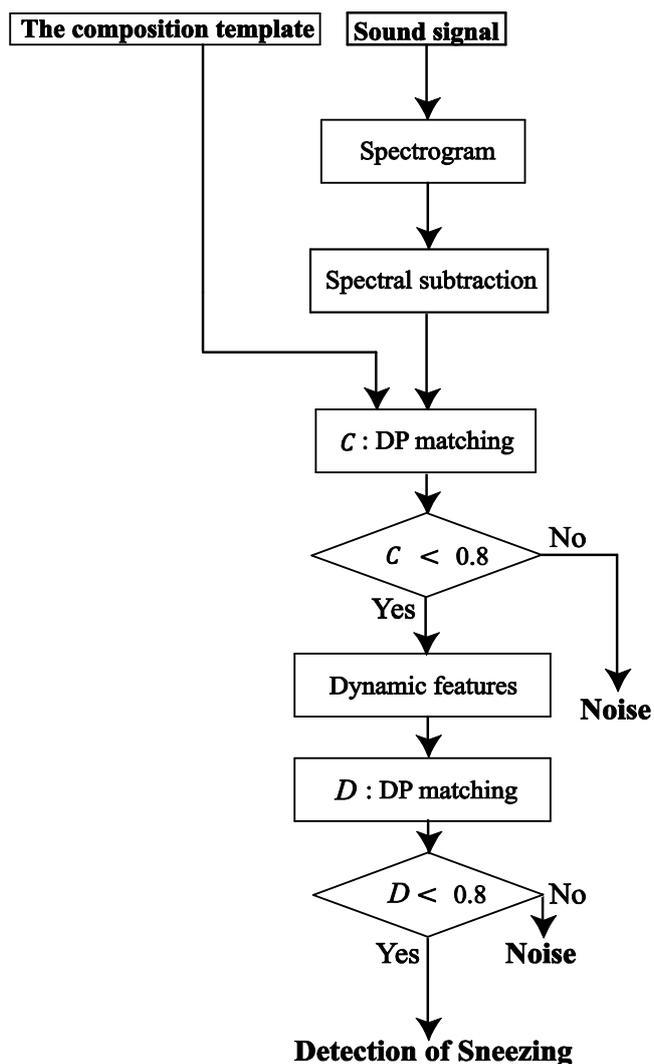


Fig. 3.21 検出法のフローチャート

**Table 3.5** および **Table 3.6** に実験結果について示す. まず **Table 3.5** では収録した区間中, くしゃみ音が含まれる区間は 14, 雑音区間が 61924 であり, 提案手法では, 識別率を従来法の 13% から 54% に大幅に改善することができた. これは, 雑音の原因となっていた金属柵音とくしゃみ音を動的特徴によって識別できたことから, 誤検出の原因を減らすことができたと考えられる. さらに前節の手法では見逃していた 1 つのくしゃみ音を検出することができ感度 93% となった. これは, 前節での見逃しが, 内積計算であったため, 切り出した時間によってずれが生じており対象のくしゃみ音で類似度が上がらなかったためだと考えられる. さらに切り出しのオーバーラップを多くさせることで対処できる可能性があるが解析時間が長くなる短所があり, 時間変化に対応可能である提案手法が有効であることが示された. **Table 3.6** も同様に識別率は 60%, 感度 93% ととなり, 識別率を 7% 向上させることができた. よって提案手法が異なる豚舎環境でも有効であり, どちらの豚舎でも感染症の陰性・陽性の判定に必要な 50% 以上の識別率および感度を満たしていることがわかった.

**Table 3.5** 畜産草地研究所のデータに対する実験結果

**Time : 3 hour**

		B		Total
		Sneezing	Noise	
A	Sneezing	(a) 13	(c) 1	14
	Noise	(b) 11	(d) 61913	61924
Total		24	61914	

A : Correct value

B : Result of value of the discrimination

Discriminative rate (  $a / ( a+b )$  ) : 54 %

Sensitivity (  $a / ( a+c )$  ) : 93 %

**Table 3.6** 全農飼料中央研究所のデータに対する実験結果

**Time : 30 minute**

		B		Total
		Sneezing	Noise	
A	Sneezing	(a) 14	(c) 1	15
	Noise	(b) 8	(d) 561	569
Total		22	564	

A : Correct value

B : Result of value of the discrimination

Discriminative rate (  $a / ( a+b )$  ) : 64 %

Sensitivity (  $a / ( a+c )$  ) : 93 %

#### 3.3.3 まとめ

本節では、前節の問題点であった、誤検出の原因となる金属柵音とくしゃみ音の識別方法として動的特徴をスペクトログラムに導入し、その識別を試みた。また僅かな時間変化によって前節の内積計算の手法では類似度が算出できない問題があったことから DP マッチングによる手法によって、時間変化に対応する類似度算出法を検討し、スペクトログラム上での類似度の算出方法を提案した。前節の複数の豚くしゃみ音からのスペクトログラムでの特徴抽出に加え、動的特徴の導入と類似度の計算方法を DP マッチングによる手法を取り入れた提案手法について、異なる環境下である畜産草地研究所及び全農飼料中央研究所で収録した 2 つのデータに対し提案手法を適応した結果、畜産草地研究所のデータでは識別率は 54%、感度が 93%であり、前節の従来手法に比べ誤検出率を 31%、見逃し率を 7%向上することができた。また全農飼料中央研究所のデータに対しても、感度 93%と識別率 64%の結果となり、誤検出の原因を低減することができた。よって提案手法はどちらの豚舎においても感染症の陰性・陽性の判定に必要な 50%以上の識別率および感度を満たしていることがわかった。

一方、問題点として動的特徴の導入にも関わらずくしゃみ音と誤検出してしまった音は動的特徴も同様であるような金属柵音や、水しぶきの音と金属柵音が同時に発生した際などの音であった。前述の金属柵音は人が聴いてもくしゃみ音か判断することができず、さらなる特徴を検討しなければ検出が難しいものと考えられる。また同時に発生しマイクロフォンに入る音についてはマイクロフォンのアレイ化によって方向推定と組み合わせることで今後分離できる可能性がある。また豚舎に設置する際の注意点として、マイクロフォンの設置場所や豚舎での雑音環境によって豚舎毎にパラメータを調整することが必要であるとともに、今回の手法で用いた相乗平均によって算出された見本となるテンプレートの種類や収録された環境によっては結果が異なる可能性には十分注意が必要である。

### 3.4 まとめ

本章では、異なる雑音環境下における特定音の検出として畜産草地研究所と全農飼料中央研究所の2か所の豚舎において収録した音からくしゃみ音を検出する手法について検討した。検出に使用するマイクロフォンを1つとしてビームフォーミングなどの手法を用いず雑音環境下で収録された見本となるデータから特徴を抽出し、異なる豚舎でも特定の音を検出できる手法を目指し2つのことを実施した。まず複数の豚くしゃみ音を用いる音源特徴の抽出と検出法として定常的な雑音の低減方法及びくしゃみ音の特徴である高周波の強調を行う手法を用い、複数の豚くしゃみ音からスペクトログラムの相乗平均によって特徴を抽出する手法を提案した。また検出法として特徴を抽出したスペクトログラムの画像を見本となるテンプレートとして用い、内積計算に基づく類似度の計算方法を提案した。しかし、内積計算に基づく手法では時間変化を一致させない場合、類似度を算出することができなかった。また、スペクトログラムの画像を見本データとして使用したところ、スペクトログラムの形状が類似している豚が柵に当たる金属音などによって誤検出が発生し、全農飼料中央研究所では識別率は53%、感度が93%であるものの、畜産草地研究所では識別率が13%、感度は86%であり、この値は先行研究における感染症の陰性・陽性の判定に必要な識別率を満たしておらず、識別率を向上させる必要があることがわかった。

そこで、誤検出の原因となる金属柵音とくしゃみ音の識別方法として動的特徴をスペクトログラムに導入し、その識別を試みた。また僅かな時間変化によって前節の内積計算の手法では類似度が算出できない問題があったことからDPマッチングによる手法によって、時間変化に対応する類似度算出法を検討し、スペクトログラム上での類似度の算出方法を提案した。同一の収録データを用い実験の結果、畜産草地研究所のデータでは識別率は54%、感度が93%であり、前節の従来手法に比べ誤検出率を31%、見逃し率を7%向上することができた。また全農飼料中央研究所のデータに対しても、感度93%と識別率64%の結果となり、誤検出の原因を低減することができた。よって提案手法はどちらの豚舎においても感染症の陰性・陽性の判定に必要な50%以上の識別率および感度を満たしていることがわかった。

今後の展望として、マイクロフォンの設置場所や豚舎での雑音環境によって豚舎毎にパラメータを調整する必要があるため、距離や環境に柔軟に対応できる検出法を検討するとともに、既存の手法であるマイクロフォンのアレイ化による音の強調や分離方法によってさらなる検出精度の向上が期待される。

## 第4章 少数マイクロフォンによる音源位置推定

### 4.1 はじめに

前章までにおいて、様々な環境下において特定の音を検知する方法について述べた。本章では検知した音に対し、どの位置で発生した音源なのかを音によって推定する音源位置推定の手法と原理について述べる。音を検出するセンサであるマイクロフォンは、1つではその場の音圧を電圧に変換する機能のみであり、どの位置で発生した音なのかを判別することは難しい。そこで複数のマイクロフォンを用い、それぞれのマイクロフォンで受信された信号の時間差や周波数上の差異を用いることで位置を推定する音源位置推定法<sup>[86]</sup>が提案されている。マイクロフォンを複数運用する際、そのセンサはマイクロフォンアレイと呼ばれマイクロフォンの数や設置間隔などによって特性が変わり、入力信号に対する処理も異なることからアレイ信号処理分野として多くの研究が実施<sup>[87]</sup>されている。音源位置推定法の代表的な手法としてビームフォーミング法<sup>[88]</sup>、MUSIC法<sup>[89,90]</sup>、相互相関法<sup>[91]</sup>、白色化相互相関法<sup>[92]</sup>、サブバンドピークホールド法<sup>[93]</sup>などがある。ビームフォーミング法は、多くのセンサを用いることで複数の到来時間差から音源方向推定精度を高めているため、設置機材の大型化が問題となる。MUSIC法は、受信信号を空間相関行列で表現し、固有ベクトルの張る空間へ変換することで雑音との直交性を用いることで高い分解能で音源の方向推定を行えるが、計算処理が多い問題がある。相互相関法は伝搬時間差を測定する手法としては最も基本的であり、相互相関関数から伝搬時間差を測定するため、計算処理は前述の手法と比較して少なく、リアルタイム処理が可能である。相互相関法の問題点としては、反射音の影響により推定誤差が大きくなる場合挙げられる。これに対して白色化相互相関法およびサブバンドピークホールド法は反射音耐性を持つとされているが、相互相関法と比較すると計算処理量が多い。そこで、本研究では相互相関法を用い到来時間差を計算することとした。

また複数のマイクロフォンによって位置推定を行う場合、従来は多くのマイクロフォンを必要としてきた。室内の壁や天井に約30個のマイクロフォンを設置し、その部屋の発話者の位置を推定する研究<sup>[94]</sup>や、12個のマイクロフォンを用い異常音の発生位置や音を強調する手法<sup>[95]</sup>、約60個のマイクロフォンを用い移動音源の音声認識<sup>[96]</sup>などが研究されている。しかし、リアルタイムでの位置推定には演算処理や機材コストから少数のマイクロフォンが適していると考えられ、少数で正確に位置推定が可能となれば上記の音声強調などの分野に貢献が可能であると感られる。そこで本研究では位置の推定に必要な最小個数のマイクロフ

オンである4つのマイクロフォンを用いカメラとマイクロフォンでの位置推定や、反射波を用いる位置推定法について検討を行う。また立体構成にした4つのマイクロフォンを持つマイクロホンアレイを作成し、2組のみ用い屋外において連続音に対する音源位置推定法について検討を行った。

## 4.2 原理

### 4.2.1 相互相関法による到来時間差の計算方法

本節では、音から位置推定するための原理となる到来時間差の計算方法及び到来時間差からの位置の計算方法について述べる。本節の原理を基に次節以降の計算方法や応用について検討を行う。

**Figure 4.1** に音源から音波が到来し、2つのマイクロフォンによって受信する様子について示す。音源とマイクロフォンが十分に離れているとすると、マイクロフォンで受信された信号は平面波と近似できる<sup>[86]</sup>ことから図のように示すことができる。今音源での信号を  $s(t)$  とするとマイク1とマイク2のセンサでの受信信号はそれぞれの到来時間  $\tau_1$ ,  $\tau_2$  を用いて

$$\begin{aligned} x_1(t) &= s(t - \tau_1), \\ x_2(t) &= s(t - \tau_2). \end{aligned} \tag{4.1}$$

と示すことができる。ここで相互相関関数  $R_{cc}$  を次式で定義する。

$$R_{cc}(\tau) = \sum_{t=0}^T x_1(t)x_2(t + \tau), \tag{4.2}$$

$$T_{DOA} = \arg \max_{\tau} (R_{cc}(\tau)), \tag{4.3}$$

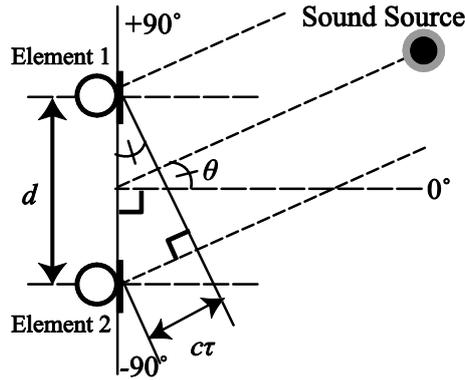


Fig. 4.1 到来時間差と音源方向

ここで  $\tau$  は時間指数であり、それぞれのマイクでの到来時間差( $\tau=|\tau_1 - \tau_2|$ )となる。  $T$  は信号長である。式(4.2)のように相互相関関数  $R_{cc}$  が最大値となる  $\tau$  とは、センサでの受信信号が一致する時間差であり、この時間指数を到来時間差  $T_{DOA}$  とする。相互相関法では受信信号の式(4.1)に含まれる音源信号  $s(t)$  が時間差をもって到来し、その位相差が一致する箇所が最大となり、信号の類似度判定にも使用されている。しかし、それぞれのセンサで受信される振幅に影響され、一方のセンサの振幅が大きい場合などは最大値が正しい値をとらず誤差の原因になることが考えられる。またセンサ間の距離  $d$  を大きくした場合や計測時間などによっては、周波数によってはセンサ間に数波長含むことで、相互相関関数のピークが複数現れるサイドローブの影響も考慮する必要がある。適切な設計やサイドローブでも目的の信号からのピークを見失わない手法などが必要となる。

#### 4.2.2 到来時間差からの方向・位置の計算原理

前述の原理によって到来時間差を計算することで、マイクロフォンアレイから音源に対する方向や位置を計算することができる。方向推定は Fig. 4.1 に音源から到来した平面波を2つのマイクロフォンで受信する場合を示す。マイク間の到来時間差を  $\tau$  とし、このときの音速を  $c$  とすると到来時間の距離差は  $c\tau$  となり幾何学的な関係により音源の方向である  $\theta$  は、

$$\theta = \sin^{-1}\left(\frac{c\tau}{d}\right), \quad (4.4)$$

によって求めることができる。この時音速  $c$  とセンサ間距離  $d$  は既知でアレイの音源の方向を計算することができる。次に Fig. 4.2 に音源から到来した球面波を4つのマイクロフォンで受信する場合について示す。マイク3を基準とした場合、音源と各マイクの距離はマイク0, マイク1およびマイク2は  $c\tau_{d0}+l$ ,  $c\tau_{d1}+l$ ,  $c\tau_{d2}+l$  と示すことができる。各音源の座標系と音源からの距離を整理すると、

$$\begin{pmatrix} 2L_A & 2L_E & -2c\tau_{d0} \\ 0 & 2L_E & -2c\tau_{d2} \\ 2L_A & 0 & -2c\tau_{d1} \end{pmatrix} \begin{pmatrix} x \\ y \\ l \end{pmatrix} = \begin{pmatrix} c\tau_{d0}^2 + L_A^2 + L_E^2 \\ c\tau_{d0}^2 + L_E^2 \\ c\tau_{d0}^2 + L_A^2 \end{pmatrix}, \quad (4.5)$$

と示すことができる. 式(4.5)の左辺の正方行列に対して両辺から逆行列の積を求めることで未知数である  $x y l$  を求めることができる. ここで,  $l$  はマイク 3 を基準としたときの音源までの距離であることから  $z$  を導くことができる. 以下の節ではこれらの原理を基に対象である音源位置の検出を行う.

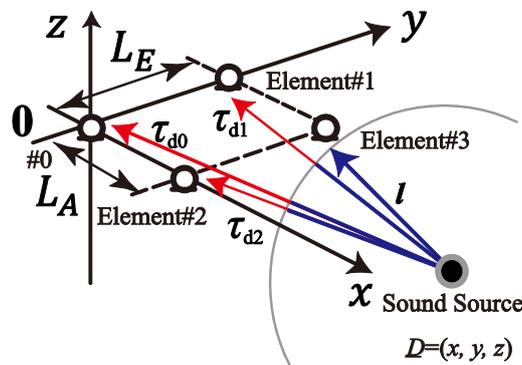


Fig. 4.2 到来時間差と音源位置推定

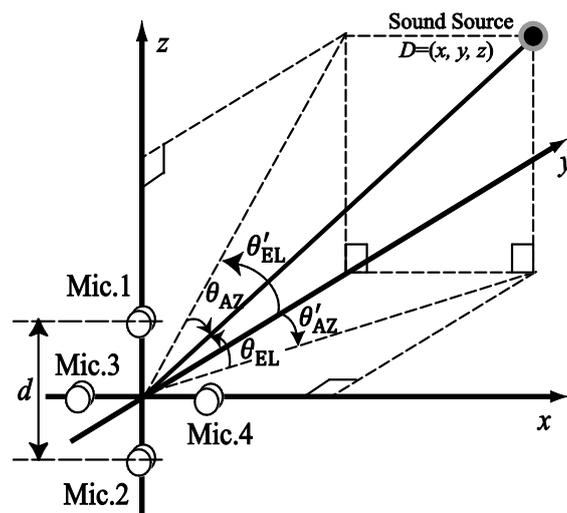


Fig. 4.3 4つのマイクロフォンとカメラによる位置推定

### 4.3 カメラとマイクロフォンを用いる豚舎内の位置推定

#### 4.3.1 カメラとマイクロフォンによる位置推定の計算手法

本節では、音源として 2,3 章における豚のくしゃみ音の位置検出することを目的として、カメラと 4 つのマイクロフォンにより実験を行った。Figure 4.2 に 4 つのマイクロフォンによる方向推定の原理について示す。式(4.4)に 2 つのマイクロフォンからの方向推定について示したが、ここでは 4 つのマイクロフォンを Fig. 4.2 のように設置し、Mic. 1, 2 を用いて仰角方向の角度  $\theta_{EL}$  deg. , Mic. 3, 4 を用いて方位角方向  $\theta_{AZ}$  deg. を求める。また、3次元であるため、

$$\begin{aligned}\theta'_{AZ} &= \sin^{-1}\left(\frac{\sin \theta_{AZ}}{\cos \theta_{EL}}\right), \\ \theta'_{EL} &= \sin^{-1}\left(\frac{\sin \theta_{EL}}{\cos \theta_{AZ}}\right),\end{aligned}\tag{4.6}$$

のように補正を行い正しい方位角  $\theta_{AZ}$  deg. , 仰角方向  $\theta_{EL}$  deg. を求める。この座標をカメラ中央に設置したカメラ上に表示させることで、くしゃみを発生させた豚の位置を特定することができる。この際、正しいくしゃみ位置の基準として、カメラ画像からの豚の頭部が 3 deg. 程度であることから前後 3 deg. を正解の範囲とした。本節において、4 つのマイクロフォンを用いる場合には、ここで定義した座標系とマイクロフォン番号、音源方向及び正解の範囲とした。

#### 4.3.2 豚舎での実験条件

位置推定を評価するため、2 章の実験で用いた畜産草地研究所のデータのうちくしゃみ音の発生位置を確認できた 10 個のくしゃみ音に対し本手法を適応する。なお位置の計算に用いるカメラは Web カメラの解像度は 640×480 画素であり、角度 1 deg. に対する画素数は 12 であった。なお実験条件は Table 2.1 に示す音速及びマイク間距離に基づくものとする。

#### 4.3.3 実験結果と考察

収録した信号のうち豚房内で発生した 10 区間のくしゃみ音に対し、4 つのマイクロフォンを用いて到来時間差を求めた。到来時間差の計算には、収録した信号で相互相関法を用いた。またくしゃみ音の特徴である 30 kHz 以上のハイパスフィルタをそれぞれに通過させたものを比較した。実験の結果を Fig. 4.3 および Table 4.1 に示す。まず方向推定に関する結果では相互相関法による手法では 50.0% で正解の範囲である 3 deg. 以内で検出できることがわかった。正しく推定できなかった No.7, 8 については豚が  $\theta'_{AZ} = 13$  deg. ,  $\theta'_{EL} = 10$  deg. の地点からマイクアレイと対称の方向にくしゃみを発していた。このため、 $\theta'_{EL} = 15$  deg. にある木の仕切り版で反射をし、その反射位置と推定される位置を検出し、誤推定の原因になった

### 4.3 カメラとマイクロフォンを用いる豚舎内の位置推定

と考えらる. No. 9 はくしゃみ音が背景雑音に埋もれており, 30 kHz 以上の成分が含まれていないものであった. このくしゃみ音は検出実験でも雑音として検出できていないものであった. ハイパスフィルタを通過させた信号に対する相互相関法での推定結果は高周波帯での反射等の影響で正しく推定できないことが考えられる.

Table 4.1 豚舎内の方向及び位置推定結果

No.	Time (h:m:s)	CC(error) (deg.)	CC <sub>HI</sub> (error) (deg.)	Location(error) (m)
1	0:31:24	○(-0.53)	×(-27.16)	×(+1.16)
2	0:39:28	○(+1.17)	○(+1.17)	○(-0.13)
3	0:39:45	○(+1.17)	○(+1.17)	○(-0.03)
4	0:40:11	○(+1.17)	×(+3.69)	○(-0.10)
5	0:40:20	○(+1.17)	×(+19.57)	○(-0.24)
6	1:42:02	×(+10.04)	×(-5.96)	×(+1.03)
7	1:57:54	×(+6.13)	×(-18.96)	×(+0.98)
8	1:57:56	×(+6.13)	×(-18.96)	×(+1.00)
9	2:04:29	×(+12.37)	×(-14.96)	×(+1.33)
10	2:42:48	×(-11.68)	×(-3.74)	×(+8.13)

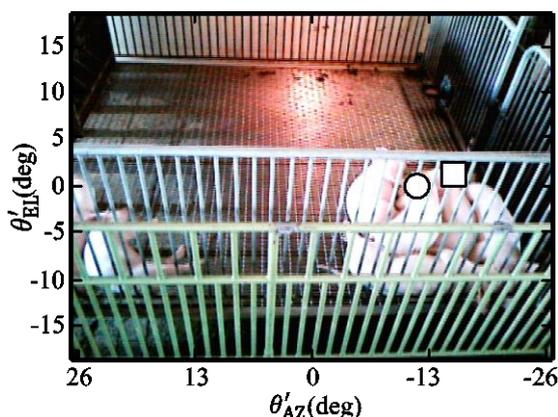


Fig. 4.4 カメラにおける豚舎内における方向及び位置推定結果の例  
(○が実際の発生位置, □が推定位置)

次に位置推定の結果について示す。4つのマイクロフォンから得られる3つの到来時間差から位置推定が可能であるか実験を行った。相互相関法を用い4つのマイクロフォンから3つの到来時間差を求めるため、それぞれの6通りの組み合わせ( $=_4C_3$ )から到来時間を求め推定結果の平均をとることで誤差の低減を行った。なお、位置の計算には式(4.5)を用い正解の範囲を豚の頭と同様の大きさである0.3 m程度とした。豚舎内で発生したくしゃみ音のうち、正解の範囲は4区間のみであり、方向推定可能であった区間に対しても位置推定が困難であることがわかった。正解したNo. 2-5については、同一箇所が発生しており、音源となる豚の口付近がカメラから確認でき、位置から逆算した到来時間差と比較し正しい到来時間差が得られていることがわかった。不正解となった原因として一つは、実際のくしゃみ音の音源がマイクロフォンに向かっていないことから、音源が原理となった球面波構造をしていない可能性がある。このため、3つの到来時間差が計算できず位置の解を推定できない。また、式(4.5)の左辺のセンサ間の距離の項が関連している。正しい到来時間であってもセンサ間の短い場合、僅かな誤差が計算結果に大きな誤差を与える。この影響をシミュレーションにより考察した結果、今回対象としたマイクロフォン間の距離がMic. 1とMic. 3などで0.14 m程度となることから、正しい時間差の場合にも1つのマイクロフォンに0.05 mの誤差を与えた場合、各座標軸の平均誤差は0.37 mとなる。マイク間の幅を10倍(1 m)とした場合、0.21 mの誤差であった。よって正確な位置を推定するためには、対象とするくしゃみ音を発する口の方向やセンサ間の距離を適切にする必要がある。

#### 4.3.4 まとめ

マイクロフォンとカメラ画像による位置推定法として、マイク2組のみの組み合わせによって方向推定をし、その結果をカメラ画像と合わせることで50.0%のくしゃみ音をカメラ上に推定でき、位置探査が可能であることが示された。

また、4つのマイクロフォンから位置推定を行った結果、音源がマイク方向に向いている場合、40.0%のくしゃみ音を位置探査可能であった。遠方場を用いる場合は、カメラから得た壁等の反射面を活用した推定を行うことで誤差の低減が可能であると考えられる。また、近傍場を用いる場合は反射等の影響に加えセンサ間の距離を考察することが必要となることがわかった。

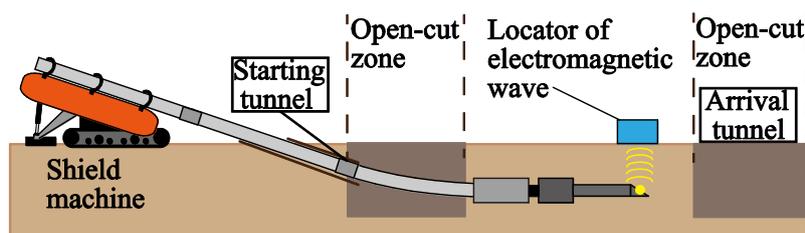


Fig. 4.5 誘導式水平ドリル工法とロケータによる位置推定の概要

## 4.4 反射音を用いる土中の位置推定

### 4.4.1 少数センサによる土中位置推定の背景

前節までにおいて少数のマイクロフォンで音源から直接マイクロフォンに到達する場合における音源位置推定について述べた。音波は空気中に限らず、電波などが通過できない土中も振動として通過できる特徴があることから、マイクロフォンを地中に埋め込み、直達音だけでなく地表面での反射波も含めた音源位置推定<sup>[97]</sup>について検討する。

現在のインフラ整備では水道管やガス管、電線の埋設が盛んに行われており、新規施工だけでなく耐久年数の過ぎた配管の交換のため埋設部分のすべての舗装や地面を開削する工法が一般的となっている。しかし、この工法は作業中の道路封鎖が必要であり、開削から埋設まで工期が長くなることから、新たな工法として誘導式水平ドリル工法が提案されている<sup>[98]</sup>。この工法は、地上にドリル本体を設置し、ドリル先端で配管の埋設予定の地中のみの掘削を行うため、開始地点である発進坑と終了地点である到達坑のみを開削すればよい。そのため、**Fig. 4.5**のように、埋設部のすべて舗装や地面を開削する必要がないので、道路封鎖や工期の短縮ができる。一方、この工法はドリル先端の正確な位置を管理することが重要であり、現在はドリル先端内部に電磁波発生器を設け、その電磁波を直上のロケータで受信し推定する手法が用いられている。しかし、土中で電磁波が減衰してしまう場合や直上に構造物や川がありロケータを設置できない場合には推定できない問題がある。そこで、電磁波に比べ減衰が少ない音響的手法を用いて直上以外から推定する手法が検討されている。ドリル先端位置の推定法では地表面上に複数のジオフォンを設置し、ドリル先端部で発生する破碎音を音源として、音波の到来時間差を用いて音源位置を推定するリアルタイムな手法<sup>[8]</sup>である。しかし、17個のジオフォンを約10m<sup>2</sup>に設置し、各ジオフォンの位置を正確に測定する必要があることから、少数かつ測定容易な手法が望まれている<sup>[99]</sup>。そこで本章では、ドリル先端部の直上から離れた位置に設置した2本の杭状ジオフォンアレイによってドリル先端部で発生する破碎音の位置を推定する手法を提案する。杭状ジオフォンアレイは2つのジオフォンで構成され、本手法では2本のジオフォンアレイを2箇所の縦穴に差し込み、位置推定に必要

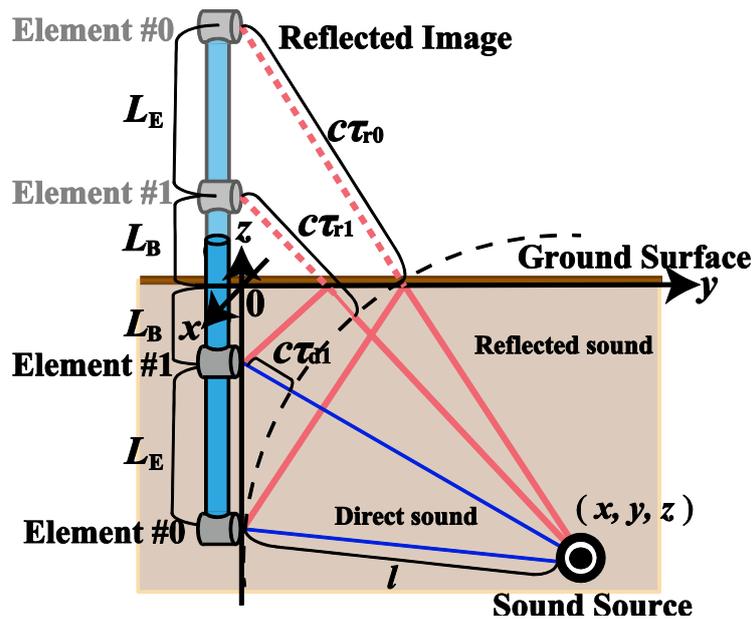


Fig. 4.6 地表面反射による音波伝搬経路の例

最小限の4つのジオフォンへの音波の到来時間差を基に音源位置を推定する。本手法では、音源からジオフォンに直接受信される直達音だけではなく、地表面からの反射音を考慮した計算を行うことにより音源位置を推定する<sup>[100,101]</sup>。なお、本節では、ベクトルを一重下線、行列を二重下線として表記する。

#### 4.4.2 反射音を用いる位置の計算手法

本節で行う実験の音波伝搬は音線理論<sup>[102]</sup>に基づくとする。また、ドリルの先端部で発生する破砕音はインパルス性であることが知られている<sup>[103]</sup>、よってインパルス信号を生成し音源信号とする。位置推定法では少素子で構成されたジオフォンアレイを作成し埋設することで、地中を伝わる音波のうち直接伝わる直達波に加え地表面で反射する反射音を利用する土中の音源位置測定法を提案する。なお、直達音および反射音の到来時間差を求めるため式(4.5)の原理を応用した手法を用いる。対象とする音源の波長が地表境界面の凹凸より十分大きい場合には、反射面を平面境界とみなし鏡面反射が起こるものとして計算できるものとし。粘土質の土中を対象とした場合、音速は約 1,500 m/s、伝搬する音波は 3 kHz 以下であった。波長は約 3 ~ 300 m であり、地表面の凹凸が約  $1.0 \times 10^{-2} \sim 3.0 \times 10^{-2}$  m とすると波長は地表面凹凸に比べ十分に長いとして扱える。

一対のジオフォンアレイを地中埋設した際の音源と素子間の音波伝搬経路をFig. 4.6に示す。地表面と素子1との距離を $L_B$ 、素子1-2間の距離を $L_E$ とする。音源の位置 $\underline{D} = (x, y, z)$ とすると、音源で発生した音波はまず素子0に到来する。この時刻を基準として、音速 $c$ とすると、同一の音波は到来時間差 $\tau_{d1}$ で素子1と $c\tau_{d1}$ の距離差で受信される。地表面で反射した音波は同様に素子1に $\tau_{r1}$ の距離差で到来し、素子0に $\tau_{d0}$ の距離差で受信される。

#### 4.4 反射音を用いる土中の位置推定

ここで、各素子への受信信号は音源の信号を  $\phi(t)$  とすると、素子0, 素子1 の受信信号  $s_0(t)$ ,  $s_1(t)$  は、式(4.7), (4.8) で示される。

$$s_0(t) = \phi(t) + \phi(t - \tau_{r0}), \quad (4.7)$$

$$s_1(t) = \phi(t - \tau_{d1}) + \phi(t - \tau_{r1}), \quad (4.8)$$

ここで、**Figure 4.7** に示すように2つの素子の受信信号の時間差を計算すると、直達音と反射音の到来時間差に相当する時間位置にピークが現れ  $\tau_{d1}$ ,  $\tau_{r1}$ ,  $\tau_{r0}$  を求めることができる<sup>[104]</sup>。ここで素子0を基準とし、もう一对のジオフォンアレイを考慮すると、素子2, 素子3に到来する受信信号  $s_2(t)$ ,  $s_3(t)$  は、式(4.9), (4.10) で示される。

$$s_2(t) = \phi(t - \tau_{d2}) + \phi(t - \tau_{r2}), \quad (4.9)$$

$$s_3(t) = \phi(t - \tau_{d3}) + \phi(t - \tau_{r3}), \quad (4.10)$$

各素子間の相互相関関数を計算することで、二対のジオフォンアレイの各素子に対する直達音、反射音の到来時間差  $\tau_{d1}$ ,  $\tau_{d2}$ ,  $\tau_{d3}$ ,  $\tau_{r0}$ ,  $\tau_{r1}$ ,  $\tau_{r2}$ ,  $\tau_{r3}$  が求められる。相互相関結果は反射による振幅の反転を考慮して絶対値を求めたピークとなり、**Figure 4.8** のように示すことができる。

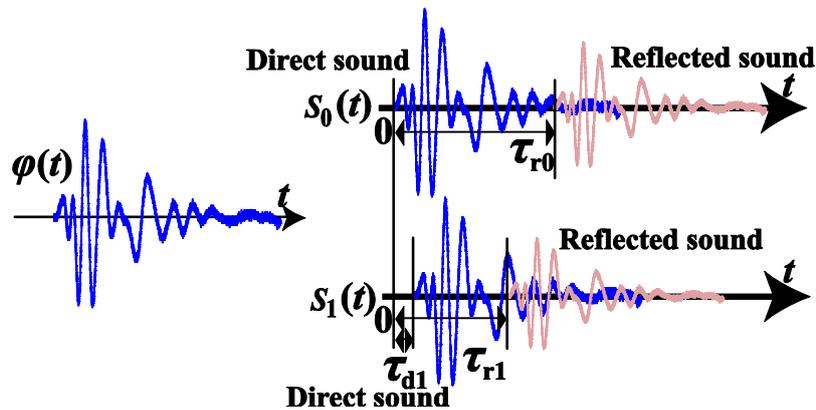


Fig. 4.7 反射波を含む受信信号の例

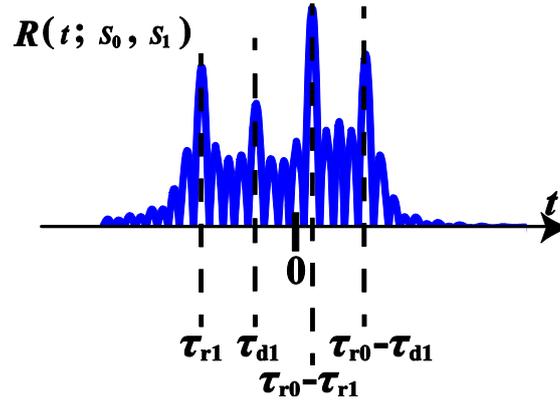


Fig. 4.8 反射波を含む相互相関波形

Figure 4.6 の一対のジオフォンアレイが  $z$  軸上 ( $x=0$ ) に位置するとき, 素子 0, 素子 1 の座標を  $\underline{m}_0 (\equiv (0, 0, -L_B - L_E))$ ,  $\underline{m}_1 (\equiv (0, 0, -L_B))$  とすると, 素子 0 と素子 1 の地表に対する虚像の座標は,  $\underline{r}_0 (\equiv (0, 0, L_B + L_E))$ ,  $\underline{r}_1 (\equiv (0, 0, L_B))$  となる. 音源の座標系と直達音, 反射音の距離差から,

$$\|\underline{m}_0 - \underline{D}\| = l, \quad (4.11)$$

$$\|\underline{r}_0 - \underline{D}\| = l + c\tau_0, \quad (4.12)$$

$$\|\underline{m}_1 - \underline{D}\| = l + c\tau_1, \quad (4.13)$$

$$\|\underline{r}_1 - \underline{D}\| = l + c\tau_1, \quad (4.14)$$

と示される ( $\|\cdot\|$  はベクトルのノルムを表す).

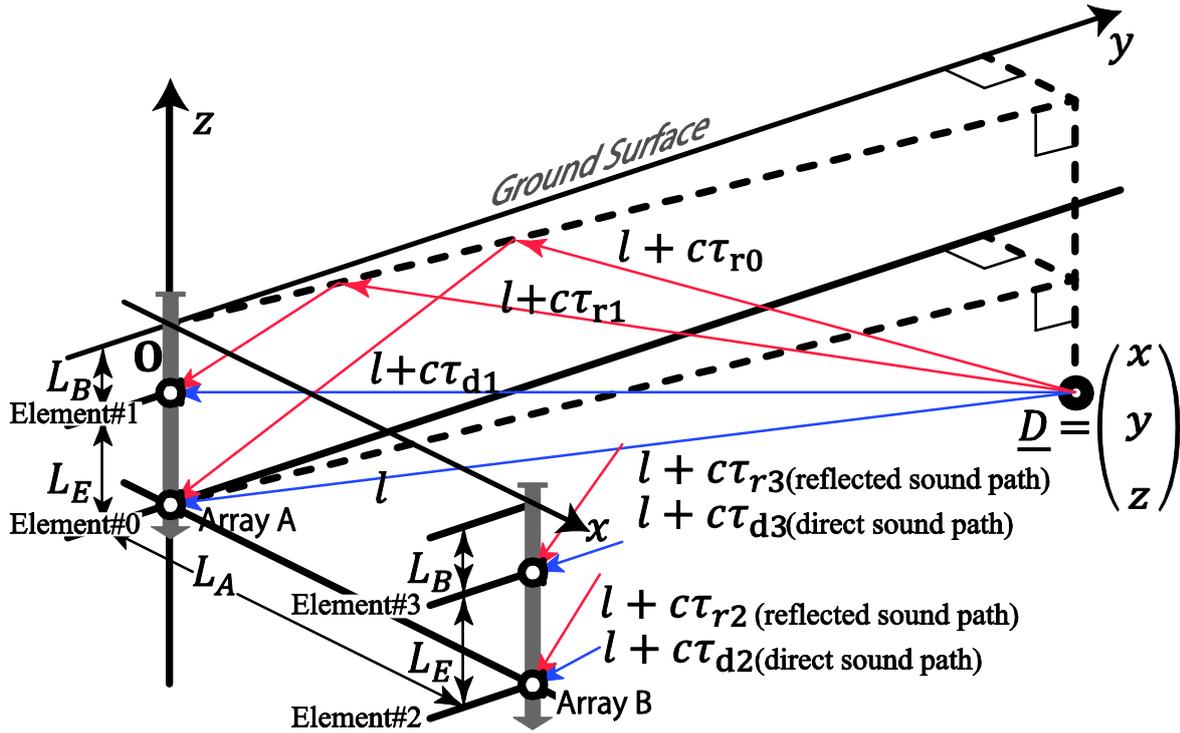


Fig. 4.9 2つのジオフォンアレイによる音波伝搬経路

ここで、Fig. 4.9 のように二対のジオフォンアレイをアレイ A ( $x=0$ ) とアレイ B ( $x=L_A$ ) とすると、アレイ A が式(4.11)-(4.14)のとき、アレイ B も同様に素子 2, 素子 3 は  $\underline{m}_2 (\equiv (L_A, 0, -L_B - L_E))$ ,  $\underline{m}_3 (\equiv (L_A, 0, -L_B))$ , 虚像の座標を  $\underline{r}_2 (\equiv (L_A, 0, L_B + L_E))$ ,  $\underline{r}_3 (\equiv (L_A, 0, L_B))$  の距離差から、

$$\|\underline{m}_2 - \underline{D}\| = l + c\tau_{d2}, \quad (4.15)$$

$$\|\underline{r}_2 - \underline{D}\| = l + c\tau_{r2}, \quad (4.16)$$

$$\|\underline{m}_3 - \underline{D}\| = l + c\tau_{d3}, \quad (4.17)$$

$$\|\underline{r}_3 - \underline{D}\| = l + c\tau_{r3}, \quad (4.18)$$

と表される。未知数を  $x, y, z, c$  として整理すると以下の関係が成り立つ。

$$T \cdot \begin{pmatrix} x \\ z \\ cl \\ c^2 \end{pmatrix} = \underline{L}, \quad (4.19)$$

ここで  $\underline{T}$ ,  $\underline{L}$  は,

$$\underline{T} = \begin{pmatrix} 0 & 2L_E & 2\tau_{d1} & \tau_{d1}^2 \\ 2L_A & 0 & 2\tau_{d2} & \tau_{d2}^2 \\ 2L_A & 2L_E & 2\tau_{d3} & \tau_{d3}^2 \\ 0 & 2(2L_B + 2L_E) & 2\tau_0 & \tau_0^2 \\ 0 & 2(2L_B + L_E) & 2\tau_1 & \tau_1^2 \\ 2L_A & 2(2L_B + 2L_E) & 2\tau_2 & \tau_2^2 \\ 2L_A & 2(2L_B + L_E) & 2\tau_3 & \tau_3^2 \end{pmatrix}, \quad (4.20)$$

$$\underline{L} = \begin{pmatrix} L_E^2 \\ L_A^2 \\ L_E^2 + L_A^2 \\ (2L_B + 2L_E)^2 \\ (2L_B + L_E)^2 \\ L_A^2 + (2L_B + 2L_E)^2 \\ L_A^2 + (2L_B + L_E)^2 \end{pmatrix}, \quad (4.21)$$

ただし,  $l^2 = (x^2 + y^2 + (z - L_B - L_E)^2)$  である. ここで式(4.21) に,  $\underline{T}$  の擬似逆行列  $\underline{T}^+$  を左からかけると,

$$(x \ z \ cl \ c^2) = \underline{T}^+ \cdot \underline{L}, \quad (4.22)$$

となり, 音源位置  $\underline{D}$  を求めることが可能となる. ここで,  $y$  は  $l^2 = (x^2 + y^2 + (z - L_B - L_E)^2)$  より導出される.

従来手法には原理で述べた計算手法を用い, 提案手法と従来手法の評価基準として, 無次元化した値を用い音源距離あたりの誤差を評価するため, 以下の式によって原点から音源の真位置までの距離  $l_{\text{True}}$  と原点から測定された位置までの距離  $l_{\text{Est}}$  より測定誤差率  $\eta$  を求めることとした.

$$\eta = \frac{|l_{\text{True}} - l_{\text{Est}}|}{l_{\text{True}}}, \quad (4.23)$$

### 4.4.3 土中における実験条件

提案手法の有効性を検証するため、実際の土中で実験を実施した。地表面に素子を配置する従来手法との比較をした。実験の概要図を**Fig. 4.10**に示す。実験環境の土質は固い粘土質で、実験付近は約10 m 四方に亘り概ね同様の地質であった。音源となる地点を掘削し、その音源位置をハンマーで加振し信号を発生させた。音源位置を $\underline{D}$  ( $\equiv (0.20, 1.00, -0.70)$  (m)) とし、埋設するジオフォンアレイは圧電式の素子 (H1a; AQUARIAN) を用いて作成したものを使用して、 $L_E = 0.20$  m,  $L_A = 0.25$  m,  $L_B = 0.15$  m の地点に埋設した。収録には、PC (Sony, VAIO PCG-791N) と Analog-Digital コンバータ (National Instruments, USB-6212) を用いた。また予備実験として、**Fig. 4.10**のように音速測定実験を行った。地表面から0.30 m の深さに2つの素子を1 m の間隔で埋設し、0.5 m 離れた側面から加振させ伝搬時間を測定した結果、音速は $c = 1,637$  m/s であった。以下、音速はこの値を用いる。予備実験により求めた音速と素子間の最小距離0.2 mに対し分解能を得るためサンプリング周波数96 kHz とした。音源位置を、10秒間に5回の加振し、これを20回繰り返し計100サンプルを得た。比較対象である従来手法では、素子0を原点とし、すべての素子を地表面状に  $L_E = 0.20$  m,  $L_A = 0.25$  m の間隔で設置した。到来時間差の推定時には、素子間の距離に応じた時間閾値を設定し、その範囲以外の値については除外するとともに、相互相関関数を加算することでピークを得た。

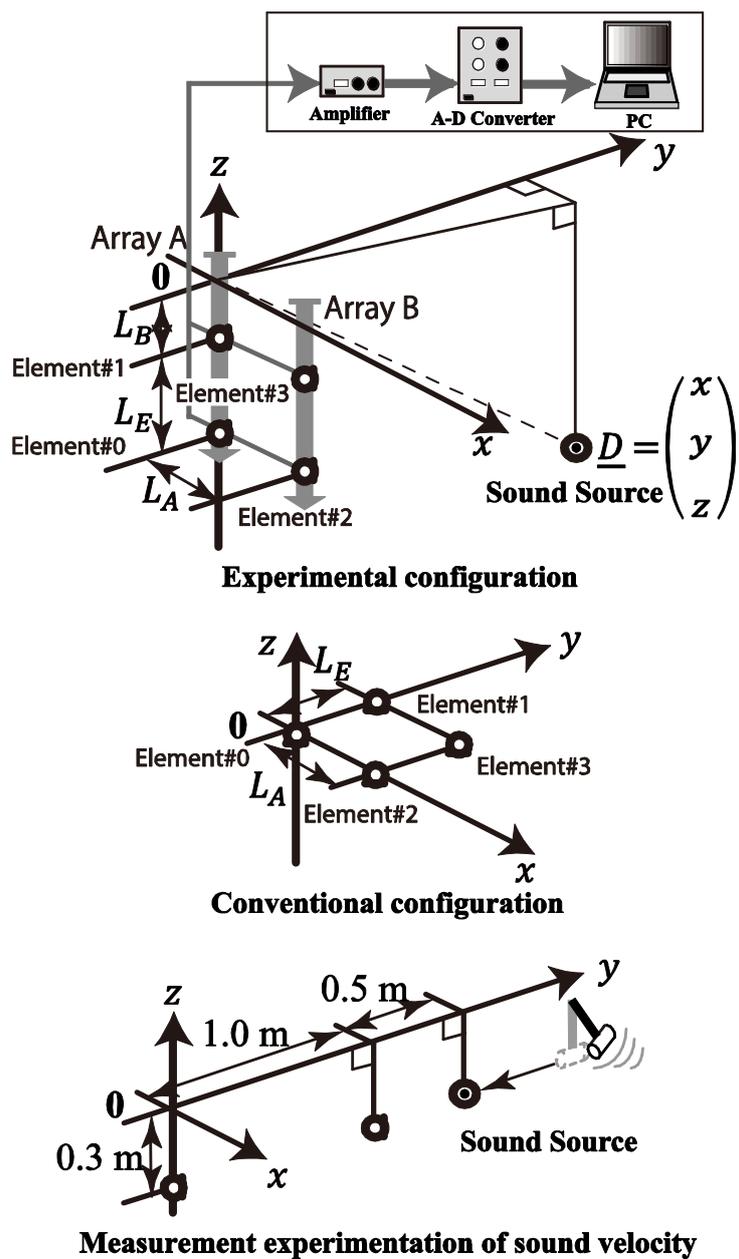


Fig. 4.10 音速の測定と実験概要

## 4.4.4 結果と考察

**Table 4.2** に、提案手法および従来手法の推定値および真値からの誤差を示す。推定の結果、提案手法では  $y$  軸方向の測定誤差が最も大きく  $+4.08$  m となり、従来手法では  $z$  軸方向の測定誤差が最も大きく  $-11.86$  m となった。また、提案手法での音速推定の結果  $1,505$  m/s となった。各軸方向に対する誤差の原因として、素子の配置が考えられる。提案手法の素子の配置面は  $x-z$  平面の配置となり、従来手法は  $x-y$  平面の配置となっている。このため音源から各面で受信する信号の到来時間差を十分に推定できていないことが原因であると考えられる。4 つの素子を同一平面上に設置する場合には、その素子面の法線軸方向の音源からの各素子の到来時間差に対応するサンプル数の間隔が広がるため位置推定精度が他軸に比べ低下すると考えられる。しかし、直達音だけでなく反射音も利用する提案手法の誤差は、従来手法に比べて小さいことが分かる。この原因として素子を地表面に設置する従来手法では、多くの素子を用いることで到来時間差の誤差を補間していたが、少数素子の場合には補間ができず、誤差がそのまま結果に影響することが考えられる。これに対して、提案手法では未知数の導出に疑似逆行列を用いて 7 つの式から 4 つの解を推定していることから  $y$  軸方向への到来時間誤差をより正確に推定が可能となっていると考えられる。また、提案手法および従来手法は共に、それぞれの素子の位置関係から音源位置の測定を行うため、これらの素子の位置を正確に配置することが重要であると考えられる。測定誤差率を比較すると、提案手法  $\eta_p=2.35$  に対し、従来手法  $\eta_c=8.03$  であり、提案手法は従来手法に比べ 3.42 倍の精度で位置測定可能であることが示された。

**Table 4.2** 土中の位置推定結果の比較

	$x$ (m)	$y$ (m)	$z$ (m)	$c$ (m/s)
	(error)	(error)	(error)	(error)
Proposed method	0.15 (-0.05)	5.08 <b>(+4.08)</b>	-0.11 (-0.69)	1505 (-132)
Conventional method	0.26 (+0.06)	1.58 (+0.58)	-11.86 <b>(-11.16)</b>	-
True value	0.20	1.00	-0.70	1637

\* Bold characters: maximum value of error

#### 4.4.5 まとめ

本節では地表面反射を利用する土中の音源位置推定法について述べた。杭状のジオフォンアレイを地中に埋設し用いることで反射波の収録を可能とし、相互相関法によって反射波の時間差から位置推定の計算方法について提案をした。

実際の土中で評価実験を行った結果、提案手法は従来手法に比べ3.42倍の測位精度を有することがわかった。また、4つの素子のみを用いた測定法では、素子で構成される面の直行する方向に対して分解能が低いことがわかった。実規模を想定した場合、素子間隔の長距離化が必要となるとともに、それぞれの構成する素子位置の設置誤差が位置測定誤差に影響することがわかった。本実験における実規模スケールのシミュレーションについては本文は割砕し、**付録B**に示すこととする。

## 4.5 立体構成マイクロフォンによる位置推定

### 4.5.1 マイクロフォンを用いるドローン位置推定の背景

本節では4つのマイクロフォンを立体的に配置したマイクロフォンアレイを2組作成し、音源位置推定について検討した。前節の反射音を用いる手法では地面以外に反射面となる場所が無い場所だったため、実地における実験でも反射音の到来時間差を算出できたものの、屋外の空気中ではどこでどのように反射したのかを判別するのは難しく、室内であれば残響として見えなくなってしまう。そこで屋外での音源位置推定の対象として小型無人飛行ロボット（以下ドローン）を用い位置推定法を検討する。

ドローンはカメラや温度計、マイクロフォンなどのセンサを取り付け、橋や道路などのインフラ点検や上空からの三次元復元図の作成、災害現場の探索などに活用するなど応用がされている<sup>[105,106]</sup>。特に人にとって危険な箇所の点検や測量など従来は膨大な時間をかけて行っていた作業を大幅に削減できる可能性があることから様々な応用の研究や実用化が進んでいる。中でも橋や高速道路などのインフラの点検は目視や打音検査などによる点検が主流であり、ドローンによって目視点検が主流となれば作業効率の向上が期待できる。一方で、定期的な検査で前回と同じ場所を点検する場合や正確な3次元立体図などを作成するにはドローンがどの位置から撮影しているのかの位置情報が重要である。しかし、建物間や橋梁下などはGPSの精度が悪い場所では、正しいドローン位置を確かめる必要がある。加速度センサなどで補正する制御方法等も研究されているが正確な位置を推定することが難しい<sup>[107]</sup>。また実際の操縦者が特定の位置でドローンから撮影する際にも目視や、ドローンに装着されたカメラからの画像で相対的に位置を判断する必要がある。これらには操縦者の経験や空間認知能力を必要とする上にドローン自体の正確な位置を記録することができない。リアルタイムに位置を計測するためにドローンにLEDなどを着け地表面に設置したカメラ等から位置推

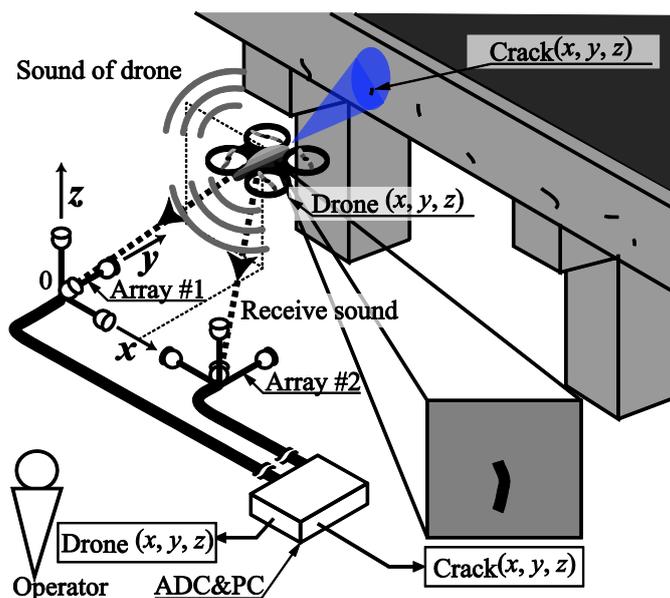


Fig. 4.11 提案手法を用いるドローン位置推定の応用例



Fig. 4.12 実験で使用した AR. Drone 2.0

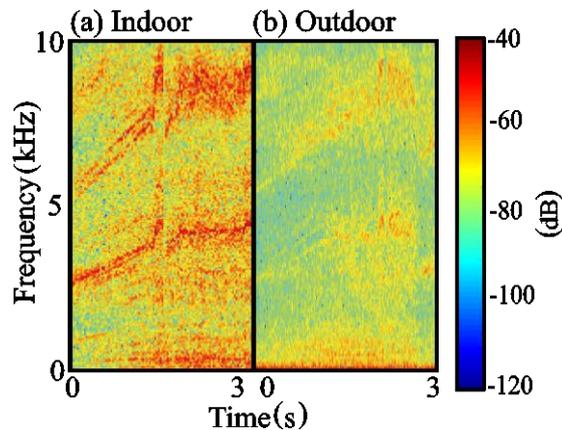


Fig. 4.13 ドローン音の無響室及び屋外の周波数スペクトログラム

定する手法も提案<sup>[108,109]</sup>されているが、ドローンに追加装備が必要であり積載重量の観点などの問題がある。そこでFig. 4.11のようにドローンが飛行する際に回転翼で発せられる音に着目し、その音の発生位置である音源位置推定を行うことにより、ドローン本体に追加装備不要なリアルタイム位置推定を検討する。

音をドローンに応用する手法として、ドローン本体にマイクロフォンを載せるものはある<sup>[110,111]</sup>が、地表面にマイクロフォンを設置し計測するものはドローンの方向を測定するに留まっており、位置推定に至ってはいない。そこで本節では、ドローンから発せられる音源位置を4つのマイクロフォンからなるマイクロフォンアレイを地表面に2つ設置することで音の到来時間差からドローンの位置推定法を提案する<sup>[112,113]</sup>。

#### 4.5.2 最適化手法を用いる連続音の位置の計算手法

本実験で使用する機体を Fig. 4.12 に示す。このドローンは全長約 0.8 m 重さ約 1 kg であり、HD カメラとオプションにより GPS のフライトレコーダを装着可能である。また Wi-Fi による制御によって屋外では 50 m 程度以内でカメラからの画像を見ながら操作可能であり、実際の点検等を想定した場合の最低限の機能を備えていると考えられる。

ドローンの飛行音の特徴を確認するため、無響室(1 m 遠方から集音)および屋外(10 m 遠方から集音)においてドローンの上昇からホバリングまでの音を収録した。Fig. 4.13 にその周

波数スペクトログラムを示す。2 秒付近まではプロペラの回転数上昇に対応し、周波数が高くなっていることがわかる。屋外では遠方にあることで飛行音自体は小さくなっており 2 kHz 以下には風音や様々な雑音を含んでいるもので占められている。しかし、無響室での実験と同様にプロペラ回転数に対応する様子が確認できる。そこで、以降では、収録信号に対し 8 kHz のハイパスフィルターを用い周波数帯域を制限することにより主要な雑音等の影響を減じた。

次にドローンの飛行音と屋外環境音から到来時間差の比較を行った。ここでの信号処理では前述のハイパスフィルタを通過させた信号から相互相関を行った。平成 27 年 9 月 29 日 0 時から 2 時の間に大学構内の広場において、精密騒音計 (NL-41: RION) と録音装置 (D-50: SONY) を設置し、原点から 100 m の距離まで徐々にドローンを離していく実験を行った。Fig. 4.14 に実験の概要を示す。三脚を用い地上 1.0 m にそれぞれ固定し、距離の計測にはレーザ測距計 (CH-9435: Leica Geosystems AG) を用いた。まず精密騒音計によりドローンがホバリング状態の際の騒音レベルと、その直前および直後の環境音の平均騒音レベルの比較を Fig. 4.15 に示す。なお騒音レベルの周波数特性はフラットで設定した。

ここで環境音の変動があるのは、ドローンの距離を動かしている間に環境音の時間変化が原因である。1 m 程度の距離では 15 dB 以上大きい 70 dB 程度であり、これは部屋で掃除機をかけている時と同程度であり、環境音の 55 dB は部屋の通常会話程度とされている。夜間に 55 dB となっていた原因は広場の虫の音や車の走行音である。ドローン音は環境音に比べ 50 m 程度までは大きいですが、それ以降は環境音とほぼ同等となった。次に録音装置によって収録した信号から到来時間差を計算させた。その結果を Fig. 4.16 に示す。計測開始点である 1 m では音源が近く他の距離の時間差からずれているものの、50 m 程度までは到来時間差が計算できることがわかった。これらの結果からドローン音が環境音より大きい際には正しい時間差が測定でき、環境音が 55 dB 程度では 50 m 程度まで時間差を計測できることがわかった。

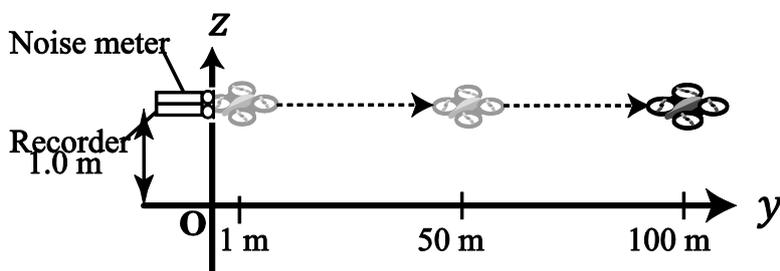


Fig. 4.14 環境音とドローン音比較の実験概要

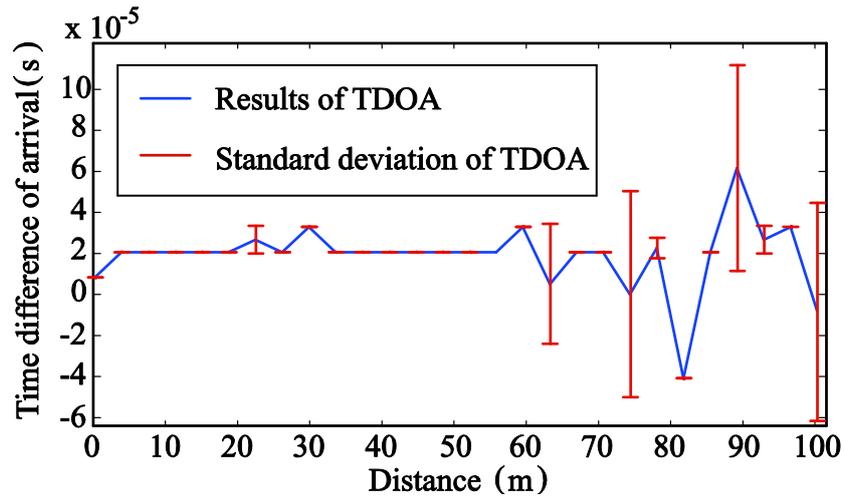


Fig. 4.15 ドローン音の到来時間差の計算結果

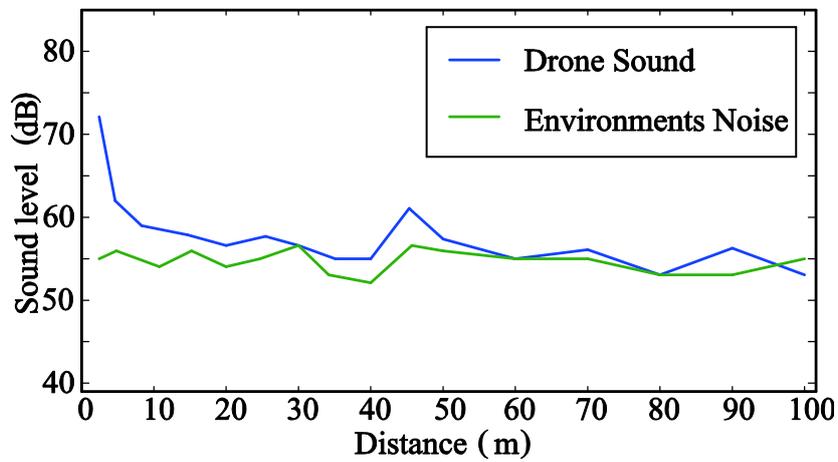


Fig. 4.16 環境音とドローン音の騒音レベルの比較

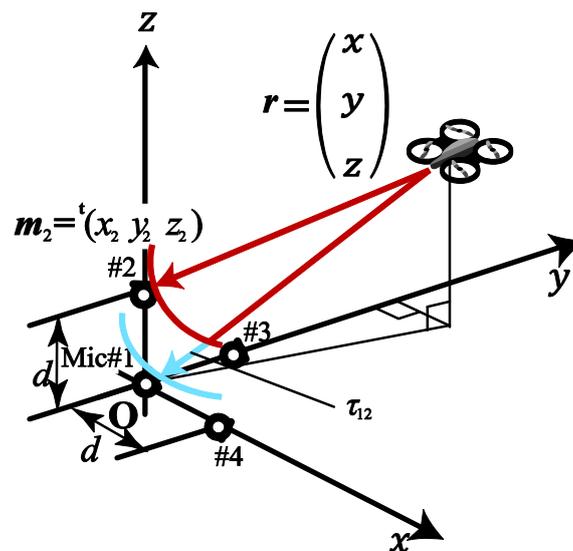


Fig. 4.17 到来時間差の原理とマイクロフォンアレイ計測

続いて計算手法について **Fig. 4.17** を基に示す, 4つのマイクロフォンが1組となったマイクロフォンアレイを2組設置した場合を考える. マイク間距離を  $d$  とし, ここで  $i$  番目のマイクの座標を  $\mathbf{m}_i = {}^t(x_i \ y_i \ z_i)$  同様に  $j$  番目を  $\mathbf{m}_j = {}^t(x_j \ y_j \ z_j)$  とした時, この2つのマイク間の到来時間差を  $\tau_{ij}$  とする. またドローンの真の位置を  $\mathbf{r}_{\text{real}} = {}^t(a \ b \ c)$  とする. これらを整理して,

$$c\tau_{ij} = \|\mathbf{r}_{\text{real}} - \mathbf{m}_i\| - \|\mathbf{r}_{\text{real}} - \mathbf{m}_j\|, \quad (4.24)$$

この式は, 理想的な音の伝搬で成り立つものであり実際の音波には雑音をはじめ風切り音など様々な影響をうける. ここで, (5)式と, 音源位置として推定するドローンの位置を  $\mathbf{r}_{\text{est}} = {}^t(x \ y \ z)$  より,

$$S_{ij} = c\tau_{ij} - \left( \|\mathbf{r}_{\text{est}} - \mathbf{m}_i\| - \|\mathbf{r}_{\text{est}} - \mathbf{m}_j\| \right), \quad (4.25)$$

$$\mathbf{F} = \arg, \min \mathbf{R} \cdot S_{ij},, \quad (4.26)$$

と式(4.26)のように定義し, この  $S_{ij}$  を 0 とするような  $\mathbf{r}_{\text{est}}$  の解を式(4.26)の最適化問題として扱い, すべてのマイクロフォンの組み合わせで計算する. 例えばマイクロフォンアレイ1組の4つであれば  ${}_4C_2$  の6通りの式が計算され, 今回使用する2組のアレイすべてのマイクロフォンでの組み合わせでは  ${}_8C_2$  の28通り計算でき冗長性をもたせることができる. 最適化問題の計算には Levenberg – Marquardt (LM) 法を用いた<sup>[114,115]</sup>. またここで,  $\mathbf{R}$  は前述の信号の類似度に起因する尤度であり, 最適化問題を解く際に類似度が高く信頼性の高い到来時間差をもつ式を重視し計算させることができる. これによって, 各マイクロフォンでの風切り音や雑音などの影響を減じることができる. 評価実験として, 実際のドローンを用いて屋外でマイクロフォンアレイ2組を作成・設置し, 提案手法による位置推定精度について評価を行った. この際それぞれのアレイで得られる6つの到来時間差から求めた推定位置と, アレイ2つのもつ8つのマイクロフォンアレイから得られる28つの到来時間差を用いる手法について比較した. なお, 目標精度としてGPSでの誤差が1m程度以上であり, 使用するドローンの大きさが0.8m程度であることから0.8mとした.

4.5.3 ドローンに対する実験条件

4つのマイクロフォン(WM-61A: Panasonic)からなるマイクロフォンアレイを、中心のマイクを基準にマイク間距離  $d = 0.4 \text{ m}$  で2つ作成し Fig. 4.7 のように設置した. この際アレイ間の距離  $l = 5.0 \text{ m}$  とした. 収録にはノート PC (B354: TOSHIBA) と Analog-Digital コンバータ (USB-6259: National Instruments) を用い, サンプリング周波数を  $100 \text{ kHz}$  とした. またドローンの座標系は  $y$  軸上の  $\mathbf{r}_1 = {}^t(0 \ 5.0 \ 1.5)$  及び  $\mathbf{r}_2 = {}^t(2.5 \ 5.0 \ 1.5)$ , 同位置で  $z$  軸を  $0.5 \text{ m}$  下げた  $\mathbf{r}_3 = {}^t(2.5 \ 5.0 \ 1.0)$  の3か所とした. 実験中の気温は  $22.4 \text{ }^\circ\text{C}$  であり風速は  $0.5 \text{ m}$  程度以下であった. **Figure 4.18** に実験構成図について示す.

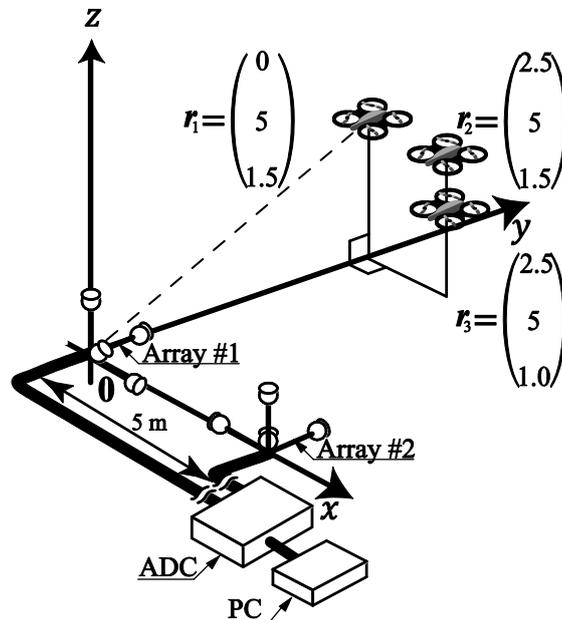


Fig. 4.18 ドローンの位置推定のための実験構成図

Table 4.3 2組のマイクロフォン毎の計測結果

	$x(\sigma)$	$y(\sigma)$	$z(\sigma)$
$\mathbf{r}_1 = {}^t(0 \ 5.0 \ 1.5)$	-0.09 (0.03)	<b>4.14</b> (0.28)	1.26 (0.08)
$\mathbf{r}_2 = {}^t(2.5 \ 5.0 \ 1.5)$	<b>1.11</b> (0.05)	4.57 (0.33)	<b>0.55</b> (0.35)
$\mathbf{r}_3 = {}^t(2.5 \ 5.0 \ 1.0)$	<b>1.07</b> (0.04)	<b>2.00</b> (0.03)	<b>0.42</b> (0.12)

Unit: (m)

## 4.5.4 結果と考察

Table 4.3 に、1組のアレイで得られる6つの到来時間差を2組分（計12個）用い、最適化問題を解いた結果について示す。表は位置推定結果と括弧内にその標準偏差を示す。なお太字については目標精度以下になったところを示している。結果をみると音源とマイクロフォンアレイ1の距離に近い $\mathbf{r}_1$ では概ね位置推定できているものの、それ以外の値は精度外の結果となった。 $\mathbf{r}_1$ でアレイ1に近い結果となった原因として、この類似度である $\mathbf{R}$ の値がアレイ2より高く、アレイ2の到来時間差が考慮されなかった可能性がある。 $\mathbf{r}_2$ 、 $\mathbf{r}_3$ では $\mathbf{R}$ の値はほぼ同程度だったものの、アレイ1・アレイ2の計算結果をそれぞれ別に見てみると実際の音源位置とそれぞれのアレイを線分で結んだ線分上のアレイに近い付近に推定していた。この原因として、今回作成したマイク間距離 $d$ のみで得られる到来時間差と最適化問題の特性が考えられる。例としてFig. 4.18の $\mathbf{r}_1$ にアレイ1から音源までの線分を引く。この線分上である程度離れた音源位置からそれぞれのマイクロフォンで測定される到来時間差と、その音源位置から線分上に少し前後した到来時間差はマイクロフォン近傍に音源位置がある場合と比べると変化がとても少ない。このような場合、本手法の最適化問題で解の初期値は0としているため、線分上の変化が少ないような0に近い側の解付近で最適化されてしまうことが考えられる。しかし一つのアレイでのマイク間距離を数mなどに大きくすることは、設置の簡易化を目標とする本研究に不適である。そこで、それぞれのアレイ間距離を用いることを検討した。この手法を行うためには2組のアレイで時刻を同期させておく必要がある。

Table 4.4 すべてのマイクロフォンの組み合わせによる計算結果

	$x(\sigma)$	$y(\sigma)$	$z(\sigma)$
$\mathbf{r}_1 = {}^t(0 \ 5.0 \ 1.5)$	-0.17 (0.09)	5.36 (0.77)	1.60 (0.04)
$\mathbf{r}_2 = {}^t(2.5 \ 5.0 \ 1.5)$	2.52 (0.04)	4.57 (0.33)	0.85 (0.35)
$\mathbf{r}_3 = {}^t(2.5 \ 5.0 \ 1.0)$	2.46 (0.03)	4.24 (0.29)	0.60 (0.15)

Unit: (m)

Table 4.4 に、それぞれのアレイのマイクロフォン合計 8 個からなる 28 通りの到来時間差を用い計算した結果について示す。結果はすべての音源位置で目標精度以内となり、アレイ 1 組ずつから計算する手法に比べ推定精度を向上させることができた。これは、アレイ内での 4 つのマイクロフォンでの到来時間差に加え、アレイ間距離 5 m での組み合わせで到来時間差を計算していることから、前述した線分上の変化が少ない範囲を拡大できたことが原因であると考えられる。つまり、このアレイ間距離を広くすることで前述の影響を減らし精度を向上させることができる。本結果に対し、式(4.23)を用い 3 つの場所の平均測定誤差率を比較すると、提案手法  $\eta_p=0.10$  に対し、類似度  $\mathbf{R}$  を用いない従来手法  $\eta_c=0.32$  であり、提案手法は従来手法に比べ 3.2 倍の精度で位置測定可能であることが示された。しかしながら、この 2 組のアレイを広げていくと雑音や音源からの音が到達せずアレイをまたがる組み合わせのマイクロフォン間での類似度  $\mathbf{R}$  は下がり推定できない可能性がある。実際の利用を考える際にはこのトレードオフを考慮した最適な配置を検討する必要がある。

### 4.5.5 まとめ

本節では、立体構成をしたマイクロフォンアレイ 2 組を作成し、実際のドローン位置を音のみを使用する手法により推定した。位置の計算には相互相関の類似度を重みづけとした最適化計算による音源位置推定法を提案し、実験によって位置推定精度について評価した。その結果、提案手法は従来手法に比べ 3.2 倍の精度で位置測定可能であることが示された。しかしながら、この 2 組のアレイを広げていくと雑音や音源からの音が到達せずアレイをまたがる組み合わせのマイクロフォン間での類似度  $R$  は下がり推定できない可能性がある。実際の利用を考える際にはこのトレードオフを考慮した最適な配置を検討する必要がある。

## 4.6 まとめ

本章では音源位置推定の手法として 4 つのマイクロフォンアレイとカメラを用いる手法、4 つのマイクロフォンで反射波を用いる手法、8 つのマイクロフォンから 26 通りの到来時間差を用い最適化計算によって位置推定する手法について提案した。カメラを用いる手法では豚くしゃみ音が高周波成分をもつ特徴から相互相関法とハイパスフィルタを施した相互相関法を用い、相互相関法のみ手法では 50% 程度の正解率を得ることができた。またカメラを用いない音のみの位置推定では 20% 程度の正解率となった。カメラを用いることができない土中において音源位置推定を実施し、反射音を計算に用いることで従来手法よりも 3.4 倍位置推定精度が向上することがわかった。最後に、8 つのマイクロフォンを用い到来時間差のすべてを用い相互相関値を類似度として重みづけを行った最適化計算による位置推定手法によって、従来法に比べ 3.2 倍程度精度向上することがわかった。以上により、従来数十のマイクロフォンを用い音源位置推定を行っていた問題に対し様々なアプローチによって少数マイクロフォンでも音源位置推定ができることが示唆された。

## 第5章 結 論

本研究の目的は、雑音を有する様々な環境下で少数のマイクロフォンを用いて特定の音の検出及びその音源の位置推定を行うことであった。

音の検出のためには雑音の影響の少ない見本となる音を大量に用意する必要があり、用意が難しい音や見本となる音に雑音を含む場合への有効的な手法は提案されていない。そこで本研究では単一のマイクロフォンを用い、畜産業のひとつである養豚豚舎内で豚が呼吸器感染症になった際に生じるくしゃみ音の検出方法について検討を行った。豚舎内は、常に排気ファンやヒータ、冷房などの音に加え豚の鳴き声や柵などに当たる活動騒音、餌作業や清掃作業などに伴う作業音など多くの雑音が発生している環境であり、対象とするくしゃみ音は音声認識で用いられる見本データとなる音声コーパスがなくこの雑音環境から収集する必要がある。また、このくしゃみ音は常時発生していることがなく収集が難しく、従来の主要な音声認識法である大量学習によるモデル化ができない。そこで、このような雑音環境で発生回数が少ない音に対する音源の特徴抽出法について検討を行った。また特徴抽出法の汎用性を評価するため異なる環境で収録された音に対し実験を行い提案法の評価をした。

特定の音を検出する手法として、くしゃみ音に含まれる 30 kHz 以上にもなる高周波成分の特徴や周波数毎の継続時間から 5 つのテンプレートでくしゃみ音を検出し、誤検出の少なさを示す識別率は 63%、見逃しにくさを示す感度は 85% となり先行研究における感染症の陰陽性の識別に有効であることが示された。一方、他の豚房では本手法で検出することが難しいことがわかった。そこで、それぞれの豚房で収録できた 14 個の音のスペクトログラムによる相乗平均で特徴を抽出し、見本となるデータとして用いる手法を提案し、定常雑音の低減や周波数の時間方向の変動を特徴とする手法を提案した。この際類似度の計算手法として収録信号とくしゃみ音特徴を抽出した見本データをスペクトログラムの類似度を DP マッチングによって算出する手法を提案した。本手法によって、2 つの異なる豚舎においても識別率 54% 以上、感度 93% 以上となり、音による豚呼吸器感染症の発見法として提案手法の有効性が示された。

次に検出した音の発生位置を推定する手法について検討した。従来、マイクロフォンを用い音の発生位置を求めるためには多くのマイクロフォンを同時に収録する必要があった。本研究では上記の検出対象であるくしゃみ音に限らず、様々な環境で音の発生位置を推定するため、3 つの音源を対象に少数マイクロフォンでの位置推定の検討を行った。豚のくしゃみ音に対しマイクロフォンとカメラを組み合わせた位置推定法や、カメラや電磁波の届かない土中のドリル先端位置を模した音源を検出対象とし、土中の音源が地表面での反射音する特

---

徴を用い位置推定精度を向上させる方法について検討した。また屋外で飛行するドローンの飛行音から位置推定を行う手法について検討を行った。

音の発生位置の特定のため複数のマイクロフォンによって音の到達時間差が生じる原理を用い、4つまたは8つの少数のマイクロフォンによって音源の位置を推定する手法について検討をした。豚のくしゃみ音に対してカメラとマイクロフォンを組み合わせ角度推定から位置を推定する手法と従来法によるマイクロフォンのみでの位置推定についてその推定精度を検証し、カメラでは50%、マイクロフォンのみでは20%程度で発生位置を特定した。また、土中にマイクロフォンを差し込むジオフォンアレイを作成し、地表面反射によって少数でも位置推定精度を向上させる方法について検討し、実際の土中において提案手法の評価を行った。その結果、地表面に置く従来方法に対し、提案手法は3.4倍の精度向上が期待できることがわかった。最後に音の到来時間差の発見が難しい連続音に対し、立体構成マイクロフォンを作成し、ドローンを用い音の類似度によって重みづけされた最適化計算による手法を提案した。実際のドローンへ適応したところ、約6m離れたドローンに対し0.5m程度で位置推定できることがわかった。また類似度を用いない手法に比べ3.2倍の位置推定精度向上が可能であることが示された。

本研究によって従来は多くの見本となるデータや統計的な解析が必要であった特定の音を検出する手法に対し、雑音環境下かつ発生回数が少ない特定の音に対する特徴の抽出及びその検出法を提案することができた。本手法は雑音の排除が難しい環境での収録や、単一のマイクロフォンによって音源との距離が数mの比較的遠い環境でも有効であることから様々な環境での特定の音の検出法として応用が期待される。さらに、その発生位置を少数のマイクロフォンでも高精度に特定が可能であることから、特定の音の検出とその音源位置の情報によって今後の音を使ったセンシング技術に大きく貢献できると考えられる。

## 謝 辞

筑波大学システム情報系知能機能工学域 教授 水谷孝一先生には指導教員として研究の進め方や生活, 将来, 家族についてなど公私共々ご指導を頂きました. 特に学会発表や論文執筆に関して積極的に機会を与えて頂くとともに DDP (デュアルディグリープログラム) での副専攻履修についても勧めて頂き, 博士課程での深い洞察力と問題解決能力に加え, 様々な知識や広い視野, 多面的な物の考え方と同時に履修するための計画性及び胆力を身に着けることができました. 大学を選んでいた際, 準備生として研究室見学に来て本当によかったと改めて感謝致します. 日頃からのご指導, ご鞭撻本当にありがとうございました.

筑波大学システム情報系知能機能工学域 准教授 若槻尚斗先生には, 実験や論文執筆などを丁寧にご指導して頂き論文採用などの成果を上げることができました. また共通の趣味である自動車について, 研究相談から脱線し車の話をすることも多々ありましたが, 若槻先生のおかげできつく苦しいときも楽しく過ごせたことがこの研究室の居心地の良さの寄り心だと思っております. ご指導, ご鞭撻厚く御礼申し上げます. 筑波大学システム情報系知能機能工学域 准教授 海老原格先生には, 研究に関するアドバイスや海外出張の同行など多くのご指導をして頂きました. とても一つ年上の先輩には思えないほど鋭い洞察力や研究遂行能力, さらに研究室の学生指導に至るまで思料なさるお姿は私の目標です. 4年間のご指導, ご鞭撻厚く御礼申し上げます. 筑波大学システム情報系知能機能工学域 助教 善甫啓一先生には, 実験や論文執筆への多大なるご協力, 公私に渡る相談など本当にお世話になりました. 恐らく今後ともお世話になりますが, 善甫さんのおかげでこの論文の大半が出来たと言っても過言でないほどご指導頂きました. 感謝致します.

また, 本論文を執筆するにあたり, 筑波大学 システム情報系 構造エネルギー工学域 教授 京藤敏達先生, 農研機構動物衛生研究部門 領域長 兼 岐阜大学 客員教授 西藤岳彦先生, 筑波大学 システム情報系 知能機能工学域 助教 前田祐佳先生には, 専門的な見地から大変有益なご指摘をいただきました. 本論文内容及び発表内容に関しまして, 丁寧にご指導いただきましたことを深く感謝申し上げます. 特に農研機構動物衛生研究部門 領域長 西藤岳彦 様及び同所属 主任研究員 竹前喜洋 様には, 本論文では加筆できなかったインフルエンザへの感染させた豚の収録実験において多くの知見と研究成果を得ることができました. この収録実験によって本論文執筆時で 500 サンプル以上の感染した豚のくしゃみ音を得ることができました. 今後これらのデータを用いさらなる豚インフルエンザの研究の進展ができると確信しております. 引き続き音響研豚グループをよろしく願い申し上げます. 2012年当時, (独) 農研機構 畜産草地研究所において上席研究員 池口厚男 様 (現

宇都宮大学 教授) 及び業務 2 科の皆さまにはお忙しい中で実験場所を提供して頂き誠にありがとうございました。2014 年に収録を行わせて頂いた全農飼料中央研究所 養豚研究室の皆様は、お忙しい中実験場所を提供して頂き誠にありがとうございました。

豚舎内の音や画像の収録を行う音響研チーム豚の岩田氏、味藤氏、富樫氏には臭い、汚い、苦しい環境における収録のため、準備や設置、解析に渡り多くの苦勞を掛けました。最後まで自分にしっかりついてきてくれたおかげで、礎を築くことができました。本当に感謝しております。特に、スーパーアドバイザーである青木拓也氏によって強靱かつ正確なシステムを制作して頂いたおかげで豚舎の音を収録することができました。本当に感謝しております。

4 年間の研究室生活で出会った全ての先輩、同期、後輩に深く感謝しております。そんな研究室の学生の中でも、博士後期課程に在籍する、青木拓也氏、小木曾里樹氏、宮本隆典氏、には様々な相談に乗っていただいたことに加え、公私において様々なことを楽しく実施できたのは、間違いなくこの仲間の存在があったからだと思っております。本当にありがとうございました。趣味を押さえきれず先生方を含む 3 年で計 60 名近くの学生を巻き込み音響研自動車部（フジキンレンタカー）を立ち上げ、さらにこの仲間達とともに車両作成から全員初心者での耐久レース参加をし 2 年で大学ランキングトップに立つなど大きな成果を残すことができたことは一生の思い出です。趣味の集まりにも関わらず、私の厳しい開発オーダーや練習メニューに早朝から夜遅くまで研究と平行して付き合ってくれた多くの方々に感謝致します。他にも川岸ゼミに所属した多くの後輩達、研究室で私の愚痴に付き合ってくれた多くの方々など書ききれない人数の研究室の方々に感謝しております。皆さまのおかげで厳しくも楽しい研究生生活を送ることができました。

最後に、大学院に進学する機会を与えてくれた所属組織に感謝するとともに、何よりも大切な日頃の食事や家事、3 人の育児と家庭と私を支えてくれた妻 川岸志都香氏に深く感謝致します。

## 参考文献

- [1] K. Mizutani, N. Wakatsuki and T. Ebihara, "Introduction of measurement techniques in ultrasonic electronics :Basic principles and recent trends", Jpn. J. Appl. Phys., 55, 07KA02, 2016.
- [2] 宮崎裕道, 近藤高弘, 蜂谷弘之, “土中音波による位置探査システムの開発”, 土木学会第 64 回年次学術講演会, VI-145, pp.289-290, 2009.
- [3] 西村寛, “水産音響の研究と技術開発を回顧して”, 海洋音響学会誌, Vol. 18, No. 2, pp.9-13, 1991.
- [4] 竹内康人, 工藤雄一, “ビデオ信号級の水中通信における超音波と電磁波の使い分けについて”, 電子情報通信学会技術研究報告. US, 超音波, Vol. 105, No. 239, pp.1-5, 2005.
- [5] M. Yoshioka, S. Sato, T. Kikuchi, and Y. Matsuda, "Influence of Ultrasonic Nonlinear Propagation on Hydrophone Calibration Using Two-Transducer Reciprocity Method", Jpn. J. Appl. Phys.45, 4547, 2006.
- [6] Y. Wang, H. Hachiya, T. Nakamura, and H. Fujimori, "Estimation of Correlation of Acoustic Reciprocal Transmissions Using Central Equatorial Pacific Tomography Data", Jpn. J. Appl. Phys.44, 4729, 2005.
- [7] 松本三千緒, 山上順民, 蜂谷弘之, “土中音波による位置探査フィールド実験”, 土木学会第 64 回年次学術講演会, VI-146, pp.291-292, 2009.
- [8] J. Tamura, Y. Kawamura, H. Mochiji, N. Sasaki, K. Mizutani and H. Okawa, “A Method for Estimating the Location of the Drill-Bit During Horizontal Directional Drilling Using a Giant Magnetostriuctive Vibrator”, Jpn.J.Appl.Phys.50, 07HC15, 2011.
- [9] S. Furui, "Cepstral analysis technique for automatic speaker verification", IEEE Trans. Acoust. Speech Signal Processing, vol. 29, no. 2, pp. 254-272, 1981.
- [10] A. Mohamed, T. N. Sainath, G. E. Dahl, B. Ramabhadran, G. E. Hinton, M. Picheny, "Deep belief networks using discriminative features for phone recognition", Proc. ICASSP, pp. 5060-5063, 2011.

- 
- [11] 古井貞熙, "人と対話するコンピュータを作っています-音声認識の最前線-", 角川学芸出版, pp.17-64, 2009.
- [12] 中川聖, "ロバストな音声認識のための音響信号処理", 日本音響学会誌, Vol.53, No.11, pp.864-871, 1997.
- [13] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, B. Kingsbury, "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups", *IEEE Signal Processing*, Vol.29, No.6, pp.82-97, 2012.
- [14] A. Mohamed, G. Dahl, G. Hinton, "Acoustic modeling using deep belief networks", *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 1, pp. 14-22, Jan. 2012.
- [15] 大瀧淳司, 小菅太, 三澤勉, 植原孝之, 藤江龍一, "異音を定量的に検出する手法の開発", 富士通テン技報, Vol.23, No.2, pp.37-46, 2005.
- [16] 服部公央亮, 近江泰志, 田口亮, 梅澤太造, "CHLAC と連続 DP マッチングを用いた機械駆動音異常の自動検出", 電気学会論文誌 C, 電子・情報・システム部門誌, Vol. 131, No.2, pp.367-374, 2011.
- [17] 迫孝司, 徳重虎太郎, 吉江修, "週数変調解析によるすべり軸受ラビング以上早期検出技術", 電気学会論文誌 C, 電子・情報・システム部門誌, Vol. 132, No.9, pp.1454-1464, 2012.
- [18] Akira Saso, "Acoustic surveillance based on Higher-Order Local Auto-Correlation", *IEEE International Workshop on Machine Learning for Signal Processing*, 2011.
- [19] Akira Saso and Nyamerdene Odontsenel "Acoustic Novelty Detection Based on AHLAC and NMF", *IEEE International Symposium on Intelligent Signal Processing and Communication Systems*, pp872-874, 2012.
- [20] 佐々木洋子, 吉井和佳, 加賀美聡, "無限混合ガウスモデルを用いた未知クラスに対応可能な実環境音分類法", 人口知能学会 AI チャレンジ研究会, Vol. 36, pp.39-44, 2009.
- [21] 鳥谷部 達也, 上野 貴博, 森田 登, 中西 悠二, 渡邊 尚利, "電動機電磁騒音のケプストラム解析による絶縁診断の各種事例の紹介", 電気学会研究会資料, 回転機研究会, Vol.49, pp.7-12, 2012.
- [22] H. Yoshimura, H. Hazama, J. Aoki, M. Toyoda, Y. Naito, and K. Awazu, "Evaluation of a Delay-Line Detector Combined with Analog-to-Digital Converters as the Ion Detection System for Stigmatic Imaging Mass Spectrometry", *Jpn. J. Appl. Phys.*, Vol.50, 056701, 2011.

- [23] T. Motegi, K. Mizutani, and N. Wakatsuki, "Simultaneous measurement of air temperature and humidity based on sound velocity and attenuation using ultrasonic probe", *Jpn. J. Appl. Phys.*, Vol.52, No. 7, 07HC05, 2013.
- [24] 大賀寿郎, 山崎芳男, 金田豊, 音響システムとデジタル信号処理, コロナ出版, pp.47-48, 1995.
- [25] 古井貞熙, "音声情報処理," 森北出版, pp.2-24, 1998.
- [26] 前川喜久雄, 籠宮隆之, 小磯花絵, 小椋秀樹, 菊池英明, "日本語話し言葉コーパスの設計", 日本音声学会, 音声研究 Vo.4, No.2, pp.51-61, 2000.
- [27] 大石康智, "あらゆる音の検出・識別を目指して: 音響イベント検出研究の現在と未来", 日本音響学会 2014 年秋季研究発表会, 3-8-1, pp.1521-1524, 2014.
- [28] 中村哲, "実音響環境に頑健な音声認識を目指して", 電子情報通信学会技術研究報告. EA, 102(33), pp.31-36, 2002.
- [29] 松本弦, "音声認識における環境適応技術", 電子情報通信学会, 技術報告, SP99-111, 1999.
- [30] 農林水産省, 農林水産大臣官房統計部, 2015.
- [31] A. Frost, C. Schoeld, S. Beulah, T. Mottram, J. Lines, C. Wathes, "A review of livestock monitoring and the need for integrated systems", *Computers and Electronics in Agriculture*, Volume 17, Issue 2, 139-159, 1997.
- [32] M. White, "Porcine respiratory disease complex (PRDC)", *Livestock*, Vol.16, No.2, p.40, 2011.
- [33] M Shimizu, S Yamada, Y Murakami, T Morozumi, H Kobayashi, K Mitani, N Ito, M Kubo, K Kimura, M Kobayashi, et al. "Isolation of porcine reproductive and respiratory syndrome (PRRS) virus from Heko-Heko disease of swines", *J Vet Med Sci.* 56, 389-391, 1994.
- [34] W. Zhu, F. Zhong, X. Li, "Automated Monitoring System of Pig Behavior Based on RFID and ARM-LINUX", *IITSI, 2010 Third International Symposium*, 2-4 April, 431-434, 2010.
- [35] 仙名和浩, "呼吸器感染症の浸潤と咳・くしゃみ回数", 日本 SPF 豚研究会, *All About Swine*, 19, 9-11, 2001.
- [36] S. Ferrari, M. Silva, M. Guarino, J. M. Aerts, D. Berckmans, "Real-time recognition of sick swine cough sounds, *Computers and Electronics in Agriculture*", Volume 64, Issue 2, 318-325, 2008.
- [37] 宮原強, "豚の飼育管理上の主な留意点と今後の技術的検討課題", 日本豚病研究会, No.16, 1-5, 1989.

- 
- [38] 大臣官房統計部, 平成 27 年度肥育豚生産費(平成 28 年 12 月 9 日発表)資料, 農林水産統計, 2016.
- [39] 高木道浩, "豚繁殖呼吸障害症候群: 感染とワクチン", 日本 SPF 豚研究会, All about SWINE, Vol.47, pp.2-9, 2015.
- [40] 山根逸郎, 呉克昌, 石川弘道, 高木道浩, 宮崎綾子, 鈴木孝子, 芝原友幸, 久保正法, 小林秀樹, 國保健浩, 恒光裕, "PRRS の発生に関わる呼吸器疾患および繁殖障害などによる経済的な損失調査-アンケートを用いた疫学調査と全国の被害損額の推定-", 豚病会報, No.55, pp.33-39, 2009.
- [41] 経済産業省, 平成 27 年畜産物流統計-2.1.1 食肉卸売市場別月別取引成立頭数・価格-, 畜産物流通調査, 2016.
- [42] 塩原広之, "養豚経営-生き残るための手がかり(3)-", セミナー経営技術, (社)群馬県畜産協会・総括畜産コンサルタント  
( < <http://jlia.lin.gr.jp/cali/manage/128/k-semina/128ss2.htm> > 最終アクセス日: 2017.1.4)
- [43] 馬上齊, 今井杏子, 佐藤菜摘美, 仲山美樹子, 小見清, 中田稔, "事故率低減のための豚の施設管理・飼育管理・衛生管理改善指導", 新潟県広報誌, vol.39, No. 917, pp.1-5, 2016.
- [44] 山根逸郎, 呉克昌, 石川弘道, 高木道浩, 宮崎綾子, 鈴木孝子, 芝原友幸, 久保正法, 小林秀樹, 國保健浩, 恒光裕, "PRRS の発生に関わる呼吸器疾患および繁殖障害などによる経済的な損失調査-6 農場を対象にした調査結果-", 豚病会報, No.54, pp.8-13, 2009.
- [45] 佐々木 逸士, 佐々木 健, "非音声の環境音認識に関する研究", 精密工学会, 2015 年度精密工学会春季大会, pp. 973-974, 2015.
- [46] クグレ マウリシオ, トッサヴァイネン テーム, 中津 美紅, 岩田 彰, "環境音認識支援装置「サウンドウォッチャー」", 生体医工学, Vol. 53, No. Supplement, pp.S185\_02, 2015.
- [47] 渡辺 惇郎, 黒 柳奨, "パルスニューラルネットワークを用いた聴覚情報処理システムにおける混合音認識手法に関する研究", 電子情報通信学会技術研究報告. NC, ニューロコンピューティング 112(480), pp.273-278, 2013.
- [48] Christophe Couvreur, "Environmental Sound Recognition: A Statistical Approach", Facult'e Polytechnique de Mons, Phd Thesis, 1997.
- [49] K.S.R. Murty, B. Yegnanarayana, "Combining evidence from residual phase and MFCC features for speaker recognition", IEEE Signal Processing Letters, Vol.13, Issue 1, pp.52-55, 2006.

- [50] J. Darch, B. Milner, S. Vaseghi, "MAP prediction of formant frequencies and voicing class from MFCC vectors in noise," *Speech Communication*, Vol.48, Issue.11, pp.1556-1572, 2006.
- [51] J. C. Junqua, "The Lombard Reflex and its Role on Human Listener and Automatic Speech Recognizers", *J. of Acoust. Soc of Am*, Vol.1, pp510-524, 1993.
- [52] Takuji Kawagishi, Keiichi Zempo, Koichi Mizutani, Naoto Wakatsuki, Tadashi Ebihara, and Youhei Kawamura: "Detection of ship screw deflection using small element number of hydrophone array," *Proceedings of the 34th Symposium on Ultrasonic Electronics*, Vol. 34, 2P6-2, pp. 357-358, 2013.
- [53] 鳥谷部 達也, 上野 貴博, 森田 登, 中西 悠二, 渡邊 尚利, "電動機電磁騒音のケプストラム解析による絶縁診断の各種事例の紹介", *電気学会研究会資料, 回転機研究会*, Vol.49, pp.7-12, 2012.
- [54] T. Ahmed, M. Uppal, A. Muhammad, "Improving efficiency and reliability of gunshot detection systems", *IEEE Int. National Conf., Speech and Signal Processing*,(DOI: 10.1109/ICASSP.2013.6637700) 2013.
- [55] A. Mesaros, T. Heittola, A. Eronen, T. Virtanen, "Acoustic event detection in real life recordings", *IEEE, Signal Processing Conference, 18th European*, 2010.
- [56] 朴玄信, 滝口哲也, 有木康雄, "制約付き非負行列因子分解を用いた音声特徴抽出の検討", *電子情報通信学会, 言語理解とコミュニケーション*, Vol.108(337), pp.43-48, 2008.
- [57] 澤田宏, "非負値行列因子分解 NMF の基礎とデータ/信号解析への応用", *電子情報通信学会*, Vol.95, No.9, pp.829-834,2012.
- [58] 正田備也, 喜安千弥, 宮原末治, "肺音分類のための独立成分分析とクラスタリングによる特徴量抽出", *情報処理学会, データベースと Web 情報システムに関するシンポジウム*, pp.1-8,2007.
- [59] 牧野昭二, 荒木章子, 向井良, 澤田宏, "独立成分分析に基づくブラインド音源分離", *電子情報通信学会. EA, 応用音響*, Vol.103, No.129, pp.17-24, 2003.
- [60] 藤本雅清, 有木康雄, "GMMに基づく音声信号推定法を用いた雑音下音声認識", *情報処理学会研究報告音声言語情報処理*, Vol.121, pp.25-30, 2002.
- [61] A. Tjahyanto, D. P. Wulandari, Y. K. Suprpto, M. H. Purnomo, "Gamelan instrument sound recognition using spectral and facial features of the first harmonic frequency", *Acoustical Science and Technology*, Vol.36, No.1, pp.12-23, 2015.
- [62] 川野憲市郎, 榎本崇宏, 芥川正武, 木内陽介, 小中信典, 川田育二, 陣内自治, ウダインタアビラトナ, "マイクロフォンにより録音された呼吸音のニューラルネットワークに基づく高速検出法", *生体医工学会*, Vol. 53, No. Supplement, pp.S239\_01, 2015.

- 
- [63] 63 J. Carletta, "Announcing the AMI Meeting Corpus", The ELRA Newsletter, Vol. 11, No.1, pp. 3-5, 2006.
- [64] 64 The British Broadcasting Corporation, BBC SOUND EFFECTS LIBRARY, Corp. SOUND IDEAS, CD/CD-ROM(40),2000.
- [65] 65 国立情報学研究所 音声資源コンソーシアム(National Institute of Informatics - Speech Resources Consortium)  
(<http://research.nii.ac.jp/src/index.html>>最終アクセス日: 2017.1.4)
- [66] X Zhou, X Zhuang, M Liu, H. Tang, M. Hasegawa, J. T. Huang, "HMM-Based Acoustic Event Detection with AdaBoost Feature Selection", Multimodal Technologies for Perception of Humans, vol. 4625, pp. 345-353, 2008.
- [67] 山田眞吾, "コストロスモデルに基づいた天気用法の評価指数の提案", 日本気象学会, 短報 401:503, pp.759-766, 2001.
- [68] 川久保清, "スクリーニング検査でみられた運動誘発 ST 偏位と冠危険因子", 日本産業衛生学会, Vol.36, No.6, pp.440-445, 1994.
- [69] A. H. Katherine, "Wiley-Blackwell", 4th edition, p.31, 2004.
- [70] 川岸卓司, 水谷孝一: "マイクロフォン (音響的手法) を用いる豚呼気感染症の検出法," JOYO INDUSTRIAL RESEARCH 常陽産研 NEWS, Vol. 295, pp. 16-19, 2015.
- [71] 佐藤洋, 長友宗重, 吉野博, 矢島吉紀, "残響・騒音の音声聴取に及ぼす影響の評価に関する実験的検討", 日本建築学会計画系論文集, Vol.484, pp.1-8, 1996.
- [72] 日本工業規格, 音の伝搬における空気吸収の計算, JIS Z, 8738, 1993.
- [73] T. Kawagishi, K. Mizutani, K. Zempo, and N. Wakatsuki: "Monitoring of swine cough and sneezing sounds using time-frequency analysis for detecting respiratory tract infection," Proceedings of the 22nd International Congress on Sound & Vibration (ICSV22), Florence, Italy, CD-ROM, T04.RS07-0785(R), 8 pages, 2015.
- [74] T. Kawagishi, K. Mizutani, K. Zempo, and N. Wakatsuki: "Monitoring of Swine Sneezing Using Multi Templates Matching for Detecting Respiratory Tract Infection," 7th International Conference on Sustainable Agriculture for Food, Energy and Industry in Regional and Global Context (ICSAFEI2015), Serdang, Malaysia, USB, ICSAFEI-134(R), 7pages, 2015.
- [75] 北岡教英, 赤堀一郎, 中川聖一, "スペクトルサブトラクションと時間方向スムージングを用いた雑音環境下音声認識", 電子情報通信学会論文誌 D, Vol.J83-D2, No.2, pp.500-508, 2000.
- [76] T. Lotter, P. Vary, "Speech enhancement by map spectral amplitude estimation using a super-Gaussian speech model", EURASIP Journal on Applied Signal Processing, Vol. 2005, pp.1110-1126, 2005.

- [77] X. Lu, Y. Tsao, P. Shen and C. Hori, "Spectral patch based sparse coding for acoustic event detection," Proc. ISCSLP 2014, pp. 317-320, Sept. 2014.
- [78] 東京大学教養学部統計学教室, 基礎統計学 I 統計学入門, 東京大学出版会, pp.30-32,1991.
- [79] K. Briechle, U. D. Hanebeck, "Template matching using fast normalized cross correlation", Proc. SPIE 4387, Optical Pattern Recognition XII, 95, 2001.
- [80] 迫江博昭, "音声認識と DP", 日本オペレーションリサーチ学会, Vol24, pp.324-330, 1985.
- [81] 益子貴史, 徳田恵一, 小林隆夫, 今井聖, "動的特徴を用いた HMM に基づく音声合成", 電子情報通信学会論文誌 D, Vol.J79-D2, No.12, pp.2184-2190, 1996.
- [82] S. Furui, "Speaker-independent isolated word recognition using dynamic features of speech spectrum", IEEE Signal Processing Society, Vol.32, Issue 1, pp.52-59, 1986.
- [83] R. Bellman, An introduction to the theory of dynamic programming, The Rand Corporation, Santa Monica, Calif., 1953
- [84] R. Turetsky, and D. Ellis, "Ground-Truth Transcriptions of Real Music from Force-Aligned MIDI Syntheses", 4th International Symposium on Music Information Retrieval ISMIR-03, pp.135-141, 2003.
- [85] D. Ellis (2003). Dynamic Time Warp (DTW) in Matlab  
( < <http://www.ee.columbia.edu/~dpwe/resources/matlab/dtw/> > 最終アクセス日: 2017.1.4)
- [86] 浅野太, 音のレイ信号処理-音源定位・追跡と分離, コロナ社, pp.69-106,2011.
- [87] 安部正人, "多数センサによる音源推定", 日本音響学会誌, Vol.51, No.5, pp.384-389, 1995.
- [88] 宝珠山治, 杉山昭彦, "音響ビームフォーミングと多次元信号処理," 電子情報通信学会技術研究報告, 97(518), pp.1-8, 1998.
- [89] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas and Propagation, Volume AP-34, No. 3, pp. 276-280, Mar. 1986.
- [90] 大井堂史昌, 陶山健仁, "実環境における MUSIC 法による到来方向推定の改善," 電子情報通信学会技術研究報告, 106(66), pp.11-16, 2006.
- [91] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. ASSP, Volume 24, Issue 4, pp. 320-327, Aug. 1976.

- 
- [92] 西浦敬信, 山田武志, 中村哲, 鹿野清宏, “マイクロホンアレーを用いた CSP 法に基づく複数音源位置推定,” 電子情報通信学会技術研究報告, J83-D-II(8), pp.1713-1721, 2000.
- [93] 鈴木敬, 金田豊, “サブバンドピークホールド処理を用いた音源方向推定,” 日本音響学会誌 65 巻 10 号, pp.513-522, 2009.
- [94] 城戸 健一, 劉 家祥, 野戸 広之, 安倍 正人, “異常音の発生位置とそのスペクトルの推定”, 電気学会論文誌 C, 電子・情報・システム部門誌, Vol. 110, No. 9, pp.547-553, 1990.
- [95] A. Y. Nakano, L. Wang, K. Yamamoto, S. Nakagawa, "Acoustic source localization based on distributed microphone arrays in a living room", Proc. of the 2008 autumn meeting of ASJ, 2-6-18, pp.703-706, 2008.
- [96] 山田 武志, 中村 哲, 鹿野 清宏, “マイクロホンアレーによる 3 次元ビタビ探索に基づく移動話者の音声認識”, 電子情報通信学会技術研究報告, 音声, Vol.97, No.115, pp.31-38, 1997.
- [97] 川岸卓司, 善甫啓一, 水谷孝一, 若槻尚斗, 川村洋平: “直達音と地面反射音を利用する少素子ジオフォンアレイによる土中音源位置推定”, 土木学会, 土木学会論文集 F3(土木情報学), Vol.69, No. 2, pp. I\_130-I\_138, 2014.
- [98] HDD 推進委員会, “HDD(誘導式水平ドリル)工法の適用標準とそれを超える施工事例仕様書”, No-Dig Today, No.73, pp.81-84, 2010.
- [99] 和田洋, “非開削による配管埋設技術の高度化”, 建設の施工企画, No. 658, pp. 15-20, 2004.
- [100] 谷口慶祐, “地震波を用いた地下構造の推定に関する実験の実践例”, 京都教育大学環境教育研究年法, Vol. 13, pp. 131-144, 2005.
- [101] 川岸卓司, 善甫啓一, 水谷孝一, 若槻尚斗, 川村洋平: “地面反射を利用した土中の音源位置推定,” 土木学会, 2013 年度土木情報学シンポジウム講演集, Vol. 38, (12), pp.39-40, 2013.
- [102] 久野和宏, 野呂雄一, 井研治, “室内音響における波動論と音線理論の関係”, 日本音響学会, Vol. 49, No. 2, pp. 77-83, 1993.
- [103] J. W. Rector and B. P. Marion, “The use of drill-bit energy as a downhole seismic source”, GEOPHYSICS, Vol. 56, No. 5, pp. 628-634, 1991.
- [104] K. Zempo, T. Ebihara, and K. Mizutani, “Direction of arrival Estimation Based on Delayed-Sum Method in Reverberation Environment”, Jpn. J. Appl. Phys., 51, 07GB09, 2012.

- [105] F. Caballero, L. Merino, J. Ferruz and A. Ollero : Vision-Based Odometry and SLAM for Medium and High Altitude Flying UAVs, *Journal of Intelligent & Robotic Systems*, Vol. 54, Issue 1, pp 137-161, 2009.
- [106] 奥谷啓太, 吉田尚水, 中村圭佑, 中臺一博 : クロドロコプタ搭載のマイクロホンアレイを用いた屋外音環境理解の逐次雑音推定による向上, *日本ロボット学会誌*, Vol. 31, pp.676-683, 2013.
- [107] 鳥羽晃平, 今城直人, 高田洋吾 : スピーカとマイクロホンを用いた位置検出システムによるロボットの位置推定, *ロボティクス・メカトロニクス講演会 2015*, 2A2-M03, 2015.
- [108] J. Baek, S. Park, B. Cho and M. Lee : Position tracking system using single RGB-D Camera for evaluation of multi-rotor UAV control and self-localization, *IEEE AIM International Conference*, pp. 1283-1288, 2015.
- [109] H. Oh, D. Won, S. Huh, D. H. Shim and M. Tahk : Indoor UAV Control Using Multi-Camera Visual Feedback, *Journal of Intelligent & Robotic Systems*, Vol. 61, Issue 1-4, pp 57-84, 2011.
- [110] K. Okutani, T. Yoshida, K. Nakamura and K. Nakadai : Outdoor Auditory Scene Analysis Using a Moving Micro-phone Array Embedded in a Quadcopter, *IEEE RSJ International Conference*, pp. 3288-3293, 2012.
- [111] S. C. Lee, W. R. Lee, K.H. You : TDoA Based UAV Localization Using Dual-EKF Algorithm, *Control and Automation*, Vol. 65, pp.47-54, 2009.
- [112] 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: "立体構成マイクロフォンアレイを用いる UAV 飛行音の位置推定," *土木学会, 2015 年度土木情報学シンポジウム講演集*, Vol. 40, pp.199-200, 2015.
- [113] 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: "風影響下における立体構成マイクロフォンアレイでのドローン位置推定," *日本音響学会, 2016 年春季研究発表会講演論文集, CD-ROM, 1-Q-35*, pp.1285-1286, 2016.
- [114] Y. T. Chan and K. C. Ho : A Simple and Efficient Estimator for Hyperbolic Location, *IEEE Trans. SP*, Vol. 42, No. 8, pp. 1905-1915, 1994.
- [115] C. Li, Y. Li and Y. Shen, L. Liu and Q. Cao: An Optimization Algorithm for Wireless Sensor Networks Localization Using Multiplier Method, *IEEE CSO Conf.*, Vol. 2, pp. 337-341, 2010.

## 付録 A

# 市販コンデンサマイクロフォンによる 高周波測定のための計測実験

### A.1 通常のマイクロフォンで高周波を測定するために

本研究では 20 – 50 kHz もの高周波信号を市販のコンデンサマイクロフォンによって計測している。通常マイクロフォンは人の声を収録する目的で作成されているため、収録に対応している周波数は高い場合でも 20 kHz 程度のみデータシートしか付属していない場合が多い。Fig. A-1 に本研究で使用するマイクロフォンのデータシート<sup>[A-1]</sup>を示す。

そこで、本研究で使用するコンデンサマイクロフォンでは使用の前に高周波計測用のマイクロフォンとともに計測を行い、その周波数特性を計測することによって安価かつ高感度で高周波まで対応できるマイクロフォンを市販品の中から探し使用した。本付録では、その計測方法及び結果について示す。

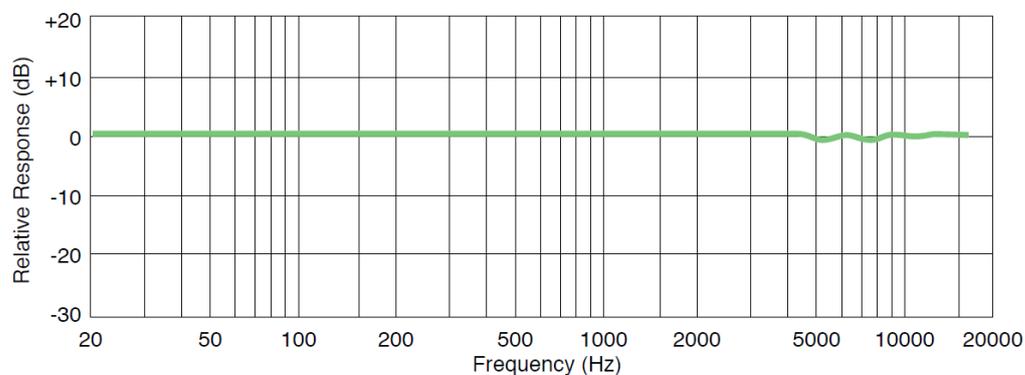


Fig. A-1 WM-62PC の周波数特性.

## A.2 計測方法及び計測結果

100 kHz まで集音可能な高周波対応の計測マイクロフォン（Brüel & Kjær, Type 4939-A-011）と実験に使用したコンデンサマイクロフォン（Panasonic, WM62PC）及びアンプ（SONY:PCM-D50）を用いて同時に収録し、正弦波スイープ信号をスピーカーから発生させ 0 - 45 kHz の周波数帯域の計測マイクとの比較を行った。サンプリング周波数は論文と同様に 96 kHz とした。その結果を Fig. A-2 に周波数スペクトルを示す。

比較の結果、計測マイクに対して 20 kHz 付近まで同等の周波数特性を有し 25 - 35 kHz 付近まで -10 dB 以内、45 kHz まで -20 dB 以内の特性で有している、無音状態でのノイズレベルを測定し、ノイズに対してして 40 - 45 kHz 付近で +20 dB のマージンがあることから、本計測実験では感度は低下するもののノイズレベルより十分大きく収録できることが確認できた。

同様に、研究で使用した他のマイクロフォンにおいても計測を行うことで、ノイズレベルよりも十分大きく高周波が測定できるマイクロフォンを使用することができた。なお、マイクロフォンの製造過程のバラつきによって同じ品名のマイクロフォンであっても高周波の特性が担保されない点には注意が必要である。

また実験で使用を繰り返し使用した結果、温度や湿度によって感度が低下<sup>[A-2]</sup>することから定期的な交換や感度のチェックなどが必要だと考えられる。

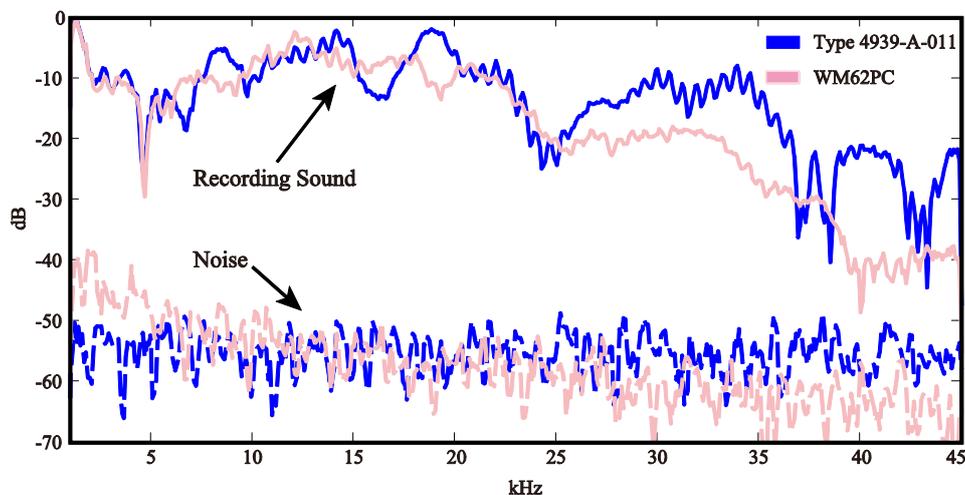


Fig. A-2 WM-62PC 及び Type 4939-A-011 の周波数特性

参考文献等

- [A-1] WM-62PC/62PK Omnidirectional Back Electret Condenser Microphone Cartridge,  
Panasonic .
- [A-2] 瀧浪弘章,” 騒音計の性能に影響を与える様々な要因”, 日本騒音制御工学会, 騒  
音制御, Vol.29, No.5, pp.332-338, 2005.

## 付録 B

### 地表面反射音による音源位置推定法

### 実規模シミュレーション

#### B.1 実規模シミュレーションの条件

本研究の4章4節における土中の位置推定を実規模を想定した設置条件にした場合の設置場所における計測誤差の影響や、各センサにおける誤差について検討を行った。4章4節における評価実験より、4つの素子による手法では、提案手法および従来手法ともに素子の構成する面（素子構成面）に対して測位分解能が低いことがわかった。また、素子設置時の誤差が測定結果に影響を与えることから、音源の深さをHDD工法でのドリル先端深さを考慮して10 m として2つの実験を行った。素子構成面の測位誤差を考察するため、音源の方向を正面方向から側面方向まで音源の角度を変化させ位置測定結果の比較を行った。また、提案手法および従来手法ともに最も誤差が小さくなる方向に対し、素子の1つの位置を移動させた場合における、素子の位置誤差が与える音源位置の測定精度への影響について比較を行った。

位置測定シミュレーションのシミュレーション条件について **Table B.1** に示す。音源には約 4 m 離れた地点から加振した際に受信した音源の周波数特性と近似したガウシアンパルス音源を使用した。**Fig. B.1** に受信した信号周波数及びシミュレーションで用いた音源を

Table B.1 シミュレーション条件

$L_E$	2.0 m
$L_A$	2.5 m
$L_B$	2.0 m
Radius	2.0 m
Depth of the sound source	10.0 m
Sampling frequency	96 kHz
Sound velocity	1,500 m/s

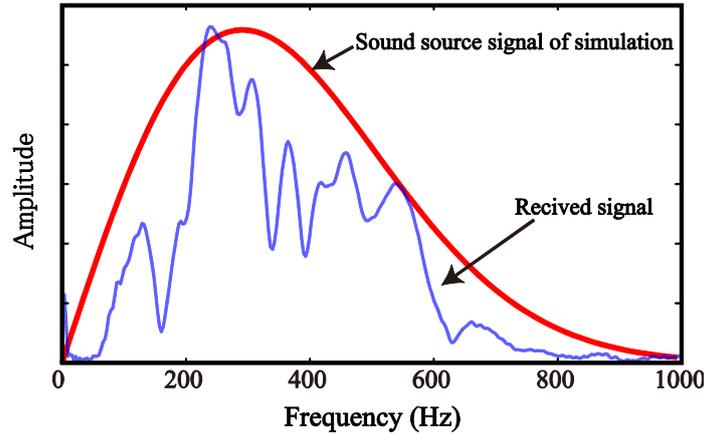


Fig. B-1 計測信号とシミュレーションで使用した音源

示す。なお、振幅については、表示のために正規化した値を用いた。音源には中心周波数を 320 Hz のガウシアンパルス信号を用いた。

## B.2 センサ構成面に対する実験と考察

実験の概略図を Fig. B.2 に示す。評価実験と同様の配置として、提案手法を  $x-z$  平面の配置として従来手法を  $x-y$  平面とした。この際、各素子を評価実験の 10 倍の間隔とし、音源位置を深さ 10 m のまま  $\theta$  を 0 から 90 度まで変化させた。各角度に対して受信信号から音源位置を測定し、提案手法と従来手法の各座標軸における真値との誤差を求めた。Fig. B.3 にその結果を示す。

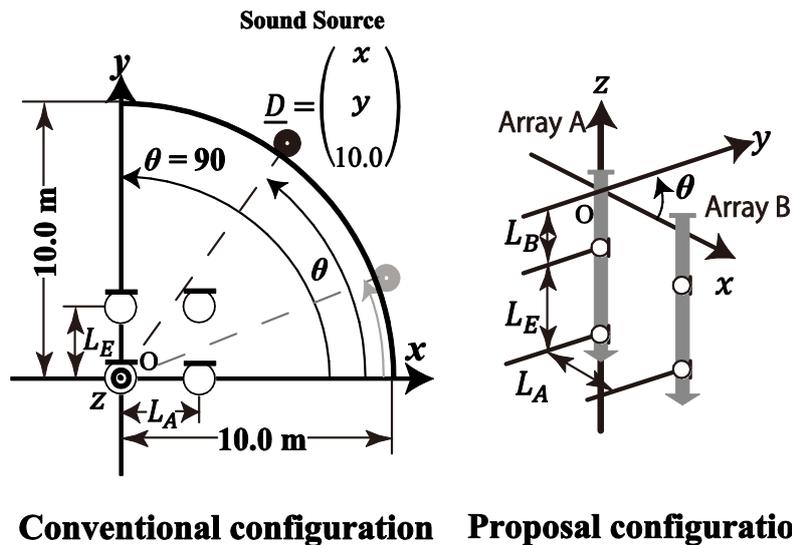


Fig. B-2 提案手法と従来手法によるセンサ構成とシミュレーション条件

a) 各軸方向成分の誤差の考察

Figure B.3 の  $x$  軸誤差については、提案手法の誤差は従来手法に比べ少なく一定であるが、従来手法については 75-90 度までが最大誤差 10 m 程度となり、75 度以前については 0.5 m 以下の誤差となっている。誤差が最大となった 83 度付近について、 $y$  座標は -0.25 m と誤差は小さいものであった。その要因として、 $y$  軸方向に 2 つの素子が音源に対し正面となることで、3 つの到来時間差から座標を推定する従来手法において 2 つの到来時間差が同一の値となり、従来手法の行列式で位置を計算する解が不安定となったことが考えられる。同様の理由から、 $x$  軸方向に 2 つの素子が音源に対し正面となる 7 度付近で誤差が大きくなった。提案手法については、 $y$  軸方向に 2 つの素子が音源に対し正面となる約 83 度で、到来時間差が同一の値となることで、解が不安定になる箇所での誤差が最大となる。しかしその誤差は  $y$  軸方向で約 0.5 m 程度であり、提案手法は従来手法に比べ解が不安定となる方向に対しても誤差が小さいことが示された。

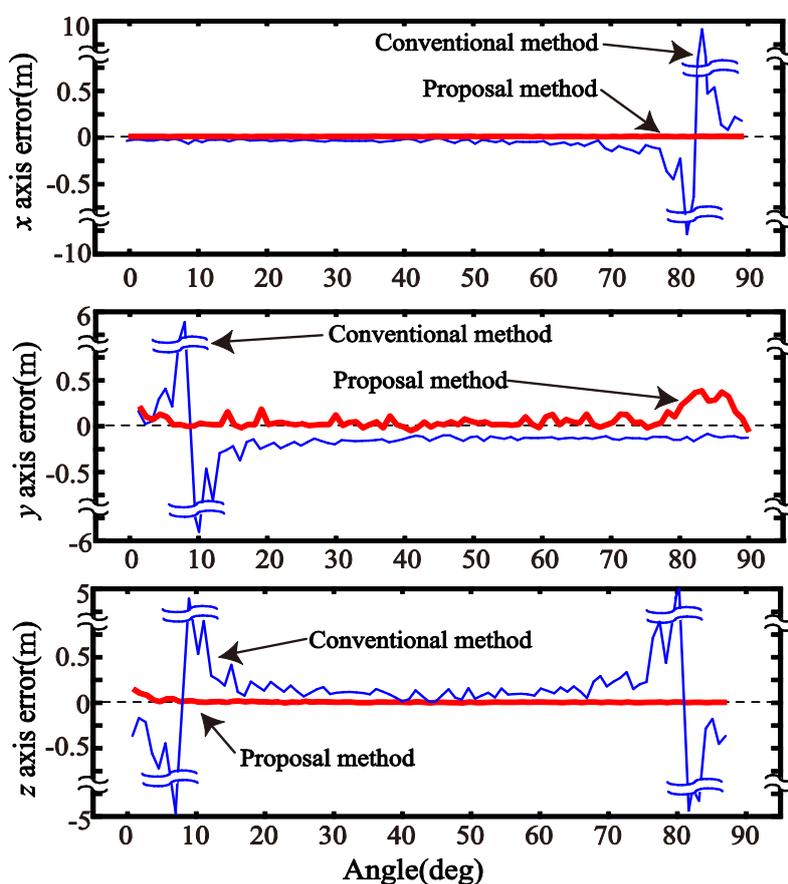
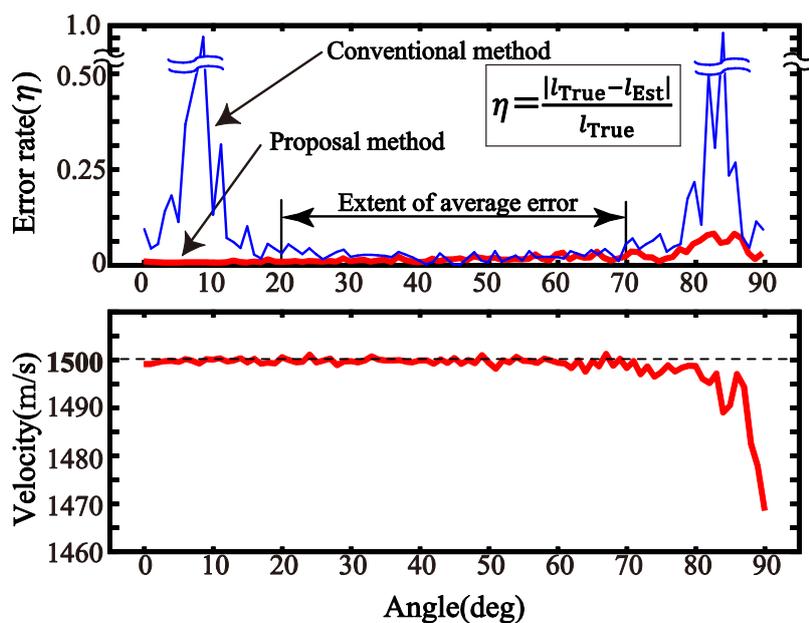


Fig. B-3 提案手法と従来手法による各軸の計測誤差

## b) 測定誤差率及び音速推定

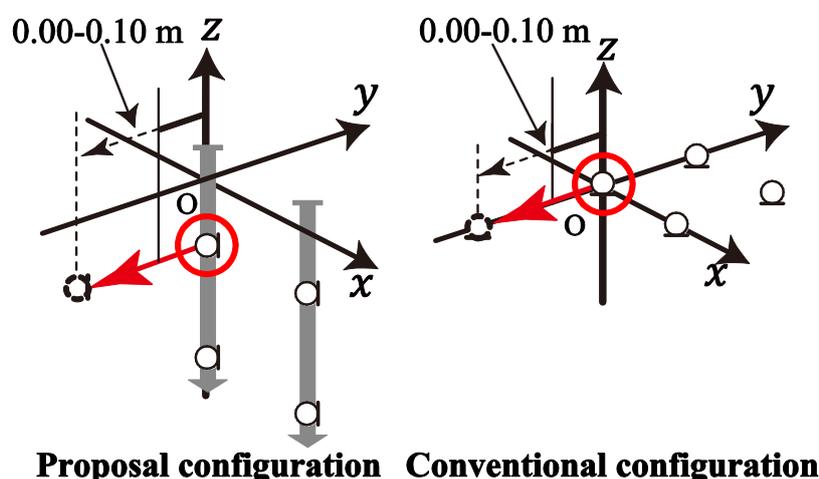
測定誤差率及び提案手法で求めた音速の値について **Fig. B.4** に示す。提案手法の測定誤差率は 0 - 75 度付近まで  $\eta_P = 0.12$  以下の誤差率であった。従来手法は 35 - 40 度付近で誤差率が  $\eta_C = 0.10$  以下となった。しかし、従来手法は素子構成面に対して 2 つの素子が音源の正面となる 7 度および 83 度付近で誤差率は最大となった。誤差率の低い角度での比較するため平均測定誤差率として、20 度から 70 度の測定誤差率の平均値を算出すると、提案手法が  $\eta_P = 0.0115$  であり、従来手法は  $\eta_C = 0.0182$  であった。よって、到来時間差が正しく得られ、解が正しく計算できる平均測定誤差率内の範囲においても、提案手法は従来手法に比べ 1.58 倍の精度で位置測定可能であることが確認された。また、提案手法による音速推定の計算結果については素子構成面に対し音源が正面に位置し、解が不安定となる 80 - 90 度付近を除き、音速を正しく推定できることがわかった。



**Fig. B-4** 音速の計算結果及び測定誤差率の比較

### B.3 素子の設置誤差に対する実験と考察

前項の実験と同様の配置で従来手法と提案手法における素子の設置誤差について比較した。この際、提案手法と従来手法ともに誤差が少ない 35 度において、素子 0 を  $-y$  軸方向に 0 から 0.10 m まで 0.01 m 刻みに移動させた。これにより、1 つの素子の設置誤差が音源位置測定に与える誤差について考察した。実験の概略図を **Fig. B.5** に示す。この際、提案手法では素子 1 つの直達音及び反射音の受信信号が遅延することとなり、他の素子への受信信号は前項の実験と同様の値を受信信号として計算した。なお、それぞれの素子の配置および 35 度における音源の位置、音速については前項の実験と同様の値を用いた。



**Fig. B-5** 提案手法と従来手法によるセンサ構成とシミュレーション誤差の条件

#### a) 各軸方向成分の誤差の考察

**Figure B.6** に各座標軸に対する音源位置との測位誤差について示す。提案手法では、各軸ともに最大で 0.7 m 程度であった。従来手法では、素子位置の誤差が音源位置の測定精度に与える影響が大きく、 $x$  軸方向に 12.2 m、 $y$  軸方向に 8.1 m、 $z$  軸方向に 13.4 m の音源測位誤差が素子設置位置誤差に対し、ほぼ線形に増加した。これは、従来手法が 4 つの素子に対して発生する 3 つの到来時間差をもとに位置測定 ( $x, y, z$ ) を行うため、1 つの素子の到来時間差の誤差が位置推定に直接影響することが原因と考えられる。提案手法では、7 つの到来時間差から 4 つの未知数 ( $x, y, z, c$ ) を求めるため、1 つの素子の位置ずれによって起こる 2 つの到来時間差の誤差を平均化し、推定結果の誤差を抑えることになる。

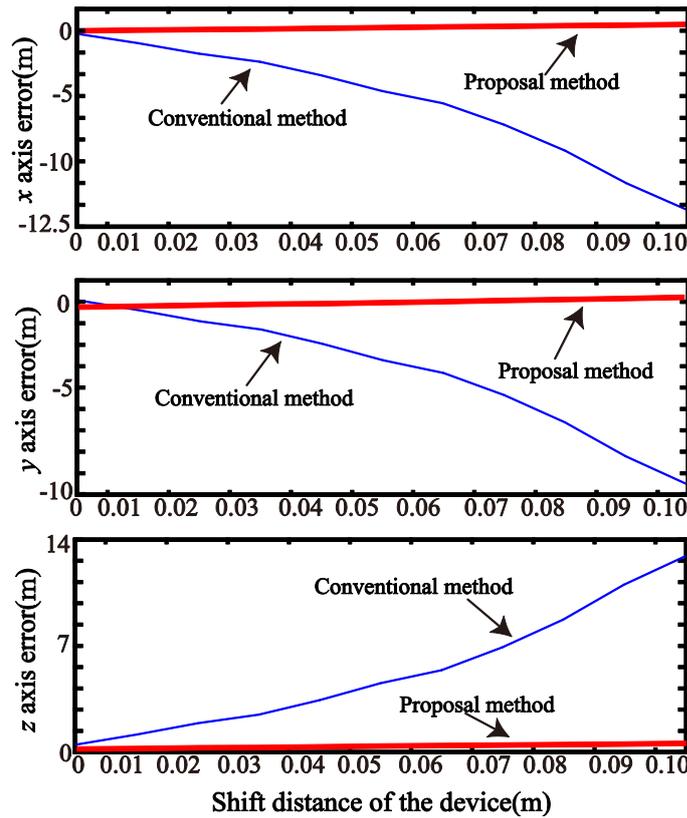


Fig. B-6 提案手法と従来手法による各軸の計測誤差

b) 測定誤差率及び音速推定

Figure B.7 に素子の設置位置に誤差を与えた場合の測定誤差率および音速の推定結果について示す。従来手法では素子の設置位置誤差 0 m のとき測定誤差率は  $\eta_c = 0.13$  であり、設置位置誤差 0.1 m のとき測定誤差率  $\eta_c = 1.32$  であり、音源測位誤差は 10 倍以上大きくなった。これに対し提案手法の音源測位誤差は、誤差 0 m のときに測定誤差率は  $\eta_p = 0.12$  であり、素子の設置位置誤差が最大の 0.10 m のときに最小値の  $\eta_p = 0.09$  となり、提案手法の位置推定精度は従来手法に比べ 14.6 倍向上することがわかった。提案手法において測定誤差率が減少した理由として、Fig. B.6 の y 軸方向の誤差は負方向から正方向に増加しておりその絶対値は誤差 0 m のときから減少していることから結果として測位誤差を低減していることがわかった。また、音速の結果については、誤差の最大値が約 +30 m/s となり Fig. B.4 の音速最大誤差と同程度の誤差が推定された。よって 1 つの素子を -y 軸方向に位置誤差を与えた場合、音速に与える影響が大きくなることがわかった。

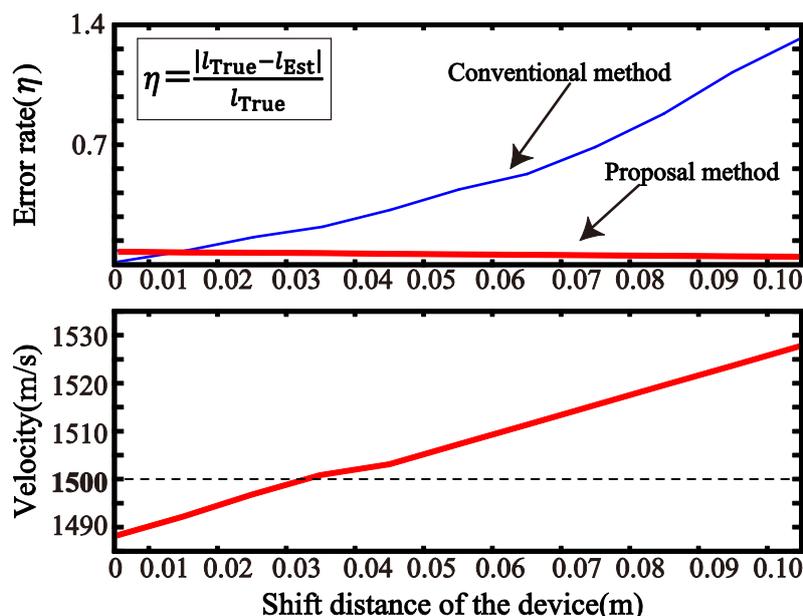


Fig. B-7 音速の計算結果及び測定誤差率の比較

## B.4 まとめ

素子構成面に対する音源方向の測定位置誤差について検証し、提案手法は0-75 度付近で測位誤差が小さく、従来手法では35-40 度付近で最も測位誤差が小さかった。これは、素子構成面が音源に対して法線軸方向に位置する場合を除き、各素子に対する音源からの伝搬経路長が長くなる分だけ、到来時間差を比較的正確に計測できることから誤差が低減されていると考えられる。この際、提案手法は従来手法に比べ1.58 倍の精度で位置測定可能であることが確認された。よって発生する音源に対し、適切な素子構成面を考慮することで、位置測定ができることが示された。また提案手法では、音速の推定も80-90 度を除き可能であることがわかった設置誤差の影響と測定結果のシミュレーションでは、提案手法と従来手法ともに誤差が少ない35 度において、素子0 を-y 軸方向に0 から0.10 m まで0.01 m 刻みに移動させ、1つの素子の位置誤差が音源測位に与える誤差について考察した。実験の結果、従来手法では、素子の位置誤差に比例して音源位置の測定精度が低下し誤差が10 倍程度となった。これに対し、提案手法では-y 軸方向への素子位置誤差に対して音源測定精度に大きな影響はない一方、音速の推定値に与える影響が最も大きくなりその誤差が+30 m/s であることがわかったその結果、提案手法は従来手法に比べ素子の設置位置誤差が音源の位置測定に与える誤差が小さいことがわかった。

本実験によって、規模を大きくした場合にも提案手法が有効であることが示された。しかし、距離が大きくなることによる減衰や連続音に対する処理、他の反射環境がある場合などについてはさらなる検討が必要であると考えられる。

## 著者論文

### 公表済み論文（学位請求申請に関連のある論文）

#### 査読付雑誌論文

- (J-1) 川岸卓司, 松梨夏季, 水谷孝一, 若槻尚斗: “豚の咳・くしゃみ監視システムに関する研究”, 農業施設学会, , vol45, Vol. 45, No. 1, pp. 1-7 (March 2014).
- (J-2) 川岸卓司, 善甫啓一, 水谷孝一, 若槻尚斗, 川村洋平: “直達音と地面反射音を利用する少素子ジオフォンアレイによる土中音源位置測定”, 土木学会, Vol.69, No. 2, 特集号, pp. I\_130-I\_138, (March 2014).

#### 査読付国際会議論文

- (C-1) Takuji Kawagishi, Koichi Mizutani, Keiichi Zempo, and Naoto Wakatsuki: "Monitoring of swine cough and sneezing sounds using time-frequency analysis for detecting respiratory tract infection," Proceedings of the 22nd International Congress on Sound & Vibration (ICSV22), Florence, Italy, CD-ROM, T04.RS07-0785(R) (8 pages) (14 July, 2015).
- (C-2) Takuji Kawagishi, Koichi Mizutani, Keiichi Zempo, and Naoto Wakatsuki: "Monitoring of Swine Sneezing Using Multi Templates Matching for Detecting Respiratory Tract Infection," 7th International Conference on Sustainable Agriculture for Food, Energy and Industry in Regional and Global Context (ICSAFEI2015), Serdang, Malaysia, USB, ICSAFEI-134(R) (7 pages) (26 August, 2015).

---

## 公表済み論文（学位請求申請に関連のない論文）

### 査読付雑誌論文

- (J-1) Satoki Ogiso, **Takuji Kawagishi**, Koichi Mizutani, Naoto Wakatsuki, and Keiichi Zempo: " Self-localization method for mobile robot using acoustic beacons," ROBOMECH Journal, doi:10.1186/s40648-015-0034-y, (2 September 2015).
- (J-2) **川岸卓司**, 川口明子, 小林隆史, 大澤義明: “訪日外国人の訪問地集中指標と一極集中是正に向けた分析-北海道の宿泊地を事例として-”, 計画行政学会, Vol. 38, No. 4, pp.50-59, 2015.

### 査読付国内会議論文

- (JC-1) **川岸卓司**, 善甫啓一, 水谷孝一, 若槻尚斗, 川村洋平: “地面反射を利用した土中の音源位置推定,” 土木学会, 2013 年度土木情報学シンポジウム講演集, Vol. 38, (12), pp.39-40 (17 September, 2013).
- (JC-2) **川岸卓司**, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: "立体構成マイクロフォンアレイを用いる UAV 飛行音の位置推定," 土木学会, 2015 年度土木情報学シンポジウム講演集, Vol. 40, pp.199-200 (9 October, 2015).

### 報 告

- (D-1) **川岸卓司**, 松梨夏季, 水谷孝一, 若槻尚斗: “複数テンプレートを用いる豚のくしゃみ監視,” 2013 年度農業施設学会次大会講演要旨集, P-22, pp.165-166 (29 August, 2013).
- (D-2) **川岸卓司**, 松梨夏季, 水谷孝一, 若槻尚斗: “豚のくしゃみ音を用いた健康監視,” 大学院生学際研究フォーラム院生プレゼンバトル 2013, ポスター部門 (4 November, 2013).
- (D-3) **川岸卓司**, 松梨夏季, 水谷孝一, 若槻尚斗: “豚の健康監視のためのくし

- やみ音の自動検出技術,”SATテクノロジーショーケース 2014, No.44, pp.46, (24 January, 2014).
- (D-4) 川岸卓司, 小木曾里樹, 水谷孝一, 若槻尚斗: “少数マイクロフォンによる音源方向推定を用いる移動ロボットの自己位置推定法,” 日本機械学会 No. 14-2, ロボティクス・メカトロニクス講演会 2014 講演論文集, 2A2-C08, 4-pages, (27 May, 2014).
- (D-5) 川岸卓司, 水谷孝一, 若槻尚斗, 善甫啓一: “時間周波数特徴量を用いる豚呼吸器感染症の検出,” 2014 年度農業施設学会大会講演要旨集, P-24, pp.147-148 (28 August, 2014).
- (D-6) Takuji Kawagishi: “Detection of respiratory tract infection in swine using sound spectrogram”, The 11th IEEE Transdisciplinary-Oriented Workshop for Entrant Researchers, p. 21 (29 November, 2014).
- (D-7) 川岸卓司, 善甫啓一, 水谷孝一, 若槻尚斗, 川村洋平: “少数ジオフォンによる土中音源探査,” SAT テクノロジーショーケース 2015, No. 72, pp.75 (21 January, 2015).
- (D-8) 川岸卓司, 水谷孝一, 若槻尚斗, 善甫啓一: “豚舎内におけるパターンマッチングを用いる咳音検出,” 日本音響学会, 2015 年春季研究発表会講演論文集, CD-ROM, 3-P-1, pp.1357-1360 (5 September, 2014).
- (D-9) 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: “少素子マイクロフォンアレイを有する音源追尾ロボットの複数音源環境下における音源分離特性,” 日本機械学会 No. 15-2, ロボティクス・メカトロニクス講演会 2015 講演論文集, 2P1-V09, 4-pages, (19 May, 2015).
- (D-10) 川岸卓司, 水谷孝一, 善甫啓一, 若槻尚斗: “豚舎内における豚くしゃみ音特徴量を用いる検出,” 農業環境工学関連 5 学会 2015 年合同大会講演要旨集, P-241, pp.68 (17 September, 2015).
- (D-11) Takuji Kawagishi: “Sound localization of underground position using geophone array”, The 12th IEEE Transdisciplinary-Oriented Workshop for Entrant Researchers, p. 45 (28 October, 2015).
- (D-12) 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: “マイクロフォンアレイによるドローン飛行音の位置推定,” SAT テクノロジーショーケース 2016, P-95, pp.97 (4 January, 2016).
- (D-13) 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: “風影響下における立体構成マイクロフォンアレイでのドローン位置推定,” 日本音響学会, 2016 年春季研究発表会講演論文集, CD-ROM, 1-Q-35, pp.1285-1286 (9 March 2016).
- (D-14) 川岸卓司, 水谷孝一, 善甫啓一, 若槻尚斗, 竹前喜洋, 西藤岳彦: “豚舎内における環境騒音と活動騒音,” 2016 年度農業施設学会大会講演要旨集, P-28, pp.149-150 (29 August, 2016).
- (D-15) 川岸卓司, 小木曾里樹, 水谷孝一, 善甫啓一, 若槻尚斗: “ドローンプロペラによる下方風速と等価音源位置の関係,” 日本音響学会, 2016 年秋季研究発表会講演論文集, CD-ROM, 1-Q-4, pp. 1305-1306(14 September 2016).

---

## 受賞歴等

- (A-1) 川岸卓司: 2013 年度農業施設学会年次大会, 最優秀ポスター賞, 大会委員長 後藤 清和, (2013. 8).
- (A-2) 川岸卓司: 2013 年度土木情報学シンポジウム, 優秀講演者賞, 土木情報学委員会 委員長 矢吹 信喜, (2013. 9).
- (A-3) 川岸卓司: つくば院生ネットワーク主催大学院生学際研究フォーラム院生プレゼンバトル 2013, 奨励賞, 筑波大学 学長 永田恭介, (2013.11).
- (A-4) 川岸卓司: The 10th IEEE Tokyo Young Researchers Workshop, Best Discussion Award, 実行委員会委員長 長谷川孔明 (2013.11).
- (A-5) 川岸卓司: SAT テクノロジーショーケース 2014, ベストアイデア賞, つくばサイエンス・アカデミー会長 江崎 玲於奈, (2014.1).
- (A-6) 川岸卓司, システム情報工学研究科長表彰 (筑波大学): "少数センサを用いるパッシブ音源探査に関する研究," 研究科長 稲垣敏之 (2014.3).
- (A-7) 川岸卓司, 知能機能システム専攻優秀修士論文賞 (筑波大学): "少数センサを用いるパッシブ音源探査に関する研究," 専攻長 岩田洋夫 (2014.3).
- (A-8) 川岸卓司: 2014 年度農業施設学会年次大会, プレゼンテーション賞, 農業施設学会会長 佐竹隆顕, (2014. 8).
- (A-9) Aclab team a (小木曾里樹, 川岸卓司, 青木拓也、古川聖也), IBM Bluemix Challenge 優秀賞, "sleeff", 審査委員長 ヴィヴェック・マハジャン(専務執行役員 ソフトウェア事業本部長) (2014.9).
- (A-10) 川岸卓司: SAT テクノロジーショーケース 2015, ベストアイデア賞, つくばサイエンス・アカデミー会長 江崎 玲於奈, (2015.1).
- (A-11) Aclab team a (川岸卓司, 小木曾里樹, 青木拓也): Ceartech Japan 2015, Qi ワイヤレスパワーアイデアコンテスト, フィリップス賞, Wireless Power Consortium, (2015.10).
- (A-12) 川岸卓司: SAT テクノロジーショーケース 2016, ベストアイデア賞, つくばサイエンス・アカデミー会長 江崎 玲於奈, (2016.2).

- (A-13) 川岸卓司: 知能機能システム専攻長表彰 (筑波大学), 専攻長 葛岡英明 (2016.3).
- (A-14) 川岸卓司: 茗溪賞 (筑波大学), 理事長 江田昌佑(2016.3).
- (A-15) 川岸卓司, 国際技術交流援助 (海外渡航) ,公益財団法人 東電記念財団:"時間周波数解析を用いる豚呼気感染症発見のためのくしゃみ音監視", Acoustical Society of America and Acoustical Society of Japan Joint meeting(2016.11)
- (A-16) 川岸卓司: 2016年度農業施設学会年次大会, プレゼンテーション賞, 農業施設学会会長 池口忠夫, (2016. 8).
- (A-17) Aclab team a (青木 拓也、小木曾里樹, 松雪俊, 川岸 卓司),: GUGEN 2016, quadcept 賞(2016.12).
- (A-18) Aclab team a (青木 拓也、小木曾里樹, 松雪俊, 川岸 卓司),: 電子情報通信学会, 2016年度知的環境アプリケーションアイデアコンテスト, 最優秀賞 (2017.1).
- (A-19) Takuji KAWAGISHI, "4pABc15. Monitoring of swine sneezing using time-frequency analysis for detecting diseases", Acoustical Society of America and Acoustical Society of Japan Joint meeting, Best Student Paper Award, (2017.1).
- (A-20) 川岸卓司: SATテクノロジーショーケース 2017, 総合得点賞, つくばサイエンス・アカデミー会長 江崎 玲於奈, (2017.1).
- (A-21) 川岸卓司, システム情報工学研究科長表彰 (筑波大学) : "少数マイクロフォンを用いる環境雑音を考慮した音源の特徴抽出及び位置推定に関する研究," 研究科長 北川博之 (2017.3).