

# 屋外音環境理解における音源検出の性能評価と可視化

## Visualization of Sound Detection for Outdoor Scene Analysis

長峰 諒英<sup>†</sup>, 大畑 琢磨<sup>‡</sup>, 上村 知史<sup>‡</sup>, 小島 諒介<sup>‡</sup>, 杉山 治<sup>‡</sup>, 中村 圭佑<sup>\*</sup>, 中臺 一博<sup>‡,\*</sup>

Akihide Nagamine<sup>†</sup>, Takuma Ohata<sup>‡</sup>, Satoshi Uemura<sup>‡</sup>,

Ryosuke Kojima<sup>‡</sup>, Osamu Sugiyama<sup>‡</sup>, Keisuke Nakamura<sup>\*</sup>, Kazuhiro Nakadai<sup>‡,\*</sup>

<sup>†</sup>東京工業大学 工学部 電気電子工学科, <sup>‡</sup>東京工業大学 大学院 情報理工学研究科,

<sup>\*</sup>(株)ホンダ・リサーチ・インスティテュート・ジャパン

<sup>†</sup>Department of Electric and Electrical Engineering, Tokyo Institute of Technology,

<sup>‡</sup>Graduate School of Information Science and Engineering, Tokyo Institute of Technology,

<sup>\*</sup> Honda Research Institute Japan Co., Ltd.

### Abstract

本稿では、屋外での音環境理解を目指して、クアドロコプターに搭載したマイクロホンを用いた音源定位を扱う。これまで、プロペラ音や風切り音が存在する環境下で音源を定位する手法を開発したが、1) 方位角のみを扱っていた、2) 音源検出結果を表示するビューアがなく直感的に状況がわかりにくかったという問題があった。本稿ではこれらの問題の解決を図るため、1つ目の問題に対しては、仰角の定位を行うことができるように拡張するとともに、音源が地上付近にあることを仮定して、音源までの距離推定を行う。これによって、方位角、仰角、距離情報からなる3次元定位を可能にした。2つ目の問題については、クアドロコプターのセンサから得られる3次元位置データ、および3次元音源定位結果を用いてマイクロホンアレイが3次元的に移動する場合でも、これらを3次元マップ上に表示するツールの開発を行った。これらを実装したプロトタイプシステムを構築し、3種類の実機を用いて、実際に屋外で21種類の音源を用いた収録を行った。提案する3次元定位手法を、実機ベース、および音源ベースの指標で評価し、その有効性を示すとともに、ケーススタディベースで音源の直感的な可視化が実現できることを示した。

## 1 はじめに

屋外環境での音環境理解は、災害地での救助活動や異常音検出など様々な応用が期待できる有用な分野である。内閣府の革新的研究開発推進プログラム (ImPACT) では、極限災害環境でもタフに仕事ができる遠隔自律ロボットの実現を目指す「タフ・ロボティクス・チャレンジ」がプロジェクトとして採択され<sup>1</sup>、屋外ロボットの基盤技術への重要性が認知されてきている。屋外環境での音環境理解は、タフ・ロボティクス・チャレンジでも、極限音響という重要なテーマとして位置づけられている。

我々は、こうしたプロジェクトに先駆け、これまでに培ってきたロボット聴覚技術を用いて、屋外環境理解実現に向け

た研究を行っている。ロボット聴覚は、主に屋内のロボットを対象にして、人とのインタラクションをロボットに備えた耳を用いて実現することを目的とした日本発の研究分野である[Nakadai 00]。ロボットの耳で音を聞く場合は、スマートホンの場合とは異なり、遠隔からの発話を認識する必要があるため、様々な雑音を扱う必要がある。そこで、マイクロホンアレイ処理を導入して、音源定位・音源分離・音声認識といった機能に着目した研究を行ってきた[Nakamura 09, Nakajima 10, Yamamoto 07]。また、ロボット聴覚で培ってきた技術をロボット聴覚のオープンソースソフトウェア HARK (HRI-JP Audition for Robot with Kyoto University) として、一般公開を行っている。

### 1.1 屋内と屋外音環境理解の違い

屋内と屋外では、前述の雑音問題の性質が異なるため、同じ雑音抑圧技術でもそのフォーカスは異なる。屋内では、周囲の騒音と共に、残響が存在する(もしくは、残響を考慮する必要がある)ことが大きな特徴である。特に、音声認識が残響に対しての頑健性が低いという特徴を持っていることから、音声認識では残響が大きな問題である。一般的な屋内では壁、天井、床など音を反射するものに囲まれていることから残響を避けることは難しく、国際学会でも Reverb Challenge のような残響抑圧技術を競うコンペティションが行われている<sup>2</sup>。一方で、残響は、屋内の音響環境に関する情報が含まれている。例えば、方位角や仰角推定と比較すれば、音源距離推定の精度は低いものの、残響情報を積極的に利用することで音源距離推定が可能であることが報告されている[丹羽 14]。

屋外では、特殊な状況を除けば、一般に残響を考慮する必要はないといえる。これは、残響を扱う必要がない反面、屋外での音源距離推定が難しいことを示している。また、周囲の雑音が大きなダイナミックレンジで、動的に変化する。風、湿度、温度の変化があるため、音速自体が一様ではないばかりか、時間的にも変動する。点音源を仮定できない雑音源も多く存在し、そのモデル化も困難であるといった厄介な特徴を持っている。

<sup>1</sup><http://www.jst.go.jp/impact/program07.html>

<sup>2</sup><http://reverb2014.dereverberation.com/>

## 1.2 屋外音環境理解の関連研究

我々は、これまで、屋外音環境特有の問題を解決するため、音源定位にフォーカスして研究を行っている。例えば、奥谷らは、屋内の音源定位用に開発した一般化固有値展開に基づく GEVD-MUSIC (MULTiple SIGNAL Classification based on Generalized EigenValue Decomposition) 法 [Nakamura 09] を時間的に変動する雑音に対応するように拡張した iGEVD-MUSIC (incremental GEVD-MUSIC) 法を報告した [奥谷 13]。ベースとなった GEVD-MUSIC 法は、クアドロコプタで事前収録した音響信号を用いて、雑音に関する知識である雑音相関行列の推定を行うため、上述のようにモデル化が難しい雑音源であっても精度よく推定できたが、動的に変化する雑音に対応することは難しかった。iGEVD-MUSIC 法における雑音相関行列の推定は、短時間での雑音は定常であるという仮定の下、対象区間より、時間的に少し前の時刻の音響信号を用いて雑音相関行列の推定を行うため、雑音相関行列を動的に推定することができ、屋外での音源検出性能を著しく向上できる。古川らは、この考え方をさらに発展させて、クアドロコプタ自身が作り出す雑音の変化に対応するため、クアドロコプタのステータス情報に対してガウス過程を用いることにより、雑音相関行列を動的に推定する手法を報告している [Furukawa 13]。大畑らは、GEVD の計算コストを削減するために、一般化特異値展開 GSVD (Generalized Singular Value Decomposition, GSVD) を導入した iGSVD-MUSIC 法を提案した [Ohata 14]。また、さらに雑音相関行列の推定誤差に対応するため、雑音相関行列の大きさをスケールリングできる CMS (Correlataion Matrix Scaling) 法を併せて用いることを提案した [大畑 14]。これらの手法を用いた結果、音声では 15 m 程度、ホイッスルなど検出しやすい音源では 20 m 程度遠方の音源でも精度良く検出をできることを示した。このように、要素技術としては、屋外環境に耐えうる音源定位技術が構築されつつある。

## 1.3 課題とアプローチ

しかし、こうした技術の実用化を考えた場合、以下のような課題を解決する必要がある。

1. 屋外は三次元環境であるにもかかわらず、一次元（方位角）のみの音源定位を扱っていた。
2. 音源検出結果を表示するビューアがなく直感的に状況がわかりにくかった。

本稿では、これらの問題の解決を図るため、1 つ目の問題に対しては、仰角の定位を行うことができるように拡張する。また、方位角と仰角平面上に音源探索を頑健に行うことができる音源探索法を提案する。さらに、音源が地上付近にあることを仮定して、音源までの距離推定を行う。これによって、方位角、仰角、距離情報からなる 3 次元定位を可能にした。2 つ目の問題については、クアドロコプタのセンサから得られる 3 次元位置データ、および 3 次元音源定位結果を用いてマイクロホンアレイが 3 次元的に移動する場合でも、これらを 3 次元マップ上に表示するツールの開発を行った。

## 2 音源定位手法

本稿では、オフラインでの評価を前提としていることから、MUSIC 法の中で性能がもっともよい iGEVD-MUSIC

法をベースに定位を行う。

### 2.1 iGEVD-MUSIC 法

iGEVD-MUSIC 法は、GEVD-MUSIC 法の雑音相関行列推定を逐次的に行うことができるように改良した手法である。これによって、動的な雑音が存在する環境でも頑健に音源定位を行うことが可能となる。以下に、そのアルゴリズムを説明する。

$f$  フレーム目の入力音響信号をフーリエ変換して得られる  $X(\omega, f)$  から、以下のように相関行列  $R(\omega, f)$  を定義する。

$$R(\omega, f) = \frac{1}{T_R} \sum_{\tau=f}^{f+T_R-1} X(\omega, \tau) X^*(\omega, \tau) \quad (1)$$

ただし、 $\omega$  は周波数ビン番号、 $T_R$  は相関行列の計算に用いるフレーム数である。

次に、 $f$  番目のフレームに対して、 $f_s$  前のフレームから、 $T_N$  フレーム分の信号は雑音区間であると仮定して、雑音の相関行列  $K(\omega, f)$  を求める。

$$K(\omega, f) = \frac{1}{T_N} \sum_{\tau=f-f_s-T_N}^{f-f_s} X(\omega, \tau) X^*(\omega, \tau), \quad (2)$$

GEVD-MUSIC は、雑音の相関行列には、与えられた雑音区間から事前に計算したものを使用しており、動的な雑音の変化に対応できないという問題があった。iGEVD-MUSIC 法では、フレームごとに（逐次的に）雑音が推定できるため、動的な雑音変化に対応できることが期待できる。

$K$  の逆行列を用いて、以下のように雑音成分を白色化することが出来る。こうして得られた  $K^{-\frac{1}{2}}(\omega, f) R(\omega, f) K^{-\frac{1}{2}}(\omega, f)$  を一般化固有値展開し、固有ベクトルを計算する。

$$K^{-\frac{1}{2}}(\omega, f) R(\omega, f) K^{-\frac{1}{2}}(\omega, f) = E(\omega, f) \Lambda(\omega, f) E^*(\omega, f) \quad (3)$$

ただし、 $\Lambda(\omega, f)$  は降順に並んだ固有値を対角成分に持つ行列である。 $E(\omega, f)$  は、固有値ベクトルを並べた行列である。

これと音源方向  $\psi$  に対応した伝達関数  $G(\omega, \psi)$  を用いて MUSIC 空間スペクトル  $P(\omega, \psi, f)$  を計算する。

$$P(\omega, \psi, f) = \frac{|G^*(\omega, \psi) G(\omega, \psi)|}{\sum_{m=L+1}^M |G^*(\omega, \psi) e_m(\omega, \psi)|} \quad (4)$$

ただし、 $L$  は目的音源数である。 $e_m$  は、 $E_l$  に含まれる  $m$  番目の特異値ベクトルを表す。音源方向を推定するために  $P(\omega, \psi, f)$  を以下のように  $\omega$  方向に平均する。

$$\bar{P}(\psi, f) = \frac{1}{\omega_H - \omega_L + 1} \sum_{\omega=\omega_L}^{\omega_H} P(\omega, \psi, f) \quad (5)$$

なお  $\omega_H$ 、 $\omega_L$  は使用する周波数ビンの上限と下限に対応したインデックスである。

最後に、 $\bar{P}(\psi, f)$  に対してピーク検出と閾値処理を行い、得られたピークに対する  $\psi$  を音源方向として検出する。

## 2.2 仰角推定と2次元音源探索手法

一般に、音源方向  $\psi$  としては、方位角  $\theta$  のみを扱うことが多く、屋内では、このような1次元定位でも比較的問題になることが少ない。しかし、クアドロコプタなど屋外での音源定位を前提にする場合には、仰角に対する定位能力が求められる。そこで、本稿では、以下のように2次元に拡張して、定位を行う。

$$\psi = (\theta, \phi) \quad (6)$$

このような定義を行っても、上述の MUSIC アルゴリズム自体は基本的に一般性を失わない。ピーク検出についても、 $\theta$  直線上ではなく、 $\theta-\phi$  平面上で行う必要がある。実際には、ピーク検出の問題はそれほど簡単な問題ではないが、本稿では、以下のように、音源数が、高々1であると仮定し、単純な最大値検出によって、ピークを検出し、定位を行った。

$$\Psi(f) = \begin{cases} \operatorname{argmax}_{\psi} \bar{P}(\psi, f) & (\bar{P}(\Psi, f) \geq P_{th}) \\ \emptyset & (\text{otherwise}) \end{cases} \quad (7)$$

$P_{th}$  は音源かどうかを判断するための閾値であり、実験的に求めた。

## 2.3 音源距離推定

上述のように、音源方向は極座標表現  $\Psi(f) = [\Theta(f), \Phi(f)]$  として得られる。これは、xyz 軸からなる直交座標系では3次元表現になるため、方位角と仰角からなる音源方向推定はしばしば3次元音源定位と呼ばれることがある。しかし、方位角と仰角の2次元の情報しか含まれていないため、実際には、3次元音源定位とは言えない。真に3次元音源定位を実現するためには、音源までの距離情報を推定する必要があり、屋外音環境理解では、マップ上に音源表示を行うためにも距離情報まで推定できることが望ましい。しかし、本稿の冒頭に述べたように一般に音源距離推定問題は難しい。さらに、屋外では距離推定の重要なキューとなる残響情報の利用が困難であるため、音源距離推定問題は一層難しい問題となっている。

そこで、本稿では、音源は地上付近（主に人間の口元の高さ）にあるという仮定を置くことによって、この問題の解決を試みる。

まず、得られる音源方向は、クアドロコプター座標系での値になっているため、航法データを用いて、絶対座標への変換を行い、絶対座標系での方位角と仰角のペア  $[A, E]$  を得る。

クアドロコプタの地表からの高度を  $h$ 、音源の高度を  $h_{src}$  とすれば、音源距離は、以下のようにあらわすことができる。

$$D = \left| \frac{h - h_{src}}{\sin(E)} \right| \quad (8)$$

従って、クアドロコプターの中心を原点にとれば、3次元音源位置は、以下のように表すことができる。

$$P_s = [A, E, D] \quad (\text{極座標系}) \quad (9)$$

$$= [D \cos(E) \cos(A), -D \cos(E) \sin(A), D \sin(E)] \quad (\text{直交座標系}) \quad (10)$$

## 3 音源可視化システム

得られた音源定位結果をクアドロコプタの航法データや地図データとともに可視化を行うシステムを構築した。Fig. 1 に構築した音源可視化システムの構成図を示す。我々が利用しているクアドロコプタである Asctec 社の Pelican は、ジャイロ、高度センサ、GPS、加速度センサ、磁気センサを搭載しており、位置、姿勢、速度、加速度が取得できる。これらに加えて、システムインフロンティア社の多チャンネル収録装置 RASP-24 と MEMS マイクロホンで構成される小型軽量の 16ch マイクロホンアレイを設置した (Fig. 2a 参照)。クアドロコプタ搭載センサからの情報、およびマイクロホンアレイからの音響信号は WiFi (IEEE 802.11ac) 経由でデータ収録用の端末に送信される。この際、センサデータを同期収録する必要があるため、ROS<sup>3</sup> を用いて実現した。端末側では、受信した信号のうち、音響信号は、2節で説明した音源定位手法を用いて、定位を行う。実装は HARK<sup>4</sup> を用いた。得られたクアドロコプタ極座標系での2次元の音源定位情報とクアドロコプタの情報を用いて絶対座標系での音源位置を算出し、KML (Google Earth(Keyhole) Markup Language) 形式に変換後、Google Map 上にこれらのデータを表示する。また、予め人の位置がわかっている場合には、その音源位置を登録し、その位置に人オブジェクトを表示しておくことができる。実際に、登録した人位置に音源があるとシステムが判断した場合には、これを人の発話と見なし、表示した人オブジェクトの色の変更を行う。

## 4 評価実験

構築したシステムの評価実験を行うため、実際に屋外で21種類の音源をスピーカから出力し、音源定位実験を行った。クアドロコプタには、Asctec 社の Pelican (Fig. 2a 参照)、enRoute 社の Zion (Fig. 2b 参照) を用いた。また、ヘリウムガスを入れたバルーンの周囲に 16ch マイクロホンアレイを設置して、これを浮遊させ、クアドロコプタと同様の実験を行った (Fig. 2c 参照)。

### 4.1 実験条件

実験の測定条件について、Tab. 1 にまとめる。「固定」は、屋外測定ではあるが、筐体をしっかり固定し、プロペラが回転しても動かない状態で収録を行った。ただし、バルーンは、固定しても風で流されてしまうため、完全な固定はできなかった。「移動」は、実際にクアドロコプタを浮遊させホバリングに近い動作を行った状態で収録を行った。固定条件と比べれば、風の影響が大きくなり、また、プロペラ音の動的な変化への対応が必要となる。音源の位置に関しては、大まかな方向は得られるものの正確なリファレンスを得ることは困難であった。

使用した21種類の音源、およびその音量を Fig. 3 にまとめた。音量は、wav ファイルの最大値を 0dB として算出している。音量は一つの目安ではあるが、音源毎に周波数特性が異なるため、音源定位のしやすさと完全な相関はない。MUSIC に用いる伝達関数については、実測ではなく、幾何計算で算出した。MUSIC の処理で用いる音源数  $L$  は 1 とした。

<sup>3</sup><http://www.ros.org/>

<sup>4</sup><http://www.hark.jp/>

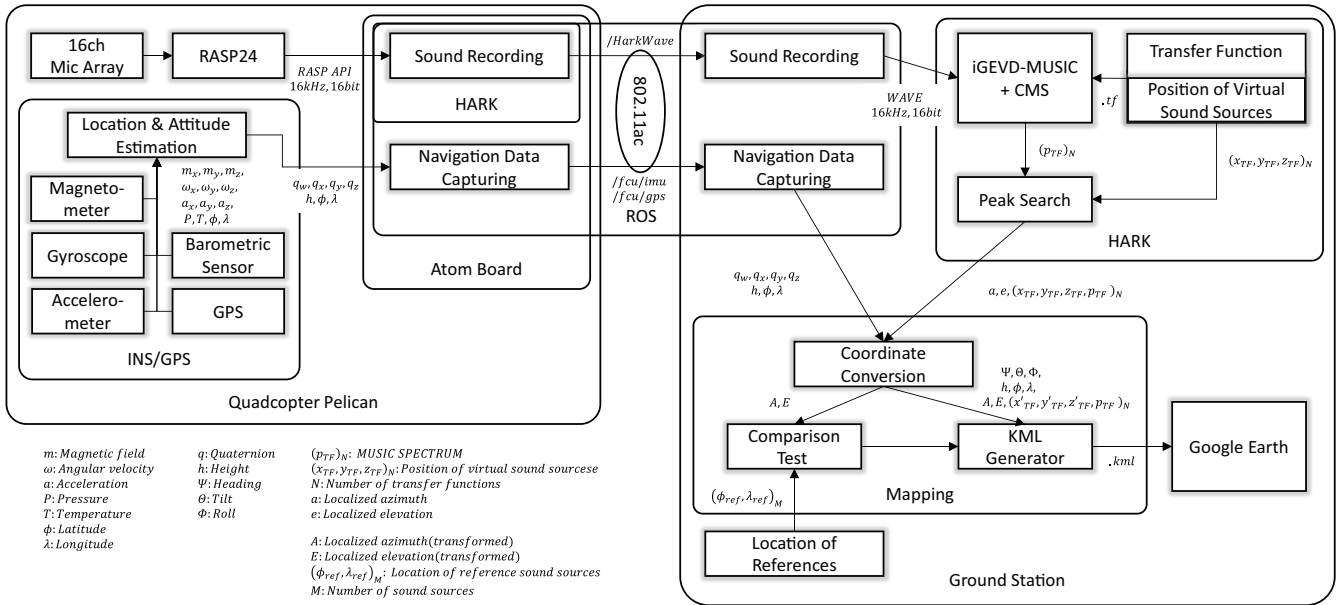


Figure 1: 可視化システム構成図

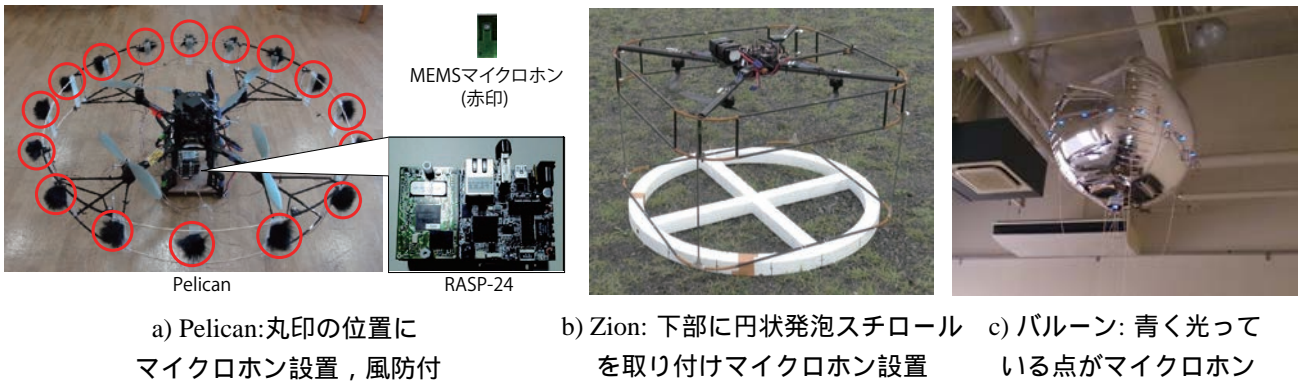


Figure 2: マイクロホンアレイ搭載クアドロコプタ

Table 1: 実験条件 (移動条件の高度，距離，角度は目安)

ラベル	クアドロコプタ		音源方向		使用音源	
	高度 [m]	水平距離 [m]	仰角 [deg]	方位角 [deg]	音源種類	測定回数 (音源毎)
バルーン固定	0	3	0	65	20	10
Pelican 固定	0	3	0	0	21	10
Zion 固定	0	3	0	0-360 (45度毎)	1	10
Pelican 移動 A	5	3	60	0	7	3-10
Pelican 移動 B	5	5	45	0	7	3-10
Pelican 移動 C	5	10	27	0	7	3-10

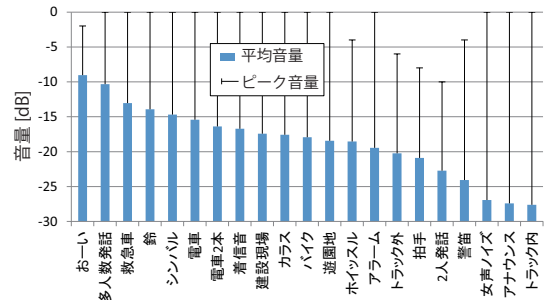


Figure 3: 使用音源の種類と音量

#### 4.2 評価指標

音源定位の評価には，以下の3つの指標を用いた．

指標 1: 軸別定位正解精度

指標 2: クアドロコプタベース定位正解率

指標 3: 音源ベース定位正解率

指標 1 は，全音源数を  $N$  とした場合に，方位角，仰角別にクアドロコプタから見て，一定の角度 ( $a_{th}$ ) 以内に定位した正解数 ( $C$ )，定位はしているものの角度が  $a_{th}$  以内でない定位誤り数 ( $S$ )，音源の定位そのものがされなかった

削除誤り数 ( $D$ )，よけいに定位をしてしまう挿入誤り数 ( $I$ ) をカウントし， $(N - S - D - I)/N$  を計算した値となる． $C = N - S - D$  であるので， $I$  が多ければこの指標は負の値を持つ．奥谷らが用いた LAR (Localization Accuracy Rate) [奥谷 13] と同様の指標である．本稿では， $a_{th}$  として，方位角に対しては， $5^\circ$ ，仰角に対しては， $10^\circ$  を用いた．

指標 2 は，クアドロコプタから見て，実際の音源位置の一定角度 ( $b_{th}$ ) 以内に定位しているかどうかを示す指標であり，指標 1 と同様，クアドロコプタと音源の距離によ

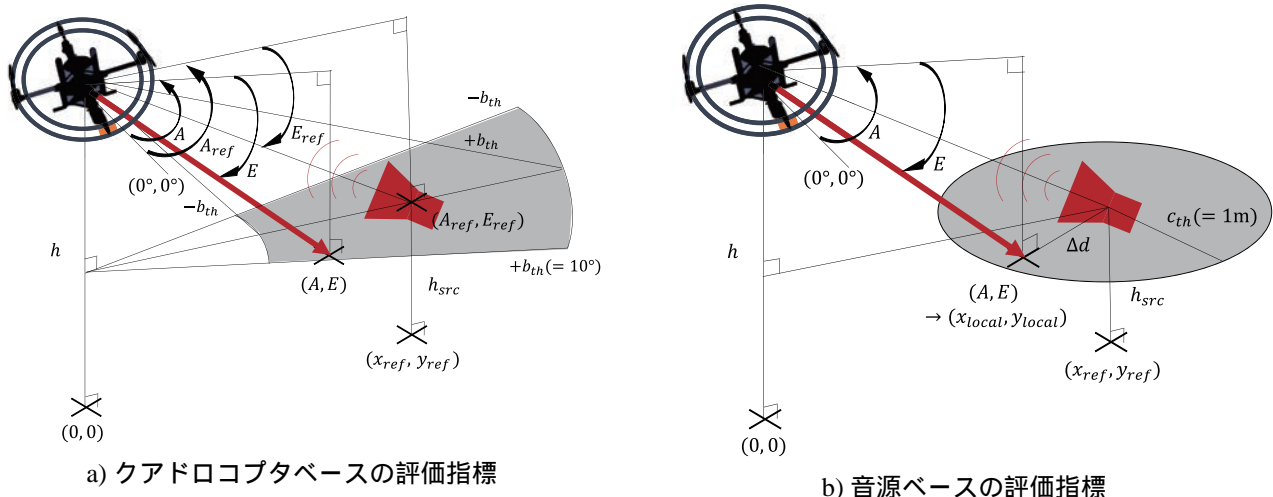


Figure 4: 実験で用いた評価指標

らず精度が変わらない指標である．奥谷らが用いた LCR (Localization Correct Rate)[奥谷 13] に倣い、指標 1 の  $C/N$  に相当する値とした．具体的な正解の判定条件は、Fig. 4a) における網掛け部分、つまり、以下の 2 つの条件を同時に満たす場合とした．

$$|A - A_{ref}| \leq b_{th} \quad (11)$$

$$|E - E_{ref}| \leq b_{th} \quad (12)$$

本稿では  $b_{th}$  として、 $10^\circ$  を用いた．

指標 3 は、音源位置から一定範囲 ( $c_{th}$ ) に定位したかどうかを示す指標である．指標 1 と同様、 $C/N$  に相当する値とした．具体的な正解の判定条件は、Fig. 4b) における網掛け部分、つまり、以下の条件を満たす場合とした．

$$\Delta d = \sqrt{(x_{ref} - x_{local})^2 + (y_{ref} - y_{local})^2} \leq c_{th} \quad (13)$$

本稿では  $c_{th}$  として、 $1\text{m}$  を用いた．この指標は、実際には、音源定位自体は極座標系で行われるので、たとえ音源定位結果が同じであっても、クワドコプタからの距離が遠い音源ほど、精度が劣化する．

また、Tab. 1 記載のリファレンスデータは、正確性に欠けるため、この値をそのまま用いず、定位結果のヒストグラムを作り、その中央値が Tab. 1 記載値から、 $\pm 20^\circ$  であれば、中央値をリファレンスの値とするようなキャリブレーションを行った．

### 4.3 実験結果

Tab. 2 に、Pelican 固定、バルーン固定条件に対する指標 1 の結果を示す．また Tab. 3 に、Zion 固定条件に対する指標 1 の結果を示す．Pelican では、プロペラ音が存在するにもかかわらず、ほぼ問題なく定位が実現できていることがわかる．着信音については、他の音源よりも性能の劣化が見られる．これは、着信音は特定の周波数のみにパワーが分布していることから、プロペラ音の周波数に埋もれやすいためではないかと考えられるが、より詳細な分析は今後の課題である．一方、バルーンはプロペラ音がないにもかかわらず、Pelican よりも定位性能が悪いという結果に終わった．特に、仰角については、音源定位結果は得られるものの、実際の方向とは大きく乖離した方向に定

位が得られた．性性能が悪い原因としては、前述のようにバルーンが風で揺れてしまいしっかり固定できなかったこと、マイクロホンバルーンの表面に貼りつけたため、風によるバルーンの変形に応じてマイクロホンアレイのレイアウトが変形してしまったこと、そもそもマイクロホンの位置を正確に計測することが難しかったことが挙げられる．また、仰角方向の定位に関しては、そもそもマイクロホンレイアウトが円状であったことから、上述の理由と合わさって性能が劣化したと考えられる．バルーンに対しては、プロペラ音が存在しないため、論理的には Pelican よりも良好な音源定位性能が得られるはずであるので、マイクロホン設置方法を工夫するなどして改良を行いたい．Zion については、音源の方向を変えながら定位性能の変化を調べた．概ね、Zion のプロペラ騒音下でも問題なく定位が可能であるといえる．方向毎の性能の変化が若干見られるが、少なくとも方位角に関しては、データ解析により、風の影響であることがわかった．

Tab. 4 に Pelican 移動条件に対する指標 2 の結果、Tab. 5 に Pelican 移動条件に対する指標 3 の結果を示す．移動条件は、固定条件と比較して、定位が難しいことがわかる．ホイッスルや警笛を見る限り、指標 2 では距離とは無関係に性能が出ていること、指標 3 では距離が大きくなる ( $A \sim C$ ) につれて、性能が劣化していることがわかる．指標 2 では、音源距離が  $10\text{m}$  離れていても検出できている．これは、音源距離が  $20\text{m}$  程度離れていてもホイッスルの定位が可能である (音声は  $12\text{m}$  程度まで) とした報告[大畑 14]を裏付けるものである．しかし、指標 3 では、ホイッスルでも定位性能が落ちており、距離まで含めた 3 次元的な音源定位性能で見るとこの距離では難しいことがわかる．これは、最初に 2 次元の音源方向だけ定位を行い、音源に近づくことで 3 次元音源定位を行う段階的なアクティブな音源定位が必要であることを示唆しているといえる．移動 A 条件では、音源までの距離が近いため、指標 2 と指標 3 で大きな差が見られない．つまり、指標によらず、定位しにくい音源は定位しにくいことがわかる．例えば、アナウンスは全く定位ができないという結果となった．これは、Fig. 3 にも示したように、アナウンスは音量が小さいために信号対雑音比が小さくなってしまっているためである．

Table 2: バルーン固定, Pelican 固定に対する軸別定位正解精度 [%] (すべて母数は 10)  
無印は 500 Hz - 2800 Hz, \* 付は 2800 Hz - 6000 Hz で評価を行った。「-」はデータ収録の失敗により算出できず。

		おーい	多人数発話	救急車	鈴	シンバル	電車	電車 2 本
バルーン	仰角	0	0	0	0	0	0	0
	方位角	100	80	100	90	60	100	80
Pelican	仰角	100	100*	100	100*	100	100*	100
	方位角	100	100*	100	100*	100	100*	100
		着信音	建設現場	カラス	バイク	遊園地	笛	アラーム
バルーン	仰角	0*	0	0	-	0	0	0
	方位角	60*	100	70	-	100	100	100
Pelican	仰角	60*	100*	100	100*	100	100*	100
	方位角	60*	100*	100	100*	100	100*	100
		トラック外	拍手	2 人発話	警笛	女声雑音	アナウンス	トラック内
バルーン	仰角	0	0	0	0	0	0	0
	方位角	70	60	80	50	70	80	70
Pelican	仰角	100	90*	100	100	100*	100	100*
	方位角	100	90*	100	100	100*	100	100*

Table 3: Zion 固定に対する軸別定位正解精度 [%] (母数は 10): 仰角範囲を 0-45° に制限

	Elevation	Azimuth
-90°	90	90
-45°	100	100
0°	100	100
45°	70	100
90°	90	100
180°	80	90

Table 4: Pelican 移動条件に対する指標 2 の結果 [%]

	おーい	救急車	鈴	カラス	ホイッスル	警笛	アナウンス
Pelican 移動 A	-	86	100	100	90	60	0
Pelican 移動 B	70	-	-	-	100	100	-
Pelican 移動 C	80	-	-	-	90	60	-

Table 5: Pelican 移動条件に対する指標 3 の結果 [%]

	おーい	救急車	鈴	カラス	ホイッスル	警笛	アナウンス
Pelican 移動 A	-	86	100	100	70	80	0
Pelican 移動 B	40	-	-	-	90	50	-
Pelican 移動 C	20	-	-	-	0	20	-

#### 4.4 可視化の例

可視化については, ケーススタディとして, 飛行実験の一例に処理を行ったケースを示す. Fig. 5 に実験の際の飛行データ, また Fig. 6 に収録した音響信号に対する MUSIC スペクトログラムを示す. Fig. 5 からは, クワドロコプタの 3 次元的な動きを把握することは難しい. また, クワドロコプタの向きが変化するため, Fig. 6 からだけでは, 音源が検出できていそうであることは見て取ることができるものの, いつどこに音源があったのかまでを把握することは難しい.

Fig. 7 は, これらのデータをすべて統合し, Google マップ上に表示した結果のスナップショットとなっている. 図は, 上下 2 枚の画像のペアが時系列で並んでおり, 各ペアの上の写真は, 実際にカメラで現場を収録したデータからキャプチャしたものであり, 下の画像は, カメラの画像と同じ視点に合わせて再合成した (Google マップ上に表示した) 結果である. 再合成画像は, 実際にカメラで収録したデータに近い結果が得られており, 直感的に状況の把握が可能な可視化が実現できたといえる. また, さらにカメラ画像からでは音源がどこにあるかまではわからないが, 再合成画像では音源の位置や発話時刻まで可視化することができ, より多くの情報をわかりやすく表示できていることがわかる.

#### 5 おわりに

本稿では, 2 種類のクワドロコプタ, およびバルーンに搭載したマイクロホンアレイを用いて, 様々な音源を実際に屋外環境で収録し, 2 次元, および 3 次元音源定位の性

能評価を行った. また, 定位結果の可視化について報告した. 結果として, 方位角, 仰角からなる 2 次元の音源定位はプロペラ音や風が存在する環境下でもロバストに動作することが示された. また, 距離の推定は音源までの距離が近ければ有効であることを示唆する結果を得た. 一方で, 実環境では正確なリファレンスデータの取得が難しく, 評価を行う際には, リファレンスデータの誤差も考慮に入れる必要があることが分かった. 今後は, システムのオンライン化, 音源同定の導入を行う予定である.

#### 謝辞

本研究は科研費基盤 (S) No.24220006 の支援を受けた.

#### 参考文献

- [Furukawa 13] Furukawa, K., Okutani, K., Nagira, K., Otsuka, T., itoyama, K., Nakadai, K., and Okuno, H. G.: Noise Correlation Matrix Estimation for Improving Sound Source Localization by Multirotor UAV, in *Proc. of the IEEE/RSJ International Conference on Robots and Intelligent Systems (IROS)*, pp. 3943-3948, IEEE (2013)
- [Nakadai 00] Nakadai, K., Lourens, T., Okuno, H. G., and Kitano, H.: Active Audition for Humanoid, in *Proc. of 17th National Conference on Artificial Intelligence (AAAI-2000)*, pp. 832-839, AAAI (2000)
- [Nakajima 10] Nakajima, H., Nakadai, K., Hasegawa, Y., and Tsujino, H.: Blind Source Separation with parameter-free adaptive step-size method for Robot Audition, *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 18, No. 6, pp. 1476-1485 (2010)
- [Nakamura 09] Nakamura, K., Nakadai, K., Asano, F., Hasegawa, Y., and Tsujino, H.: Intelligent Sound Source Localization for Dynamic Environments, in *Proc. of the IEEE/RSJ International Conference on Robots and Intelligent Systems (IROS)*, pp. 664-669, IEEE/RSJ (2009)

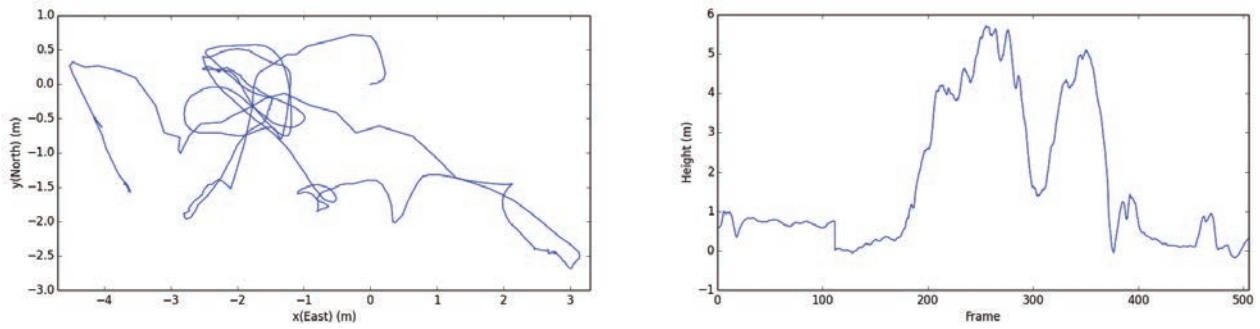


Figure 5: 航法データ (左: x-y 平面, 右: 高度)

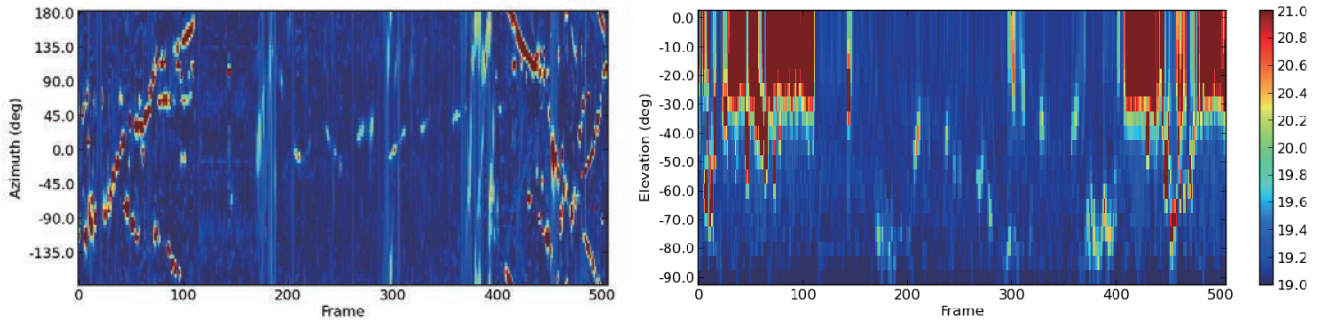


Figure 6: MUSIC スペクトル (左: 方位角, 右: 仰角)

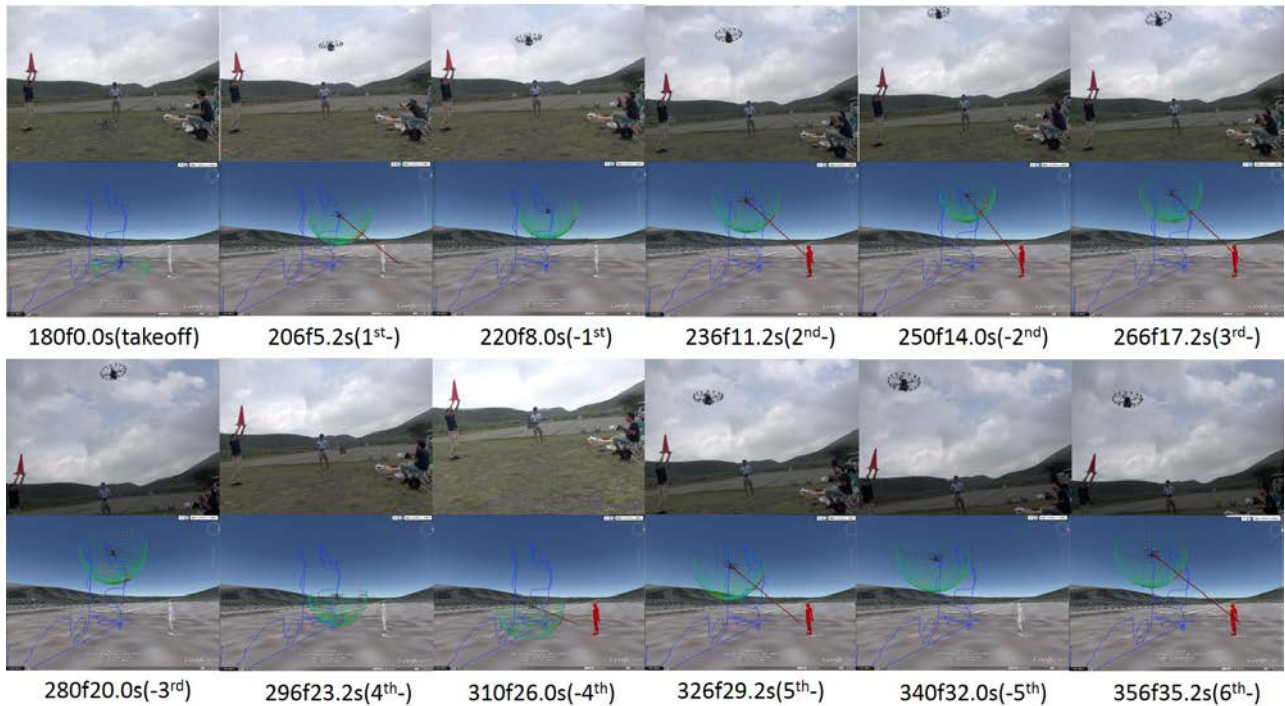


Figure 7: スナップショット (上: 実際の画像, 下: 可視化システムでの表示, 数字はフレーム数 f 時刻 s)

[Ohata 14] Ohata, T., Nakamura, K., Mizumoto, T., Tezuka, T., and Nakadai, K.: Improvement in Outdoor Sound Source Detection Using a Quadrotor-Embedded Microphone Array, in *Proc. of the IEEE/RSJ International Conference on Robots and Intelligent Systems (IROS)*, IEEE Press (2014)

[Yamamoto 07] Yamamoto, S., Nakadai, K., Nakano, M., Tsujino, H., Valin, J.-M., Komatani, K., Ogata, T., and Okuno, H. G.: Design And Implementation Of A Robot Audition System For Automatic Speech Recognition Of Simultaneous Speech, in *Proc. of the 2007 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU-2007)*, pp.

111-116, IEEE (2007)

[奥谷 13] 奥谷 啓太, 吉田 尚水, 中村 圭佑, 中臺 一博: クワドロボ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, *ロボット学会誌*, Vol. 31, No. 7, pp. 38-45 (2013)

[大畑 14] 大畑 琢磨, 長峰 諒英, 中村 圭佑, 水本 武志, 中臺 一博: 相関行列スケーリングを用いた iGSVD-MUSIC 法による屋外環境音源探索の向上, *日本ロボット学会第 32 回学術講演会*, pp. 111-03 (2014)

[丹羽 14] 丹羽 健太, 江崎 知, 日岡 裕輔, 西野 隆典, 武田 一哉: 空間相関行列の固有値分布に着目した音源別距離推定, *電子情報通信学会論文誌 A*, Vol. J97-A, No. 2, pp. 68-76 (2014)