

## クアドロコプターを用いた飛行雑音に頑健な音源定位

古川 孝太郎<sup>1</sup>, 奥谷 啓太<sup>2</sup>, 柳楽 浩平<sup>3</sup>, 大塚 琢馬<sup>3</sup>, 中臺 一博<sup>2,4</sup>, 奥乃 博<sup>3</sup>

<sup>1</sup> 京都大学 工学部情報学科 <sup>2</sup> 東京工業大学大学院 情報理工学研究科

<sup>3</sup> 京都大学大学院 情報学研究科 <sup>4</sup> (株) ホンダ・リサーチ・インスティテュート・ジャパン

### 1. 序論

多数の回転翼を持つ無人航空機 (Multirotor UAV) はホバリングや垂直離着陸といった高い機動性を持ち、リモートセンシングにおけるプラットフォームとして近年特に注目されている。実際、画像取得による応用は多数報告されており、音情報の利用への期待も高い。本研究では、四翼の Multirotor UAV であるクアドロコプターから周囲の音源を発見し定位する手法を開発する。

Multirotor UAV 上で収録した音響信号を扱う上での大きな課題は、回転翼の駆動や風などに起因する非定常な飛行雑音の存在であり、これらは定位性能の劣化要因となる。一般に雑音環境下で音源定位を行う場合、雑音の空間相関行列 (以下、雑音相関行列) を推定し GEVD-MUSIC 法 [1] に与えることでその影響を緩和できる。例えば iGEVD-MUSIC 法 [2] では、注目音源が雑音に対して比較的短時間だけ継続するという性質を利用し、注目する時間フレームの直前を雑音とみなす。この手法は非定常な雑音にも頑健な定位を期待できるが、音源が長く継続する場合は音源を含めた観測すべてを雑音とするため、その先頭部分しか検出されないなどの欠点がある。

本稿では、機体センサのモニタ情報を用いた飛行雑音に頑健な音源定位手法を提案する。本手法はまずモニタ情報を利用して雑音相関行列を Gauss 過程 [3] により回帰し、得られた雑音相関行列と観測信号を GEVD-MUSIC 法に与えて音源を定位する。入力が多チャンネル音響信号、モニタ情報、出力は各時刻での音源方向であり、ステアリングベクトルをマイクアレイについての事前知識として利用する。また問題の簡単化のため音源数は既知であると、各音源位置は移動しないと仮定する。

### 2. 雑音相関行列の回帰による音源定位

本手法は Gauss 過程による雑音相関行列の回帰と、GEVD-MUSIC 法による音源定位という 2 つの段階からなる。Gauss 過程は連続空間上の確率過程であり、回帰問題などに応用される。また GEVD-MUSIC 法は、観測中の雑音を白色化した上で、静音や無指向性雑音環境を仮定した音源定位手法である MUSIC (Multiple Signal Classification) 法を適用する手法である。

#### 2.1 Gauss 過程による雑音相関行列回帰

具体的な手順は次の通りである。まず音源の無い状態でクアドロコプターを飛行させ、特徴量  $\mathbf{u}$  と目的変数  $\mathbf{v}$  の対応が付けられたデータ  $\mathcal{D} = \{(\mathbf{u}_i, \mathbf{v}_i)\}_{i=1, \dots, d}$  を収録する。 $\mathbf{u}$  は xyz の各方向についての傾きと速度、及び 4 個の各モーターへの供給電力を表す PWM (Pulse Width Modulation) 信号からなる 10 次元ベクトルである。また  $\mathbf{v}$  は空間相関行列  $Q$  を、その Hermite 性と正定値性を考慮し次の手順でベクトル化したもので、マイク数を  $M$  とすると  $M(M+1)/2$  次元ベクトルとなる。

$$\mathbf{v} = \text{vec}(\text{diag}(L)^{\frac{1}{2}} + L - \text{diag}(L))$$

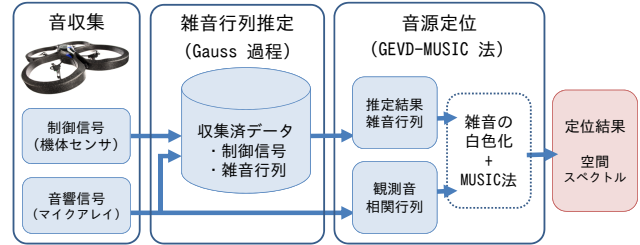


図 1: 本手法の概要

ここで  $\text{vec}$  は行列の上三角部分の列ベクトルを連結する操作、 $\text{diag}$  は対角成分からなる対角行列を得る操作とし、 $L$  は  $Q$  の Cholesky 分解によって得た上三角行列とする。これらの学習データを用意した上で、音源がある状態で新たに観測した特徴量  $\mathbf{u}_{d+1}$  から条件付分布  $p(\mathbf{v}|\mathcal{D}, \mathbf{u}_{d+1})$  を求め、その平均  $m(\mathbf{u}_{d+1})$  を  $\mathbf{v}_{d+1}$  の予測値とする。

Gauss 過程は平均関数  $m$  と共分散関数 (カーネル)  $k$  によって特徴付けられる。カーネルは、広く利用される RBF カーネルをもとにしており、入力についてスケール不変な Mahalanobis カーネル [4] を次のように変形して用いるものとする。

$$k(\mathbf{u}_i, \mathbf{u}_j) = \exp\left(-\frac{\gamma(\mathbf{u}_i - \mathbf{u}_j)^T \Sigma^{-1}(\mathbf{u}_i - \mathbf{u}_j)}{\text{dim}(\mathbf{u}_i)}\right) \quad (1)$$

$\Sigma$  は共分散行列の対角成分とする。 $m(\mathbf{u}_{d+1})$  は条件付き Gauss 分布の平均として次のように求められる。

$$m(\mathbf{u}_{d+1}) = \mathbf{k} (K + \rho^{-2}I)^{-1} V^T \quad (2)$$

$$\mathbf{k} = (k(\mathbf{u}_1, \mathbf{u}_{d+1}), \dots, k(\mathbf{u}_d, \mathbf{u}_{d+1})) \quad (3)$$

$$V = (\mathbf{v}_1, \dots, \mathbf{v}_d) \quad (4)$$

ここで  $\rho^2$  は加法雑音の分散、 $I$  は  $d$  次元単位行列、 $K$  は  $(i, j)$  要素に  $k(\mathbf{u}_i, \mathbf{u}_j)$   $i, j = 1, \dots, d$  を持つ Gram 行列である。求めた  $\mathbf{v}_{d+1}$  の予測値を上記のベクトル化の逆の手順で行列に戻し、雑音相関行列とする。

#### 2.2 GEVD-MUSIC 法による音源定位

GEVD-MUSIC 法の基礎となる MUSIC 法は、雑音部分空間と信号部分空間の直交性を利用して音源の到来方向にピークを持つ空間スペクトルを得る手法である。以下では行列  $A$  の共役転置を  $A^H$ 、その逆行列を  $A^{-H}$  と表す。また空間相関行列は信号のスペクトルについてマイク間の相関を取ることで求める。

まずマイク数  $M$ 、音源数  $N$  とするとき、原信号の空間相関行列  $S$  の固有値分解により  $N$  個の正固有値と  $M - N$  個の零固有値を得られる。したがって分散  $\sigma^2$  の白色雑音を仮定する場合、観測信号の空間相関行列  $R$  は  $S$  の固有値を  $\lambda_{S,i}$  として次式で表せる。

$$R = S + \sigma^2 I \quad (5)$$

$$= \sum_{i=1}^N (\lambda_{S,i} + \sigma^2) \mathbf{e}_i \mathbf{e}_i^H + \sum_{i=N+1}^M \sigma^2 \mathbf{e}_i \mathbf{e}_i^H \quad (6)$$

ただし固有ベクトルは固有値を降順とするよう並べる。

謝辞: 本研究の一部は科研費 (基盤 S) 24220006 の補助を受けた。

このとき  $N$  個の主固有値が原信号に対応するため、その固有ベクトル  $\mathbf{e}_i$  ( $i = 1, \dots, N$ ) の張る空間を信号部分空間、残りの固有ベクトルの張る空間を雑音部分空間とする。ここで空間相関行列の性質より两部分空間は直交する。故にステアリングベクトルを  $\mathbf{a}(\theta)$  として次式のように空間スペクトル  $P$  を定義すると、 $\theta$  が音源方向となるときのピーク値を取る。

$$P(\theta) = \frac{\|\mathbf{a}(\theta)\|^2}{\sum_{i=N+1}^M |\mathbf{a}^H(\theta)\mathbf{e}_i|^2} \quad (7)$$

これは単一の周波数ビンについての値であるから、更に各周波数成分の和を取り所望の結果を得る。

GEVD-MUSIC 法では上記の標準固有値分解に代え、雑音相関行列  $Q$  による一般化固有値分解 (GEVD, Generalized Eigenvalue Decomposition) を用いる。

$$R\mathbf{e}_i = \lambda_{R,i}Q\mathbf{e}_i \quad (i = 1, \dots, M) \quad (8)$$

これは  $Q = \Phi^H\Phi$  とする  $\Phi$  によって観測信号中の雑音を白色化した上で (標準) 固有値分解を行うことに等しい。

$$R' = \Phi^{-H}R\Phi^{-1} = \sum_{i=1}^M \lambda_{R',i}\mathbf{f}_i\mathbf{f}_i^H \quad (9)$$

したがって同様の議論により空間スペクトルを得る。

### 3. 実験と評価

本手法を用いた音源定位性能を評価するため、実験を行った。収録したデータについて、標準固有値分解を用いた MUSIC 法、観測の一部から得た一定の雑音相関行列を用いる GEVD-MUSIC 法、iGEVD-MUSIC 法 [2] 及び本手法を適用し、その結果を比較する。

使用したデータは、クアドロコプター AR.Drone<sup>†</sup>に図2のように搭載したマイクアレイにより無響室で収録した、飛行中の雑音の8チャンネル音響信号及び機体センサのモニタ情報である。音響信号は標準化周波数 16 [kHz] として 400 [s] 程度、モニタ情報は約 50 [ms] 毎に取得し、これをおよそ 7:1 の割合で学習データと評価データに分けた。また周囲の 5 度間隔 72 方向について TSP 応答を収録して得たステアリングベクトルを用いた。音源として 500 [Hz] の純音と白色雑音を用い、様々な SN 比で飛行雑音をシミュレーション混合したデータによる定位実験を行った。

本手法より図3のような空間スペクトルを得た。各空間スペクトルに対して閾値を設定し、これを超える値を持つ方向を音源方向とみなす場合の適合率と再現率から F 値を計算した。F 値が最大となるよう閾値を調整した場合の各手法の性能を表1にまとめる。ここで SEVD は標準の MUSIC 法、const. は一定の雑音相関行列による GEVD-MUSIC 法、iGEVD は iGEVD-MUSIC 法である。const. は音源の無い先頭フレームを雑音とみなし、また iGEVD では評価データを除くデータ上で最大の F 値を最大とするパラメータの値を用いた。

表中の下線を引いた値が各条件における最良値である。これより本手法は純音を音源に用いた場合や SN 比の高い場合に、既存手法よりも音源定位性能が高いとわかる。一方で白色雑音を音源に用い、かつ SN 比が低い場合には iGEVD-MUSIC 法を超える結果は得られなかった。これはおそらく本手法が、音源の細かな変化を補足できる程度まで正確に雑音相関行列を推定できていないためであると考えられる。

<sup>†</sup>Parrot SA: <http://ardrone.parrot.com/parrot-ar-drone/>

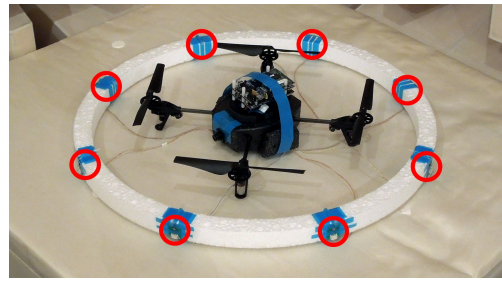


図2: クアドロコプター (丸印の位置にマイク)

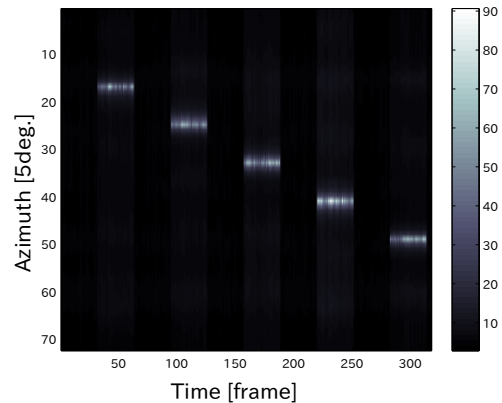


図3: 空間スペクトル (白色雑音, SN 比:10 [dB])

表1: 最大 F 値の評価結果 (下線が最良値)

固有値分解	SEVD	GEVD			
		雑音相関行列	一定	動的	
手法	-	const.	iGEVD	本手法	
純音	-2.5 [dB]	0.702	0.768	0.593	<u>0.799</u>
	0 [dB]	0.817	0.860	0.651	<u>0.872</u>
	2.5 [dB]	0.901	0.923	0.631	<u>0.935</u>
白色雑音	4 [dB]	0.456	0.454	<u>0.695</u>	0.457
	10 [dB]	0.624	0.651	<u>0.886</u>	0.690
	16 [dB]	0.928	0.886	0.928	<u>0.933</u>

### 4. 結論

本研究ではクアドロコプターに搭載したマイクアレイにより周囲の音源を定位する上で、飛行中に発生する飛行雑音による性能劣化を抑圧するため、雑音相関行列の Gauss 過程による回帰と GEVD-MUSIC 法への適用を行った。この結果、本手法は音源の種類によらず安定的な定位が可能であった。今後の展開として、特に SN 比が低い場合の性能向上のために、特徴量の選択とパラメータの最適化、また本手法と既存手法との組合せなどが考えられる。

### 参考文献

- [1] K. Nakamura *et al.*, Intelligent sound source localization for dynamic environments, in *Proc. of IEEE/RSJ IROS*, IEEE Press, 2009, pp. 664-669.
- [2] K. Okutani *et al.*, Outdoor auditory scene analysis using a moving microphone array embedded in a quadcopter, in *Proc. of IEEE/RSJ IROS*, IEEE Press, 2012, pp. 3288-3293.
- [3] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning, The MIT Press, 2006.
- [4] S. Abe, Training of support vector machines with mahalanobis kernels, in *Artificial Neural Networks: Formal Models and Their Applications*, Vol. 3697 of Lecture Notes in Computer Science, 2005, pp. 571-576.