

マイクロホンアレイ搭載 UAV を用いた屋外実環境実時間音源探査

干場功太郎¹, 若林瑞保², 鷲崎海², 石木隆洋², 公文誠², Daniel Gabriel¹, 中臺一博^{1,3}, 奥乃博⁴
 1 東京工業大学 2 熊本大学 3 (株) ホンダ・リサーチ・インスティテュート・ジャパン 4 早稲田大学

1 はじめに

われわれは、被災地等での要救助者の探索を目的に、マイクロホンアレイ搭載 UAV (Unmanned Aerial Vehicle) を用いた音源探査の研究を行っている。こうした手法を確立できれば、上空から広範囲かつ迅速に被災者の音声を探査する有効な手段ができると考えられる。これまで、MUSIC (Multiple Signal Classification) 法を拡張させたさまざまな音源定位手法を提案してきた [1]。2016年11月、提案手法を用いたオンラインシステムを実装し、タフロボティクスチャレンジ評議会 屋外フィールドにおいて、実環境・実時間音源探査の公開デモンストレーションを行った。本稿では、デモンストレーションで用いた実環境・実時間音源探査システムの設計・開発について述べる。

2 音源探査システムの設計

屋外の実環境においてリアルタイムで音源探査を行うためのシステムの設計について述べる。

2.1 機体の組み立て、準備の効率化

Fig. 1 に構築したマイクロホンアレイ搭載 UAV を示す。ベースとなる機体には ACSL 社製 MS-06LA を用いた。Fig. 1 左下に示す、球形の 12 ch マイクロホンアレイを制作した。実環境で使用するため、空気を通し、かつ耐水性の素材でマイクロホンアレイの表面を覆った。また、本マイクロホンアレイは本体と支柱 1 箇所、ケーブルはコネクタ 1 つで接続でき、組み立ての効率化や断線リスクを低減できる。Fig. 1 右図は、機体上方から見た簡略図である機体前方の方位角を 0° とし、30° の箇所にマイクロホンアレイ、-90°、150° の箇所に機体のバランスを取るためのダミーウェイトを設置した。さらに、多チャンネル音響信号収録装置 RASP-ZX^{*}、GNSS/IMU モジュール、無線カメラ、通信モジュール、機体の駆動とは独立になった音源探査システム用バッテリーを、Fig. 1 中赤矢印に示すケースに封入し、機体へ固定し、準備の効率化を図った。その結果、組み立て準備にかかる時間が 2 時間から 40 分に短縮され、電源投入後 2 分程度で離陸可能となった。

2.2 システムの安定性の確保

本システムの構成を Fig. 2 に示す。本システムは、UAV-基地局間で、(1) 収録した音響データの伝送、(2) GNSS/IMU のログデータの伝送、(3) 無線カメラによる映像の伝送、(4) コントローラとの通信の 4 系統の無線通信を行う必要がある。音響データはサンプリングレート 16 kHz、量子化ビット数 24 bit で収録されるため、(1) は約 5 Mbps の帯域が必要である。そこで、システムの安定性を確保するため、それぞれの通信周波数を、(1) 2.4 GHz、(2) 900 MHz、(3) 5.6 GHz、(4) 72 MHz とし、混線を防いだ。また、(1) に関しては、より

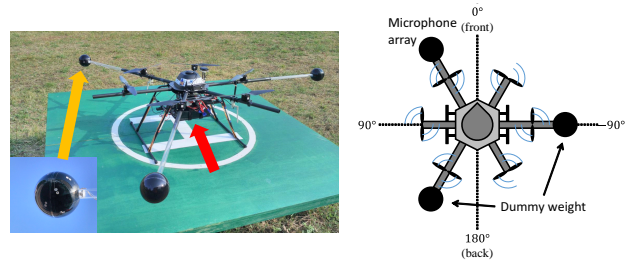


Fig. 1 UAV with microphone. Yellow arrow: Microphone array. Red arrow: Packaged system.

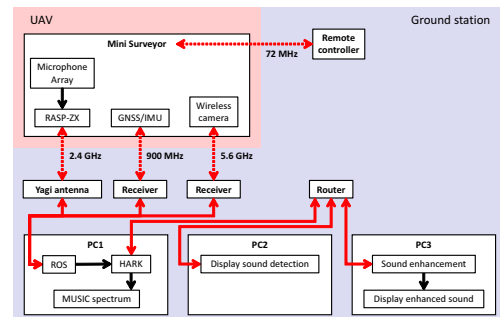


Fig. 2 Framework of online sound source detection and localization system.

確実な通信を行うため、受信側に八木アンテナを使用した。さらに、通信が安定していても数%の欠損は発生してしまうため、パケット損失が起こっても通信を継続できるように、通信プロトコルを UDP (User Datagram Protocol) とした。その結果、約 75 m 離れた地点でも 6 Mbps 以上の速度で無線通信が可能となった。

2.3 遅延の低減

Fig. 2 に示すように、UAV から取得したデータは ROS で統合して処理を行う。音響データは、PC1 にて、ロボット聴覚オープンソフトウェア HARK (Honda research institute japan Audition for Robots with Kyoto university)[†] [2] を用いて音源定位を行う。音源定位手法として、一般的な MUSIC (MULTiple Signal Classification) 法である SEVD-MUSIC (MUSIC based on Standard Eigen Value Decomposition) [3] や、これまでわれわれが提案した iGSVD-MUSIC (MUSIC based on incremental Generalized Singular Value Decomposition) [1] などが挙げられる。前者は、処理コストは小さいがノイズ耐性が低い。また、後者は、時間変化のあるノイズに対する耐性はあるが、処理コストは大きい。処理によるシステムの遅延はさまざまなパラメータに依存するが、本稿では使用する周波数帯域 (ピン数) に着目する。Fig. 3 に、使用する周波数帯域を変化させた場合の、それぞれの手法を用いたシステムの遅延の変化を示す。どちらの手法も、使用する周波数帯域が小さいほど、遅延も小さくなるのがわかる。また、iGSVD-MUSIC では、使用

Real time outdoor sound source detection and localization using a UAV with a microphone array

Kotaro Hoshiba¹, Mizuho Wakabayashi², Kai Washizaki², Takahiro Ishiki², Makoto Kumon², Daniel Gabriel¹, Kazuhiro Nakadai^{1,3}, Hiroshi G. Okuno⁴

1 Tokyo Institute of Technology 2 Kumamoto University
 3 Honda Research Institute Japan Co., Ltd. 4 Waseda University

^{*}http://www.sifi.co.jp/system/modules/pico/index.php?content_id=36

[†]<http://www.hark.jp/>

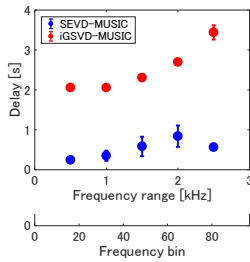


Fig. 3 Processing delay using SEVD-MUSIC and iGSVD-MUSIC.

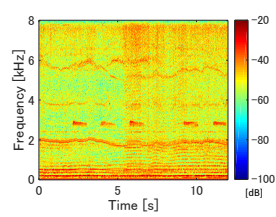


Fig. 4 Spectrogram of whistle sound recorded by UAV-embedded microphone.

する周波数帯域の変化に関わらず、SEVD-MUSICと比較して2秒程度多く遅延が発生していることがわかる。本システムでは、遅延の小さいSEVD-MUSICを選択し、さらに周波数帯域を制限することによりリアルタイム性を追求する。Fig. 4に、対象音源としてUAV搭載マイクロホンで収録した笛の音のスペクトログラムを示す。0~2 kHzに存在する成分がUAVのローター音によるもの、3 kHz付近に点在する成分が笛によるものである。そこで、本システムでは使用する周波数帯域を2.5~3 kHzに限定することにより、処理を軽減し、ローター音の影響を低減させた。その結果、システムの遅延は約0.2秒となり、リアルタイム性を確保することができた。

Fig. 5に、周波数帯域を制限した場合のMUSICスペクトルを示す。(a)がSEVD-MUSIC、(b)がiGSVD-MUSICを使用した場合の結果である。円周方向が方位角、半径方向が水平方向を 0° とした場合の仰角($-90^\circ \sim 0^\circ$)、カラーバーで各方向からの音のパワーを示している。目的音源の方向は方位角 75° 、仰角 -5° 付近に存在しており、どちらの手法においても検出できていることがわかる。また、iGSVD-MUSICではノイズが除去できているが、SEVD-MUSICでは方位角 -150° 付近にローター音が検出されている。ローター音は常に同じ方向に存在するため、検出範囲を方位角 $-90^\circ \sim 170^\circ$ に限定することで、目的音のみを検出することができるよう調整を行った。将来的にはFig. 1に示すダミーウェイトもマイクロホンアレイに置き換える予定であり、それぞれのマイクロホンアレイでの検出範囲を制限しても、すべてのマイクロホンアレイを使用することで全方向の音源の検出が可能になると考えられる。

2.4 直感的な可視化ツールの開発

デモンストレーションを行うにあたり、閲覧者が直感的にわかりやすい可視化を行う必要がある。そこで、可視化ツールの開発を行い、閲覧者にFig. 6に示すような情報の提示を行った。(a)は地上カメラからのライブビデオ、(b)はFig. 5のようなMUSICスペクトル、(c)は(b)を直感的に検出結果を示すため、UAVと音源方向をイラストで示した可視化ツール、(d)はVB-MRNMF (Variational Bayesian Multi-channel Robust Nonnegative Matrix Factorization) による音源強調[4]の結果を表示した。このような情報提示を行うことで、閲覧者は飛行場所から離れた場所においても、直感的に探索結果を確認することができる。

3 デモンストレーション

設計・開発したシステムを用いて、屋外実環境でのデモンストレーションを行った。ローター音や屋外に

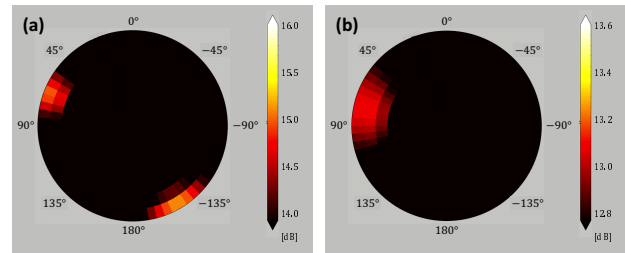


Fig. 5 MUSIC spectrum using (a) SEVD-MUSIC, (b) iGSVD-MUSIC.

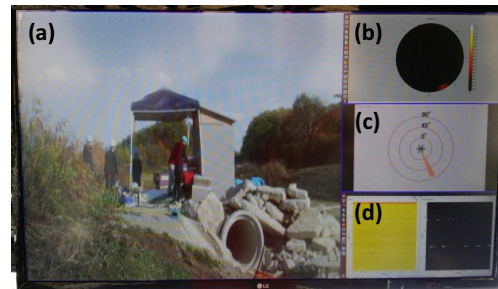


Fig. 6 Displayed information for participants. (a) Live video by ground camera, (b) MUSIC spectrum, (c) sound direction, (d) enhanced sound.

存在するノイズ(小雨が降っていたため、雨音がノイズとして存在)が多い中、誤検出することなく、実時間で目的音源を検出することができた。また、閲覧者は実際の飛行場所から離れていたが、可視化ツールにより、探索結果をリアルタイムに確認することができ、設計・開発したシステムの有用性を確認できた。

4 まとめ

本稿では、被災地等での被災者の探索を目的に、マイクロホンアレイ搭載UAVを用いた屋外実環境・実時間音源探索システムの設計・開発について述べた。機体の組み立て・準備の効率化、システムの安定性の確保、遅延の低減、直感的な可視化ツールの開発について検討を行い、屋外実環境・実時間での音源探索デモンストレーションを行うことができ、システムの有用性を確認できた。

謝辞 本稿をまとめるにあたり、協力をいただいた(株)スカイリモート 岡部和夫氏、京都大学 坂東宜昭氏に感謝する。本研究は、JSPS 科研費 24220006, 16H02884, 16K00294 および、JST ImPACT タフ・ロボティクス・チャレンジの助成をうけた。

参考文献

- [1] T. Ohata *et al.*, "Improvement in outdoor sound source detection using a quadrotor-embedded microphone array", IROS 2014, pp. 1902-1907, 2014.
- [2] K. Nakadai *et al.*, "Design and Implementation of Robot Audition System 'HARK' - Open Source Software for Listening to Three Simultaneous Speakers", Advanced Robotics, 24 (5-6), pp. 739-761, 2010.
- [3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation", IEEE Transactions on Antennas and Propagation, 34 (3), pp. 276-280, 1986.
- [4] Y. Bando *et al.*, "Variational Bayesian Multi-channel Robust NMF for Human-voice Enhancement with a Deformable and Partially-occluded Microphone Array", EUSIPCO 2016, pp. 1018-1022, 2016.