

プログラム名：タフ・ロボティクス・チャレンジ

PM名：田所 諭

プロジェクト名：ロボットインテリジェンス

委 託 研 究 開 発

実 施 状 況 報 告 書 (成 果)

平 成 2 8 年 度

研究開発課題名：

UAV 搭載マイクロホンアレイを用いた音源探索・同定

研究開発機関名：

東京工業大学 工学院 システム制御系 システム制御コース

研究開発責任者

中臺 一博

# I 当該年度における計画と成果

## 1. 当該年度の担当研究開発課題の目標と計画

UAV (Unmanned Aerial Vehicle) を用いた音源探索技術を実現するために以下の課題を設定する。

1. UAVに搭載したマイクロホンアレイを用いた音源の時間・位置の検出技術（音源探索技術）。
2. 検出した音源に対する、音源の聞き分け技術（音源同定技術）。
3. オペレータにわかりやすいユーザインタフェース技術（U I 技術）

また、本技術の展開の一環として動物サイボーグへの協力を進める。具体的には、救助犬にマイクロホン装着し鳴き声を聞き分ける技術に関して連携して推進する。

このうち、H28年度は、1の音源探索技術については熊本大学、早稲田大学との協力の下、実機デモを構築し、フィールド評価会でのデモ実施を通じて、技術実証を行う。2については、H27に構築した手法をベースに実データ収集と実用的な手法の構築を行う。3については、音源探索のオンライン可視化を目指し、マニュアルが不要なわかりやすいU I 構築を行う。救助犬への適応については、東北大学、早稲田大学とともに救助犬で収録した音響信号を解析し、音源識別に向けた課題を明らかにする。

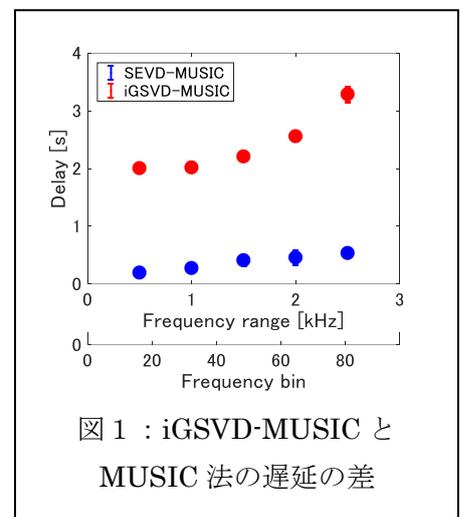
## 2. 当該年度の担当研究開発課題の進捗状況と成果

### 2-1 進捗状況

音源探索技術に関しては、音源探索の遅延低減と性能の維持というトレードオフ関係にある問題をアルゴリズムの変更とマイクロホンレイアウトを工夫することにより解決し、遅延の少ない音源探索を実機上で動作できるレベルまで達成した。音源同定に関しては、深層学習に必要な学習データの Annotations コストを低減できるより実用的な手法を提案し、有効性を示した。U I 技術に関しては、専門家でも一目で状況がわかるU I の開発を行い、音源探索実時間システムのU I として採用した。救助犬の声の聞き分けについては、動物サイボーグチームでのマイクロホン付きスーツの開発、およびデータ収集計画の遅れにより、本年度はアルゴリズムの準備を行った。次年度以降で実施予定である。

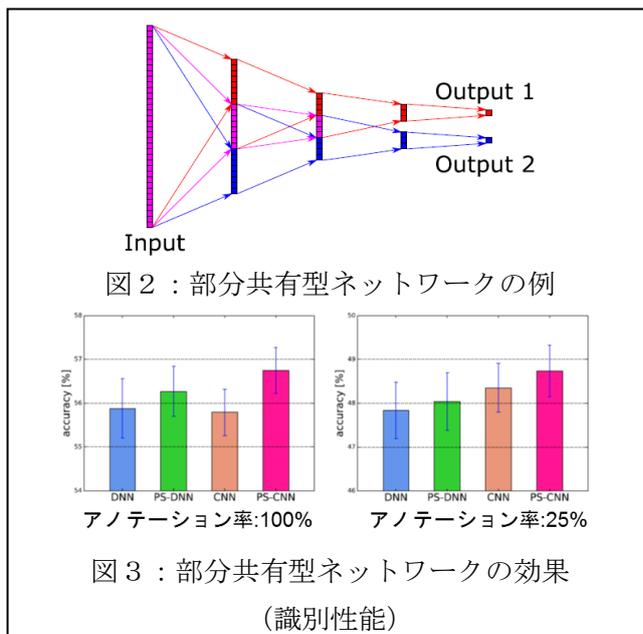
### 2-2 成果

音源探索技術の性能については、これまで用いてきた iGSVD-MUSIC (multiple signal classification based on incremental generalized singular vector decomposition) 法をオンライン処理化した場合、アルゴリズム的な制約から、遅延が大きくなり、オペレータの直観的な認識が難しいという問題が発覚した。そこで、遅延の少ないオリジナルの MUSIC 法を利用できないか検討を行った (図1)。しかし、オリジナルの MUSIC 法は UAV の騒音があるような環境下での定位は難しくなる。これまででは、マイクロホンアレイのサイズを大きくして音源定位性能を追求するアプローチで研究を進めていたため、UAV の外側にフレームを設置して、UAV を取り囲むようなレイアウトでマイクロホン



を設置していた。このアプローチでは、UAVの騒音がマイクロホンアレイの内側から全方向的に到来するため、高精度な雑音抑圧法が必要だった。このレイアウトを変更し、マイクロホンアレイを直径10cmの球体の表面に配置し、UAVのアームを延長して50cmほど離れた場所に設置した。この結果、UAVの騒音はマイクロホンアレイから見て特定の方向から到来するように見えるため、それ以外の方向から到来する音源を抽出すべき音源として検出できるようになった。つまり、マイクロホンアレイのレイアウトと位置を工夫することにより、遅延と性能のトレードオフ問題を解決することができた。このブレークスルーが熊本大学、早稲田大学とともに挑んだ2016年11月のフィールド評価会でのデモ実現につながった。

音源同定は、H27年度に信号処理による音源分離と深層学習による音源識別の組み合わせが有効であることを示した。H28年は、このうち、深層学習による音源同定において、大量の人手によるアノテーション済みの学習データが必要であるという問題にフォーカスし、この解決を図ることで、深層学習ベースの音源同定をより使いやすい実用的なものにする研究開発を行った。この問題を解決するために提案したネットワークが部分共有型ニューラルネットワークである。このネットワークは2つのネットワークを一部共有する形で接続したネットワークである(図2)。実際には一方のネットワークには回帰問題の一種である音源分離、もう一方のネッ



トワークには識別問題の一種である音源同定を学習させる。音源分離部分は信号処理的な手法により、入力とそれに対応するラベルである分離音を自動的に生成できるため、自動的にネットワークの学習が可能である。一方の音源同定には入力に対応するラベルとして音源名が必要であり、これは一般に人が耳で聞いてアノテーションをする必要があるため、データを準備するコストが大きい。部分共有型ニューラルネットワークを用いる場合、音源分離部分を学習すると音源同定部分と共有している部分も同時に学習できるため、音源同定部分の学習に必要な学習データが少なくても済むことが期待できる。実際に評価を行った結果、全データの25%しか音源名のアノテーションが行われていない学習データを使った場合でも、提案手法がDNN、CNNの両方に対して有効であることを示すことができた(図3)。

UIに関しては、水平角、方位角を同時に表示するUI(図4右上)、自己雑音の影響を排除し、目的音の方向のみを可視化できるUI(図4右中)を開発した。また、音源強調処理を併用して、時間的にどのタイミングで音イベントが検出されたかを示すUI(図4右下)も開発した。これらにより音響の専門家でなくても一目でわかるUIを開発することができた。



### 2-3 新たな課題など

音源探索技術については、飛行ロボの他グループとの技術統合を図るため、ドローン搭載の組み込みボードで処理することが肝要であるとの結論に至った。このため H29 年度に新たに組み込み処理化を行う。音源同定については、オンライン化を行い音源探索技術との統合を図る。救助犬の声の聞き分けについては、東北大学のグループより、H29 年度にデータ収録を開始するとの連絡を受けたので議論を深めながら研究開発を進めていく予定である。その他については、予定通り進捗している。

### 3. アウトリーチ活動報告

- ・ロボット聴覚のオープンソースソフトウェア HARK の無料講習会およびハッカソンの開催

2016 年 12 月 6,7 日にこれらの連続開催を行った。毎年行っている活動（講習会は 13 回目、ハッカソンは 3 回目）であり、昨年度同様、TRC も主催者である。無料講習会は 50 名（満員御礼）、チュートリアルは 20 名程度の参加者があり、盛況の成功裏に終わった。

- ・国際ドローン展での技術展示

2016 年 4 月幕張メッセにて開催された国際ドローン展に極限音響技術をポスター・機器展示で紹介した。用意した 350 部のポスターがすべてなくなり、本技術に対する関心の高さが肌で感じることができた。また、展示したマイクロホンアレイを作成した機器メーカーには、すでに問い合わせが数件あるとの報告を受けている。