

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 7 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)

このうち、H27 年度は、1 の音源探索技術については UAV の水平方向 15m 程度離れた距離の音源に対して 80% 程度の検出率を達成し要素技術としての完成を目指す。2 については、小規模ベンチマークテストで 80% の音声識別器を達成する。3 については、音源探索に関して、一目で操作できる可視化U Iを実現する。

2. 当該年度の担当研究開発課題の進捗状況と成果

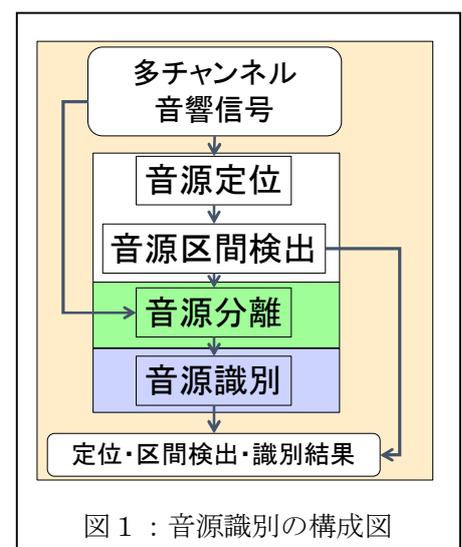
2-1 進捗状況

音源探索技術に関しては、昨年度までにオフライン処理では目標性能を達成したので、当該技術のオンライン処理化を行った。実際に、この技術を用いて実時間で動作するシステムを H28 年度に構築し、公開シンポジウムで技術展示を行う予定である。音源同定に関しては、マイクロホンアレイ処理による音源分離と深層学習による音源同定を組みわせることにより、小規模なデータセットを用いた評価で目標性能を達成することができ、要素技術として技術構築の目途が立った。U I 技術に関しては、音源探索実時間システムの構築に合わせて一目で状況がわかるU I の開発を行っており、処理時間を含めて技術的な目途が立った状態である。音源探索術と合わせて H28 年度の公開シンポジウムでの技術展示を予定している。

2-2 成果

音源探索技術の性能については昨年度報告済みである。U I 技術と合わせて公開シンポジウムに向けた開発を継続している。

音源同定は、H27 年度に検討を開始した。図 1 に構成図を示す。入力音にはプロペラ音や風雑音がこれを除去するため、音源分離を適用した。分離手法は、ロボット聴覚ソフトウェア HARK に含まれる GHDSS(Geometric High-order Decorrelation based Source Separation) 法を用いた。この手法は、ビームフォーミングとブラインド分離のハイブリッド法であり、音源分離アルゴリズムの中でも性能が高いことが知られている。しかし、音源分離を用いても完全に雑音を除去することはできないので、雑音に頑健な音源識別手法として、近年注目されている深層学習で、画像識別問題に効果が高いといわれている畳み込みニューラルネットワーク



(CNN, Convolutional Neural Network) を併せて適用した。CNN は画像識別用の手法であるが、図 2 に示すように音はスペクトログラムとして表現することにより、画像と見なすことができる。つまり音源識別問題を画像識別問題として扱うことができる。最後にシステム全体を最適化するようなパラメータチューニングを行った。これは音源検出に用いる音源検出閾値は、音源検出の際にはできるだけ音源を検出しやすいように低い値が好ましいが、音源同定の際には良好な信号対雑音比得られるように高い値が好ましいというトレードオフ関係に起因している。実際に音声 1 種類、非音声 7 種類の計 8 種類の音源を屋外で UAV のプロペラを回した状態で 3m 離れた場所から収録し、識別実験を行った。

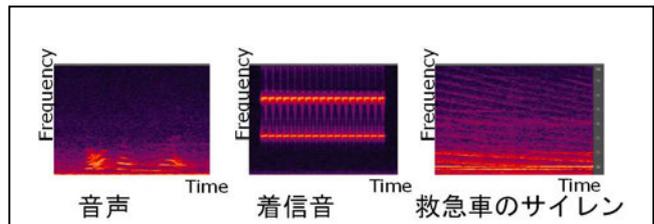


図 2：スペクトログラムの例

表 1：システム性能（識別結果）

	定位性能	識別性能	全体識別性能
定位最適化閾値	85.1	75.7	64.5
識別最適化閾値	80.2	96.9	77.7
全体最適化	83.5	95.0	79.3

収録信号の信号対雑音比は-5~-15dB と音源よりも雑音が多い状態であった (0 dB の時に信号と雑音が同等の大きさである)。結果を表 1 に示す。パラメータの最適化を行うことによりシステム全体の音源識別性能は約 80% となり、目標性能を達成することができた。

UI に関しては、オフライン処理での可視化例を図 3 に示す。左上の窓内が実際の飛行風景 (左上に UAV、右下に話者) であり、処理したデータを用いて再構成した画面が右の窓内に示されている。GoogleEarth 上に UAV と音源の位置が同時に表示されるようになっていたため、一目で状況がわかる可視化が可能となっている。実時間オンライン処理の目途は立っており、実際に実時間処理システム版を構築して、H28 年度に技術展示予定である。

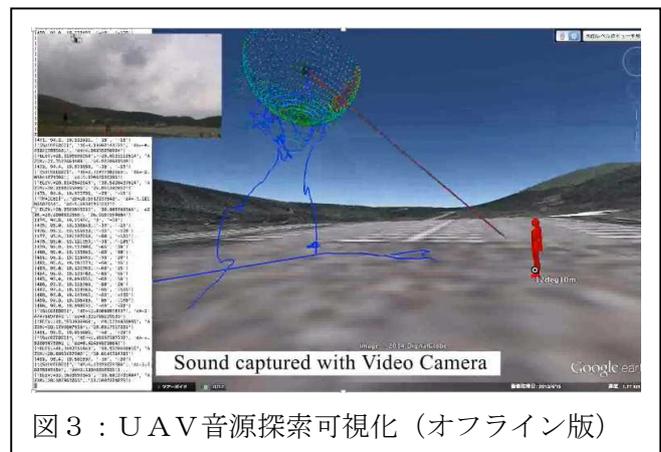


図 3：UAV 音源探索可視化（オフライン版）

2-3 新たな課題など

音源探索技術については、H27 年度で開発を完了する予定であったが、公開シンポジウムでのデモを行うため、H28 年度も継続するものとした。構築済みの技術を用いて実時間システム構築を行うことが主な活動である。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

2015 年 11 月 10,11 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 12 回目、ハッカソンは 2 回目) であるが、本年度から、ImPACT タフ・ロボティクス・チャレンジも主催者として加わった。無料講習会は 45 名、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。