

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 9 年 度

研究開発課題名：

柔軟ロボット音響センシングにおけるブラインド音源分離処理の高精度化

研究開発機関名：

首都大学東京

研究開発責任者

小野 順貴

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

本研究開発は、これまで研究開発者が国立情報学研究所に所属している期間、東京大学、筑波大学らとともに進めてきた一連の研究に継続するものであるため、まず、これまでの研究開発方針・内容を述べ、次に今回の研究開発方針・内容を述べる。

近年、本研究開発責任者らによって、高次統計量の値の変化が人間の聴覚的印象と強い相関があることが明らかになった。端的に言うと、4次統計量を大きく変化させるような信号処理は、概して瞬間的に音のトーン性（音程感）を付与することに等しく、それゆえに雑音の中から音声成分を取り出すブラインド音源分離にて用いることが出来るし、また過度にトーン性を強調するような非線形信号処理はしばしば自然界の聞きなれた雑音を人工的な雑音に変えてしまい、結果として人が聞いた時に不快感や了解性の低下を引き起こす。そこで、本研究開発では、高次統計量追跡による音質定量化及び制御理論を遠隔発話かつ劣悪環境下の音声強調・認識システムと融合し、真に自律的な音声処理システム最適化理論の構築を目標として研究開発を行う。これは、音響センシングシステム自体が環境に合わせて自律的に処理を変えることによって、音声強調品質・音声認識精度を常に保つというものである。

想定されるロボットシステムは、「(A) 柔軟ロボットに設置された位置不定マイクロホンセンサ及びブラインド音源分離ブロック」、「(B) 高次統計量に基づく音声強調処理ブロック」、及び「(C) 極限状況テレコミュニケーション・遠隔オペレーション支援ブロック」から構成される。特に (C) のブロックでは、すなわち(A)(B)により強調された被災者音声を対象にし、極限状況におけるテレコミュニケーションを実現するための一課題として、空間情報の復元、すなわち目的音声方向の定位感を残したまま、強調音声をオペレータに提示する手法の開発に取り組んできた。

しかしながら、空間情報の復元だけでは (C) の実現には不十分である。常時、オペレータに雑音抑圧音を聞き続けさせることは大きな負荷を与えるし、オフライン処理の場合にも全時間を聞きとおさなければならぬため、膨大な時間を要することになってしまう。これを解決するために本研究開発では、東京大学、筑波大学らのチームと連携しながら、発話区間の自動推定に取り組んだ。

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

索状ロボット上に実装したマイクアレーに、(A)ブロック：IVA, ILRMA (IVA + NMF)、(B)ブロック：ノイズキャンセラ、定位保持型 MOSIE に加え、(C)ブロックとして、発話区間検出、空間情報復元処理を実装し、音源定位・空間知覚品質に関しても性能を評価した。そして、システム全体の評価を29年度の11月に完了した。

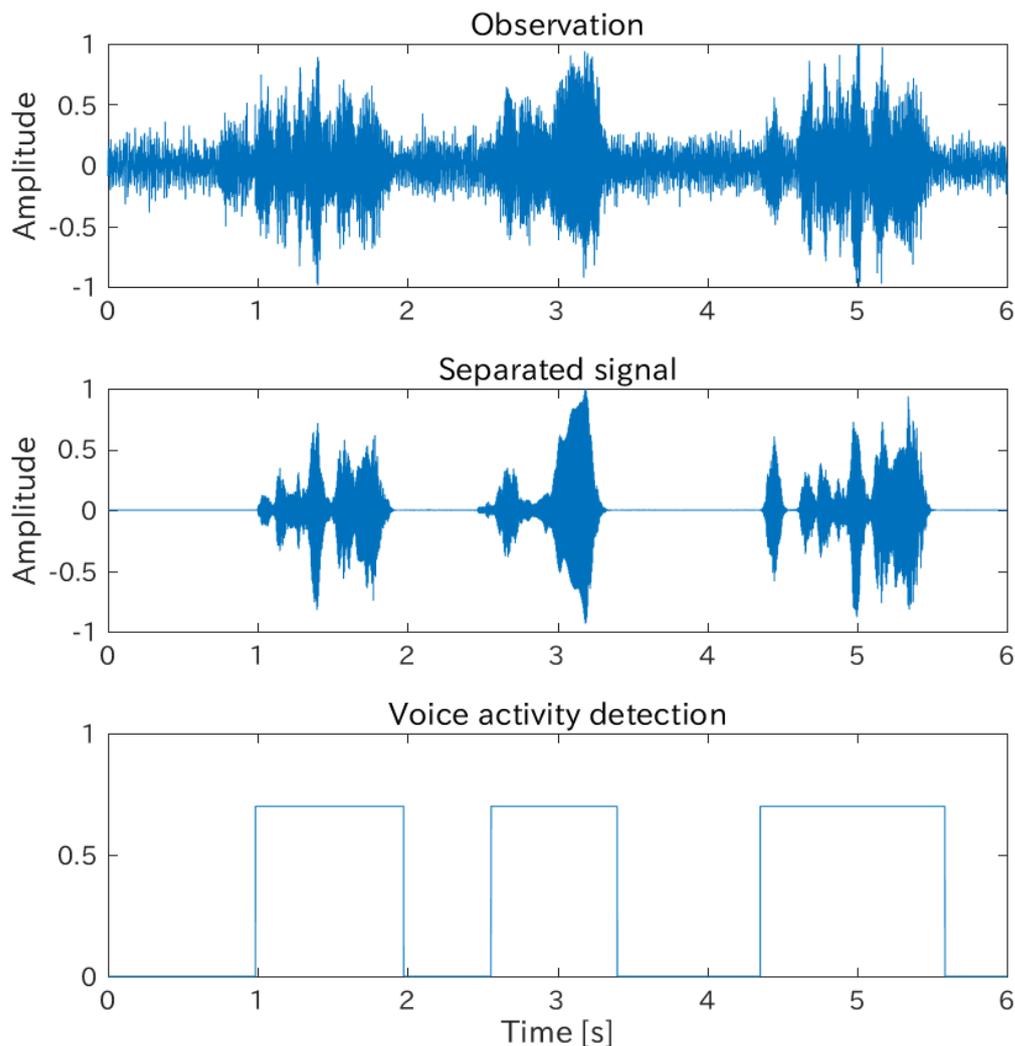
2-2 成果

音声区間の自動推定を目的に、(A) ブロック、(B) ブロックを通した雑音除去信号に対し、フレーム毎のエネルギーに対して閾値処理を行い、その後ポスト処理を行う手法を実装した。閾値は、事前に定めた固定値を用いる方法、観測信号のRMS値を用いて動的に定める方法の両方を試した。動的に信号強度が変化する状況では、後者のように観測信号自体を用いて閾値を変化させる方が妥当であると考えられる。また、ポスト処理としては、

1. 音声区間に挟まれた短い非音声区間（デモシステムでは 250ms）は音声であると判定する
2. 非音声区間に挟まれたごく短い音声区間（デモシステムでは 125ms）は非音声であると判定する
3. 音声区間を前後に伸ばす（デモシステムでは前に 37.5ms, 後ろに 125ms）

といったスムージング処理を適用することによって、音声区間、非音声区間が不必要に短時間に变化してしまうことを防ぎ、安定して音声区間が検出できることを確認した。検出例を以下の図に示す。

関連研究成果3件を、査読付き国際会議で発表した。



3. アウトリーチ活動報告

特になし。