

(続紙 1)

京都大学	博士 (情報学)	氏名	武田 龍
論文題目	A Unified Framework of Blind Separation, Blind Dereverberation and Self-voice Cancellation for Real-time Robot Audition (実時間ロボット聴覚のためのブラインド分離, ブラインド残響除去, 自己発話抑制の統一的枠組み)		
(論文内容の要旨)			
<p>本論文は, 1秒程度の残響の強い実環境で, 複数の音源からの音やロボット自身の発話があっても, 目的話者の声を聞き分け, 実時間で認識するロボット聴覚に関する研究をまとめたものである. 複数音源やロボット自身の発話, さらにはそれらの残響は自動音声認識の性能を著しく劣化させてしまうので, 音源分離と残響除去を同時に実現するロボット聴覚技術は, ロボットと人との音声コミュニケーションを実用化する上で不可欠の技術である. 本論文では, 独立成分分析 (Independent Component Analysis, ICA) に基づくアプローチが述べられている.</p> <p>第1章は序論で, 本研究の背景として, 残響下でのロボット聴覚研究の動機について述べ, 実時間処理, 最小限の事前知識, 音声認識等の後続処理との接続性という課題を指摘し, ブラインド音源分離の必要性, 及び, ロボット発話中に話者が割り込み発話をして自己発話を抑制し相手発話を分離強調する自己発話抑制機能の必要性を述べている.</p> <p>第2章では, ロボット聴覚とそのための信号処理に関する文献, 特に, ブラインド音源分離, ブラインド残響除去, 自己発話抑制についての従来研究を概観し, そこで得られた知見を基に, ブラインド音源分離の1手法であるICAに基づいたアプローチが妥当であることを述べ, ロボット聴覚における残響処理や実時間処理を実現するための統一的枠組みの必要性を指摘している.</p> <p>第3章では, ロボットの自己発話抑制と残響処理を統合したICAについて述べている. 本手法では, 観測信号とその反射音との独立性に基づく分離モデルに注目し, 新たな分離フィルタ推定アルゴリズムをICAの拡張として定式化し, その処理法を開発している. 実験によって, 本手法が従来のICAより, 残響音声に対する認識率を10~20ポイント改善することを示している.</p> <p>第4章では, ICAに基づく音源分離手法の高速化について述べている. 残響の減衰特性を考慮することにより, 分離過程の再帰表現, 分離フィルタステップサイズの区分線形モデル化, sub-ICA-filterによるステップサイズ適応という3種類の高速化法を考案し, 音源分離過程の演算量削減を実現している. 評価実験により, 本手法が固定ステップサイズよりも演算量を約30%削減し, 音源分離手法の高速化に寄与することを示している.</p> <p>第5章では, 前章までに開発してきたICAに基づく音源分離の性能評価を, 音声認識の観点から行っている. 異なるマイク数, 音源数, 音源間隔の組合せに対して, 分離音の音声認識を行っており, 本手法が従来のICAよりも約30ポイント認識率を改善することを確認している.</p> <p>第6章では, 残響下での音源分離の性能に大きな影響を与える残響時間の推定方法について述べている. ロボットが自分自身の発話した音声を用いて, 環境の残響時間</p>			

を推定する手法であり，その残響時間の指標は，ICAのエコーキャンセルフィルタの減衰率に基づいている．実験によって，本手法で得られる残響時間を基にICAのフィルタ長を適切に設定することができ，音源分離が音声認識精度で5ポイント向上し，演算量が最大60%削減されることを示している．

第7章では，ICAに基づく音源分離の実装技術として，複数スタックに基づく並列実装とリサンプルに基づくオーバーラップ・ブロック分離に基づいた方法について述べている．評価実験により，8コアCPUを用いて，リアルタイムファクタが0.5未満に抑えられることを確認するとともに，リサンプルなしの単純並列実装と比較し，約2～10ポイント音声認識率が改善することを示している．

第8章では，本研究のまとめを行い，実環境下でのブラインド音源分離に関する関連研究に対する，本研究の主要な貢献について述べるとともに，残された課題や今後の方向性についても考察を行っている．

第9章では，本論文の結論を述べている．

(論文審査の結果の要旨)

本論文は、1秒程度の残響の強い実環境で、複数の音源からの音やロボット自身の発話があっても、目的話者の声を聞き分け、実時間で認識するロボット聴覚に関する研究をまとめたものである。得られた主な成果は次の通りである。

1. ブラインド音源分離、ブラインド残響除去、及び、自己発話抑制に対する統一的な枠組みを、独立成分分析 (Independent Component Analysis, ICA) の中に組み込み、観測信号とその反射音の独立性に基づく分離手法を考案し、分離フィルタとして、実装しており、残響音声に対する認識率が10~20ポイント改善することを確認していること。
2. 自己生成音を抑制し、相手発話だけを聞き分ける自己発話抑制機能をセミブラインド分離としてモデル化し、ICAによる実現法を考案し、時間遅れの少ない実装法を開発していること。また、本手法を、ロボットやシステムの発話中にユーザが割り込み発話するようなバージイン発話に対して、自己生成音を抑制し、相手発話だけを分離し、音声認識システムに応用していること。
3. ICAに基づく音源分離手法が、残響の減衰特性を考慮すると分離過程が再帰表現できることに着目し、分離フィルタステップサイズの区分線形モデルを採用し、sub-ICA-filter によるステップサイズ適応を行う、という高速化法を考案し、音源分離過程の演算量削減を実現していること。また、評価実験により、本手法が固定ステップサイズよりも演算量を約30%削減し、音源分離手法を高速化することを確認していること。特に、マイク数、音源数、音源間隔の多様な組合せに対して、分離音の音声認識実験を行っており、本手法の優位性を確認していること。
4. ICAの分離性能を大きく左右する残響時間推定法として、ロボット自身が発した音声を用いて、ICAのエコーキャンセルフィルタの減衰率に基づいた環境の残響時間を推定する方法を考案し、推定された残響時間を用いることで、ICAの演算量が最大60%削減でき、音声分離性能が音声認識精度で最大約6ポイント向上することを実証していること。

以上本論文は、比較的強い残響環境におけるブラインド音源分離、ブラインド残響処理、自己発話抑制という3つの問題に対する統一的な枠組みを示し、時間遅れの小さい実時間処理の処理系を開発したことに意義がある。特に、残響環境における音源分離は長年の未解決問題であり、この解決策を与えたことは重要な成果と考えられる。さらに、音声対話や音楽合奏に応用し、その有効性を具体的に示しており、学術上、實際上寄与するところが少なくない。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。

また、平成23年1月24日実施した論文内容とそれに関連した口頭試問を行った結果合格と認めた。