

論文内容の要旨

博士論文題目 Blind Source Separation Based on Multistage Independent Component Analysis
(多段独立成分分析に基づくブラインド音源分離)

氏名 西川 剛樹

(論文内容の要旨)

音声ユーザーに優しいヒューマンマシンインタフェースとして、ハンズフリー音声認識やハンズフリー通話システムの実現が期待されている。しかし、実環境下では、種々の背景雑音や残響が存在するため、通常の接話型マイクロホンを用いた場合に比べユーザーの音声を高い信号対雑音比で收音できなくなる。そのため高精度な音声認識・通話システムを実現することは困難とされてきた。

本論文では、目的音の到来方位や適応処理のための非発話区間情報といった事前情報が不要であるという利点を有するブラインド音源分離(BSS)を導入することで、上記の問題の解決を図っている。BSSとはマイクロホンアレーで受信された観測信号のみから目的音を分離する技術であり、音源信号同士の独立性を用いた独立成分分析(ICA)に基づく手法が近年盛んに研究されている。しかし、従来ICAは、そのフィルタ学習の複雑さより、通常室内環境下にて十分な性能を発揮することが困難であった。そこで本論文では、以下で述べる幾つかの新規理論提案・検討を行っている。

BSSは分離フィルタを推定する領域の違いで、周波数領域ICA(FDICA)と時間領域ICA(TDICA)とに分類されるが、実環境下における実験結果より、FDICAでは、帯域分割数を過度に増やすと狭帯域信号間の独立性の仮定が成立しなくなり分離が困難になるという問題が新たに発見された。一方、TDICAにおいては、分離フィルタの反復学習における低収束性により、長い残響時間を有する混合系へ適用することは非常に困難であることが確認された。そこで本論文においては、まず、FDICAによって分離された信号をTDICAの入力とみなし、FDICAにおける残留クロストーク成分をTDICAによって分離することにより上記の問題を解決する手法(多段独立成分分析;MSICA)の提案を行っている。実環境下での音源分離実験より、提案手法はTDICA及びFDICAに比べ高い音源分離性能及び音声認識性能を有していることが実証された。

次に、MSICA の多素子化に関して拡張アルゴリズム提案および検討を行っている。先に提案された基本 MSICA では、マイクロホンと音源は同数であるというモデルを仮定していた。一般に残響環境下において、さらに高精度な音源分離を実現するためにはマイクロホン素子数を増やす必要があるが、素子数を増やすことで FDICA 部における音源の入れ替わり問題が複雑になる。この問題を解決するために、音源と同数のマイクロホンにより構成されたサブアレーを用いた MSICA に基づく優決定 BSS を提案している。実環境下における音源分離実験より、提案法はマイクロホン数を増やすことで分離性能が向上するということが確認された。

最後に、分離音質品質を保証した安定アルゴリズムへの拡張に関して検討を行っている。一般に、音声信号のように時間的に相関のある信号に対しては、ホロノミック拘束による無相関化を避けるために非ホロノミック拘束 TDICA を適用しなければならないが、その学習安定性は保証されない。この問題を解決するために、まず MSICA と線形予測器を組み合わせた手法を提案し、更に、ホロノミック拘束の TDICA による分離信号の白色化に寄与する成分を推定し、推定された処理歪み成分を用いて分離信号の音質を改善する手法も提案している。残響環境下における実験結果より、提案手法は従来のホロノミック拘束やノンホロノミック拘束に基づく TDICA に比べ高い安定性を有し、高い分離性能を実現できることが示された。

以上の理論提案およびその実験的検証により、残響環境下におけるより実用的な BSS の可能性が示された。

(論文審査結果の要旨)

音声ユーザーに優しい柔軟なヒューマンマシンインタフェースとして、ハンズフリー音声認識やハンズフリー通話システムが望まれている。しかし実環境下では、種々の背景雑音等が混入してしまうため、高い音声認識率や SN 比を達成することは困難とされてきた。本研究では、ユーザー音声の発話位置や区間等の事前情報が不要であるという利点を有するブラインド音源分離(BSS)を導入することで、上記の問題の解決を図っている。BSS は、観測信号のみから目的音を分離する技術であり、音源信号同士の独立性を用いた独立成分分析(ICA)に基づく手法が近年精力的に研究されている。しかし、従来の ICA は、特に残響環境下においてその性能の劣化が顕著であった。そこで本論文では、以下で述べる幾つかの新規理論提案を行い、その有効性を実環境データによって検証している。

- (1) 周波数領域 ICA の性能限界を、狭帯域信号の独立性仮説とフィルタの自由度の観点より説明し、両者にトレードオフが存在することを世界で初めて示した。また、この考察に基づき、周波数領域 ICA と時間領域 ICA とを組み合わせた多段 ICA を提案し、その優位性を実機実験（通常室内および車室内）により実証した。
- (2) 多段 ICA の多素子化拡張を行うため、サブアレー信号処理の観点を導入し、素子数に応じて適切な音源分離性能が達成されるシステムを考案した。
- (3) 分離後音声の品質にも注目し、スペクトル歪みの少ない高品質な BSS を達成するために MSICA と線形予測器を組み合わせたアルゴリズムを提案した。また、その考え方を更に発展させ、安定なホロノミック学習の上で歪み成分を減算補正する手法も考案し、その優位性を実環境データにより検証した。

これらの手法は、従来よりあまり実用的ではなかった実残響環境下での BSS 問題の解決に関して、有用かつ画期的なものであったと高く評価できる。本研究成果は、3 編の英文原著論文や 10 件の国際会議・21 件の国内会議にて発表され、3 件の論文賞・奨励賞を受賞したことから、その研究レベルの高さが内外より評価されていることが分かる。

以上述べたように、本論文の音響・音声信号処理分野への貢献は極めて大きい。平成 16 年 1 月 24 日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は博士論文（工学）の学位論文として十分な価値があるものと判断した。