

報告番号	※甲	第	号
------	----	---	---

主 論 文 の 要 旨

論文題目 Blind source separation with the low computational costs for the mobile and portable speech equipment
(小型音声処理装置のための低演算量で実現可能な音声音源分離に関する研究)

氏 名 近藤多伸

論 文 内 容 の 要 旨

本論文では、周波数ビンの選択により計算量を削減した、周波数領域独立成分分析 (Frequency Domain Independent Component Analysis; FDICA) に基づくブラインド音源分離 (Blind Source Separation; BSS) を提案する。遠隔会議や携帯電話といった音声通話装置では、明瞭な音声品質が求められる。通常、小型の音声処理装置は、デジタルシグナルプロセッサ (Digital Signal Processor; DSP) といった組み込みプロセッサにより設計・製造される。組み込みプロセッサは低消費電力とリアルタイム処理を同時に実現するため必要となる。BSS は高品質な音声強調への応用も目的として広く研究されきており、中でも FDICA は最も代表的な手法のひとつである。通常、音響的には残響が不可避であり、FDICA は逆フィルタ法に相当することから残響の影響を最小限にすることが期待される。しかしながら、FDICA の計算コストは極めて高い。FDICA は高次統計量を用いた反復最適化法で実現されることが多く、また逆フィルタの係数は周波数ビン毎に反復推定されるためである。現在の DSP には高い性能を示すものもあるが、適切な DSP の選択は利用シーンに応じた装置の性能によって決定される。モバイル機器に対してはバッテリーが連続稼働時間に影響し、通常は低消費電力 DSP が使われる。ICA は高次統計量による最適化手法であるため、長時間の観測信号を記憶させる必要がある。通常、大容量化の容易なダイナミック・ランダム・アクセス・メモリ (Dynamic Random Access Memory; DRAM) が信号の記憶に使われる。しかし、DRAM には無駄な電力消費に繋がるメモリアクセスのための待ち状態が、無視できない課題として存在する。DSP 自体の電力消費と DRAM アクセスの課題を併せ、高い性能の DSP を使い FDICA を組み込みプロセッサで実現することは困難な課題であると解る。提案法は ICA が動作する際に周波数ビンの数を減らすことで、計算コス

トを減らすことを目的とする．計算コストを減らすことで，FDICA を組み込みプロセッサにて実現することが容易となる．提案する周波数ビン選択法は，空間相関を利用し，特に本論文では空間相関行列の行列式と，2つのマイク間での振幅二乗コヒーレンス (Magnitude Squared Coherence; MSC) を用いる．

FDICA に対する二つの周波数ビン選択方法のひとつは携帯電話などのモバイル機器に適したもので，もうひとつは可搬性のあるテレビ会議システムなどに適したものである．モバイル機器向けには空間相関行列の行列式を周波数ビン選択に用いる．行列式は理論的に解析され，音源方向に関する情報と音源パワーの情報を同時に評価できることを示す．言い換えると，信号の分離性を行列式により評価できることに相当する．選択された周波数ビンではICA の分離行列を学習し，非選択の周波数ビンでは死角型のビームフォーマ (Null-Beamformer; NBF) の出力を仮の分離信号としウィナーフィルタを得る．実験により分離性能の向上を確認した．マイクロホンアレイ信号処理の音響的な限界から示される信号歪がキャンセルされ性能向上に寄与したと考えられる．セグメンタル SNR が明らかな改善を示す一方，ケプストラム歪は周波数ビンの選択数を減らす程に悪化する傾向が見られた．従って，二つの評価値の間にはトレードオフが見られるが，これは周波数ビンの選択数の決定基準に役立つ．本論文では実験的に 64 ビンとした．周波数ビン選択を実施しない FDICA の従来法と比べたとき，提案法は計算コストを 80 % 以上削減する．提案法は効果的な計算コスト削減を示しつつ，セグメンタル SNR は 2dB 程度の向上を見せ，一方ケプストラム歪の低下は 1dB 程度に抑えられている．

可搬性の音声処理装置として，小型かつ凝集型の收音システムである正十二面体マイクロホンアレイ (Dodecahedral Microphone Array; DHMA) が以前に提案された．DHMA では 10 面にそれぞれ 6 個のマイクロホンを実装した．DHMA を使った BSS では，FDICA が優決定条件下で動作するため，マイク数が音源数を上回る．FDICA 分離行列の大きさは，DHMA を使った BSS では非常に大きくなり計算コストを押し上げる．パーミュテーション解決は特に計算コストを必要とするが，これはクラスタリング手法が多数の類似度計算を含むことに由来する．ここでは MSC を周波数ビンの選択に用いる．DHMA の形状は球状マイクロホンアレイに似ているが，音響的な特徴は多数の平面の組み合わせから成るため異なる．MSC の計算には 2 つの任意のマイクロホンが選ばれ，MSC の値が小さなほど分離の良い周波数領域であるという基準で評価した．DHMA の形状は MSC を小さくすることを実験的に確認した．提案法では，周波数ビンの選択を不等間隔とする．68 % の周波数ビンを用いる場合で，計算コストの削減結果は 80 % を超える一方，信号対妨害音比 (Signal-to-Interference Ratio; SIR) による分離性能は周波数ビン選択により低下した．しかし，等間隔に周波数ビンを選択した場合に比べて，その低下は 1dB 以下に抑えられた．分離処理に伴う歪成分は，セグメンタル SNR が若

干低下するも、ケプストラム歪は改善した。これら歪に対する客観評価指標についても、等間隔に周波数ビンを選択した場合と比べて提案法による改善が見られた。

いずれの周波数ビン選択方法も、2つのマイクロホンを用いた空間相関を計算する。また空間相関は2次統計量に相当する。これは高次統計量に基づくFDICAによるBSSが、2次の統計量を組み合わせることで信号歪を同程度に保ちながら、計算コストを削減できることを示している。音声処理装置がDSPにより実現される場合には、DRAMアクセスのための待ち状態が実現上の重要な課題となることを述べた。FDICAは長時間の信号をDRAMに蓄えなければならない。DSPには高速な内部メモリもあるが、これは単にデータを保持するだけに使うべきではない。提案法は周波数ビンの数を減らし同時にメモリ消費量も減らす。これは同時にメモリアccessの回数も減らし、DRAM利用時の課題を緩和する。本論文ではFDICAアルゴリズムの計算コストを明らかにした。装置に内在されるべきDSPを想定し、目標とする演算量はモバイル機器で200メガの演算回数、可搬性の装置で160ギガの演算回数と見積もった。提案法による計算コストは、目標計算コストを下回ることが確認できた。またFDICAアルゴリズムの中で、計算コストに支配的な要因が周波数ビンの数で変わることを示した。これはエンジニアがソフトウェアを最適化する際の指針となる。

現在、世の中で使われている音声強調分野での客観的評価指標を考慮すると、10dBから20dBの絶対的なSIRが許容される性能と考えて差し支えない。提案法により分離性能が低下しても、現実的な市場応用に対しては許容される範囲内に収まった。一方、歪に関する指標は、80%以上と大幅な計算コスト削減を実施しても、若干の低下は見られつつも削減前と同等と言える。つまり提案法は、一方では分離性能と計算コストの現実的なトレードオフを示し、他方で歪と計算コストのトレードオフは改善を示した。提案するブラインド音源分離法が現実的でかつ、組み込みプロセッサを用いた音声処理装置へ供するのに適していると言える。今後の課題は、非選択周波数ビンの分離手法が提案法に留まらず検討の余地がある点、またオンライン手法への拡張も挙げられる。