

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 2 日現在

機関番号：34419

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25420423

研究課題名(和文) 即時稼働可能なリアルタイム雑音除去システムの開発

研究課題名(英文) Development of immediately operational real-time noise reduction systems

研究代表者

五反田 博 (GOTANDA, Hiromu)

近畿大学・工学部・教授

研究者番号：10153751

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：音声のスパース性を利用して音源の到来方向(DOA: Direction of Arrival)をフレーム単位に推定する方法と、その推定値をもとにソフトマスク処理を行って音源をリアルタイムに分離する方法を提案した。その結果、固定音源については、その方位が $\pm 30^\circ$ の範囲、残響時間が200msec以下、SN比が15dB以上あれば、DOAが精度良く推定できて目的音源を高品質に分離できることが確認された。また移動音源については、これと同一環境下において、音源を様々な運動パターンで32～96cm/sの速さで移動しても、誤差 4° 未満の精度で移動音源をリアルタイムに追跡できることを確認した。

研究成果の概要(英文)：In the research project, we have proposed a frame-wise DOA (Direction of Arrival) estimation using the sparsity of speech sound sources and a real-time sound source separation using soft-mask filters designed based on the DOA estimate. For fixed sound sources, our approach has been confirmed that the DOAs can be estimated with high accuracy and that the target sound can be separated in real time in high quality under the condition that the target sources are in a range of $\pm 30^\circ$, the reverberation time is not more than 200msec and the SN ratio is not less than 15dB. Under the same condition, it has been confirmed that the moving sound source can be tracked in real time with an accuracy of less than error 4° even if the sound moves in a variety of movement patterns at a speed of 32-96cm/s.

研究分野：工学

キーワード：ブラインド信号処理 音声のスパース性 音源フレーム検出 フレーム単位のDOA推定 リアルタイム音源分離 移動音源追尾 低消費電力型雑音除去

1. 研究開始当初の背景

タブレット端末や携帯電話が世界的に普及する中、これらの機器に内蔵されたビデオ通話アプリ(例えば FaceTime)や音声認識アプリ(例えば Siri)によるビデオ通話サービスや音声認識サービスを利用するユーザが増えている。しかし、これらのサービスが一般ユーザにまで行き渡るには、音環境(室内の音響特性や個々の音源の発生位置等)に関する事前情報が無くても、人間のように目的音声のみを即時リアルタイムに抽出できる電池駆動型雑音除去システムを構築する必要がある。このようなシステムの実現は、音声信号(音波)だけでなく、超音波や電波、脳波等を信号とする分野でも望まれており、ブラインド信号分離として世界的に研究が進められている。

音声信号については、話者や雑音等の音源を固定した残響の少ない環境で、話者音声を明瞭に抽出することに国内外の研究機関で成功している。これらの場合、予め収録した1~2秒間の音声データにICA(Independent Component Analysis)法を適用して、成分置換とスケールを補正する学習過程(1~2秒)を経て、音源分離(雑音除去)処理が開始される。しかし、通信回線で100ms程度の遅延を想定したビデオ通話サービスにとって、このように音声データの収録と学習に数秒間要することは致命的であることから、即時リアルタイムに稼働する雑音除去システムの開発が望まれていた。

2. 研究の目的

ICA原理に基づく音源分離の場合、「数秒間の学習データが必要」、「ICA法は計算負荷が大きい」、「音源数がマイク数より多いとき適用できない」、「音源の移動や音源数の増減により音環境が変動する場合、適用できない」等の問題があるため、その適用は限定的な環境に限定されていた。そこで本研究では、マイクロホンデータの取得から音声の分離までを10ms程度で処理できる即時稼働型リアルタイム雑音除去システムのプロトタイプの開発を試みた。

3. 研究の方法

1) フレーム単位のDOA推定法の開発

ICA法では数秒間の収録音声をもとに学習が展開されるため、収録開始から数秒遅れて音声は分離されることになる。そこで、音源のDOAを32~64msのフレーム単位に推定するとともに、推定されたDOAに基づいて目的音源のみに指向性を向けるフィルター(ソフトマスク)を設計して、目的音源を分離する方法の開発を試みた。この場合、フレームを32msの間隔でずらすことでほぼリアルタイムに音源が分離されることになる。

2) 代表的DOA推定法MUSICとの比較

上記で開発を試みたフレーム単位のDOA推

定法の有効性や発展性を調べるため、既存の代表的なDOA推定法であるMUSICと比較実験を行った。提案法とMUSICは、前者が断続的な広帯域信号/後者が連続的な狭帯域信号を対象にしている点、前者が単一フレーム/後者が複数フレームを利用する点異なる。

3) DOA推定値に基づく分離フィルター設計

提案法によるDOA推定値をもとに、マイクロホンで観測された混合信号から目的音のみを分離するフィルター(ソフトマスク)の設計を試みた。また、その有効性を検証するため、3音源が活性する状況下で目的音を分離抽出し、分離音声の品質を評価した。

4) 提案法の移動音源に対する有効性の検討
単一の音源が様々な運動パターンで移動する状況で、その音源の方位を追跡できるか、シミュレーションにより検討した。

4. 研究成果

1) フレーム単位のDOA推定法の開発

2個のマイクロホンで観測した音声データを32~64msのフレームに分け、そのフレームを短時間離散フーリエ変換(STFT)した。また、そのとき得られる複素スペクトルの位相差をもとに、フレーム内の各時間周波数点で局所DOAを求めた。さらに、音声のスパース性を利用して、局所DOAの頻度分布をHoyer尺度でフレーム単位に評価し、当該フレームが無音/複数音/単一音のいずれであるかを弁別し、単一音と判定されたフレームについては局所DOAの頻度分布のピークを探索し、ピークを採るとき方位を目的音源のDOA推定値と採択する方法を提案した。

この提案法について、2つの妨害音が存在するもとの、シミュレーションを行った結果、目的音源がブロードサイド方位から $\pm 30^\circ$ の範囲にあれば、残響時間が200msec以下でSN比が15dB以上のもとの、目的音源のDOAを精度良く推定できることを確認した。

2) 代表的DOA推定法MUSICとの比較

提案法について、既存の代表的なDOA推定法であるMUSICとの比較実験を行った。その結果、MUSICは提案法に比べてより広角な範囲のDOA推定が可能であるが、音声のように断続的な広帯域信号のDOAを推定した場合、推定値のバラツキが大きくリアルタイム性に欠けることが明らかになった。また、MUSIC法による推定値のバラツキの原因をCramer-Rao Boundの観点から明確にした。

3) DOA推定値に基づく分離フィルター設計

2つの妨害音が存在するもとの得られた目的音源のDOA推定値をもとに、目的音源のみに指向性を向けるフィルター(ソフトマスク)を設計した。このフィルターにより、マイクロホンでの観測値(混合信号のSTFT版)から目的音がリアルタイムに分離できるこ

とを確認した。

さらに分離音声の品質を3つの音質評価指数 ISR (Source Image to Spatial distortion Ratio), SAR (Signal to Artifacts Ratio), SDR (Signal to Distortion Ratio) で評価した。その結果、残響時間が200msでSN比が20dBのとき、ISR = 4.70dB, SAR=1.98dB, SDR=3.06dBと高品位に目的音を分離できることが確認された。

4) 提案法の移動音源に対する有効性の検討

提案法を移動音源のDOA推定(音源追跡)に拡張し適用することを検討した。具体的には、目的音源を方位30°から-30°の間を秒速8cmで直線的に移動させた場合、±30°の方位区間を秒速48cmで往復移動させた場合、移動速度を32cm/sと96cm/sと変化させて往復移動させた場合、音源が直線的に動く途中で方位を0°から30°もしくは0°から-30°にジャンプ移動させた場合等について、シミュレーションを行った。その結果、残響時間が200msec以下でSN比が20dB以上であれば、誤差4°未満の精度でリアルタイムに音源を追跡できることを確認した。

従来法による音源追跡の場合、拡張Kalman filterやParticle filter等にみられるように、マイクに入る音声データの収録だけでなく、音源の動きを予測する運動モデルの導入が必須である。このように運動モデルを導入する点は、固定音源の定位にはない特徴であるが、そこでの計算量は一般に膨大となるためリアルタイムに音源を追跡することは困難である。一方、提案法では何らの運動モデルも必要としない。

以上のことから、提案法は、固定音源のDOA推定だけでなく移動音源の追跡に対しても新たな道筋をつけたものと考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計19件)

N. Iwasaki, M. Tamaki, S. Fukase, K. Inoue, H. Gotanda: "A study on sound source tracking based on a frame-wise DOA estimation," Proc. the 47th ISCIE Int'l Symp. on Stochastic Systems Theory and Its Application (SSS15) pp.1-7 [to be published] (査読有)
Y. Yoshino, K. Kiyota, T. Ishibashi and H. Gotanda: "Effect of Blind Source Separation for Analog Modulation," ICIC Express Letters, Vol. 10, No. 1, pp.15-20, Jan. 2016. (査読有)
T. Ishibashi, K. Fujimori, H. Shintani, C. Okuma, K. Hayama and H. Gotanda: "Blind Source Separation without Scaling Indeterminacy Using Amplitude Ratio of Observed Signals," ICIC

Express Letters, Vol. 10, No. 1, pp. 191-196, Jan. 2016. (査読有)

K. Hayama, T. Ishibashi, C. Okuma and H. Gotanda: "Implementation of Directional Characteristics by Real-Time Processing of Sounds Observed by Two Microphones," ICIC Express Letters, Vol. 10, No. 1, pp. 251-254, Jan. 2016. (査読有)

H. Shiratsuchi and H. Gotanda: "Frequency domain blind channel estimation without phase ambiguity for QAM-OFDM systems," Proc. the 9th Int'l Conference Signal Processing and Communication Systems (ICSPCS2015), pp. 1-8, Dec. 2015. (査読有)

K. Hayama, T. Ishibashi, H. Shintani, M. Motoki and H. Gotanda: "Implementation of directional characteristics by real-time processing of sounds observed by three microphones," Proc. the 3rd IIAE Int'l Conference on Intelligent Systems and Image Processing, pp.305-308, Sep. 2015. (査読有)

C. Okuma, T. Ishibashi, K. Hayama and H. Gotanda: "Variable arbitrary directional characteristic pattern and its application to two-channel microphone system," Proc. Life Engineering Symposium 2015, pp. 239-241, Sep. 2015. (査読有)

石橋孝昭: "音声コミュニケーション装置と今後の課題," 月刊自動認識, Vol. 28, No. 6, pp.32-36, 2015 (査読無)

H. Sogo, H. Tanaka, N. Iwasaki, T. Matsuzaki, H. Shiratsuchi and H. Gotanda: "Blind channel estimation for QAM-OFDM without phase ambiguity," ICIC Express Letters, Part B: Applications, Vol. 6, No.3, pp. 627-635, March 2015. (査読有)

N. Iwasaki, S. Nishimura, K. Inoue and H. Gotanda: "Blind source separation using frame-wise DOA estimates based on DUET," ICIC Express Letters, Part B: Applications, Vol. 6, No.3, pp. 877-886, March 2015. (査読有)

N. Iwasaki, K. Inoue and H. Gotanda: "A Real Time Oriented Sound Source DOA Estimation Based on Sparseness," Trans. ISCIE, Vol. 27, No. 12, pp. 493-500, Dec. 2014. (査読有)

S. Nishimura, N. Iwasaki, K. Inoue and H. Gotanda, "A new blind sound source separation based on frame-wisely obtained DOA estimates," Extended Abstract of the 46th ISCIE Int'l Symp. on Stochastic Systems Theory and Its Application (SSS14), pp. 23-24, Nov.

2014.

N. Iwasaki, T. Matsuzaki, G. Hirano, H. Shiratsuchi, K. Inoue and H. Gotanda:

“Studies on real time DOA estimation based on DUET,” ICIC Express Letters, Part B: Applications, Vol. 5, No. 2, pp.377-386, April 2014. (査読有)

田中宏典、十川勇人、岩崎宣生、松崎隆哲、白土浩、五反田博：“QAM-OFDM システムのブラインドチャンネル推定”、信号処理、Vol. 18, No. 2, pp. 77-88, March 2014. (査読有)

N. Iwasaki, S. Nishimura, K. Inoue and H. Gotanda: “Frame-wise Estimation for Direction-of-Arrival of Sound Sources,” Proc. the 45th ISCIE Int’l Symp. on Stochastic Systems Theory and Its Application (SSS13), pp.373-380, Aug. 2014. (査読有)

T. Ishibashi, Y. Tajiri, K. Inoue and H. Gotanda: “An Approach to Blind Source Separation Based on Rotation of Joint Distribution of Observed Mixture Signals,” Proc. RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing, pp. 353-356, Nov. 2013. (査読有)

H. Sogo, H. Tanaka, T. Matsuzaki, H. Shiratsuchi and H. Gotanda: “Semi-blind channel estimation for QAM-OFDM,” Proc. the 45th ISCIE Int’l Symp. on Stochastic Systems Theory and Its Application (SSS13), pp.337-345, Nov. 2013. (査読有)

T. Ishibashi, Y. Tajiri, K. Inoue and H. Gotanda: “A noise reduction method using joint distribution of observed signals,” Proc. the 45th ISCIE Int’l Symp. on Stochastic Systems Theory and Its Application (SSS13), pp.367-372, Nov. 2013. (査読有)

岩崎宣生、井上勝裕、五反田博：“DUETに基づくフレーム単位での音源到来方向推定”、電子情報通信学会技術研究報告、Vol.113, No.118, pp.119-124, July 2013. (査読無)

[学会発表](計24件)

杉本尚之、松崎隆哲：“FPGA実験ボード上のオーディオコーデックを用いたA/D・D/A変換時に関する研究”、情報処理学会第78回全国大会、2016/3/10、慶應義塾大学(横浜市)

岩崎宣生、玉置将人、深瀬駿、井上勝裕、五反田博：“フレーム毎のDOA推定に基づく移動音源追尾に関する検討”、第58回自動制御連合講演会、2015/11/14、神戸大学(神戸市)

杉本尚之、松崎隆哲：“FPGA実験ボード上のオーディオコーデックを用いた

A/D・D/A変換時における動作解析及び信号遅延に関する研究”、第68回電気・情報関係学会九州支部連合大会、2015/09/26、福岡大学(福岡市)

樋口佳奈、石橋孝昭、大隈千春、葉山清輝、五反田博：“無音区間を利用した騒音環境下での話者音声の抽出”、第23回電子情報通信学会九州支部学生会講演会、2015/09/04、福岡大学(福岡市)

藤森光平、石橋孝昭、大隈千春、葉山清輝、五反田博：“実環境下での観測信号の同時分布の回転による音源分離”、第23回電子情報通信学会九州支部学生会講演会、2015/09/04、福岡大学(福岡市)

岩崎宣生、玉置将人、深瀬駿、井上勝裕、五反田博：“リアルタイムDOA推定に基づく移動音源追尾に関する検討”、計測自動制御学会ライフエンジニアリング部門シンポジウム、2015/09/03、九州工業大学(飯塚市)

杉本尚之、松崎隆哲：“FPGAボード上を用いたA/D・D/A変換における遅延と位相歪みに関する研究”、近畿大学サイエンスネットワーク第5回院生サミット、2015/07/11、近畿大学(紀の川市)

石橋孝昭、大隈千春、葉山清輝、五反田博：“複数の目的音抽出可能なマイクシステムの開発”、第59回システム制御情報学会研究発表講演会、2015/05/20、中央電気倶楽部(大阪市)

石橋孝昭、大隈千春、葉山清輝、五反田博：“2チャンネルマイクによる任意指向特性の形成”、第33回計測自動制御学会九州支部学術講演会予稿集、2014/12/06、九州工業大学(北九州市)

西村俊哉、岩崎宣生、五反田博：“リアルタイムDOA推定に基づく音源分離に関する検討”、第33回計測自動制御学会九州支部学術講演会予稿集、2014/12/06、九州工業大学(北九州市)

西村俊哉、岩崎宣生、五反田博：“リアルタイムDOA推定とソフトマスクに基づく音源分離と評価”、第22回電子情報通信学会九州支部学生会講演会、2014/09/20、鹿児島大学(鹿児島市)

清田奈那、石橋孝昭、芳野裕樹、五反田博：“独立成分分析を用いた振幅変調信号に対する信号分離の試み”、第22回電子情報通信学会九州支部学生会講演会、2014/09/20、鹿児島大学(鹿児島市)

田代皓太、石橋孝昭、大隈千春、葉山清輝、五反田博：“観測信号の同時分布の回転に基づくブラインド音源分離”、第22回電子情報通信学会九州支部学生会講演会、2014/09/20、鹿児島大学(鹿児島市)

杉本尚之、松崎隆哲：“FPGA実験ボード上でのA/D・D/A変換に関する検討”、第67回電気・情報関係学会九州支部連合大会、2014/09/18、鹿児島大学(鹿児島市)

石橋孝昭、大隈千春、葉山清輝、五反田博：“観測信号の振幅比を利用したブライント音源分離”、第67回電気・情報関係学会九州支部連合大会、2014/09/18、鹿児島大学（鹿児島市）

葉山清輝、石橋孝昭、大隈千春、五反田博：“2マイクのリアルタイム処理による指向特性の実現”、第67回電気・情報関係学会九州支部連合大会、2014/09/18、鹿児島大学（鹿児島市）

大隈千春、石橋孝昭、葉山清輝、五反田博：“プログラミング制御による可変指向性マイクシステムの開発”、第67回電気・情報関係学会九州支部連合大会、2014/09/18、鹿児島大学（鹿児島市）

正木範朗、松崎隆哲：“FPGAボード上における音声データの取り扱いに関する検討”、情報処理学会第76回全国大会、2014/3/12、東京電機大学(東京都足立区)

西村俊哉、岩崎宣生、五反田博：“DOA推定に基づく音源分離に関する検討”、第32回計測自動制御学会九州支部学術講演会、2013/11/30、長崎大（長崎市）

西村俊哉、五反田博：“音声到来方向のリアルタイム推定”、近畿大学サイエンスネットワーク 第4回院生サミット、2013/11/23、近畿大学（飯塚市）

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

近畿大学学術情報リポジトリ

<https://kindai.repo.nii.ac.jp/>

6. 研究組織

(1)研究代表者

五反田 博 (GOTANDA Hiromu)

近畿大学・産業理工学部・教授

研究者番号：10153751

(2)研究分担者

白土 浩 (SHIRATSUCHI Hiroshi)

近畿大学・産業理工学部・准教授

研究者番号：30315460

石橋 孝昭 (ISHIBASHI Taka-aki)

熊本高等専門学校・情報通信エレクトロニクス工学科・准教授

研究者番号：60455178

松崎 隆哲 (MATSUZAKI Takanori)

近畿大学・産業理工学部・准教授

研究者番号：20363385