

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 1 日現在

機関番号：17501

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25350670

研究課題名(和文) 補聴システムの高性能化に向けた話者移動に頑健な残響抑圧処理の研究

研究課題名(英文) Robust speech enhancement against speaker's movements in noisy reverberant environments for hearing aid

研究代表者

古家 賢一 (Furuya, Ken'ichi)

大分大学・工学部・教授

研究者番号：10643611

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：本研究では、補聴システムの性能向上を目指し、まず、画像情報と音響情報を統合した音声強調手法を提案し、画像情報を組み合わせることでより音声強調性能が向上することを確認した。さらに、マルチチャネル非負値因子分解を用いて環境を学習し、音声と雑音・残響とを行列分解アルゴリズムで分離できないか実験的に検討を行った。その結果、初期値の設定法として位相変化に基づく初期値設定法を空間相関行列に与えると有効であることを確認した。

研究成果の概要(英文)：We propose two speech enhancement methods for hearing aid. The first is the method combining image information captured by camera and audio information captured by microphone array. Experiment results show that the performance with the proposed method using microphone array and camera is superior to the system only with microphone array. The second method uses the Multichannel Nonnegative Matrix Factorization (MNMF) for speech enhancement. The problem of MNMF is initial-value dependency for the performance. In experimental results, the initial-value settings based on the phase differences between the microphones was more effective than the random settings

研究分野：音響信号処理

キーワード：マイクロホンアレイ 画像処理 非負値行列因子分解 音響モデル

1. 研究開始当初の背景

(1) 加齢に伴う聴力低下を補う補聴システムにおいて、マイクロホンで収音された音声を劣化させる妨害音として、雑音と部屋の残響がある。雑音については多くの研究がなされ、市販の高性能補聴器に実装されている。しかし、残響は雑音と性質が大きく異なり音声と相関が非常に高く通常の雑音抑圧技術では抑圧できない。これまでに残響による聞き取りにくさの多くの報告がされているにも関わらず残響抑圧処理についてはまだ研究途上であり、急速な高齢化が進む社会においては急務な研究課題である。

(2) 実用化に向け残る大きな課題のひとつは、日常生活における話者位置の移動に対して残響抑圧性能が大きく劣化してしまう点である。現状の方法では、話者が僅か 10 cm 移動するだけで効果がなくなってしまい、改めて残響特性を学習するのに約 60 秒かかる。この課題が解決すれば、話者が動きながら話す日常生活においても、音声強調を実現でき音声の劣化を回復できる。

2. 研究の目的

(1) 本研究では、従来の音声強調で課題であった話者位置の変動に頑健な手法を確立し、その応用として実際の残響・雑音環境下において補聴システム性能を大幅に向上させることを目的とする。

3. 研究の方法

(1) 話者位置を画像情報から推定するためにカメラ・マイクロホンアレイ型デバイスを用いる。構築する残響モデルが妥当かどうかを検証するためのデータとして、通常の会議室、講義室の残響時間 (0.5~1 秒程度) の部屋において、同一話者が移動する場合と複数話者が交代して発話する場合を想定して測定を行う。1m 四方の中心を基準点として、その前後左右を 10 cm 間隔でインパルス応答を複数の条件の部屋において TSP 測定法で測定し、複数話者が交代して発話する場合に対応して、基準点を 3m 以上離れた位置に移動して同様に測定する。部屋の条件に関しても最低 3 つの残響時間の異なる部屋において測定する。画像センサー、距離センサーによる音源位置の推定技術の調査のため、データとしてインパルス応答の測定と同時にカメラによって音源位置の画像も記録しておく。

(2) 最適な基底関数の組み合わせ探索と残響モデルの構築を行う。音声は物理的には音の波であり波動的な性質を持つ。これを利用すれば、話者位置をパラメータとする残響モデルが構築できる。ただし、部屋の形や壁の吸音率などの境界条件は未知であり、また部屋によって異なるため一般的には解析的な波動解を求めることは困難である。したがって、本研究では試行錯誤的に複数の基底関数を選択し、パラメータフィッティングによ

てその中から話者位置パラメータの変化に対して誤差が最も少ない最適な基底関数の組み合わせを探索する。探索された最適な基底関数の組み合わせを用いて残響特性を表現する残響モデルを構築する。音声、残響、雑音を分解するアルゴリズムとしてマルチチャンネル非負値行列因子分解を用いて基底関数による音声の分解の検討をする。

4. 研究成果

(1) カメラ画像から話者位置を検出し雑音・残響を抑圧する MV 法の測定および評価した。残響で音源位置が正確に検出できない場合でも動作するように、カメラと距離センサーを利用して音源位置を検出し、その画像から得られた情報とマイクロホンアレイで得た音情報を統合して雑音・残響抑圧を行った。図 1 にカメラによる画像情報とマイクロホンアレイによる音情報を統合したシステム構成図を示す。実験では、口元にスピーカを埋め込んだダミーヘッドを話者の代わりに用い、カメラが組み込まれたマイクロホンアレイ前方に設置し収録を行った。ここではカメラ一体型マイクロホンアレイとしてマイクロソフト社の Kinect を使用する。まず Kinect のカメラ画像から話者の頭部位置座標を検出し、その検出した頭部位置座標を用いて Kinect から話者の頭部までの方向角度を推定する。次に推定した方向角度を従来の指向性制御手法 (MV 法) に適用して処理を行う。この手法をここでは画像 MV 手法と呼ぶ。MV 法は目的音声に歪を許さない指向性制御法である。

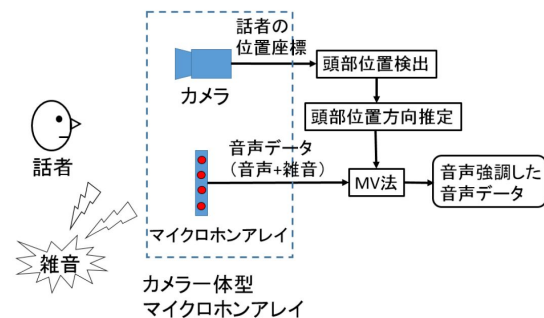


図1 システム構成図

(2) 実験を行い性能を評価した。図 2 に実験での話者と雑音、カメラ一体型マイクロホンアレイ Kinect の配置図を示し、図 3 に実際の実験の様子を示す。Kinect から目的音 (話者) が 0° 方向に存在し、雑音 (ホワイトノイズ) が +30° 方向に存在する場合と、それに加えて雑音 (音声) が 30° 方向に存在する場合を評価した。目的音は Kinect から 1m と 3m の地点、雑音は Kinect から 1m の地点に配置した。目的音源として口の位置にスピーカを内蔵したダミーヘッドを使用した。従来の MV 法については観測信号

から相関法を用いて音源位置を検出した。実験の結果、図4には平均SN比改善量を示しており、エラーバーは10回測定したときの標準偏差である。図から分かるように、音情報のみで音源位置を検出する場合(MV法)に比べ、提案法(画像MV法)の方がより高精度で位置を特定でき、音声強調の性能も向上することが確認できた。

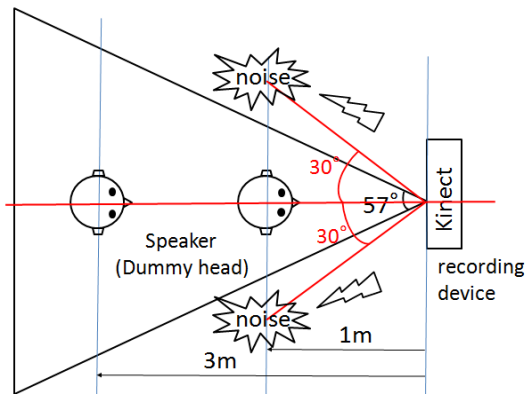


図2 実験配置図



図3 実際の実験風景

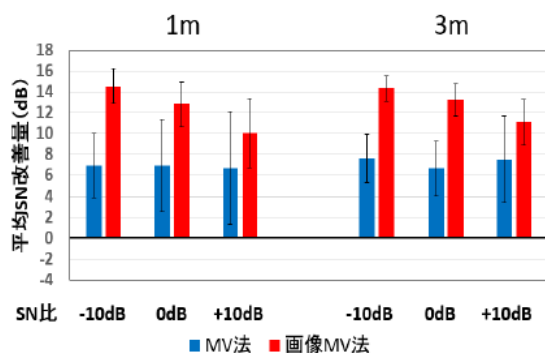


図4 SN比改善の実験結果

(3)次に、非負値行列因子分解を用いた最適な基底関数の反復的計算法における初期値の検討をした。音声信号を原音声、雑音、残響に分解するための基底関数について、分解するアルゴリズムとしてマルチチャンネル非負値行列因子分解(マルチチャンネルNMF)を用いて基底関数による音声の分解の検討を行った。図5は、マルチチャンネルNMFアルゴリズムの処理の流れを示す。マルチチャンネルNMFでの基底関数は、特に初期値に対するばらつきが大きく、適切な初期値の設定が課題であり、様々な初期値について検討した

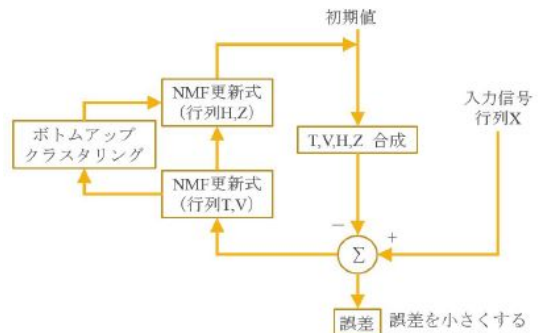


図5 マルチチャンネルNMFアルゴリズム

(4)適切な初期値の設定について、実験的に調査を行った。図6の環境で測定したインパルス応答に畳み込み作成した。この場合、複数の話者が周りで話している特定の話者の音声のみ分離して強調する状況を想定している。インパルス応答長300、サンプリング周波数16kHz、STFTのフレームサイズ1024、シフトサイズ256とし、基底関数 $K=30$ 、音源数 $L=3$ 、分離性能を計算するための更新回数500とした。様々な初期値選択法について実験した結果、位相変化に基づく初期値選択と空間相関行列に基づく初期値選択が性能向上に効果的であることが確認できた。図7は、位相変化に基づく初期設定法の効果を示し、ランダムで初期値を設定するのに比べ、平均的に性能が改善している。

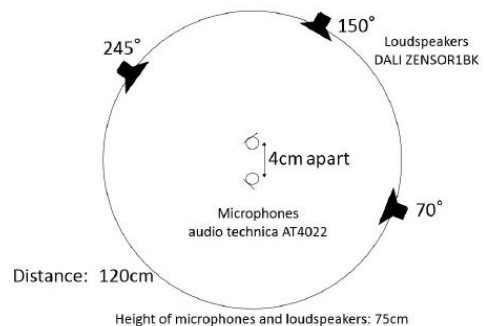


図6 マルチチャンネルNMFの実験環境

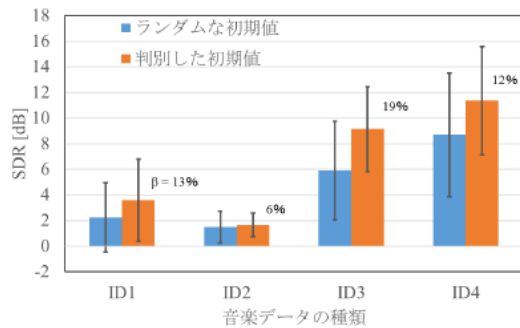


図6 初期値選択法の効果

5. 主な発表論文等

〔雑誌論文〕(計2件)

YOSHIYAMA, F., UENOHARA, S., NISHIJIMA, K., HIOKA, Y. and FURUYA, K., Log-likelihood method to select initial values of multichannel non-negative matrix factorization, 2015 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2015, 2015, pp. 126-127. 査読有

DOI: 10.1109/ICCE-TW.2015.7216813

NISHIJIMA, K., UENOHARA, S. and FURUYA, K., 2015. Snore activity detection using smartphone sensors, IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2015, 2015, pp. 128-129. 査読有

DOI: 10.1109/ICCE-TW.2015.7216814

〔学会発表〕(計13件)

西島 恵介, 上ノ原 進吾, 古家 賢一, SVMを用いたいびき区間検出における学習データ数に関する検討, 日本音響学会, 2016年3月9日, 桐蔭横浜大学(神奈川県横浜市)

黒田 翔馬, 上ノ原 進吾, 西島 恵介, 古家 賢一, 多チャンネルインパルス応答測定における多重M系列信号と多重Log-SS信号の巡回性について, 応用音響研究会, 2015年11月13日, 熊本大学(熊本県熊本市)

工藤 晋也, 西島 恵介, 上ノ原 進吾, 古家 賢一, ウェーブレット変換を用いた心音分析における雑音抑圧のパラメータについて, 日本音響学会, 2015年9月17日, 会津大学(福島県会津若松市)

黒田 翔馬, 上ノ原 進吾, 西島 恵介, 古家 賢一, 多チャンネルインパルス応答測定における多重M系列信号と多重Log-SS信号の比較, 日本音響学会, 2015年9月16日, 会津大学(福島県会津若松市)

西島 恵介, 上ノ原 進吾, 古家 賢一, 雑音抑圧を用いたいびき区間検出の性能改善, 日本音響学会, 2015年9月17日, 会津大学(福島県会津若松市)

吉山 文教, 上ノ原 進吾, 西島 恵介, 古家 賢一, 平面波モデルを用いたマルチチャネル非負値行列因子分解アルゴリズム, 日本音響学会, 2015年3月16日, 中央大学(東京都文京区)

西島 恵介, 上ノ原 進吾, 古家 賢一, 雑音特性の事前学習を用いたいびき区間検出, 日本音響学会, 2015年3月16日, 中央大学(東京都文京区)

K. NISHIJIMA, S. UENOHARA, and K. FURUYA SNORE ACTIVITY DETECTION USING SMARTPHONE IN AMBIENT NOISE ENVIRONMENT, Kyushu-Youngnam Joint Conference on Acoustics 2015, 2015年1月10日, 長崎大学(長崎県長崎市)

F. YOSHIYAMA, S. UENOHARA, K. NISHIJIMA, and K. FURUYA, PHASE DIFFERENCE AND LOG-LIKELIHOOD METHODS TO SELECT INITIAL VALUES FOR MULTICHANNEL NON-NEGATIVE MATRIX FACTORIZATION IN SOUND SOURCE SEPARATION, Kyushu-Youngnam Joint Conference on Acoustics 2015, 2015年1月10日, 長崎大学(長崎県長崎市)

吉山 文教, 上ノ原 進吾, 西島 恵介, 古家 賢一, 音源分離における直交化制約を用いたマルチチャネルNMFの検討, 応用音響研究会, 2014年11月21日, 九州大学(福岡県福岡市)

吉山 文教, 上ノ原 進吾, 西島 恵介, 古家 賢一, マルチチャネル非負値行列因子分解における分離性能の高い初期値の判別法, 日本音響学会, 2014年3月16日, 日本大学(東京都千代田区)

有満 大輝, 上ノ原 進吾, 西島 恵介, 古家 賢一, 画像処理による頭部位置検出を用いた音声強調の評価, 日本音響学会, 2014年3月16日, 日本大学(東京都千代田区)

西島 恵介, 古家 賢一, 兼板 佳孝, 周波数領域におけるいびき音の特徴, 日本音響学会, 2014年3月16日, 日本大学(東京都千代田区)

6. 研究組織

(1) 研究代表者

古家 賢一 (FURUYA, Ken'ichi)

大分大学・工学部・教授

研究者番号: 10643611

(2)連携研究者

西島 恵介 (NISHIJIMA, Keisuke)

大分大学・工学部・助教

研究者番号：30237698