

令和元年5月14日現在

機関番号：57403

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K21583

研究課題名(和文)超狭空間指向性マイクロホンを用いた音声指令による電子機器制御

研究課題名(英文)Control of electronic device by voice using ultra-narrow space directional microphone

研究代表者

石橋 孝昭 (Ishibashi, Takaaki)

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

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：騒音環境下で電子機器を安定して音声制御をさせるための高雑音環境における雑音除去の実用化を目指す。荷物や資料などを持って両手がふさがっているときに電子機器を操作したい状況を想定し、電子機器に取り付けた複数のマイクロホンを用いて、目的とする発話者に対して方向だけでなく距離に対する指向性を強化形成する手法を新規に提案する。周囲雑音の影響を従来のものから大幅に低減し、独自に開発した目的話者音声を短時間処理で安定してリアルタイムに抽出するアルゴリズムを適用する。今回は電動車椅子の利用者を想定し、ジョイスティックの操作なしで音声指令により電動車椅子を操作できる制御システムを構築する。

研究成果の学術的意義や社会的意義

音声による電子機器の制御の利用を考えると、屋内のプライベート空間で使用されると思われる。周囲に他人がいる屋外で、スマートフォンなどの音声認識機能を利用する人を見かけないからである。したがって、発話者とマイクロホンの距離は1mよりも近く、騒音レベルや残響時間も、室内を想定して研究を進めた。そのため、これらの条件に基づいて雑音を抑制する関数を見つける本研究は、これからのIoTに関連した技術である。

研究成果の概要(英文)：It aims at the practical application of the noise reduction in the high noise environment to control the electronic device by voice under a noisy environment. Assume a condition in which you want to operate the electronic device when both hands are closed with luggage or materials. A method is proposed to create a beamform for direction and distance to a target speaker using microphones attached to an electronic device. An algorithm is applied to reduce the influence of ambient noise and extract the target speaker speech originally developed in a short time processing. Assuming a user of the electric wheelchair, the control system which can operate the electric wheelchair by voice command without joystick operation was created.

研究分野：信号処理

キーワード：雑音除去 信号処理

## 様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

### 1. 研究開始当初の背景

本研究を計画したとき、ブラインド信号分離と呼ばれる観測された混合音声から目的とする話者音声の抽出方法に関して、世界中で研究が進められていた。また、研究代表者もこのブラインド信号分離について、伝達関数のゲイン比や位相差を推定して音源の復元法を提案していた。また、ブラインド信号分離では音源の情報未知であるため、音源数も分からない問題がある。従来では、この問題は無視され、マイクロホン数と同じ数の音源があると仮定して解かれていた。この問題に対して、分離信号の相関を求めることで音源数を推定する方法が提案されているものの、分離された信号数以上の音源が存在するときには理論的に推定できない。すなわち、音源数がマイクロホン数よりも多いときに機能しないという問題があった。音源数推定問題に対して研究代表者は、2本のマイクロホンをから得られる観測信号に対して、マイクロホン数以上の音源が存在するときにも音源数推定可能なアルゴリズムを提案し、その推定結果に基づいて音源分離して目的音だけを出力できる方法を考案していた。

このように、ブラインド信号分離の研究は世界中で進められていた。そのため、実用化に向けた取り組みが進められてもよさそうであるが、製品化されたものは、ほとんどない状況であった。エンジン音などを抑制できるノイズキャンセリング機能を持つ製品が多数存在するのに対して、ブラインド信号分離の実用化は困難だと考えられた。その理由に、パラメータの決定のために数秒間の音声データが必要であり、大きなメモリが必要となることや、逐次更新アルゴリズムであるため、収束までに時間を要することが挙げられた。そこで、アルゴリズムの単純化によって実用化が可能となるブラインド信号分離のアルゴリズムの開発が重要であると考えられた。

### 2. 研究の目的

本研究では、騒音環境下で電子機器を安定して音声制御をさせるための高雑音環境における雑音除去の実用化を目指す。高齢者や障害のある人が荷物や資料などを持って両手がふさがっているときに電子機器を操作したい状況を想定し、電子機器に取り付けた複数のマイクロホンを用いて、目的とする発話者に対して方向だけでなく距離に対する指向性を強化形成する手法を新規に提案する。周囲雑音の影響を従来のものから大幅に低減し、独自に開発した目的話者音声を短時間処理で安定してリアルタイムに抽出するアルゴリズムを適用する。今回は特に高雑音環境下での電動車椅子の使用者が荷物を持ってタブレット端末を操作することを想定し、ジョイスティックの操作なしで音声指令により電動車椅子を操作できる制御システム機構を実現する。

### 3. 研究の方法

音声は8kHzのサンプリング周波数でサンプリングされることが多い。16kHzでのサンプリングされることもあり、音響信号は44.1kHzでサンプリングされる。本研究の音声信号処理を考えたところ、処理コストを減らすために8kHzサンプリングが望ましい。高いサンプリング周波数と比較すると情報は低下する。すなわち、処理単位フレーム長については、8kHzサンプリングの場合、10ミリ秒のデータで処理すると考えると80ポイントしか存在しないことになる。しかしながら、本研究で用いる統計データはビッグデータを必要としないために、十分なデータ量であると考えられる。また、データ量を必要とするために、過去に取得したデータの利用も考えなければならないが、時間分解能が低下する。そこで、50~70dBの騒音下で、サンプリング周波数と処理単位フレーム長の組合せを変えて分離性能を評価する。

マイコンへのモノラル音声の入出力と処理はこれまでに終えている。本研究では、2チャンネル以上の複数の音声を扱うため、複数のマイクロホンの入力を同時に処理できるよう、プログラムを書き換える。同時に、入力のための周辺回路の素子を選定する。その後、パソコンで開発した雑音抑制アルゴリズムをマイコンに実装し、マイクの数や配置に対する雑音抑制能力を調査する。さらに、音声による制御が可能な電動車椅子を開発する。具体的には、マイクに向かって発声した音声を本研究で作製した機器で雑音を除去し、既存の音声認識で認識した後、その言葉の通りに移動する車椅子のプロトタイプを作る。

### 4. 研究成果

(1) 可変指向性パターンを形成することで雑音を抑制して目的音声を強調する方法を開発した。複数の原信号が存在するとき二つのマイクロホンで観測された混合信号に対して、マイクロホン1で観測された信号の振幅をグラフの横軸に、マイクロホン2で観測された信号の振幅をグラフの縦軸としてプロットすると同時分布を作ることができる。この同時分布は、それぞれの音源が単体で存在するときに現れる同時分布の直線が重なってできる。そのため、音源数と同じ数の直線成分を持つ。この分布から雑音成分を削除して目的音声の成分だけを強調する。すなわち、同時分布を利用して目的音声を抽出することは、雑音成分が作る直線成分を削除して、目的信号の成分が作る直線成分を抽出することを意味している。その事実に基づいて、目的音声の方向に対する分布に空間的な指向性を与えることで目的音声を抽出することが可能となった。この指向性は、プログラムで指向性の方向や幅を変更することができる。また、このアルゴリズムは、1つの目的音声に対して1つの方向の指向性を持たせることもできるし、複数の音源に対して複数の指向性を持たせることも可能であるという特徴を持っている。さら

に、目的方向に指向性を向けることもできるし、雑音方向に死角を持たせることもできるので、歪みの少ない目的音声を抽出することを可能とした方法である。

(2) 同時分布を回転させることで雑音を削除して目的音声だけを抽出する方法を開発した。目的音声と雑音の二つが存在するとき作成された同時分布に対して、雑音成分の振幅が0となるように分布を回転させることで、雑音の無い目的音声を出力させるアルゴリズムを開発した。この方法は、データ数があまり多くなくても機能する特徴を持っていて、抽出される目的音声の歪みも少ない。また、短時間フレーム処理を可能としており、現在の取得データだけでは分布の回転による目的音抽出が不可能となる場合、過去のデータを利用して相関や尖度を有効な分離データを持っているかどうかの判定尺度として用いることで、処理フレーム長が0.01秒となっても話者音声抽出できることを確認した。このとき、処理フレームごとに目的音声と雑音の出力チャンネルが変化しないことや、目的音声の振幅の変動を無くしてフレーム処理後に自然な音声を復元できることも確認できた。

(3) 電動車椅子は歩行の不自由なときの移動手段として多くの場面で利用されており、その高機能化に関する研究開発を行った。電動車椅子を操作する方法には、ジョイスティック操作やハンドル操作などが利用される。一般的な状況であれば問題ないものの、荷物を運ぶときや作業をしているときに、音声によって電動車椅子を操作したいという要求もある。音声認識技術は急速な進歩によって実用化が進んでおり、スマートフォンやPCを搭載した音声認識機能付きの電動車椅子の検討も進められている。スマートフォンやPCを用いた音声認識は、認識率が高いものの、装置自体が高価になる。さらに、機器の起動に時間がかかる問題がある。スマートフォンやPCはOSで動作するため、OSが起動しソフトウェアが動作を開始するまでには数分の時間を要する。電動車椅子を制御するシステムが起動するまで移動ができないことは大きな問題となる。そこで本研究では、マイコンを利用した音声認識によって電動車椅子を制御できるシステムを開発した。具体的な処理は、これまでの目的音声抽出技術をマイコンに実装して、マイコンと連携可能な市販されている音声認識機を用いて音声認識した結果から、従来のジョイスティックによって作り出していた電圧変化をマイコンによって作成し、その作成した電圧をモータードライバに伝える。本システムは安価で手軽に実装でき、容易にセンサや他の装置と連携できる特長を持っている。また、使用者に合わせてカスタマイズ可能である。

## 5. 主な発表論文等

### 〔雑誌論文〕(計3件)

C. Okuma, K. Hayama and T. Ishibashi, "Two-channel microphone system with variable arbitrary directional pattern," ICIC Express Letters, Vol. 12, No. 3, pp. 229-236, 2018.

H. Shintani, Y. Oshiro, K. Nagata and T. Ishibashi, "Low-cost and Small-size EMG Signal Measurement Device for Electric Wheelchair Control," ICIC Express Letters, Part B: Applications, Vol. 9, No. 1, pp. 31-35, 2018.

T. Ishibashi, H. Shintani and K. Nagata, "Fast Blind Source Separation and Target Human Speech Extraction Method for Acoustic Signals," ICIC Express Letters, Vol. 11, No. 12, pp. 1715-1721, 2017.

### 〔学会発表〕(計17件)

M. Iikawa and T. Ishibashi, "Blind Source Separation Based on Rotation of Joint Distribution Without Inversion of Positive and Negative Sign," Proceedings of the 7th IIAE International Conference on Industrial Application Engineering 2019, pp. 229-232, 2019.

Y. Kawakami and T. Ishibashi, "Estimation for the Number of Sound Sources Using Directional Histogram of Observed Mixture Signals," Proceedings of the 7th IIAE International Conference on Industrial Application Engineering 2019, pp. 225-228, 2019.

Y. Oshiro, T. Ishibashi, K. Hayama, C. Okuma and K. Eguchi, "Extraction of Multiple Sound Signals using Two-Channel Microphone System with Variable Arbitrary Multi-Directional Pattern," 2019 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP'19), pp. 268-271, 2019.

飯川聖人, 石橋孝昭, "同時分布の回転による音源分離とスケーリングの解消," 第26回電子情報通信学会九州支部学生会講演会論文集, D-19, 2018.

川上雄大, 石橋孝昭, "観測信号の方位ヒストグラムを用いた音源数の推定," 第26回電子情報通信学会九州支部学生会講演会論文集, D-18, 2018.

井上公友, 石橋孝昭, 大隈千春, 江口啓, "任意可変指向性パターン形成法に基づく複数の目的音声の抽出," 第26回電子情報通信学会九州支部学生会講演会論文集, D-3, 2018.

T. Ishibashi, K. Inoue, C. Okuma and K. Eguchi, "Low Distortion Target Speech

Extraction Method Using Two-Channel Microphone System, ” The 6th IIAE International Conference on Intelligent Systems and Image Processing 2018, pp. 428-431, 2018.

井上公友, 石橋孝昭, 大隈千春, 江口啓, “ 低歪み目的音声抽出マイクロフォンシステムの開発, ” 産業応用工学会全国大会 2018, pp. 95-96, 2018.

樋口佳奈, 大隈千春, 石橋孝昭, “ 短時間フレーム処理を利用した 3 音源に対する音源分離, ” 第 25 回電子情報通信学会九州支部学生会講演会論文集, D-36, 2017.

K. Higuchi, C. Okuma and T. Ishibashi, “ Blind Source Separation and Human Speech Extraction for Three Sound Sources Using Silent Interval, ” The 5th IIAE International Conference on Intelligent Systems and Image Processing 2017, pp. 205-208, 2017.

C. Okuma, K. Hayama and T. Ishibashi, “ Two-Channel Microphone System with Variable Arbitrary Directional Pattern, ” Proceedings of 12th International Conference on Innovative Computing, Information and Control (ICICIC2017), C5-6, 2017.

K. Higuchi, C. Okuma and T. Ishibashi, “ Silent Interval Detection and Target Human Speech Extraction, ” Proceedings of 12th International Conference on Innovative Computing, Information and Control (ICICIC2017), C5-5, 2017.

T. Ishibashi, S. Nakao and K. Nagata, “ Blind Source Separation and Target Speech Extraction for Acoustic Signals, ” Proceedings of 12th International Conference on Innovative Computing, Information and Control (ICICIC2017), C5-4, 2017.

T. Ishibashi, K. Higuchi, C. Okuma and K. Hayama, “ Noise reduction method for multiple sound source signals and its application to two-channel microphone system, ” Proceedings of the 5th IIAE International Conference on Industrial Application Engineering 2017, pp. 336-339, 2017.

K. Higuchi, C. Okuma and T. Ishibashi, “ Target sound extraction using silent interval of human speech, ” International Symposium on Artificial Life and Robotics, AROB 22nd 2017, pp. 900-903, 2017.

樋口佳奈, 大隈千春, 石橋孝昭, “ 騒音の同時分布を利用した話者音声の抽出, ” 第 24 回電子情報通信学会九州支部学生会講演会論文集, D-39, 2016.

樋口佳奈, 大隈千春, 石橋孝昭, “ 観測信号の相関係数に基づく無音区間を利用した騒音環境下での話者音声抽出, ” 産業応用工学会全国大会 2016, pp. 46-47, 2016.

〔その他〕

ホームページ

<https://kumamoto-nct.ac.jp/gyouseki/1000029.html>

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。