

令和元年6月11日現在

機関番号：34315

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16094

研究課題名(和文)音環境推定に基づく危機検知システムの開発

研究課題名(英文) System development of critical situation detection based on sound environment estimation

研究代表者

福森 隆寛 (Fukumori, Takahiro)

立命館大学・情報理工学部・助教

研究者番号：60755817

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：本研究課題では、家庭内の危機的状況を検知するために、音環境をリアルタイムで自動推定できる(音声、生活音、異常音を理解できる)アルゴリズム、および音響システムを開発した。このシステムは、はじめにマイクロホンに入力された音から人間の聴覚特性と相性が良いとされている音声・音響特徴量を抽出する。そして、この特徴量に基づいて構築したディープニューラルネットワークを使って、危機的状況の度合いを評価し、その結果をコンピュータのディスプレイに表示する機能を有している。

研究成果の学術的意義や社会的意義

本研究は、叫び声、異常音、報知音を危機として捉えて、音環境を推定できる音響システムを開発しており、高齢者にとどまらず、全ての単身生活者を犯罪や災害などの危機的状況から守る一助となる。特に核家族や単身生活を想定すると、単身生活者と離れて暮らす家族が、本人の生活状況を遠隔地から随時確認することが可能となり、安全安心な生活基盤の形成に貢献できる。さらに、危機的状況を検知することで重大な犯罪を未然に防ぐ効果も期待できる。

研究成果の概要(英文)：In this study, for detecting critical situation in the house, I have developed an algorithm and an acoustic system that automatically estimate acoustic environments (recognize speech, daily-life sound and abnormal sound) in real time. This system firstly extracts acoustic features that well represent human auditory characteristics from a sound measured by a microphone. After that, the critical degree is evaluated by a deep neural network that is trained with these acoustic features. Then, the result is shown on the computer display.

研究分野：音響情報処理

キーワード：音環境推定 危機検知

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

1. 研究開始当初の背景

少子高齢化・核家族化が急進している日本では、高齢者の日常生活を支援できるロボットを家庭内へ導入する動きがある。これまでは家事・雑事・移動などの生活支援を中心とする研究が取り組まれていた(例:トヨタ自動車株式会社 パートナーロボット)。しかし、核家族や単身生活を想定すると、このようなロボットは犯罪や災害などの危機的状況から生活者を守る役割も担わなければならない。特に高齢者の聴力は加齢によって低下するため、音環境を理解できる範囲は極めて限定的となるため、不法侵入者や火災報知音などに気付かず、高齢者が重大な犯罪や災害に巻き込まれる危険性がある。この課題を解決するために、本研究では図1のように聴力が低下した高齢者に代わって、音環境を推定できる(音声、生活音、異常音を理解できる)ロボット聴覚の開発に挑戦した。

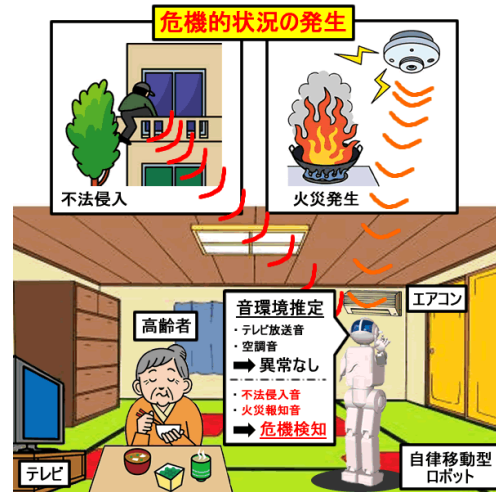


図1：音環境推定可能なロボット聴覚

2. 研究の目的

昨今の日本では、家庭用ロボットが国民の日常生活を支援するための取り組みが推進されている。このようなロボットは、核家族や単身生活を想定した場合、家事・移動などの生活支援に加えて、単身生活者を犯罪や災害などの危機的状況から守る役割も担う必要がある。そこで家庭内の危機的状況を検知するために、音環境をリアルタイムで自動推定できる(音声、生活音、異常音を理解できる)ロボット聴覚を開発することを目的として本研究課題を遂行した。この研究によって、単身生活者(特に高齢者)を危機的状況から守ることが可能となり、安全安心な生活基盤の形成に大きく貢献できる。

音環境推定に基づくロボット聴覚の開発に向けて、研究代表者は研究期間内に以下3点を明らかにした。そして、これら3点の技術的要素を用いて、音環境に存在する音声・生活音・異常音をリアルタイムで自動推定できるロボット聴覚を開発し、単身生活者(特に高齢者)を危機的状況から守るための基盤技術を確立させることを目指した。

(1) 家庭内に存在する不要雑音の除去

正確な音環境推定を実現するにあたって、家庭内に存在する不要な雑音を事前に除去する必要がある。これまでに研究代表者は、工場やオフィスなどに存在する雑音成分から自動推定した最適な雑音抑圧パラメータを用いて雑音を正確に抑圧する技術を開発した実績を有している。本研究では、この技術の枠組みを応用して、家庭内騒音の周波数特性や定常性などから最適な雑音抑圧パラメータを特定し、そのパラメータを用いて不要な家庭内騒音の抑圧を行った。

(2) 音環境推定に適切な音響特徴量の特定

家庭内における音情報(音声・生活音・異常音)の推定に用いる音響特徴量を明らかにした。具体的には、周波数特徴量(メルケプストラム係数など)や時間特徴量(音の大きさ、定常性など)を組み合わせながら音環境推定実験を実施して、それぞれの音情報に対して高い推定性能を達成できる音響特徴量を特定した。

(3) 音情報の構造化と危機的状況の検知

(2)で明らかにした音響特徴量を用いて推定した音情報(音声・生活音・異常音)に基づいて、それぞれの音情報を詳細に構造化する(例:音声を平静音声や叫び声に分類する)技術開発に挑戦した。さらに構造化された音情報を用いて、その音情報の危機的要因の有無を検知できるシステムを構築した。

3. 研究の方法

本研究では、家庭内の危機的状況を検知することを目指して、音環境をリアルタイムで自動推定できるロボット聴覚を開発した。具体的には、本研究を以下に示す4つのサブテーマに分けて研究を実施した(1)-(3)で音環境推定から危機的状況の検知までの一連のアルゴリズムを開発し、(4)でこれらの研究成果を統合したロボット聴覚のプロトタイプを開発した。

(1) 研究1：家庭内に存在する不要な雑音成分の除去

利用環境において正確な音情報の推定を行うために、事前に不要な雑音成分を除去する手法を確立した。研究代表者は、これまでに工場やオフィスなどに存在する雑音を自動推定し、その雑音に最も適切なパラメータで雑音除去できる改良型スペクトルサブトラクション法を確立した。本研究では、この技術と同じ枠組みで家庭内騒音の周波数特性や定常性を解析して最適な雑音抑圧パラメータを特定し、そのパラメータを用いて不要な家庭内騒音の抑圧を行った。

(2) 研究 2：音環境推定に適切な音響特徴量の特定

音声・生活音・異常音、それぞれの音情報を適切に表現できる音響特徴量を明らかにした。具体的には、これまで提案されている音響特徴量（メルケプストラム係数や音のパワーや定常性など）を用いて音環境推定実験を実施して、それぞれの音情報に対して高い推定性能を達成できる音響特徴量を特定した。研究代表者は、これまで利用環境に存在する外乱成分（雑音や残響など）を推定することで外乱に頑健な音声認識システムを構築した実績を有している。本研究の枠組みを音声・生活音・異常音などへ展開して、高精度な音情報の推定に挑戦した。

(3) 研究 3：音情報の構造化と危機的状況の検知

音情報（音声・生活音・異常音）の推定結果に基づき、それぞれの音情報を詳細に構造化できる（具体例：DNN などを用いて音声を平静音声や叫び声に分類する）技術開発を行った。さらに事前に主観的に定義した音情報と危機的状況の対応関係に基づいて、構造化された音情報の危機的要因の有無を判定する。これまでに研究代表者は環境音を用いた音場の構造化に関する研究実績を有しており、この研究成果を音声・生活音・報知音にも応用させた。

(4) 研究 4：音環境推定に基づくロボット聴覚の開発

(1)-(3)の研究成果を統合し、雑音除去、音環境推定、危機的状況検知の一連の処理をリアルタイムで実現可能なロボット聴覚のプロトタイプを開発した。

4. 研究成果

下記の通り、本研究課題の各年度の研究成果を説明する。

(1) 平成 28 年度

「3. 研究の方法」に示した「(1) 研究 1：家庭内に存在する不要な雑音成分の除去」と「(2) 研究 2：音環境推定に適切な音声・音響特徴量の特定」に取り組んだ。研究 1 では不要な雑音（室内の騒音や壁・床からの反射など）を除去するためのアルゴリズムを開発した。具体的には「騒音環境で発話された音声の特徴量（特に位相成分）を雑音が含まれない音声の特徴量に変換することで、観測音声に含まれる騒音を除去する手法」と「人間の声の高さに相当する基本周波数を用いて室内で発話された音声から床や壁の反響成分を取り除く手法」のアルゴリズムを開発し、評価実験を通してこれらの有効性を確認した。研究 2 では音環境推定の中でも音声の発話様式に着目し、平静音声と危機的音声を表す叫び声を切り分けるアルゴリズムを開発した。具体的には、喉元の声帯と声道の運動が平静音声と叫び声で異なることを明らかにし、この知見に基づいて叫び声と平静音声を切り分けるための音声特徴量を提案した。評価実験の結果、提案した音声特徴量の有効性を確認できただけでなく、この特徴量が騒音にも頑健であることを示すことができた。

(2) 平成 29 年度

「3. 研究の方法」に示した「(3) 研究 3：音情報の構造化と危機的状況の検知」と「(4) 研究 4：音環境推定に適切な音声・音響特徴量の特定」の一部に取り組んだ。研究 3 では、生活環境音の中でも特に室内で発生した音の内容を認識するためのアルゴリズムを開発した。具体的には、観測された環境音から人間の聴覚特性との相性が良いとされている音響特徴量（メルフィルタバンク出力）を抽出し、それらを用いて深層学習に基づき、環境音を自動識別するためのディープニューラルネットワークを構築した。そして、評価実験を通してこのアルゴリズムの有効性を確認した。加えて、研究 4 にも一部関係するが、危機的状況を表す叫び声をリアルタイムで検知できる音響システムを開発した。このシステムには、マイクロホンに入力された音声を研究 2 で明らかにした音声・音響特徴量に基づいて叫び声と平静音声のいずれかに分類し、さらに分類時に用いる統計的情報から叫び声らしさを評定する機能を有している。

(3) 平成 30 年度

「3. 研究の方法」に示した「(4) 研究 4：音環境推定に基づく危機検知システムの開発」を重点的に取り組んだ。具体的には、コンピュータ上で危機的状況をリアルタイムで検知できる音響システムを開発した。このシステムは、マイクロホンに入力された音を研究 2 で特定した音声・音響特徴量に基づいて危機的状況の度合いを評定し、その結果をコンピュータのディスプレイに表示する機能を有している。また、この研究 4 で開発したシステムの運用試験を進める中で、危機的状況の推定精度がマイクロホンと音源の位置関係に依存することが判明した。そこで、本最終年度において、音が発生する環境（特に発生した音がどれほど認識しにくい）を自動的に理解するために、危機的状況の推定精度を入力音から推定するアルゴリズムを開発した。具体的には、様々な環境で発生した大量の音データを使って、入力音から認識率を予測するディープニューラルネットワークを構築した。これにより、認識しにくい環境で発生した音に対して、事前に認識精度を改善するための信号処理を適用できるようになったため、より頑健に動作する危機検知システムを開発することができた。

5. 主な発表論文等

〔雑誌論文〕(計1件)

- [1] Yukoh Wakabayashi, Takahiro Fukumori, Masato Nakayama, Takanobu Nishiura, and Yoich Yamashita, "Single-Channel Speech Enhancement with Phase Reconstruction Based on Phase Distortion Averaging," IEEE Transactions on Audio, Speech and Language Processing, vol. 26, no. 9, pp. 1559-1569, 2018, DOI:10.1109/TASLP.2018.2831632. (査読有)

〔学会発表〕(計16件)

- [1] 福森 隆寛, 西浦 敬信, "残響音声と深層ニューラルネットワークを用いた音声認識性能予測の検討," 日本音響学会 2019 年春季研究発表会, 2019.
- [2] 張 宇涛, 若林 佑幸, 福森 隆寛, 西浦 敬信, "Training data proliferation for indoor-environmental sound classification based on autoencoder," 日本音響学会 2019 年春季研究発表会, 2019.
- [3] 福森 隆寛, 西浦 敬信, "インパルス応答を用いた深層ニューラルネットワークによる残響下音声認識性能予測," 日本音響学会 2018 年秋季研究発表会, 2018.
- [4] Yutao Zhang, Sakiko Mishima, Yukoh Wakabayashi, Takahiro Fukumori, Keisuke Imoto and Takanobu Nishiura, "Unknown Sound Clustering for Indoor Environmental Sound Classification Based on Self-Generated Acoustic Model," Proc. the 13th Western Pacific Commission for Acoustics, 2018.
- [5] 美島 咲子, 若林 佑幸, 福森 隆寛, 井本 桂右, 中山 雅人, 西浦 敬信, "自動生成型音響モデルに基づく室内環境音識別のための未知音源クラスタリングの性能評価," 電子情報通信学会技術研究報告, 2018.
- [6] 福森 隆寛, 中山 雅人, 西浦 敬信, 南條 浩輝, "音声強調処理を用いた Rahmonic とメルケプストラムに基づく雑音下叫び声検出," 日本音響学会 2018 年春季研究発表会, 2018.
- [7] 美島 咲子, 若林 佑幸, 福森 隆寛, 中山 雅人, 井本 桂右, 西浦 敬信, "室内環境音識別における教師なしクラスタリングに基づく自動生成型音響モデルの検討," 日本音響学会 2018 年春季研究発表会, 2018.
- [8] 福森 隆寛, 中山 雅人, 西浦 敬信, 南條 浩輝, "Rahmonic とメルケプストラムを用いた深層ニューラルネットワークによる叫び声検出の検討," 日本音響学会 2017 年春季研究発表会, 2017.
- [9] 福森 隆寛, 中山 雅人, 西浦 敬信, 南條 浩輝, "Rahmonic とメルケプストラムに基づく雑音・残響下叫び声検出の音響モデル評価," 日本音響学会 2017 年秋季研究発表会, 2017.
- [10] 美島 咲子, 若林 佑幸, 福森 隆寛, 中山 雅人, 西浦 敬信, "非圧縮音響特徴量を用いた深層ニューラルネットワークによる室内環境音識別の性能評価," 日本音響学会 2017 年秋季研究発表会, 2017.
- [11] 福森 隆寛, 中山 雅人, 西浦 敬信, 南條 浩輝, "雑音・残響下における Rahmonic とメルケプストラムを用いた叫び声検出," 電子情報通信学会技術研究報告, 2017.
- [12] Sakiko Mishima, Yukoh Wakabayashi, Takahiro Fukumori, Masato Nakayama and Takanobu Nishiura, "Investigations on Raw Features in Deep Neural Network for Indoor-environmental Sound Classification," Proc. the 46th International Congress on Noise Control Engineering, 2017.
- [13] 太田 陸斗, 若林 佑幸, 福森 隆寛, 中山 雅人, 西浦 敬信, "調波構造に基づく Deep Auto Encoder を用いた残響下音声強調," 電子情報通信学会技術研究報告, 2017.

- [14] 福森 隆寛, 中山 雅人, 西浦 敬信, 南條 浩輝, ``Rahmonicとメルケプストラムを用いた音響モデルに基づく騒音環境下叫び声検出の性能評価,' ' 電子情報通信学会技術研究報告, 2017.
- [15] Yukoh Wakabayashi, Takahiro Fukumori, Masato Nakayama, Takanobu Nishiura, and Yoichi Yamashita, ``Phase Reconstruction Method Based on Time-frequency Domain Harmonic Structure for Speech Enhancement,' ' Proc. International Conference on Acoustics, Speech and Signal Processing, 2017.
- [16] Takahiro Fukumori, Masato Nakayama, Takanobu Nishiura, and Hiroaki Nanjo, ``Shouted Speech Detection Using Hidden Markov Model with Rahmonic and Mel-frequency Cepstrum Coefficients,' ' Proc. the 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, 2016.

〔その他〕

立命館大学 音情報処理研究室 <http://www.aspl.is.ritsumei.ac.jp/>