

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 31 日現在

機関番号： 82626
 研究種目： 若手研究 B
 研究期間： 2010～2012
 課題番号： 22700192
 研究課題名（和文） 音圧差検出と認識の双方向処理に基づく移動ロボットに適した音環境理解の研究
 研究課題名（英文） Two-way approach robot audition using intensity based detection and spectral feature identification
 研究代表者：
 佐々木 洋子 (SASAKI YOKO)
 独立行政法人産業技術総合研究所・デジタルヒューマン工学研究センター・研究員
 研究者番号： 00574013

研究成果の概要（和文）：

走行中のロボットが「知っている音には気づける」機能の実現を目指した 3 年間のプロジェクトである。初年度は方位角・仰角の 2 次元で音を定位するために、ロボットに搭載可能なサイズで全方位に誤差 6[deg]以下で音源定位可能な球形マイクロホンアレイを開発した。2 年目は、一連の音響信号からどの部分が何の音かを推定する音響ダイアライゼーション技術を開発した。最終年度は、これらの定位・分離・認識技術を統合し、「どちらから何の音がするか」、実環境中で周囲の音を理解する機能を実現した。

研究成果の概要（英文）：

The project developed: 1) sphere microphone array to localize and separate surrounding audio events in azimuth and elevation, 2) statistical method to classify given audio data into known class or recognize them as unknown class.

Then, the combined system of above localization, separation and identification functions achieved recognizing surrounding auditory scene as for a robot audition.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,500,000	450,000	1,950,000
2011年度	900,000	270,000	1,170,000
2012年度	700,000	210,000	910,000
年度			
年度			
総計	3,100,000	930,000	4,030,000

研究分野： 総合領域

科研費の分科・細目： 情報学・知覚情報処理、知能ロボティクス

キーワード： 音声情報処理、マイクロホンアレイ、環境音識別

1. 研究開始当初の背景

雑音の中から必要な音を聞き分け、また複数の音を同時に認識するためのロボット聴覚の研究が盛んになってきたが、賑やかな環

境の中で呼びかけに応える、走行中に遠くの物音に気づく、といった複雑な環境にはまだ対応できていない。

2. 研究の目的

音は環境変化の影響を受けやすく、移動によって刻々と変化する音響条件や、音源ごと、距離ごとの音圧変化など非定常な要素が多い。本研究では、広い範囲を動くロボットが刻々と変化する音響条件下でよりロバストかつ高精度に音情報を知覚するシステムの実現を目指し、(1) 動的に変化する非定常なノイズに対応でき、(2) 知っている音には小さくても気づける、機能を備えた移動ロボット聴覚の実現を目的とする。

3. 研究の方法

(1) 動的に音響環境が変化する移動ロボットから周囲の音を定位・分離可能なマイクロホンアレイを開発し、(2) 様々な条件が想定される実環境に対応可能な、音の識別方法を構築する。これらの定位・分離・認識技術を統合し、移動ロボットが周囲の音を理解する機能を実現する。

(1) 全方位に高感度なマイクロホンアレイの開発：

マイクロホンアレイの設計では、遅延和ビームフォーミング(Delay and Sum Beam Forming, DSBF)の指向特性に対する評価指標を定義し、シミュレーション上で任意のマイクロホン配置に対する指向特性の定量評価を行う。DSBFの問題点として、メインローブ幅が広いこと、およびサイドローブが多いことが知られているが、この指向特性はマイクロホンの配置によって大きく異なる。シミュレーションから、ロボットに搭載可能で高い指向特性を持つマイクロホンアレイを設計する。

(2) 実環境音分類手法の確立：

実環境中の様々な音を認識するために、音源数や各音を表現する次元数など事前情報を最少にし、これまでに学習していない音を知らない音と判断可能な音源識別器を構築する。事前知識の最少化、未知の音の検出を可能とするアプローチとして、ノンパラメトリックベイズを導入する。観測データから複雑さの異なる様々な音を一度に学習可能なモデル生成法を構築する。

4. 研究成果

(1) 車輪型ロボットに搭載可能な直径 350mm、64チャンネルの球形マイクロホンアレイ：

ビームフォーミング時の指向特性を定量評価(図 1)することで最適なマイクロホン配置を設計した(図 2. 左)。方位角・仰角の2次元方向に対し全方位に高感度な指向特性を持ち、走行中のロボットから複数の音源を高精度に定位・分離可能である。開発

したマイクロホンアレイは、サイドローブの低減を実現し、全方位に高感度な特性を得られる。同外径の円周配置アレイと比較すると、1kHzにおいて平均7dB、形成したビームに対して180°方向で15dBのS/N比向上を実現した。

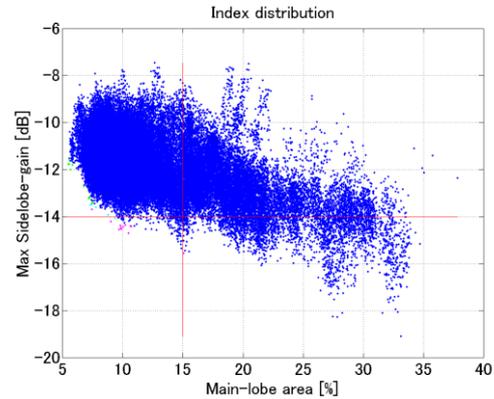


図 1. 異なるマイクロホン配置に対する指向特性の定量評価結果(横軸：メインローブ幅、縦軸：サイドローブゲイン)左下ほど高性能. ピンクの点が最適配置候補

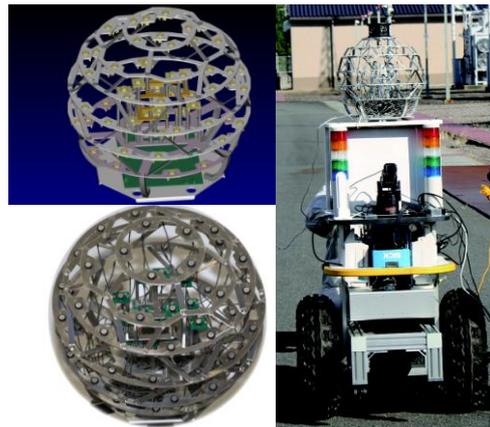


図 2. 開発した球形マイクロホンアレイとアレイを搭載した車輪移動ロボット

開発したアレイを屋外自律走行ロボット Segway RMT200 に搭載し(図 1. 右)、移動しながら同時に2音源を誤差6deg以下で定位可能であることを確認した。変電所内での走行実験結果を図 3 に示す。通路左側に低音を発する装置がある環境で、評価用に通路右側に音源位置の真値がわかるスピーカを設置して、ロボットを直進走行させた。2つのスピーカからは女声および音楽を流した。ロボットの走行速度は約1.3m/sである。スピーカ前を通過する20-40secあたりで2つのスピーカの定位結果が真値とよく一致しており、走行しながら複数の音源を同時に定位可能であることがわかる。

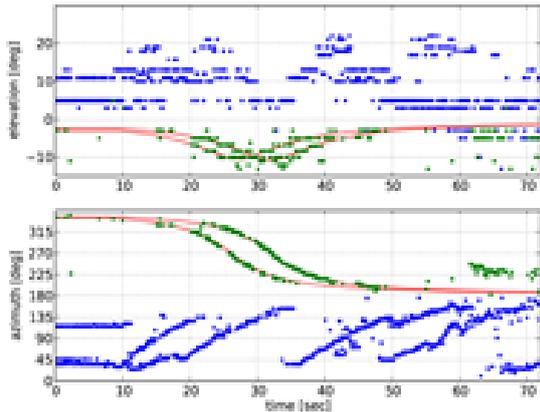


図 3. 走行中の音源定位結果
 (上段：仰角、下段：方位角)
 (赤線：スピーカ位置の真値、青点左側の定位結果、緑点：右側の定位結果)

(2) 一連の音響信号からどの部分が何の音かを推定する音響ダイアライゼーション技術：

日常環境中の音を識別するひとつのアプローチとして、ノンパラメトリックベイズを用いた音響ダイアライゼーションシステムを構築した。提案法は、実環境での様々な音に対応するために、音のモデル化に必要な事前情報を最小限にし、入力音に合わせて柔軟なモデルを構築できることが特長である。具体的には、モデルの次元数や音の種類数を入力に合わせて自動生成し、過去の学習結果にない未知の音が入力された場合は新たな音クラスとしてモデルを生成する。

複数人が会話するラジオ番組を用いた評価実験では、話者数未知の条件でも、事前に話者数を与えて話者識別を行った場合に近い精度で識別可能であることを確認した。

また実環境で録音したデータに対しても、車の走行音、人の声、扉の開閉音といった音の種類ごとに自動識別可能であることを確認した(図 4.)。本実験では、観測した音響信号に対し一部音の種類ラベルを付与し(上

段：色つき部分)、モデル生成を行った。結果は正解ラベルを付与した 3 クラスに加え、新たに 3 クラス、計 6 クラスに分類された。新クラスのひとつめ(new1)には、大通りに出る手前で遠くから聞こえる車の走行音が主に分類された。二番目(new2)に分類された 5 分 40 秒あたりと 7 分すぎの部分は、車の走行音がなく、アスファルト上を走る自転車のロードノイズが主な音源であった。三番目(new3)には、すれ違う人の話し声(女性)や、自転車が段差を越える際の金属音といった比較的高い音が含まれた。正解ラベルのない新しい音について、ほぼ適切に分類できているといえる。

未知音の検出性能評価として、7 種類の楽器音および、話し声、拍手の 9 種類の音源に対し、音源識別実験を行った。図 5. は音源識別率の評価結果である。既存手法との比較として、左端に混合数を固定し EM アルゴリズムで学習させた GMM による最尤推定結果を示す(赤)。各音源の混合数は 32 次元とし、パラメータの学習には HTK を用いた。実効的な混合数が自動調節される提案法に比べ、低い識別率となった。すべての音源を学習データに含む expA(青)では正解ラベルを持つデータの割合に関わらず高い正解率が得られており、各クラスの正解ラベルを 70% マスクした場合でも正解率 98.0% となった。学習データにない未知の音に対する識別実験である expB については、既知のクラスが完全に正解ラベルを持つ場合に正解率 73.5%、既知クラスのラベルを一部マスクした場合は 65% 前後の正解率となった。本実験を通して、既知の音に対する識別では正解ラベルが付与された割合によらず高い正解率が得られた。また正解ラベルを与えていない未知音に対しては、既知の似た音に分類されることもある一方で、既知の特徴量分布とは離れた音響イベントを新クラスと識別可能であることを確認した。時系列情報を含まない音響特徴量で可能なことを示した。

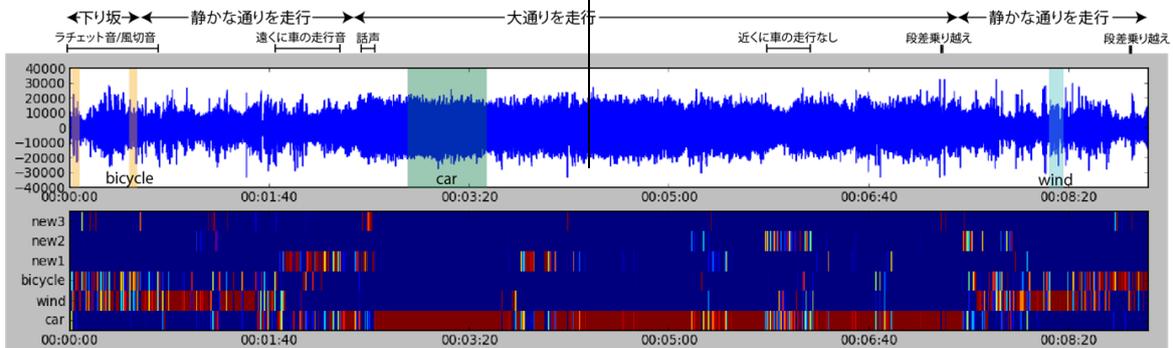


図 4. 環境音モデルの生成と識別結果

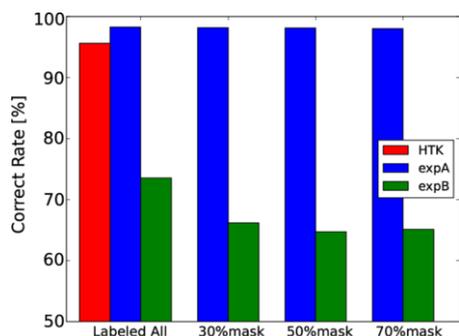


図 5. 楽器音識別結果 (expA: 未知音を含まない場合、expB: 未知音を含むモデルでの未知音の識別率)

(3) 分離音の識別、移動音の追跡 :

上記(1)(2)の定位・分離・認識技術を統合し、「どちらから何の音がするか」、実環境中で周囲の音を検出する機能を実現した。さらに各時刻で推定される音源方向と認識結果を時系列に接続する移動音の追跡機能を構築した。本追跡手法は複数の仮説を基に、複数の音を同時に扱えること、音の鳴り始めや終わりを検出し常に変化する音源数に対応できることが特長である。図 6. に実験結果を示す。断続的な2つの音源に対し、既存手法である最近棒法(上段)では音の停止時や2音の交差時に誤りがあるのに対し、提案法は、音の停止・鳴り始め、2音の交差を正しく推定で来ていることが確認できる。

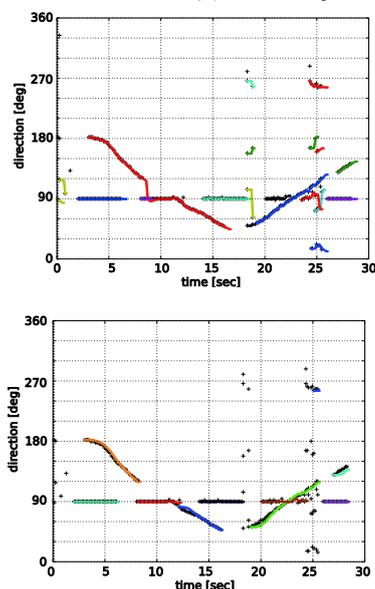


図 6. 移動音の追跡結果 (上段: 従来法、下段: 提案法) (横軸: 時刻、縦軸: 音源方向)

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 7 件)

- ① “32-channel Omni-directional Microphone Array Design and Implementation”, Yoko Sasaki, Tomoaki Fujihara, Satoshi Kagami, Hiroshi Mizoguchi, Kyoichi Oro, Journal of Robotics and Mechatronics, Vol.23-No.3, pp.378-385, 2011/06.
- ② “Map-Generation and Identification of Multiple Sound Sources from Robot in Motion”, Yoko Sasaki, Simon Thompson, Masahito Kaneyoshi, Satoshi Kagami, Proceedings of The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.437-443, 2010/10.

[学会発表] (計 7 件)

- ① “A Nested Infinite Gaussian Mixture Model for Identifying Known and Unknown Audio Events”, Yoko Sasaki, Kazuyoshi Yoshii, Satoshi Kagami, Proceedings of 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services, 2013/07 (To appear)
- ② “音響イベント認識のための未知音に対応可能な無限混合ガウスモデル”, 佐々木 洋子、吉井 和佳、加賀美 聡, 第18回ロボティクスシンポジウム, pp.245-250, 2013/03.
- ③ “Spherical Microphone Array for Spatial Sound Localization for a Mobile Robot”, Yoko Sasaki, Mitsutaka Kabasawa, Simon Thompson, Satoshi Kagami, Kyoichi Oro, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012/10.

○取得状況 (計 1 件)

名称: 点音源検出方法
 発明者: 加賀美、佐々木、溝口、榎本
 権利者: 加賀美、佐々木、溝口、榎本
 種類: 特許
 番号: 5007400
 取得年月日: 2012/6/8
 国内外の別: 国内

6. 研究組織

(1) 研究代表者

佐々木 洋子 (SASAKI YOKO)
 独立行政法人産業技術総合研究所・デジタルヒューマン工学研究センター・研究員
 研究者番号: 00574013