

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 29 日現在

機関番号：17401

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500210

研究課題名(和文) 能動耳介とロボット頭部の協調運動を用いた確率的音源定位

研究課題名(英文) Probabilistic sound source localization with active pinnae and robot head motion

研究代表者

公文 誠 (Kumon, Makoto)

熊本大学・自然科学研究科・准教授

研究者番号：70332864

交付決定額(研究期間全体)：(直接経費) 4,100,000円

研究成果の概要(和文)：能動的に動作する耳介を持つロボットシステムを用いた音源定位法を検討した。音源方向の推定には、事前測定した音響データベースとの照合による方法を考え、不確かさを確率モデルを導入した推定手法を提案した。実際の能動耳介ロボットによる実験を通じて、提案手法がロバストな音源定位が可能であることが示された。また、ロボット動作に伴う騒音の影響を検証し、基本的な動作生成法について研究を行った。

研究成果の概要(英文)：In this study, a robot system with active pinnae is considered in order to achieve sound source localization. Localization method based on pre-measured acoustic database is introduced, and uncertainty model for the system is proposed. It was verified by experiments that the proposed method succeeded to localize a sound source under noisy environment. Ego-noise that is generated during a robot motion is also considered, and a primitive action of the robot is studied.

研究分野：ロボット工学

キーワード：音源定位 能動耳介

1. 研究開始当初の背景

自転車のベルや自動車のクラクションのように、緊急・重要な情報を伝達する手段として我々は音信号を日常的に活用しており、日常生活環境で機能するロボットの実現にはロボットの聴覚機能の実現は必須である。ロボット聴覚は昨今精力的に研究される分野(例えば「計測と制御」特集号、奥乃他編, 2007, IROSでのRobot Auditionセッション 2005-2011)で、複数のマイクロホンからなるマイクアレイを利用した音響信号処理の応用などの研究が盛んである。

現実的なロボットを考えると、必ずしも音信号処理のためだけに演算装置を利用できるわけではなく、マイクの数や配置も自由には出来ないなど制限がある。その一方、ロボットは目的音源へ近づくことや、マイクロホンに手を当て、手を反射板として集音効果を得るといった、身体性の活用し受聴状況を改善できる。ロボット聴覚では、音信号処理における演算量の制限などを、身体動作を組み合わせたロボット特有の方法で補える可能性がある。

ところでこのような身体あるいは身体動作を巧みに活用しているのは生物である。例えば人間は耳介(耳たぶ)での反射など身体の音響上の影響を利用して音源方向の推定を行うこと(Shaw et al., 1967)、ネコは耳介を動かすことで音源方向の推定性能を改善する(Olmstead et al., 1980)ことが知られている。申請者はこのような生物の聴覚機能に着想を得て、耳介を用いたロボットによる音源方向へのサーボ系の設計(Shimoda, Kumon et al., 2005, 2007)、能動耳介による音源方向探査の基礎を明らかにしてきた(Kumon and Noda, 2011)。

身体の音響的影響を活用したロボットの動作生成のアプローチは、身体を音響フィルタとして活用する一方、耳介に限らず身体全体が影響してしまう大自由度で複雑なタスクである。このような複雑なタスクにおいてロボットが現実的な環境で十分な性能を実現するには、これまで示してきた原理的な方法を発展させ、大自由度と不確かさを考慮し、統合的なロボットの聴覚動作を明らかにする必要がある。そこで本研究では、ロボット(頭部)の動作と耳介の動作の協調によるロボット聴覚の理論、これに基づいて不確かさを考慮した聴覚機能のモデルを導出し、性能を改善する手法を提案することを目指すこととした。

2. 研究の目的

以上の背景の下で、本研究の開始時点では主に次の3点を明らかにすることを目的に掲げることとした。

(1) 頭部動作と耳介動作の協調によるロボット聴覚

動物のように耳介と頭部が独立して動作する系は、マイクロホンを音源方向に向けるようなタスクにおいて冗長な自由度があり、追加的なサブタスクを同時に実行可能である。聴覚機能を考えれば、自己の動作に伴いモータなどから生じるエゴノイズの抑制、身体の変形に伴う音響特性変動の活用、音源に対する配位の変化による音響特性変動を考慮するなどのサブタスクを考慮することが出来、高性能の聴覚ロボットの実現には、これらのタスクをロボットの動作に適切に割り当てる必要がある。本研究では最適な聴覚動作を与えるタスク配分法を示し、協調動作の背景にある理論を解明する。

(2) 不確かさを考慮した聴覚機能のモデルを導出

壁の反射や残響など周辺環境の影響に加え、ロボット身体自身も音響的な伝達特性を持つため、収録される音信号は目的音から大きく歪むことになる。特にロボットの身体はマイクロホン近傍で影響が大きいため、その影響を十分に考慮しなければならない。しかし、ロボットの形状は複雑で、本研究のように身体が動作する場合の影響を完全に予測することは不可能と言って良い。

そこで身体の伝達特性については相当の不確かさを含むことを許容するモデルとして**身体の位置・配位の影響による不確かさを確率的揺籃として考慮した受聴モデル**を導入し、聴覚機能における不確かさを積極的に検討することで性能改善に役立てる。

(3) 頭部と耳介の協調動作による確率的音源定位

耳介のような方向選択性の高い音収録装置は、対象音を耳介正面に捉える場合は有用であるが、耳介の横や後方など「死角」に相当する方向から目的音がある場合、対象信号が抑圧されてしまう問題がある。特に音源定位では音源が耳介の正面にあるとは限らないため、**耳介を用いた聴覚装置では積極的に頭部や耳介(あるいは双方)を動かして音源を探査する必要がある**。この音源定位動作を効率的に行うには、音源の存在する可能性の高い位置を推測しながら、得られた情報を統合した適切な動作を生成する必要がある。

本研究では上述のテーマで得られる頭部・耳介の協調動作モデル、および聴覚における状態と不確かさのモデルに基づき、**不確かさにロバストな確率的状態推定と探索手法**を援用し、新しい音源定位手法を提案する。

3. 研究の方法

(1) 基本システムの構成

本研究を遂行するにあたって、まず能動耳介を用いた音源定位の手法の基礎として、正中面での音方向(音の上下方向に対応)の推定を対象に、定位法を確認し、これを実現する装置の基本性能の確認とその改良を行うこととした。(平成24年度)

(2) 不確かさのモデル化

次いで、能動的な動作を行う耳介システムによる音源定位においてどのような不確かさが生じるかをまず調べることとした(上の目的(2)に対応)。これは、受聴された音信号がロボット周囲の環境や身体の反射などによって、原信号と異なってしまい、本質的に音信号観測に不確かさが存在することによる。特に、身体による影響は対象音源との相対的な位置関係に依存し、音源位置の情報が不正確であること、ロボット自身のセンサで得た配位情報だけでは影響を正確に見積もることは出来ないことから、モデルを得ることが難しいためである。そこで、観測時に生じる不確かさに一定の確率モデルを導入し、これが推論に及ぼす効果を考慮することで音源定位の性能にどのような影響を与えるかを検討し、改善方法を考案することとした。(平成24年度～25年度)

(3) 不確かさの下での動作生成

観測に伴う不確かさについてある程度の知見が得られた後、この不確かさを含め、本研究の目的(1)(3)で考える頭部と耳介の動作に取り組むこととしたが、今回の研究では不確かさの影響が非常に強いことが分かったため、当初の目的を一部変更し、まずは能動耳介そのものの動作生成問題を先に考えることとした。(平成25年度～26年度)

不確かな観測と事前知識の下で、より精度の高い音源推定を行うため、基本的な動作パターンを検討し、動作の有効性を実験的に検証するとともに、動作しながらの音信号の受聴において、耳介の駆動音を与える影響についても調べることとした。

(4) カメラ情報との統合

さらに、本研究の発展的な応用として、カメラ情報と音源定位情報の統合についても研究の範囲を広げ、物体の背後に隠れた移動音源の位置を推定する手法などについても検討を行うこととした。

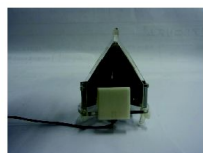


Figure 1: 正面図



Figure 2: 上面図



Figure 3: 正面図

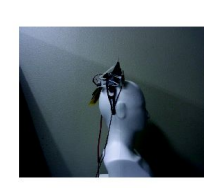


Figure 4: 側面図

4. 研究成果

(1) 能動耳介を用いた音源定位の実現

音源方向を推定するため、二つの耳で受聴した信号の間の強度差と位相差を周波数領域上での特徴量に考え、事前に音源方向との対応関係を測定したデータベースと照合することで音源方向を推定する方法を適用した。これは、耳介の反射等により音源方向が到来信号の周波数領域に埋め込まれることを利用した定位法である。

提案法の有効性を検証するため、実際の能動耳介装置を開発(上のFig.1-4)し、音源定位実験を行った所、耳介姿勢を含めて音源方向(上下・左右)を一定の精度で推定することが可能であることが確認された。

(2) 音響特徴量に基づく定位手法での不確かさのモデル化

(1)での方法は、事前に高精度に作成されたデータベースを前提にしており、また対象の音源方向がデータベース中のデータ点では正しく動作する。一方、十分な事前データが得られず、そもそもデータベース自身に不確かさが含まれる場合、この不確かさを考慮した音源方向の推定を行う必要がある。このため、特徴量の推定におけるデータ照合で、不確かさを考慮した尤度関数を導入する方法を提案した。この結果、少ない事前データ点数や雑音下で測定したデータ点から作成したデータベースを用いても、

一定のロバスト性を持った音源方向推定が可能となった。また、対象音源がデータ収録時の音源方向にはない場合、データベース点の情報を補間することで方向推定する方法があるが、この時、補間点での不確かさを推定するためにガウス回帰を用いる方法を提案した。

(3) 能動耳介の動作音の影響

能動耳介は、耳介の向きを変えることで、受聴条件を制御して、音源定位能を改善しようとするものであるが、その動作に伴って駆動雑音が生じ、音源定位の妨げになる点は考慮が必要である。本研究での調査によって、耳介動作で生じる雑音が広帯域にわたる比較的一様な雑音であることが明らかになり、この雑音の影響の及ぶ範囲を検出できるようになった。また、音源定位からこの雑音区間を除外することで、定位性能の改善が可能であることを実験を通じて示すことが出来た。

(4) 耳珠の有効性

本研究での音源方向の推定は、耳介による反射などの干渉を用いたものである。能動耳介でこれが機能することが分かったため、さらにこれを効率化する方法として、耳介開口部に突起を設けることとした。これは、人間やコウモリなどで見られる耳珠が耳介効果を強調するという報告を参考にしたもので、能動耳介システムに耳珠に相当する小板(前ページ Figure.1 の耳介下部に見える白い四角い部品)を突起として導入した。実験の結果、音源方向の特徴がより顕著に見られることが示された。

(5) 繰返しベイズ推定によるカメラ・音源情報統合によるターゲット定位

音源位置を事前に測定した音響特徴量データベースとの照合によって推定する方法の応用として、屋内での音源位置の推定が可能になる可能性がある。一方、音源の存在する領域が広い場合、異なる音源位置にあっても、音響特徴量は類似したものになる可能性が生じる。この時、音源の存在位置の確からしさは複数の地点で高い値を示すマルチモーダルな分布になり、単にその最大値を取るだけでは位置推定を達成することは出来ない。一方、カメラなどの視覚情報はターゲットの位置を正確に推定することが可能であるが、壁や障害物などが存在する環境では、視界が制限される問題がある。ここでは、これらを組み合わせることで、視界内ではカメラ情報で音源位置推定を行い、カメラで認識できない場合は音情報を用いた位置を推定する手

法を提案した。提案法では繰返しベイズ推定を用いることで音源の移動を考慮し、マルチモーダルな測定を回避することが可能となった。

実際、これを屋内で2つのマイクロホンだけを持つバイノーラルシステムで試した所、音源の移動を追跡することが確認された。

5. 主な発表論文等

[雑誌論文](計 2 件)

M. Kumon, D. Kimoto, K. Takami and T. Furukawa, Acoustic Recursive Bayesian Estimation for Non-Field-of-View Targets, Proceedings of 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services, 査読有, 2013.

M. Kumon, D. Kimoto, K. Takami and T. Furukawa, Bayesian Non-Field-of-View Target Estimation Incorporating an Acoustic Sensor, Proceedings of 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 査読有, pp. 3425-3432, 2013.

[学会発表](計 8 件)

尾堂航, 公文誠, 音源定位における能動耳介での動作の影響について, 人工知能学会 AI チャレンジ研究会, 2014 年 11 月 21 日, 慶応大学(東京都, 日吉市).
尾堂航, 公文誠, 能動耳介での音源定位に用いる周波数帯域の選択について, 第 32 回日本ロボット学会学術講演会, 2014 年 09 月 04 日~2014 年 09 月 06 日, 九州産業大学(福岡県, 福岡市).

公文誠, 尾堂航, 木元大輔, 耳珠のある能動耳介システムとその動作について, 人工知能学会 AI チャレンジ研究会, 2014 年 03 月 17 日, 京都大学(京都府, 京都市).

木元大輔, 尾堂航, 公文誠, 古川知成, バイノーラル聴覚ロボットでの特徴量マッチングに基づく音源定位における規範データの不確かさに関する考察, 第 14 回システムインテグレーション部門講演会(SI2013), 2013 年 12 月 18 日~2013 年 12 月 20 日, 神戸ポートアイランド(兵庫県, 神戸市).

木元大輔, 尾堂航, 公文誠, ガウス回帰に基づく両耳間レベル差の補間, 人工知能学会 AI チャレンジ研究会, 2013 年 12 月 06 日, 早稲田大学(東京都, 新宿区).
木元大輔, 尾堂航, 公文誠, 観測データの不確かさを考慮したバイノーラル聴覚ロボットでの音源定位手法, 第 31 回日本ロボット学会学術講演会, 2013 年 09 月 04 日~2013 年 09 月 06 日, 首都大学東京(東京都, 八王子市).

公文誠, 木元大輔, 両耳間レベル差を用いた音源定位における耳孔位置の最適化, 人工知能学会 AI チャレンジ研究会, 2012 年 11 月 15 日, 慶應義塾大学 (東京都, 日吉市) .

野田佳孝, 公文誠, 二つの能動耳介による正中面内の音源方向推定, 第 13 回システムインテグレーション部門講演会, 2012 年 12 月 18 日~2012 年 12 月 20 日, 福岡国際会議場(福岡県, 福岡市) .

6 . 研究組織

(1)研究代表者

公文 誠 (KUMON, Makoto)
熊本大学・大学院自然科学研究科・准教授
研究者番号 : 70332864

(2)研究協力者

Tomonari Furukawa
Virginia Polytechnic Institute and State University, Professor