

ロボットの持つ位置情報を利用した環境音の音源推定による家庭内作業の状態把握

State Recognition of House Chores by Environmental Sound Source Estimation Utilizing Location Information of Home Robot

学 五十嵐 大騎 (都市大) 徳永 夏帆 (都市大)
正 佐藤 大祐 (都市大) 正 金宮 好和 (都市大)

Daiki IGARASHI, Tokyo City University

Natsuho TOKUNAGA, Tokyo City University

Daisuke SATO, Tokyo City University, sato@rls.mse.tcu.ac.jp

Yoshikazu KANAMIYA, Tokyo City University

By performing sound source estimation using location information of home robot and a result of sound source localization by HARK, the robot can recognize environmental sounds based on the sound source. Thus, it is possible to distinguish even if similar sounds of feature quantities are generated, so that precision of state recognition of house chores is increased. In this paper, we developed the state recognition system and performed experiments of operation works of a microwave oven. As a result, the localization results were facing a direction of the microwave oven, and so it can be said that it is possible to recognize the sound source accurately.

Key Words: Environmental sounds recognition, Sound source estimation, Home robot

1 緒言

ロボットが家庭内作業を行うためには、作業の状態を把握する必要がある。これに対し、作業中に発生する環境音を利用することで状態把握が可能になる。例えば、家電製品から鳴る音によって、それによる作業の開始や終了の判断ができる [1]。しかし、家庭内では特徴の似た環境音が多数存在するため、環境音認識だけでは完璧な状態把握は困難である。

ロボットは様々な情報に基づいて作業を行っている。その情報の一つに、周辺環境の相対位置関係が挙げられる。ロボットの持つ環境情報と音響処理による音源情報を組み合わせることで、作業状態の認識精度の向上が期待できる。

本稿では、ロボットの持つ室内の位置情報を利用した環境音の音源推定システムを構築し、それをを用いた家庭内作業の状態認識実験の結果から、本システムの有用性を示す。

2 状態把握のための機能

2.1 ロボット聴覚用オープンソースソフトウェア

環境音認識や音源推定を行うために、音の聞き分け機能を持つ HRI-JP Audition for Robots with Kyoto university (HARK) [2] と認識エンジンである Julius [3] を導入した。以降では、状態把握に必要な各ソフトウェアの機能とロボットの持つ環境情報について述べる。

2.2 HARK による音源方向と音響特徴量

作業環境内で発生する音は、複数音源による混合音であるため、音源ごとに音を分離して目的の環境音を抽出する必要がある。そのための機能が音源定位と音源分離である。音源定位では、マイクロホンアレイによって取得した音から音源方向を推定している。また、音源分離によって音源方向に基づく分離を行い、各音源の分離音を出力する。分離音は、そのままでは認識を行うことができないため、Mel Frequency Cepstrum Coefficients (MFCC) による解

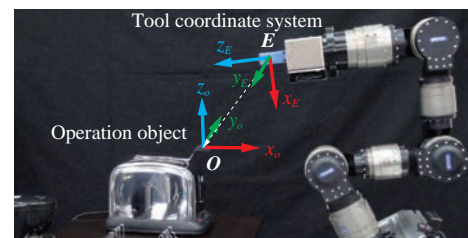


Fig. 1: Relative position between the tool coordinate system and the coordinate system of the operation object.

析を行い、音響特徴量を抽出する。

このように、HARK を利用することで環境音の音源方向と音響特徴量を得ることができる。

2.3 Julius による環境音認識

Julius では、Hidden Markov Model (HMM) による音響モデルと言語モデルに基づいて、音響特徴量との統計的な比較を行い、音の認識を行っている。また、最尤推定を採用しており、設定したモデルの中で一番近いものを認識結果としている。そのため、複数の環境音を一度に扱うことができる。その反面、モデルにない環境音にも反応するため、特徴量の近い音を誤認識してしまう問題がある。

以上のことから、Julius は環境音の認識器として利用できるが、それだけを利用した環境音認識では家庭内作業の状態把握は確実にすることはできないため、その他の情報による認識を行う必要がある。

2.4 ロボットの持つ環境情報

ロボットが物を扱うためには、操作対象物の位置を設定する必要がある。例えば、図 1 のように物を把持する場合、手先座標系から見た操作対象物の位置や姿勢を与えることで、把持動作を行うことができる。すなわち、ロボットと操作対象物の相対位置関係は、既知情報であるといえる。そこで、この情報を利用した音源推定を提案する。

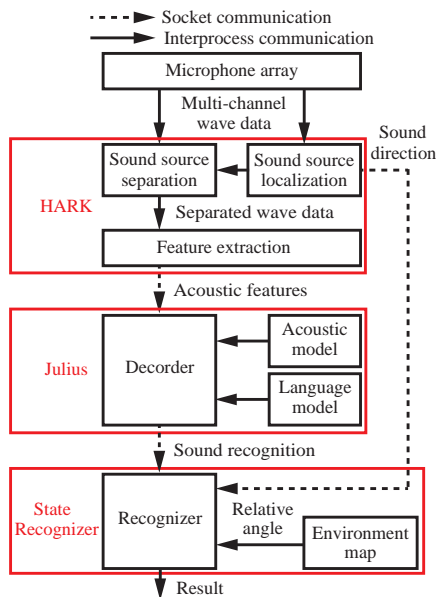


Fig. 2: State recognition system of the house chores using HARK and Julius.

3 家庭内作業の状態把握システム

3.1 システムのコンセプト

本システムでは、環境音認識と音源方向および環境情報の三つを利用して、家庭内作業の状態把握を行う。まず、ロボットと操作対象物の相対角度と音源方向から、環境音の音源推定を行う。続いて、環境音認識の結果から、音源である操作対象物や家庭内作業の状態を把握する。

音源推定によって、複数の操作対象物から特徴量の近い音が発生した場合でも、おのおのに対して状態把握を行うことができる。また、家庭内作業には作業時間が不明なものがあり、このような作業に対して環境音認識は有効的であるため、効率の良い作業が可能になる。

3.2 システムの構成

3.1 節のコンセプトを満たすシステムの構成図を図 2 に示す。本システムは、主に三つの機能で構成されており、音源推定及び状態把握を行う機能を状態把握器と呼称する。

HARK によって得られる音源の方向角は、ソケット通信によって状態把握器に送信される。状態把握器は、作業環境内に存在する操作対象物の位置データをまとめた環境地図を保有しており、ロボットの位置や姿勢から、各操作対象物との相対角度を算出している。これらの角度の差をとり、偏差角度が最小のものを、音源推定の結果とする。

一方で、HARK によって抽出された音響特徴量は、Julius のデコーダに送られる。デコーダでは、音響特徴量と各モデルの尤度を求め、環境音の認識結果を出力している。状態把握器では、音源推定と環境音認識の結果を利用して、操作対象物や作業の状態を把握している。

4 電子レンジ操作実験によるシステムの検証

4.1 実験内容

家庭内作業の操作対象物として電子レンジを利用し、状態把握のための環境音認識の検証を行った。電子レンジから発生する音は、電子レンジのワット数の操作音 (W 操作音)

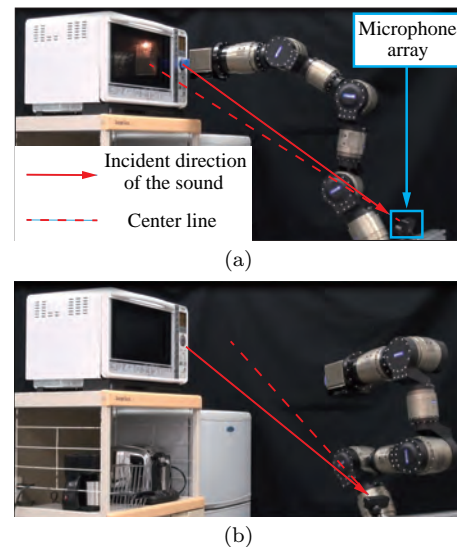


Fig. 3: The condition of the experiment using the microwave oven: (a) the robot facing the front direction pushes the start button and (b) the robot facing the direction of -30 degrees waits until the microwave oven stops.

と時間の操作音 (T 操作音) といった単発音と、メロディによって構成される開始音と終了音の 4 種類である。そこで、ロボットによる電子レンジ操作を 30 回繰り返し、それぞれの環境音の認識率を調査した。

また、電子レンジによる作業の待機中に、ロボットが移動して別の作業を行うことを想定すると、ロボットの位置や姿勢が変化することが考えられる。そこで、相対角度が変化した場合の音源推定について検証を行った。実験の様子を、図 3 に示す。このように、ロボットの回転後に電子レンジが終了した際、終了音がどのように定位されるかを調査した。なお、ロボットの回転角度を 0 deg から 90 deg の 10 deg 刻みとし、各回転角度に対し 10 回の実験を行った。

4.2 環境音の認識精度

環境音認識の実験結果を表 1 に示す。これより、W 操作音は誤認識が起こりやすいが、その他の環境音については良好な認識傾向にあることが分かる。

ここで、誤認識の原因について考察を行う。表 2 に示す W 操作音の認識結果の内訳から、二つの操作音の識別に問

Table 1: Recognition rates of the environmental sounds for the microwave oven.

Environmental sounds	Recognition rate [%]
W operation sound	73
T operation sound	100
Start sound	100
Finish sound	100

Table 2: Breakdown of the recognition of the W operation sound for the microwave oven.

Acoustic Models	recognition rate [%]
W operation sound	73
T operation sound	27
Others	0

Table 3: Sound source localization results of when the rotation angle of the robot is from -90 degrees to 0 degrees. In addition, the search range of the sound source localization is between -100 degrees and 100 degrees.

Rotation angle [deg]	Localization results [deg]							
	-40	-30	-20	-10	0	10	20	30
-90	0	1	2	0	7	0	0	0
-80	2	2	1	0	1	4	0	0
-70	0	0	1	2	2	1	4	0
-60	0	0	2	0	2	3	1	2
-50	0	0	5	2	1	0	1	1
-40	1	1	5	0	3	0	0	0
-30	1	2	3	0	3	1	0	0
-20	1	4	2	1	1	1	0	0
-10	0	2	1	2	2	1	2	0
0	0	1	4	3	1	1	0	0
Total	5	13	26	10	23	12	8	3

Table 4: Distribution probability for the error between the localization results and the relative angle.

	Error [deg]				
	0	± 10	± 20	± 30	-40
Total	10	49	25	13	3
Cumulative frequency	10	59	84	97	100
Distribution probability [%]	10	59	84	97	100

題があることが分かる．そこで、音の解析を行ったところ、W 操作音のピーク周波数が 2 kHz であるのに対し、T 操作音では 2.2 kHz であった．一方で、本システムにおける特徴量抽出時の 2 kHz 帯の周波数分解能は、480 Hz である．つまり、誤認識の原因はシステムの周波数分解能であり、誤認識を減らすためにも、分解能を調節する必要がある．

以上のことから、本システムは家電製品の開始や終了といった状態把握には有効的である．また、その他の環境音についても、パラメータの調節次第で認識精度の向上が想定されるため、多くの作業に対応した状態把握が可能になる．

4.3 音源推定による状態把握とシステムの有用性

実験の結果を表 3 に示す．また、ロボットの回転角度が 0 deg および -90 deg のときの分布を図示すると、図 4 の通りとなる．音源定位の探索範囲は ± 100 deg である．これに対し、終了音の定位結果は電子レンジの周辺に存在している．これより、音源方向と相対角度の一致は難しいが、ある程度の範囲内であれば、音源推定は可能である．

続いて、音源推定の精度について考察を行う．表 3 の結果から、定位結果は -10 deg を中心として対称的に現れているため、同程度離れた場所であれば、音源定位の傾向は同様であるといえる．そこで、定位結果と相対角度の差による分布確率を求めると、表 4 の通りとなる．音源推定の許容範

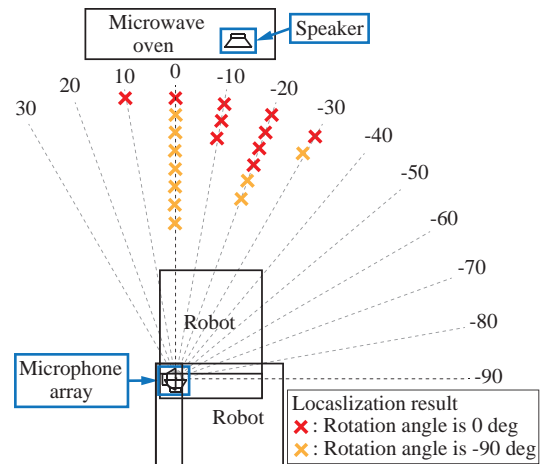


Fig. 4: Distribution of the sound source estimation results when the rotation angle of the robot is -90 degrees or 0 degrees.

囲を ± 20 deg と設定すると、84 % の確率で正確な状態把握が行うことができる．また、探索範囲が ± 100 deg であることから、五つのエリアに分けた音源推定が可能になる．

しかしながら、本システムを使ってロボットが作業を行うためには、音源推定の確実性をさらに上げる必要がある．そのための方法として、音源推定を繰り返し行い、確率的に 100 % に近づけることが挙げられる．定位結果のばらつきの原因として、音源定位の計算誤差が考えられる．そこで、表 4 の結果から事前確率を設定し、ベイズ推定に基づいて音源推定を繰り返し行う．これにより、音源定位の結果が真値に近づくため、音源推定の精度は向上し、結果として、作業の状態把握がより正確なものになることが期待できる．

5 結言と今後の課題

音源推定および環境音認識による家庭内作業の状態把握システムを構築した．本システムでは、ロボット持つの環境情報を利用して、操作対象物の位置と環境音の音源を関連付けることで、特徴量の近い音の誤認識に対応できるようにした．また、電子レンジの操作作業を対象に状態把握の実験を行った．その結果、開始音や終了音の認識精度は良好であり、電子レンジを音源とした推定が可能であることから、正確な状態把握が期待できることを示した．

今後の課題として、音源推定を繰り返し行うことが可能なシステムに変更し、状態把握の再確認をする必要がある．また、作業が複雑化するにつれて、ロボットノイズや雑音としての環境音が問題となる．そのためのノイズフィルタの設計についても考慮しなければならない．

文 献

- [1] 関戸 佐知, 黒山 佑太, 新良貴 陽平, 佐藤 大祐, 金宮 好和, “モジュラーホームロボットによる環境音を利用した動作計画に基づく朝食準備作業”, 第 32 回 日本ロボット学会 学術講演会論文集, 1N1-04, 2014.
- [2] 奥乃 博, 中臺 一博, “ロボット聴覚用オープンソースソフトウェア HARK”, 日本ロボット学会誌, vol. 28, no. 1, pp. 6-9, 2010.
- [3] 河原 達也, 李 晃伸, “連続音声認識ソフトウェア Julius”, 人工知能学会誌, vol. 20, no. 1, pp. 41-49, 2005.