

# 環境音認識を用いたモジュラーホームロボットによる朝食準備作業

Breakfast Table Setting by a Modular Home Robot With Environmental Sounds Recognition

夏目 彬弘 (都市大)      学 関戸 佐知 (都市大)  
正 佐藤 大祐 (都市大)      正 金宮 好和 (都市大)

Akihiro NATSUME, Tokyo City University

Sachi SEKIDO, Tokyo City University, sekido@rls.mse.tcu.ac.jp

Daisuke SATO, Tokyo City University

Yoshikazu KANAMIYA (D. N. Nenchev), Tokyo City University

The aim of our research is development of a modular home robot. Task of the robot is managed by timetable. Timetable needs to estimate time that is longer than each task, because there is range in end time of task. So, we focus attention on environmental sounds associated with status and changes in environment. This paper describes implementation of breakfast table setting by a modular home robot with environmental sounds recognition.

**Key Words:** Home robot, Demonstration experiment, Environmental sounds recognition

## 1 緒言

近年, iRobot 社の Roomba に代表される掃除ロボットの实用化や, Willow Garage の PR2 や複数のヒューマノイドロボットによる家事作業の実現など, 家庭内におけるロボットの利用, ホームロボットの研究開発が活発になっている。しかし, 実用化されているホームロボットは単一機能のみであり, ユーザから求められる家庭内作業を行うためには機能が不足している。多数の機構やセンサなどを冗長に搭載し, 複数の作業を実現可能なホームロボットは研究段階にあり, その価格も非常に高い。日本経済新聞社による, パートナーロボットの用途別希望購入価格調査によると, 家事支援ホームロボットは 30 万円未満であることから, 家電と同等の値段設定が望まれる [1]。すなわち, 多機能ホームロボットを実用化するためには低価格化する必要がある。

そこで我々は, 多機能で低価格なホームロボットを実現するために個々のユーザが求める最小機能でホームロボットを構成する, “Multitask” (ユーザが求める複数作業の実行), “Modularization” (ホームロボットの機能やシステムのモジュール化), “Minimalization” (機能, コストの最小化) を柱とする MMM コンセプトを提案した [2]。その MMM コンセプトに基づきモジュラーホームロボットを開発している。

モジュラーホームロボットが行う作業の各工程にはパンが焼けるまでの時間やコーヒーメーカーがコーヒーを作り終える時間などの目安となる時間が存在し, それを基に作業のタイムテーブルを作成し, 作業を実行できる。しかし, この時間には幅があるため, 完全に作業を実現するには, 実際の終了時間よりも長い作業時間を想定する必要がある。そこで, 音のカテゴリの一つである環境音に着目した。

環境音は雨音やブザー音などの周囲の状態とその変化に伴って発生する音である。そのため, 環境音は周囲の状況とその変化を把握する上で有益な情報であるといえる。すなわち, 多様な状況の変化がある家庭内で作業するホームロ

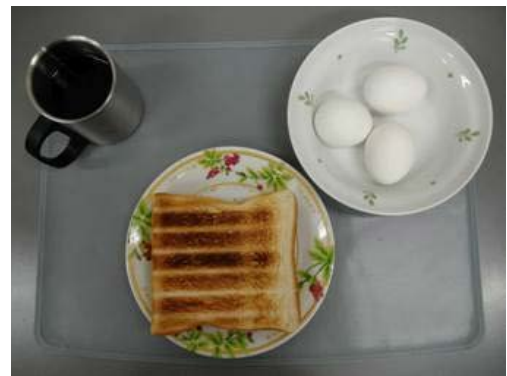


Fig. 1: Breakfast.

ボットに環境音を認識することは, 作業の状況を理解する上で有益である。よって, 環境音を認識することで各作業の終了を把握することが可能である。

本稿では, モジュラーホームロボットの有用性について述べる。また, 環境音認識の有用性を実証するため, 環境音認識を用いたモジュラーホームロボットによる朝食準備の一つであるトーストを用いたパン焼き作業を実現する。

## 2 モジュラーホームロボット

### 2.1 MMM コンセプト

多機能で低価格なホームロボットを実現するために “Multitask”, “Modularization”, “Minimalization” を柱とする MMM コンセプトを提案した。Multitask はユーザが求める複数作業を実現することを意味する。Modularization とはホームロボットを構成する要素を機能単位でモジュール化することを意味する。機能単位でモジュール化することにより, 機能を自由に選択してホームロボットを構成することができる。Minimalization は機能とコストの最小化を意味する。ユーザが必要とする機能だけでロボットを構成することにより, 低コスト化することが可能となる。すなわち, モジュラーホームロボットを構築することにより, 多機能化, 低価格化が実現可能となる。

## 2.2 モジュールの選択と作業内容

ユーザのニーズとモジュールの選択の関係を単腕と双腕を例にあげて説明する．ユーザのニーズが低価格優先で多少の作業を負担することに寛容であるならば，作業機構には単腕を選択する．ユーザのニーズが，高価であってもユーザが作業に関与しないというものであるならば，作業機構には双腕を選択する．

トースト，コーヒー，温泉卵（Fig. 1）を用意する朝食準備作業を例とすると，単腕を選択した場合は，蓋や袋を開けるという作業が困難であるため，作業を遂行するには Fig. 2 に示すようにユーザがパンやコーヒーの粉，卵のセットを行う必要がある．一方，双腕を選択した場合は蓋や袋を開けることが可能であるため，

また，ユーザから極力作業に関与せず，双腕よりもコストを抑えたいなどの要望があった際は，単腕の他に，グリッパモジュールやハンドモジュールを付加することにより，モジュラーホームロボットの作業可能な幅を拡大することが可能となる．さらに，モジュラーホームロボットの利点として，ユーザがモジュラーホームロボットに求める作業が新たに増えた際でも，モジュールを追加することで対応可能であることがあげられる．すなわち，機能追加ごとにロボット自体の買い換えが不要となる．

## 2.3 作業管理

モジュラーホームロボットに限らず，ロボットが行う作業は機器により完成する目安の時間があるため，作業のタイムテーブルを作成可能である．

先に上げた朝食準備を例にすると，ポップアップトースタがパンを焼きあげるまでの時間，コーヒーメーカーがコーヒーを淹れるまでの時間，エッグボイラが温泉卵を完成させるまでの時間などは機器により目安時間が与えられている．しかし，目安時間には幅があるため，それぞれの完成時間を冗長に設定しなければならない．これでは，トーストが冷める，温泉卵がゆで卵となる可能性がある．そこで，作業終了時に機器が発する音に着目した．

## 3 環境音認識

日常生活において，人が知覚する音には大きく分けて，音声，音楽，環境音の三つのカテゴリが存在する．音声には，人の会話やロボットに対する指令などが分類され，音楽には，歌や楽器による演奏などが分類される．環境音には，一般家庭内で発生する音に限定すると，電子レンジのブザー音やボタンの操作音，窓が開いている際の外の音などが分類される．環境音を認識することで，作業中の動作や，家庭環境の状況などを把握することが可能となる．例えば，ホームロボットが掃除機をかける作業をする際に，紙づまりの音を認識させることができれば，ホームロボットはその状況を理解し，リカバリ動作を行うことができる．

環境音認識システムにはロボット用オープンソース・ミドルウェアである Robot Operating System (ROS) 上で動作する ROS Open-source Audio Recognizer (ROAR) [3] というものがある．このシステムを用いることによりモ



Fig. 2: Preparation for task of setting the Table for breakfast by modular home robot.

ジュール化された環境音認識システムの構築が可能となる．

## 3.1 環境音認識システム ROAR

ROAR は ROS 上で動作する環境音認識システムである．機械学習の一つである one-class Support Vector Machine (OCSVM) を用いることで，認識したい環境音に対する，学習および認識が可能である．ただし，OCSVM の領域推定には超球を用いている．OCSVM の特徴は与えた情報に対して識別対象の音であるか否かを判断するシステムである．

ROAR による学習および認識の流れを Fig. 3 に示す．環境音の学習では，Fig. 3 の (a) に示すように，マイクロホンアレイから取得した環境音に対し，分類器に学習させる環境音を手動で選択する．次に，選択した環境音を分割した後に，ノイズ抑圧を行い特徴量を抽出する．その後，正規化を行い，分類器である OCSVM に学習させ，認識したい環境音のモデルを作成する．環境音の認識では，Fig. 3 の (b) に示すように，マイクロホンアレイから取得した環境音を分割した後に，取得した音のノイズ抑圧を行う．次に，あらかじめ定めた大きさに満たない音を除去した後に，特徴量を抽出する．その後，正規化を行い，分類器である OCSVM により，取得した音が，学習時に作成した環境音のモデルに含まれるか判断し，環境音の認識を行う．

## 4 環境音認識実験

ROAR による環境音認識が可能であるか判断するため，トースタによるパン焼きにおける，パンが焼けた際に発生する環境音（トースタの上がる音）の認識実験を行う．

### 4.1 方法

ホームロボットに取り付けた Kinect に向け，トースタの上がる音を 120 回学習させた後に，30 回認識させる実験を行った．実験の様子を Fig. 4 に示す．座標系はトースターの正面真ん中と，機の接点の原点とした． $y$  軸方向に  $0.5 \pm 0.02$  m 進め， $z$  軸方向に  $-0.25 \pm 0.02$  m 進めた位置に Kinect を設置した．また，ホームロボットの電源等を入れた状態で実験を行った．

### 4.2 結果と考察

実際にパンが焼けた際に発生するトースタの上がる音を 30 回認識させた結果，実験環境に左右するが，25 回正しく認識することが確認できた．また誤認識は，トースタの上がる音と近い音の場合に見受けられたが，正確な誤認識率は検証していない．以上より，30 回中 25 回正しく認識できたの

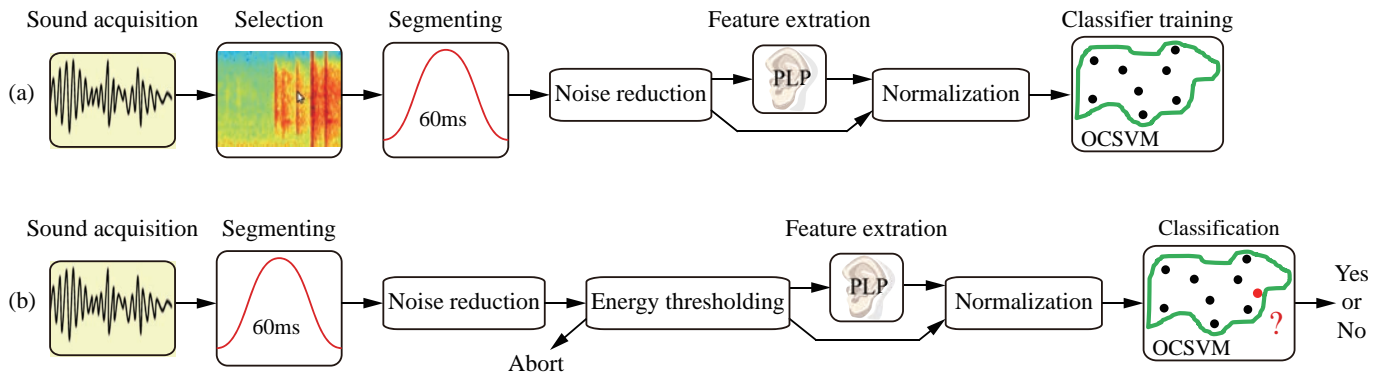


Fig. 3: The steps involved in the two ROAR pipelines. (a) To create a model of a certain audio event, many sound samples are acquired, processed, and used to train a OCSVM. (b) When the robot later needs to recognize the audio event, short-time audio signals are acquired, processed, and classified using the previously trained OCSVM.

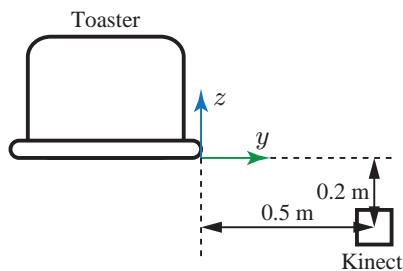


Fig. 4: State of the experiment.

で、ホームロボットへの導入が可能だと判断した。

## 5 環境音認識を用いたモジュラーホームロボットによる朝食準備動作

### 5.1 モジュラーホームロボット

我々が構築したモジュラーホームロボットを Fig. 5 に示す。このモジュラーホームロボットは主に独立二輪駆動型のモバイルベースモジュールと七自由度のアームモジュール、グリッパモジュールで構成される。グリッパモジュールには USB カメラおよびレーザセンサのセンサモジュールが搭載されており、物体の位置情報などが取得できる。この他に音を取得するにあたって、モジュラーホームロボットに安価でマイクロホンアレイだけでなく、深度センサも搭載した Microsoft 社製の Kinect for Xbox360 を導入した。

モジュラーホームロボットのシステム構成図を Fig. 6 に示す。Sounds recognition package は音認識システムのパッケージである。このパッケージ内の Control module では ROAR が動作しており、Kinect module のマイクロホンアレイより取得した音を 3.1 節で述べたように処理する。指定の環境音を認識した際には、処理結果をイーサネットを介して Modular home robot basic package 内の Control module に送信する。その結果に応じた動作をモジュラーホームロボットが実行する。

### 5.2 実験方法

環境音認識を用いたパン焼き作業の実験を行う。作業の流れは次の通りである。あらかじめ生成したパン焼き動作に、トースタの上がる音の認識を組み込むことによってパンの焼き上がりを判断して朝食準備動作を実現する。

1. アームによりパンを取得

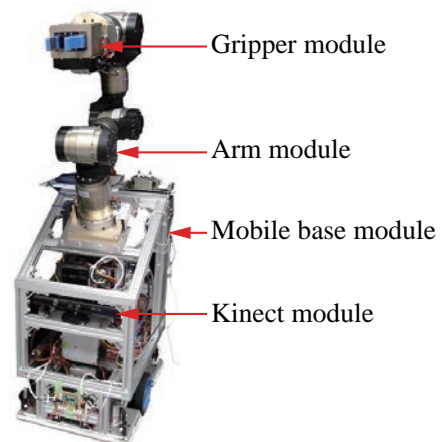


Fig. 5: Kinect on the modular home robot.

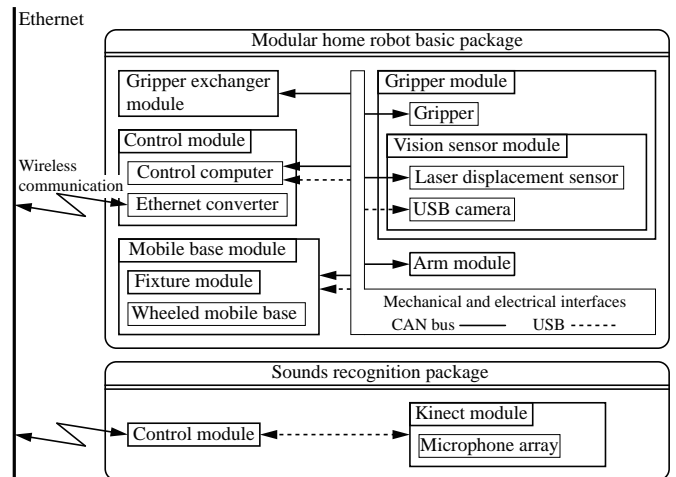


Fig. 6: Modular home robot system.

2. トースタにパンを設置
3. トースタの電源を入れる
4. トースタの上がる音を認識するまで待機
5. トースタから焼けたパンを取得
6. 焼けたパンをお皿に乗せる

### 5.3 実験結果

作業の様子を Fig. 7 に示す。Fig. 7 に示したように、作業が正しく行われていることから、環境音認識を用いたモ



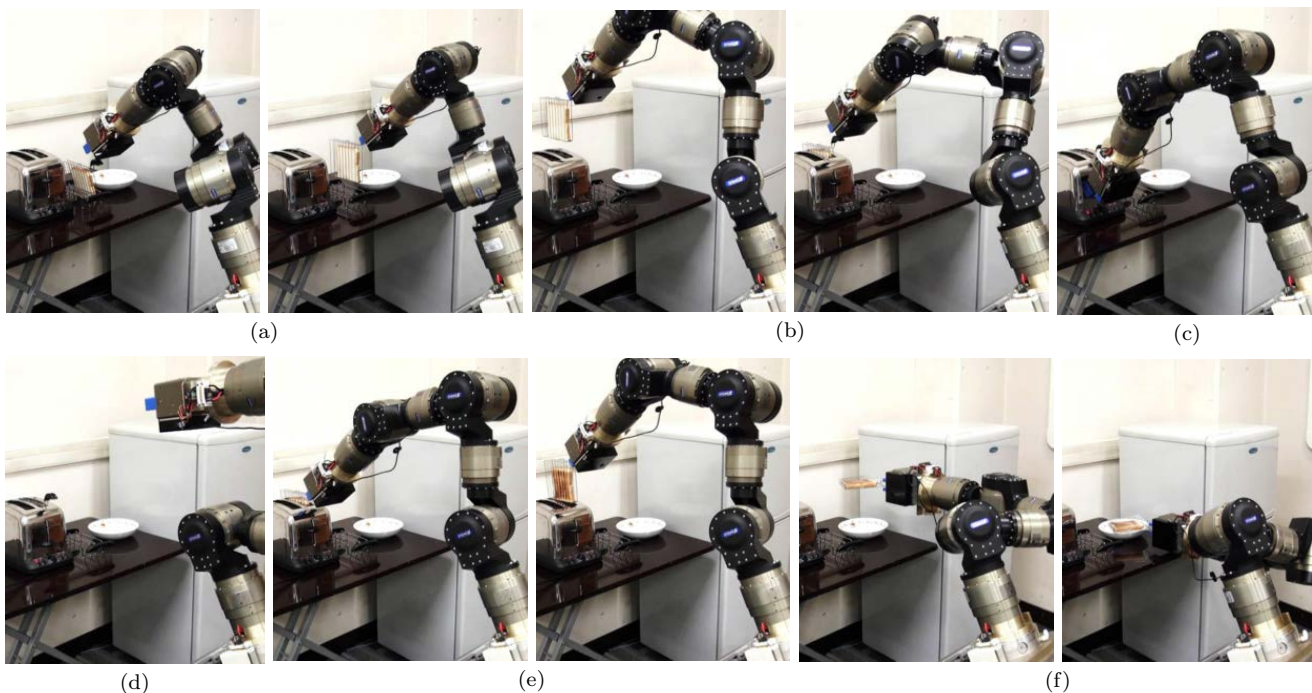


Fig. 7: Toast some bread task of (a) getting the bread, (b) setting the toaster, (c) switching on the toaster, (d) waiting for the environmental sound of the toaster when the bread was toasted, (e) getting the toasted bread, and (f) putting the bread on a plate.

ジュラーホームロボットによるパン焼き作業の動作を実現することができた。このことから、環境音を用いた作業を行えたといえるため、環境音認識を用いた作業管理は可能である。ただし、本実験ではトースタの位置などはあらかじめ指定して実験を行っている。

#### 5.4 考察

以上の結果より、新たな外界センサとしてマイクロホンアレイを用いた環境音認識による環境の状況把握を可能といえる。すなわち、環境音はトースタの上がる音以外にも様々な環境音があることから、複数の環境音をあらかじめ学習させておけば、環境音を用いた作業の管理が可能となる。つまり、環境音を認識し、モジュラーホームロボットが動作している環境の状況を瞬時に把握でき、その状況に応じた最も適した動作を動的に定めることが可能となる。

しかしながら、現状の環境音認識システムに大きく依存する。より具体的には、学習時の環境において人の話し声がしている場合や、冷暖房がついている場合など、認識時の環境が学習時と異なると認識率が下がる。

このことから、モジュラーホームロボットの動作として、学習を行う環境と認識を行う環境の変化を少なくするように動作させることや、状況によって認識に用いる学習モデルを変化させることによって、環境の変化に強いシステムの構築が可能であると考えられる。さらに、現状の環境音認識システムは誤認識率が非常に低いことから常に複数の学習モデルを用いることも、環境の変化に強いシステムの構築に有効であると考えられる。

また、識別において、学習モデルから得られる値を用いた二値化を行い認識を行なっている。このことから、学習モデ

ルから得られる二値化閾値を調節することによって、認識率をあげることが可能であると考えられる。すなわち、ロボットは現在どのタスクを行っているか把握している。このことから、曖昧な認識結果であっても、求める（想定される）環境音が認識していると、判断することによって認識率をあげることが可能であると考えられる。

#### 6 結言

環境音を用いたモジュラーホームロボットによる朝食準備の一つであるトースタを用いたパン焼き作業を実現した。しかし、現在は動作開始からトースタの上がる音を監視しているため、作業中に他の音を拾い、誤認識する可能性がある。そのため、トースタの上がる音はトースタのスイッチを入れた後にのみ発生することを考慮し、その直後から環境音認識を開始することで、誤認識を防ぐことが可能となる。

今後の課題は作業の状況に対し、どのような環境音が発生するかを関連付けることである。

#### 文 献

- [1] パートナーロボットの用途別希望購入価格調査, 日経産業新聞, 2003.
- [2] T. Tsuchiya, Y. Shiraki, S. Sekido, A. Yamamoto, D. Sato, and D. N. Nenchev, "Modular home robot system based on the MMM concept—design instance with detachable symmetric arm module," in *Proc. IEEE/ASME Int. Conf. Advanced Intelligent Mechatronics*, pp. 280–285, 2013.
- [3] Joseph M. Romano, Jordan P. Brindza, Katherine J. Kuchenbecker, "ROS open-source audio recognizer: ROAR environmental sound detection tools for robot programming," in *Autonomous Robots*, vol. 35(3), pp. 207–215, 2013.