

# 複合センサを用いた屋内環境における 行動認識に関する提案と実装

中沢 実<sup>†1,a)</sup> 小島 聖哉<sup>†1</sup> 西川 幸延<sup>†1</sup> 阿部 倫之<sup>†1</sup>

**概要:** 家電の省エネ制御といったホームオートメーションや高齢者見守りシステム等の生活を支援するサービスの実現に向けて、屋内の人物の生活行動を自動認識する研究が盛んに行われている。屋内の行動認識は様々な研究が行われており、カメラを用いる手法や、ウェアラブルデバイスを用いる手法、スマートホームを用いる手法などがある。しかしながら、プライバシーの問題やデバイスを常に装着しないといけない問題、導入コストや管理コストが高いという問題がある。この問題点に対し先行研究では10種類のセンサを1か所に設置し機械学習により38種類の行動を認識する汎用センサを構築した。しかし、センサを1箇所に設置しているため、行動がどこで起きているのか、また行動に対して複数人による行動か一人での行動であるかの区別をすることができない。本研究ではこの2つの課題を解決するために、人の位置や人数を検出できるミリ波センサの追加と1チャンネルマイクを4チャンネルマイクアレイに変更を行い2つの課題を解決することを目指した行動認識システムを提案する。

## Proposal and its Implementation of Behavioral Recognition in Indoor Environment using Composite Sensor

### 1. はじめに

家電の省エネ制御といったホームオートメーションや高齢者見守りシステム等の生活を支援するサービスの実現に向けて、屋内の人物の生活行動を自動認識する研究が盛んに行われている。本研究では低コストかつプライバシーに配慮した屋内の生活行動を認識することを目的とする。屋内の行動認識手法は様々な研究が行われており、カメラ画像を用いた画像処理により行動認識を行う手法や、ユーザの所持しているスマートフォンやウェアラブルデバイスに内蔵された加速度センサなどを用いて行動認識を行う手法、また部屋の様々な箇所にセンサが設置されたスマートホームにおいて多数のセンサ情報から行動を認識する手法などがある。しかしながら、カメラを用いるものではユーザのプライバシーを侵害する恐れがある。ウェアラブルデバイスを用いる手法ではユーザがデバイスを装着していな

いと行動を認識することができない。スマートホームを用いる手法では、多種多様なセンサを複数箇所に設置するため導入コストや管理コストが高いという問題がある。

上記で述べた問題点に対し、Laput[1]らは10種類のセンサを1か所に設置しセンサデータから機械学習により38種類の行動を認識する汎用センサを構築した。しかし、センサを1箇所に設置しているため、行動がどこで起きているのか、また行動に対して複数人による行動か一人での行動であるかの区別をすることができない。本研究では、上述した2つの課題を解決するために、人の位置や人数を検出できるミリ波センサの追加と1チャンネルマイクを4チャンネルマイクアレイに変更を行い2つの課題を解決することを目指した行動認識システムを提案する。行動が起きた位置を認識するためにミリ波センサを用いてユーザの位置を推定する。またマイクアレイを用いて音源方向を推定することで、複数人の誰が行動しているか推定する。

提案システムの評価として、2つの評価実験を行った。評価実験1では部屋の中の3箇所のうちいずれかの場所において、ハンディークリーナーを利用してもらった。人の位置や行動を推定し評価を行った。評価実験2では、被験者3人が部屋にいる状態で3人のうち1人がハンディーク

<sup>†1</sup> 金沢工業大学  
Kanazawa Institute of Technology, Nonoichi, Ishikawa 921-8501, Japan

<sup>†1</sup> 現在、金沢工業大学  
Presently with Kanazawa Institute of Technology

<sup>a)</sup> minoru.nakazawa.jp@ieee.org

リーナーを利用してもらった。人の位置や行動、音源方向を推定し複数人の場合でも行動を認識し識別できるか評価を行った。

## 2. 関連研究

屋内環境における行動認識に関して様々な研究が行われている。最近ではIoT 機器やスマートスピーカーの普及により活発に研究が行われている。行動認識を行う研究では、カメラ画像や動画を処理し行動認識を行う手法やユーザが所持しているスマートフォンやスマートウォッチ等のウェアラブルデバイスに内蔵されている加速度センサ等を用いて行動認識を行う手法、また多種多様なセンサを多数設置したスマートハウスでの行動認識の手法がある。以下の節では、それぞれの屋内環境における行動認識の手法に関する既存研究について述べる。

### 2.1 カメラを用いた行動認識

カメラを用いた手法では、Jalal ら [2] はデプスカメラを利用し画像処理により、生活行動認識手法を提案している。この研究では、人の骨格を検出することで、料理、テレビを見る、運動、手拍子、歩く、掃除の6つの行動を平均92.33%の精度で認識することができる。しかし複数人の行動は認識することができない。またカメラを利用しているためユーザが障害物などに隠れてしまうと認識することができない。Castro ら [3] は、首から下げたスマートフォンのカメラを使用し、画像をディープラーニングにより料理、食事、運転、テレビの視聴、ショッピングなどの19種類の日常生活行動を83.07%の精度で認識するシステムを構築した。この研究では認識できる行動は多いが、カメラを身に着けるためユーザのプライバシーを侵害する恐れがある。

### 2.2 ウェアラブル端末を用いた行動認識

スマートフォンには、加速度センサ、ジャイロセンサ、マイク、照度センサ、カメラなどの多数のセンサが内蔵されている。

多くの研究では加速度センサを利用し行動認識したものが多く [4]。加速度のみを使用するものでは歩く、走る、階段の上り下り、静止、運動などの身体の状態を表す行動を認識するもの [4] や、自転車に乗っているという大きな括りの行動を認識するもの [6] が多く、詳細な生活行動を認識するものは少ない。Ouchi ら [7] は加速度センサとマイクを用いて行動認識手法を提案している。

加速度センサで歩行、安静、作業の3状態を推定し作業の場合ではマイクを用いて7種類（歯磨き、皿洗い、アイロンがけ、掃除機、トイレ/手洗い、電動髭剃り、ドライヤー）の作業を推定する。歩行、安静、作業の3状態では95%以上、7種類の作業では、平均85.9%の精度で推定

することができる。Shoaib [4] らはスマートフォンの加速度センサと腕に装着した加速度センサを用いることで、歩く、座るといった行動に加えて、食事、PC作業、喫煙、コーヒーを飲むといった生活行動を認識する手法を提案した。行動認識にスマートフォンを用いることで、内蔵されている様々なセンサを利用できることと、ユーザが新たにデバイスを購入する必要がない。そのため導入・維持コストは比較的少ない。しかしユーザが常にデバイスを身に着ける必要があり、ユーザへの負担が大きいという問題がある。さらにバッテリーの交換や充電を行う必要がある。

### 2.3 多数のセンサによる行動認識

多数箇所にセンサを設置したスマートホームを用いて行動認識する手法では、異なる多数のセンサを設置し行動認識する手法 [9] や同じ種類のセンサを多数設置しセンサネットワークを構築し行動認識する手法 [10] がある。Kasteren ら [9] は、ドアや引出の開閉を測定するリードスイッチ、圧力マット、浮力センサ、温度センサなどのセンサが多数設置されたスマートホームにおいて、食事、テレビ観賞、トイレ、シャワー、洗濯、着替え、睡眠などの日常生活行動を認識するシステムを構築した。49~98%の精度で認識ができる。この研究では、認識できる行動の種類は多いが、多くのセンサが必要で導入コストが高く、行動の種類によっては認識の精度が低いという問題点がある。

Chen ら [10] は、圧力センサ、モーションセンサ、チルトセンサ、近接センサなどが多数設置されたスマートホームにおいて、お茶をいれる、歯を磨く、テレビを見る、入浴する、手を洗うといった複雑な生活行動を平均94.44%以上の精度で認識するシステムを構築した。これは、知識ベースのオントロジーにより生活行動を識別しているため、機械学習のような大量の訓練データを必要としないという利点がある。しかし、多数かつ多種のセンサを使っており、導入コストが高いという問題がある。また屋内で電化製品が使用されているかを検知する研究が多く行なわれている。

Gupta ら [11] は、家庭内の一点の消費電力を計測することで、電化製品によって生成されるEMIを用いてどの家電が使用されているかをリアルタイムに分類する手法を提案した。しかし屋内での行動では食事等の電化製品を使用しない行動もあり、使用されている電化製品のみで家庭内の行動全てを認識することは困難である。

いずれのスマートハウスを用いた手法では多数かつ多種のセンサを必要とするため導入コスト及びバッテリー交換などの維持コストがかかるという問題がある

### 2.4 先行研究

Laput ら [1] は流通している多種多様なセンサは、1つのセンサが1つの行動の認識しかできないため多くの行動

を認識しようとする多種類のセンサを複数個所に設置しなければいけない。そのためより汎用的な行動を認識するセンサを構築することを目的に、赤外線アレイセンサ、カラーセンサ、照度センサ、磁気センサ、温度センサ、湿度センサ、気圧センサ、加速度センサ、人感センサ、マイク、EMI センサの多種類のセンサを一か所に設置し行動認識を行っている。温度、湿度、気圧、照度、磁気、WiFi-RSSI、赤外線アレイ、および PIR モーションセンサは 10 Hz でサンプリングされ、3 軸加速度センサの 3 つの軸はすべて 4 kHz、マイクは 16 kHz、EMI センサは 500 kHz でサンプリングされる。10Hz でサンプリングされる低周波のセンサの場合は 7 つ（最小値、最大値、範囲、平均値、合計値、標準偏差、重心値）の統計的特徴と、高周波センサである加速度センサ、マイク、EMI センサは 100 ミリ秒ごとに FFT を行いそれぞれの特徴量に対して機械学習を用いて 38 種類の行動を平均 96% の精度で認識するシステムを構築した。センサを一か所に設置しているため検出された行動がどこで行われているかわからない。またその行動を行ったユーザが複数人なのか 1 人なのかを区別することができない。

### 3. 室内での行動認識における課題

前章でも述べたように屋内での行動認識手法にはそれぞれ以下の課題があった。カメラによるプライバシーの侵害の問題、ユーザがデバイスを常に装着することによる負担やバッテリーの問題、そして導入・維持コストが高いという問題がある。

先行研究では、カメラを使用せず代わりに多種類のセンサを使用することで、プライバシーに考慮している。そして多種類のセンサを一か所に設置することで、バッテリーの問題とユーザがセンサを装着しなければいけない問題を解決している。導入・維持コストでは多数個所にセンサを設置する手法より低コストとなる。しかし、一か所にセンサを設置したことにより、スマートハウスを用いた手法では、推定できていた、どこで何が起きたかの「どこ」が推定できなくなった。また推定された行動が複数人による行動か、一人での行動であるかの区別をすることができない。行動が起きた場所が推定できないため、例えば複数窓がある場合には、ユーザが窓を開けたことは推定できるがどの場所の窓を開けたのか推定できない。このように同じものが複数ある場合に区別することができない。また複数人での行動を区別することができないので、ユーザが 1 人の場合には問題ないが、複数人の場合にはそれぞれの行動が区別できず問題となる。そのため、先行研究を実際のサービスに適用することは難しいと考えられる。

### 4. 複合センサーを用いた行動認識

3 章で述べた先行研究の課題を解決するために、新たに

センサを導入し、ユーザの人数及び位置そして音の発生方向を検出する。先行研究では、人の位置を認識することができず、行動がどこで行われたのか推定することができない、そのため人の位置を検出するセンサを追加することで解決できるのではないかと考えた。

そして複数人での行動を区別できないという課題には先ほどの人の位置及び人数を検出するセンサに加えマイクアレイを追加し音源定位を行うことで複数人の行動を区別し行動を認識できるのではないかと考えた。人の位置や人数を検出できることで、人の人数を考慮した行動認識を行うことができ、そのためより複雑な行動を認識することができると考えられる。例えば、今まで会話していると推定されていた行動でも本システムでは、電話等を使用し通話を行い一人で会話しているのか複数人で会話しているのかを区別することができる。また人の場所を検出できるため、掃除機の使用と位置情報を用いてどこが掃除できていないかを検出することができる。応用としては、複数人が同居する場合での見守りサービスの実現や位置情報を組み合わせることで、人のいるところだけ暖房や照明を使用するといった省エネ制御が考えられる。

#### 4.1 センサーの追加・変更

先行研究で使用していたセンサに、新たにミリ波センサを追加し、マイクを従来の 1 チャンネルマイクから 4 チャンネルマイクアレイに変更した。また電化製品から出る電磁ノイズをセンシングする目的で先行研究では EMI センサを使用していた。しかし本研究では、電波を発するミリ波センサを追加したため EMI センサにミリ波センサからのノイズが混入するため EMI センサを削除した。以下の項にて新たに使用するセンサの説明をする。

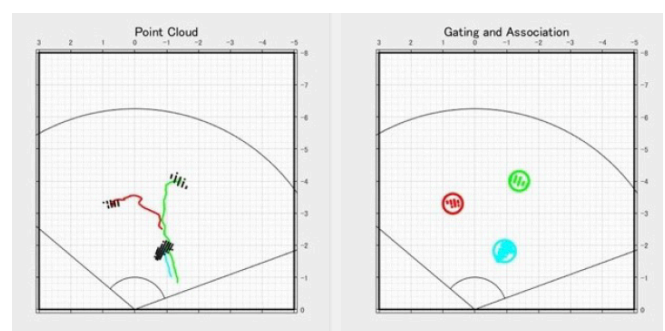


図 1 ミリ波センサによる人数カウントおよび追跡

Fig. 1 Counting and tracking the number of people using millimeter-wave sensors

##### 4.1.1 ミリ波センサ

Texas Instruments 社の AWR1642 BOOST [12] というミリ波センサは自動運転システムにも用いられているミリ波レーダーを 1 チップ、オンボード・アンテナ化し小型の開発ボードにしたものであり 76GHz~81GHz を使用する。

ミリ波センサでは、図1に示すように最大6mの範囲で人の位置および人数を検出し追跡することができる。図1の左の図が人と検出した点群及びその追跡である。右の図が検出した人の位置を表しており点群の重心を中心にして円を表示している。人の位置及び人数の検出にはTI社のリファレンスデザイン[12]及びOpenRadar[13]を参考に作成した。このリファレンスデザインでは水平面で120°最大6mの範囲で最大250個の物体を検出し最大で20人を認識し追跡することができる。またチャープ構成の変更で距離の変更ができる。ミリ波センサは内部に信号処理用のDSPとARM Cortex R4Fマイコンが内蔵されておりミリ波センサ単体で人の位置及び人数を検出することができる

#### 4.1.2 マイクアレイ

先行研究では1チャンネルのマイクADMP401を使用していたが、本研究ではseed社の開発ボードであるマイクが4チャンネル分搭載されたReSpeaker Mic Array v2.0を使用する。このマイクアレイはスマートスピーカーのような音声アシスタントシステムを想定し設計されているため内部にはXMOS XVF-3000プロセッサが搭載されておりマイクアレイ単体でDOA (Direction of Arrival) による音源方向推定やノイズやエコーの低減が行える。

#### 4.2 行動認識デバイス

上記以外のセンサは図2に示すように先行研究と同じものを使用した。デバイス構成を図3に示す。左上から右に3軸加速度センサ(MPU6500)、PIRセンサ(AMN2111)、赤外線アレイセンサ(AMG8833)左下から右に磁気センサ(MAG3110F)、照度/カラーセンサ(TCS34725)、温度センサ/湿度センサ/気圧センサ(BME280)である。これらのセンサはESP32マイコンが搭載されたM5Stackに接続され得られたセンサデータをmicroSDにCSVフォーマットで保存している。サンプリング周期は、加速度センサが100Hz、その他のPIRセンサ、赤外線アレイセンサ、磁気センサ、照度センサ、カラーセンサ、温度センサ、湿度センサ、気圧センサは10Hzとなっている。ミリ波センサは20Hz、マイクアレイは16kHzまた推定した音源方向を10Hzで取得している。WiFiを使用しLAN内のサーバーにデータを送信しておりリアルタイムにセンサデータを確認できるようになっている。

ミリ波センサでは、1.5m以上の高さに設置することが推奨されている[12]。先行研究のようにコンセントに設置する方法では高さが足りなかったため、本棚等に設置できるようなデバイスとした。図4に示すように上部にミリ波センサを搭載することで、推奨される高さを得られるようにした。最上部にはマイクアレイが搭載されている。最下部には人感センサと赤外線アレイが搭載されている。中部にはそれ以外のセンサとセンサに接続されたM5Stackが

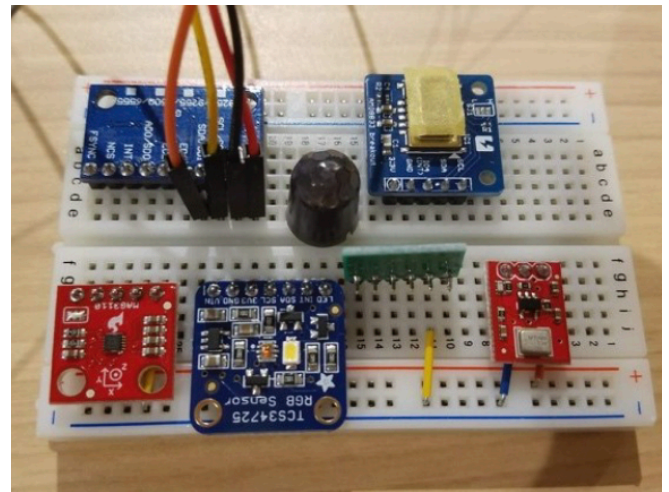


図2 搭載デバイス

Fig. 2 On-board devices

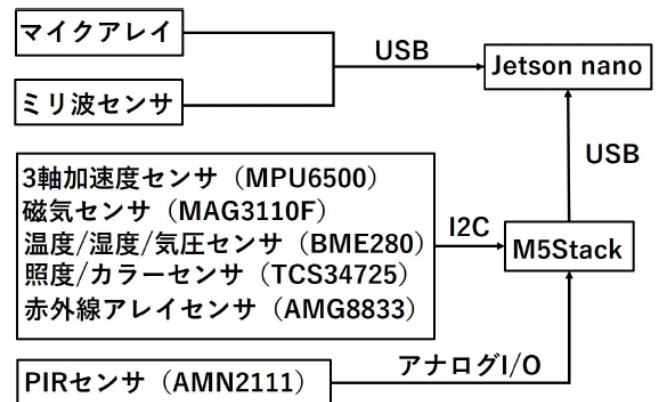


図3 デバイス構成

Fig. 3 Device Configuration

搭載されている。右下には、Jetson nanoが搭載されている。

図3に示すようにJetson nanoにマイクアレイをUSB接続しており4チャンネルでの録音及び内部プロセッサで計算された音源方向を保存している。またミリ波センサも接続しておりミリ波センサで検出された人の位置及び人数を保存している。

ミリ波はプラスチックや紙等を透過するのでセンサの周囲を覆った状態でも使用することができる。そのため、図4の右に示すようにデバイスを覆う外枠を塩化ビニル板で作成した。覆われていると正常に動作できないPIRセンサと赤外線アレイは最下部に設置し外枠の影響を受けないようになっている。こうすることで一般家庭に設置しやすいデザインにすることができた。

#### 4.3 行動認識システム

本節では、行動認識システムについて説明する。図5にシステムの概要を示す。センサ、マイクアレイのデータは特徴量の計算を行い。センサデータはSVM、マイクアレイ



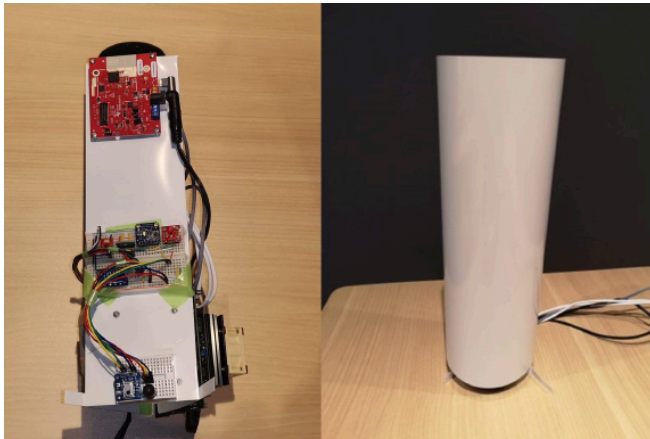


図 4 開発した行動認識デバイス

Fig. 4 Developed Device Configuration

から取得した音データは CNN を用いて行動の推定を行う。またマイクアレイから出力される音源方向とミリ波センサで検出された人の位置、人数を用いて行動が起きた場所、方向を推定する。これらの情報を組み合わせて、複数人の中から行動した人の推定を行う。

#### 4.3.1 学習データの収集

SVM, CNN の学習を行うために被験者 4 名（男子学生）から協力を得て 14 種類の行動時のセンサデータとその様子を動画で撮影した。センサデータの収集には、図 4 の開発したデバイスを使用した。デバイスを机の上に置き 1m 離れた場所で行動してもらった。ESC-50[16] のデータセットを参考に 5 秒間に必ず 1 回以上の行動が記録されるように測定を行った。1 つの行動に対して 1 人 36 回合計 3 分間の測定を行った。測定した 14 種類の行動は、ドリルの使用、ヘアドライヤー、会話、ハンディークリーナー、ドアの開閉、扇風機、ノック、電灯のオンオフ、マウス、何もしていない状態、スプレーの使用、タイピング、歩行、ファスナーの開閉である。

#### 4.3.2 特徴抽出

特徴抽出の詳細について説明する。センサデータの特徴抽出では、 $8 \times 8$  の赤外線アレイセンサは行、列ごとの平均を計算した。その他の温度、湿度、気圧、3 軸ジャイロ、輝度、色温度、3 軸加速度のセンサデータは 1 秒毎に 7 つの統計的特徴（最大、最小、範囲、平均、合計、標準偏差、重心）を計算する。マイクアレイでは、生の 4 チャンネル + マイクアレイの DSP で合成されたノイズ、エコー低減処理の 5 チャンネル、サンプリング周波数は 16kHz で録音している。チャンネルごとに分離し 960ms ごとに分割し 25ms のウィンドウで 10ms ホップ、ハン窓関数を用いて短時間フーリエ変換を行い次に 64 バンドのフィルタバンクで MFCC (メル周波数ケプストラム係数) を計算し  $96 \times 64$  のログメルスペクトログラムに変換する。

#### 4.3.3 SVM (Support Vector Machine)

センサデータの学習には scikit-learn[15] の SVM を用い

表 1 ESC-50 の分類

Table 1 Classification of ESC-50.

動物	環境音	非言語	屋内の音	都市ノイズ
犬	雨	泣き声	ノック	ヘリコプター
豚	火	拍手	タイピング	サイレン
牛	コオロギ	呼吸音	ドアの開閉	クラクション
蛙	鳥のさえずり	咳	缶の開ける音	エンジン
猫	水滴	足音	洗濯機	電車
雌鳥	風	笑い声	掃除機	鐘
昆虫	流水音	歯磨き	目覚まし時計	飛行機
羊	トイレ	いびき	時計	花火
カラス	雷雨	啜る音	ガラス	のこぎり
雄鳥	波音	くしゃみ	マウス	チェーンソー

た。RBF カーネルを使用し、ハイパーパラメータはグリッドサーチを用い求めた。学習には収集した 180 秒  $\times$  14 種類の 2520 個のデータを使用した。このデータを学習用とテスト用に 8 : 2 で分割し学習を行った結果精度は 87.55% となった。

#### 4.3.4 CNN (Convolutional Neural Network)

マイクアレイを用いて録音したデータを使用し行動認識を行う CNN について説明する。CNN モデルは AlexNet[14] を利用した。Nvidia Digits 6 の Model Store を用い ILSVRC2012 データセットで学習済みのモデルを使用しファインチューニングを行った。

CNN14 は、収集したデータのみを用いて 14 クラス分類を行うものである。CNN49 は、収集したデータに加えて ESC-50 データセットを用いて 49 クラス分類を行うものである。追加した ESC-50 データセットは、環境音分類のデータセットであり、ラベル付けされた 2000 個の環境音録音になっている。表 1 に示すように 5 つのメジャーなカテゴリーに分けられており、更に 50 のクラスに分類されている。この 50 個の中から動物カテゴリーの犬、猫以外を削除し、ドアの開閉や、ノックなどの収集した行動と同じものは統合し 49 クラスとした。

収集したデータを使用した CNN14 では Top-1 accuracy が 85.97%、Top-5 accuracy が 98.44% となった。混同行列を図 6 に示す。

ESC-50 データセットを追加した CNN49 では Top-1 accuracy が 78.17%、Top-5 accuracy が 93.53% となった。

Top-1 accuracy は、予測したラベルが実際のラベルである割合を示している。Top-5 accuracy は、予測した上位 5 つにラベルが含まれている割合を示している。

## 5. 評価

提案手法で開発した行動認識デバイスを用いた行動認識システムによって認識された行動及び行動の位置の精度を検証するために、評価実験を行った。実験環境は図 7 に示す部屋を使い、被験者 3 名（大学院生）で行った。その行

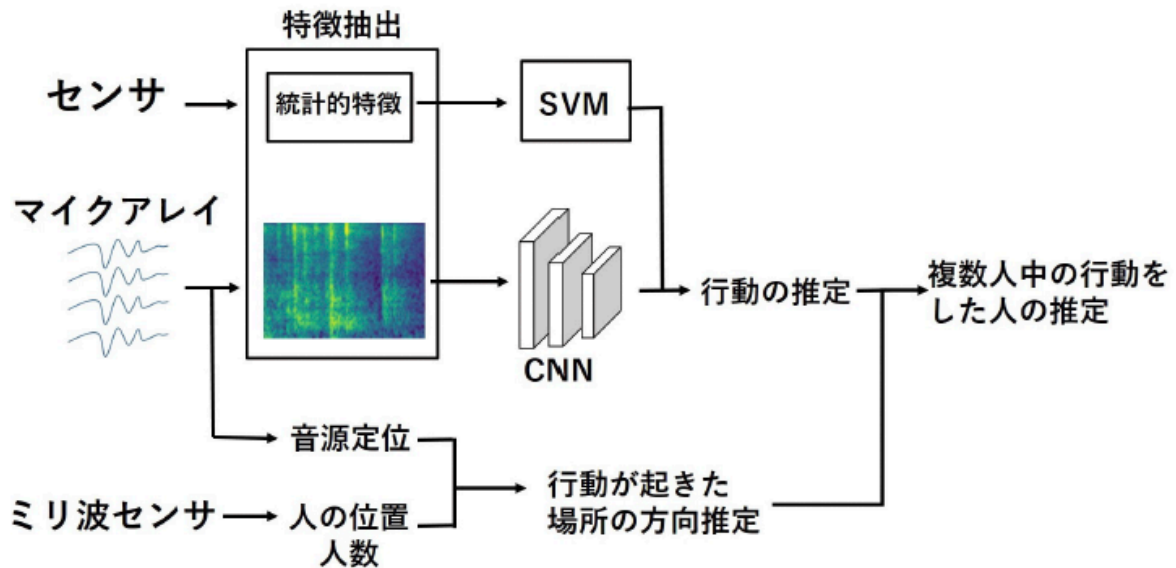


図 5 システム概要

Fig. 5 Summary of System

True \ Predict	Drill	Hairdryer	Talk	Vacuum cleaner	dor	fan	knock	light	mouse	non	spray	typing	walk	zipper
Drill	100	0	0	0	0	0	0	0	0	0	0	0	0	0
Hairdryer	0	100	0	0	0	0	0	0	0	0	0	0	0	0
Talk	0	0	87.5	0	0	0	10	0	0	2.5	0	0	0	0
Vacuum cleaner	0	0	1.333	98.67	0	0	0	0	0	0	0	0	0	0
dor	0	0	0	0	65	0	0	2.5	7.5	12.5	0	2.5	2.5	7.5
fan	0	0	0	0	0	92	0	0	0	8	0	0	0	0
knock	0	0	0	0	2.5	0	87.5	10	0	0	0	0	0	0
light	0	0	0	0	2.5	0	0	87.5	5	7.5	0	2.5	0	15
mouse	0	0	0	0	0	0	0	12.5	77.5	10	0	0	0	0
non	0	0	0.625	0	1.25	0	0	0	0.625	98.88	0	0	0	0.625
spray	0	0	0	0	0	0	0	0	5	2.5	0	92.5	0	0
typing	0	0	0	0	0	0	0	0	32.5	0	0	62.5	0	5
walk	0	0	10	0	0	0	2.5	5	10	0	0	0	72.5	0
zipper	0	0	10	2	0	0	2	12	10	8	0	0	2	54

図 6 14 クラス分類の混同行列

Fig. 6 Confusion matrix of 14 classes

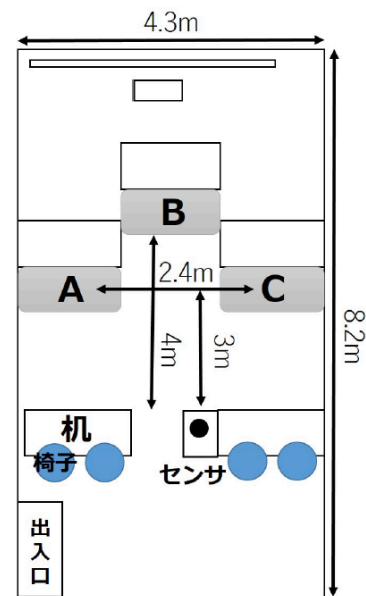


図 7 実験環境

Fig. 7 Experiment environment

動の様子を開発した行動認識デバイスとカメラを用いて記録した。検証実験では、CNN14 を用いて認識精度の高かったハンディークリーナーの行動で評価を行う。

### 5.1 評価実験の方法

検証実験 1 の目的は 1 人での行動の認識と行動場所の推定に対する精度を検証することである。検証実験 1 では図 7 に示す A,B,C のいずれかの場所に立ち机の上をハンディークリーナー用いて 30 秒間掃除してもらった。この流れを A,B,C の 3 か所で 5 回ずつ行った。

検証実験 2 の目的は、複数人が部屋にいる状態で行動を認識できるかの検証、及び複数人の中から機器を使用して

いる人を音源方向を用いての識別精度の検証することである。検証実験 2 では、被験者 3 人が部屋の A,B,C の位置に立っている状態で 3 人のうち 1 人がハンディークリーナーを用いて 30 秒間掃除してもらった。A,B,C の 3 か所で 5 回ずつ計測した。

### 5.2 実験結果と考察

評価実験 1 で測定した A,B,C の 3 箇所で 30 秒× 5 回の 150 秒× 3 人で 450 秒分のデータに対し提案したシステムで行動を推定した結果 A が 98.4%, B が 98.9%, C が 99.1% になり平均 98.8% でハンディークリーナーを使用しているこ

表 2 評価実験 2: 推定結果

Table 2 Evaluation Result 2: Estimation.

	1名	2名	3名
A	0%	20.3%	79.7%
B	15.7%	32.2%	52.1%
C	6.0%	31.8%	62.2%
平均	7.23%	28.1%	64.7%

とを推定できた。間違えて推定された行動はスプレアの使用とヘアドライヤーであった。ハンディクリーナーの電源を付けた直後の定常状態になる間に間違えて推定されていた。そのため電源のオンオフといった過渡現象のデータを増やすことにより精度の向上が見込める。

行動場所の推定として、ミリ波センサでの人の検出と場所の推定をした。人が検出され、なおかつ推定された人の座標が A,B,C の範囲内であった割合である。A が 98.2%, B が 97.7%, 96.3% となり平均 97.4% となった。手を大きく伸ばした時などに一時的に範囲外に検出されてしまうことがあった。

評価実験 1 の結果から 1 人の場合では、部屋の 3 箇所のいずれにいる場合でも高い精度で行動と場所を推定できることが確認できた。

評価実験 2 の行動の推定の結果 A が 98.6%, B が 99.3%, C が 100% となり平均 99.3% となった。部屋に 3 人いる状態であっても行動の推定が行えている。

評価実験 2 の人数の推定結果を表 2 に示す。3 人を同時に検出できた割合は、平均 64.7% となった。1 人のときに比べ精度が落ちた理由として、ミリ波センサでは、検出された動的な点群からクラスタリングを行い人数を検出しているため人同士の間隔が狭くなり点群が重なり合うと分離できなくなるためである。特に B の場合で精度が低いのは B の人が範囲内で動くことで、A または C の人と近づき 1 人として認識されているためである。

マイクアレイで推定された音源方向は平均 98.7% の精度となった。ミリ波センサで検出された人の座標と組み合わせることで平均 63.8% の精度で 3 人の中から行動している人を識別することができた。

## 6. まとめ

本論文では、一か所に設置した複数のセンサを組み合わせた複合センサを用いて、屋内での行動を認識するデバイスの開発とシステムの提案を行い 2 つの検証実験を行った。先行研究で使用していたセンサに加えてミリ波センサとマイクアレイを用いて、複数人の位置と人数、音源方向を推定し認識された行動の位置の推定及び行動している人を識別するシステムを提案した。検証実験の結果、3 人の場合でも行動を平均 99.3% 人の位置を平均 64.7% の精度で認識し 3 人の中から行動している人を平均 63.8% の精度

で 3 人の中から行動している人を識別することができた。

今後の課題として、行動認識の精度の向上を行う必要がある。マウスの使用やタイピングなどの認識精度が低い。これらの行動の精度を向上させるためには、様々な使用パターンでのデータを取得しデータを増やす必要がある。

複数人の検出精度では、3 人をそれぞれ検出できた割合が 64.7% となり精度が低い。これは 2 人を 1 人として認識してしまうためである。そのためアルゴリズムの改善が必要である。提案手法では、音源方向を用いて機器を使用している人の識別をするため同角度に複数人いる場合では、識別することができない。また、複数の行動を同時に認識することができない問題もあり、これらを解決する手法の検討を行う。

## 参考文献

- [1] Laput, G., Zhang, Y. and Harrison, C.: Synthetic sensors: Towards general-purpose sensing, Proc. CHI 17, pp.3986 - 3999, ACM(2017).
- [2] Jalal, A., Kamal, S. and Kim, D.: A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments, Sensors, Vol.14, No.7, pp.11735-11759 (2014).
- [3] Castro, D., Hickson, S., Bettadapura, V., Thomaz, E., Abowd, G., Christensen, H., Essa, I.: Predicting daily activities from egocentric images using deep learning, proceedings of the 2015 ACM International symposium on Wearable Computers pp.75-82, ACM(2015).
- [4] Shoaib, M., Bosch, S., Incel, O., Scholten, H., Havinga, P.: A survey of online activity recognition using mobile phones, Sensors, Vol.15, No.1, pp.2059-2085 (2015).
- [5] Lu, Y., Wei, Y., Liu, L., Zhong, J., Sun, L., Liu, Y.: Towards unsupervised physical activity recognition using smartphone accelerometers, Multimedia Tools and Applications, Vol.76, No.8, pp.10701-10719 (2017).
- [6] Siirtola, P., and Rönning, J.: Recognizing human activities user-independently on smartphones based on accelerometer data, IJIMAI, Vol.1, No.5, pp.38-45 (2012).
- [7] Ouchi, K. and Doi, M.: Smartphone-based monitoring system for activities of daily living for elderly people and their relatives etc, Proceedings of UbiComp 2013, pp. 103-106, ACM (2013).
- [8] Shoaib, M., Bosch, S., Incel, O., Scholten, H. and Havinga, P.: Complex human activity recognition using smartphone and wrist-worn motion sensors, Sensors, Vol.16, No.4, pp.426 (2016).
- [9] Kasteren, T. L., Englebienne, G. and Kröse, B. J.: An activity monitoring system for elderly care using generative and discriminative models, Personal and ubiquitous computing, Vol.14, No.6, pp.489-498 (2010).
- [10] Chen, L., Nugent, C. D. and Wang, H.: A knowledge-driven approach to activity recognition in smart homes, IEEE Transactions on Knowledge and Data Engineering, Vol.24, No.6, pp.961-974 (2012).
- [11] Gupta, S., Reynolds, M. S. and Patel, S. N.: ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home, Proceedings of the 12th ACM international conference on Ubiquitous computing, pp.139-148, ACM (2010).
- [12] Texas Instruments : TIDEP-01000 ミリ波レーダー・セ

ンサを使用する人数追跡/計測のリファレンス・デザイン—TIJ.co.jp, Texas Instruments (オンライン), 入手先 <http://www.tij.co.jp/tool/jp/TIDEP-01000> (参照 2019-5-12)

- [13] Pan, Edwin and Tang, Jingning and Kosaka, Dash and Yao, Ruihao and Gupta, Arjun :OpenRadar, 入手先 <https://github.com/presenseradar/openradar> (参照 2019-11-22)
- [14] Alex Krizhevsky and Sutskever, Ilya and Hinton, Geoffrey E.: ImageNet Classification with Deep Convolutional Neural Networks, NIPS2012 4824, Advances in Neural Information Processing Systems 25, pp.1097-1105 (2012).
- [15] scikit-learn, <https://scikit-learn.org/stable/> (参照 2019-10-12)
- [16] Piczak, Karol J.:ESC: Dataset for Environmental Sound Classification, Proceedings of the 23rd Annual ACM Conference on Multimedia, pp. 1025-1018 (2015)