

複合センサを用いた屋内環境における 行動認識に関する実装と評価

中沢 実^{1,a)} 西川 幸延¹ 阿部 倫之¹

受付日 2020年5月9日, 採録日 2020年11月5日

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や介護者の見守りシステムなど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さかんに行われている. 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, プライバシの問題やデバイスをつねに装着しないといけないという問題, 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 カ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ている. しかし, センサを 1 カ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 1 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.

キーワード: 行動認識, 位置認識, 音認識

Implementation and Evaluation of Behavior Recognition Indoor Environment Using Composite Sensor

MINORU NAKAZAWA^{1,a)} YUKINOBU NISHIKAWA¹ NORIYUKI ABE¹

Received: May 9, 2020, Accepted: November 5, 2020

Abstract: In order to realize services that life support such as home automation such as energy-saving control of home appliances and monitoring systems for the elderly, research on automatic recognition of living behaviors of indoor persons has been actively conducted. Various studies have been conducted on indoor behavior recognition, including methods using cameras, methods using wearable devices, and methods using smart homes. However, there are privacy issues and the need to always wear the device. There is a problem that introduction cost and management cost are high. In order to solve this problem, in the previous research, 10 kinds of sensors were installed in one place, and a general-purpose sensor that recognizes 38 kinds of actions by machine learning was constructed. However, since the sensor is installed in one place, it is not possible to distinguish where the action is taking place and whether the action is an action by multiple persons or an action by one person. In this research, to solve these two problems, we added a millimeter-wave sensor that can detect the position and number of people, and changed the 1-channel microphone to a 4-channel microphone array to solve the two problems. We propose a recognition system.

Keywords: behave recognition, position recognition, sound recognition

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など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さ

¹ 金沢工業大学
Kanazawa Institute of Technology, Nonoichi, Ishikawa 921-8501, Japan

^{a)} minoru.nakazawa.jp@ieee.org

かに行われている。さらに、これらの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が前提とされているケースが多い。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、多種多様の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ら [1] は 10 種類のセンサを 1 カ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システムを構築している。しかし、センサを 1 カ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 1 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、複数人の中からどこでどのような行動をしているかを推定する。

提案システムの評価として、2 つの検証実験を行った。検証実験 1 では部屋の中の 3 カ所のうちいずれかの場所において、ハンディークリーナを利用してもらい人の位置や行動を推定し評価を行った。検証実験 2 では、被験者 3 人が部屋にいる状態で、3 人のうち 1 人にハンディークリーナを利用してもらい、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など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など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] はデブ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、画像の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、カメラを身につ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数存在する [4], [5], [6]。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、スマートフォンの加速度センサを用いてウォーキングやバスケットボールなどの身体活動を認識する教師なし手法を提案するもの [5]、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た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）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。

Shoaiら [8] 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ている。

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、ユーザが新たにデバイスを購入する必要がないなど導入・維持コストは比較的少ないと考えられる。しかし、ユーザがつねにデバイスを身につける必要があり、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、ドアや引出し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、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、49～98%の精度で認識ができている。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われている。Gupta ら [11] は、家庭内の 1 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など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すべ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およ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箇所に設置しなければならないとし、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 1 カ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。10 Hz 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 を行っており、それぞれの特徴量に対して機械学習を用いて 38 種類の行動の識別を、平均 96%の精度で認識するシステムを構築した。

しかしながら、センサを 1 カ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分からない。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など、複数人の中からどこでどのような行動しているかを推定するには不十分であった。

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つね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、プライバシーを考慮している。そして多種類のセンサを 1 カ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ならないという問題を解決している。導入・維持コストでは多数箇所にセンサを設置する手法より低コストとなる。しかし、1 カ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において推定できていた、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、1 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、たとえば複数窓がある場合では、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 人の場合では問題ないが、複数人の中からどこでどのような行動をしているかを推定できず問題となる。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、ユーザの人数およ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およ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。たとえば、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、電話などを使用し通話を行い1人で話しているのか複数人で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。次に、センサの追加・変更にとまなうメリット・デメリット・トレードオフを述べる。
 メリット ミリ波レーダにより人の位置を推定し、マイクアレイによりどの位置からの音であるのかの判断が可能となる。

デメリット 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。

トレードオフ アクティブセンサであるミリ波センサを追加したためパッシブセンサである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センサを除去する必要がある。

本研究の1つの目的が「複数人の中からどこでどのような行動しているかを推定する」であるため、電化製品のON/OFFを用いて人の行動認識として実現するより、メリットで記載したミリ波レーダとマイクアレイによる認識の重要度を採用した。

以下の項で新たに使用するセンサの説明を行う。

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を使用している。今回は室内での利用という点を考慮にいととも、消費電力もそれほど過大でないミリ波センサ [12] を選択した。

当該センサは、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およびその追跡である。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。人の位置および人数の検出はTI社のリファレンスデザイン [12] および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

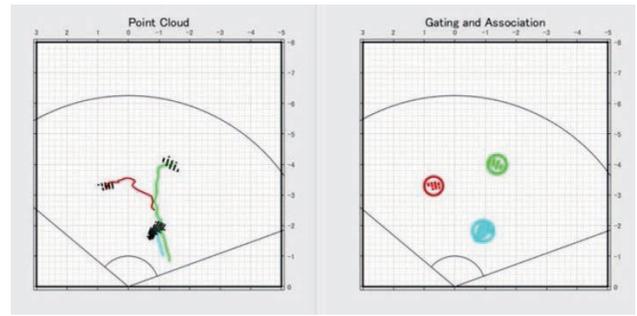


図1 人数カウントおよび追跡

Fig. 1 Counting and tracking the number of people.

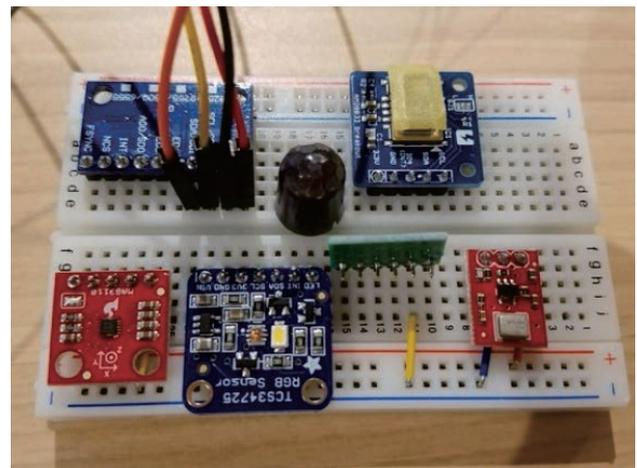


図2 搭載デバイス

Fig. 2 On-board devices.

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^{*1}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^{*2}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およ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XMOSE 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を行うことができるため、これを選定した。

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),

*1 周波数が時間とともに連続して変化すること

*2 Digital Signal Processor

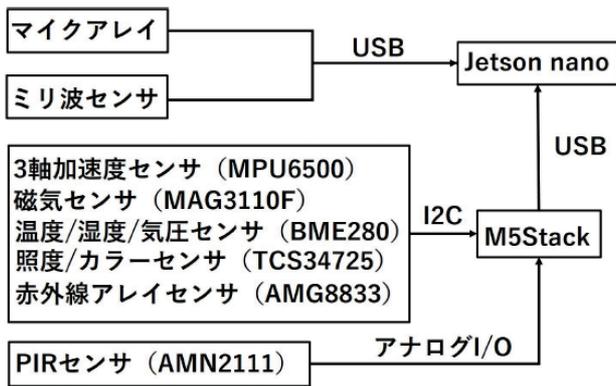


図 3 デバイス構成
Fig. 3 Device configuration.

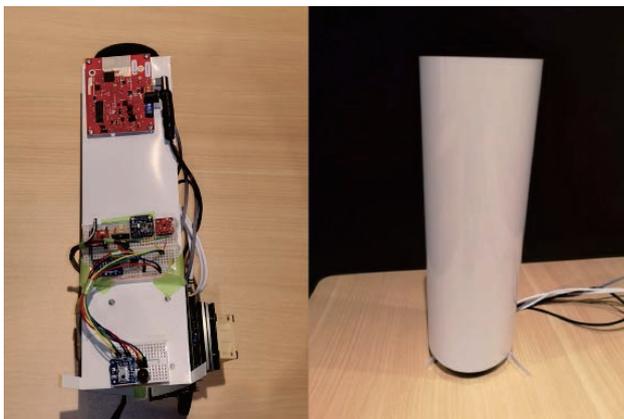


図 4 開発した行動認識デバイス
Fig. 4 Developed device configuration.

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 100 Hz, 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 10 Hz となっている。ミリ波センサは 20 Hz, マイクアレイは 16 kHz また推定した音源方向を 10 Hz で取得している。WiFi を使用し LAN 内のサーバ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 [1] 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, 本棚など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

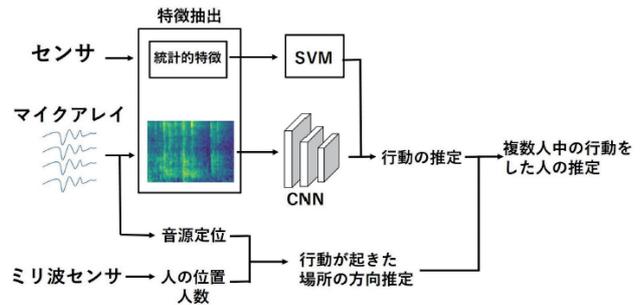


図 5 システム概要
Fig. 5 Summary of system.

が搭載されている。右下部には, 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およ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および人数を保存している。

ミリ波はプラスチックや紙など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 (高さ: 30 cm, 幅: 5 cm) にすることができた。

なお, 本行動認識デバイスは, エンジンとなる Jetson Nano のスタータキットや I2C やアナログ I/O をまとめる M5Stack を含めて, すべて既成部品で構成されており, 原価として 1 台あたり 8 万 5 千円程度となっており, 低コストな構成としている。しかしながら, ミリ波レーダの価格が構成器材の半分の価格を占めている点や Jetson Nano についても内部ソフトウェア構成を考慮し, Raspberry Pi などの安価なマイコンを活用することで 6 万円程度まで低減できるものと考えている。

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、室内において行動してもらった。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名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、会話、ハンディークリーナ、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、ファスナの開閉である。

なお、14種類の行動とした理由として、対象学生が研究や日常的に室内で行う作業とし、センサデータや環境音からラベル付け可能なものに絞った結果である。

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、8×8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秒ごとに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96×64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×14種類の2,520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2,000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に分けられており、さらに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の犬、猫以外

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 |
| 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 |

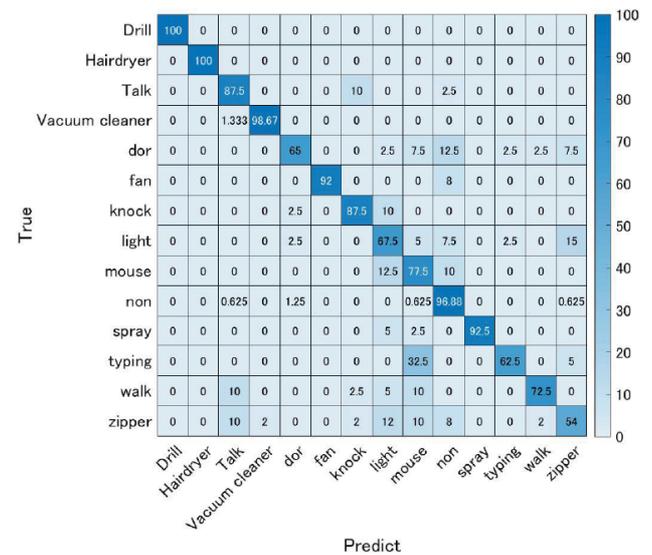


図6 14クラス分類の混同行列

Fig. 6 Confusion matrix of 14 classes.

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。

4.3.5 行動推定の統合

上述したとおり、センサ群から得られた情報はSVMを用いての行動認識を行うとともに、マイクアレイから得られた音声情報よりCNNを用いて行動認識をそれぞれ実施している。これらの結果については、暫定的ではあるが、以下のように統合を行っている。

SVMとCNNで異なる結果が出てくる場合は、CNNのTop-5 accuracy内にSVMの結果があれば、それを採用し、

それがない場合は、CNNのTop-1 accuracyを採用している。これにより、14種類の行動については、91.5%となり、90%を超える結果精度を得た。しかしながら、この手法の最適化については今後さらに検討する必要があると考える。

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および行動の位置の精度を検証するために、検証実験を行った。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の行動で評価を行う。

5.1 検証実験の方法

検証実験1の目的は、1名の行動に対して、その行動場所の推定に対する識別精度を検証することである。検証実験1では被験者1名が、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を用いて30秒間掃除してもらった。この行動をA, B, Cの3カ所で5回ずつ行った。

検証実験2の目的は、複数人の行動に対して、その行動場所の推定のための識別精度を検証することである。検証実験2では、被験者3名が部屋のA, B, Cの位置に立っている状態で3名のうち1名が机の上をハンディークリーナを用いて30秒間掃除してもらった。この行動をA, B, Cの3カ所で5回ずつ計測した。

5.2 検証結果と考察

検証実験1で測定したA, B, Cの3カ所で30秒×5回の150秒×3名で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を使用していることを推定できた。間違えて推定された行動はスプレーの使用とヘアドライヤであった。ハンディークリーナ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%, Cが96.3%となり平均97.4%となった。手を大きく伸ばしたとき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。

検証実験1の結果から1名の場合では、部屋の3カ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。

検証実験2の行動推定の結果については、Aが98.6%,

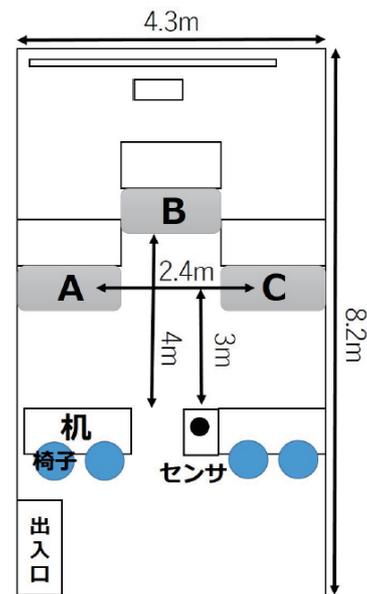


図7 実験環境

Fig. 7 Experiment environment.

表2 評価実験2: 3名中の推定結果

Table 2 Evaluation result 2: Estimated results out of three person.

| | 1名 (誤認識) | 2名 (誤認識) | 3名 (正認識) |
|----|----------|----------|----------|
| 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% |

Bが99.3%, Cが100%となり平均99.3%となった。部屋に3名いる状態であっても行動の推定が行えている。

検証実験2の人数の推定結果を表2に示す。3名を同時に検出できた割合は、平均64.7%となった。1名しか認識できなかった割合は平均7.23%, 2名しか認識できなかった割合は平均28.1%となった。1名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名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名の中から行動している人物を識別することができた。

6. まとめ

本論文では、1カ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および行動している人を識別するシステムを提案した。

本研究は、「複数人の中からどこでどのような行動をしているかを推定する」を最終目標としており、行動推定と、行動が行われた場所推定を複数人の中から認識する必要がある。そのため、本論文では、行動推定の精度と位置推定の精度をそれぞれ求めて、提案手法の有効性を明らかにした。

先行研究論文の Laput ら [1] の 38 種類の行動を平均 96% の精度で認識可能なシステムに比べて、本論文では、14 種類の行動ではあるが 90% 以上と匹敵するものとなっており、ある程度の精度は確保できている。これに加えて、本論文では、行動場所の推定については、検証実験の結果、3 名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 名の中から行動している人物を平均 63.8% の精度で識別することができている。

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。特に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。また、行動場所の推定精度の向上については、室内での行動認識場所であり、この精度を向上させることは、どこで問題やその行動が生じているのかを明示することが可能となり、重要な項目と考えている。今回の固定座標位置に対する推定精度だけでなく、今後はどの程度の解像度（分解能）まで達成できるのかを検討していく必要があると考えている。

また、今回の実験（データ取得・評価実験）については、障害物を含むノイズが少ないレベルの環境で実施している。現在、これらノイズに対する対処方法として、Yamashita ら [17] が提案している主成分分析（PCA）+独立成分分析（ICA）+経験値モード分解（EMD）を組み合わせて、主要な周波数成分を取り出して、再評価することも考えられる。しかし、今回実施したマイクアレイを用いた録音データは CNN を用いて学習されており、ノイズと CNN の関係性をいまだ見いだせていない状態である。今後、各種センサからのデータや音響データに対する結果の寄与率をベースとした学習手法を検討する必要があると考える。

さらに、複数人の検出精度では、3 名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 名を 1 名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名いる場合では、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という問題もあり、今後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. and Essa, I.: Predicting daily activities from egocentric images using deep learning, *Proc. 2015 ACM International Symposium on Wearable Computers*, pp.75–82, ACM (2015).
- [4] Shoib, M., Bosch, S., Incel, O., Scholten, H. and 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. and 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, *Proc. UbiComp 2013*, pp.103–106, ACM (2013).
- [8] Shoib, 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, p.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 Trans. 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, *Proc. 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 2019-05-12).
- [13] Pan, E., Tang, J., Kosaka, D., Yao, R. and Gupta, A.: OpenRadar, available from (<https://github.com/presenseradar/openradar>) (accessed 2019-11-22).
- [14] Krizhevsky, A., Sutskever, I. and Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks, *NIPS2012 4824, Advances in Neural Information Processing Systems 25*, pp.1097–1105 (2012).
- [15] scikit-learn, available from (<https://scikit-learn.org/stable/>) (accessed 2019-10-12).
- [16] Piczak, K.J.: ESC: Dataset for Environmental Sound Classification, *Proc. 23rd Annual ACM Conference on Multimedia*, pp.1015–1018 (2015).
- [17] Yamashita, M., Nakazawa, M., Nishikawa, Y. and Abe, N.: Examination and It's Evaluation of Preprocessing Method for Individual Identification in EEG, *Journal of*

Information Processing, Vol.28, pp.239–246 (2020).



中沢 実 (正会員)

1993年金沢工業大学大学院工学研究科情報工学専攻博士前期課程修了。同年富士通研究所入社。1996年金沢工業大学助手。2011年同大学教授となり、現在に至る。博士(工学)。自律分散システム、ロボティクス、信号処理等に関する研究に従事。2016年情報処理学会山下記念研究賞受賞。電子情報通信学会、IEEE各会員。



西川 幸延 (正会員)

1981年大阪大学基礎工学部制御工学科卒業。北陸日本電気ソフトウェア(株)入社。各種製品の組み込みソフト開発に従事。2014年より金沢工業大学基礎教育部基礎実技教育課程教授。PBL教育、組み込みシステム・IoTシステム教育を担当。



阿部 倫之 (正会員)

1989年金沢工業大学大学院工学研究科情報工学専攻博士後期課程満期退学。同年(株)日立製作所入社。1991年金沢工業大学情報工学科助手。1998年同大学助教授。2006年起業。2012年同大学准教授となり、現在に至る。工学博士。自律分散システム、SNS解析等に関する研究に従事。電子情報通信学会会員。