

# 人動作認識および物体認識を用いた置き忘れ支援

梁泳成<sup>1</sup> 谷本鯛介<sup>2</sup> 佐野睦夫<sup>1</sup> 安田清<sup>3</sup>

<sup>1</sup>大阪工業大学情報科学部

<sup>2</sup>大阪工業大学大学院情報科学研究科

<sup>3</sup>千葉労災病院

**概要：** 認知症ケアにおける忘れ物防止支援サービスを実現するうえで、置き忘れの検出は重要な要素である。本研究では、認知症ケアサービスの例として、ロボットを用いた忘れ物防止支援を行うための置き忘れ検出を挙げる。本稿では、忘れ物防止支援サービスにおける置き忘れ検出に着目し、研究を進めた。本研究では、置き忘れ検出において、人動作認識および物体認識を用いて、置き忘れ検出・人動作認識・物体認識それぞれに対して評価を行う。

## 1. はじめに

現在の日本では高齢者人口の増加による社会的問題が顕在化している。日本の総人口は平成 27 (2015) 年 10 月 1 日の時点において、1 億 2,711 万人で、その内の 65 歳以上の高齢者数は、日本の総人口の 26.7%となる 3,392 万人となった[1]。高齢者数の増加に伴い、介護者不足が深刻な問題となっている。2015 年、厚生労働省は 2025 年度には介護職員が全国で約 38 万人不足するという推計を発表した[2]。

また、2015 年 1 月 7 日、厚生労働省によると、2025 年には認知症患者数が約 700 万人に達すると推定された。この数値は、2025 年における 65 歳以上の推定高齢者数の 5 人に 1 人が認知症を患っていることになる。認知症患者に対する介護者が不足するだけでなく、介護者の負担も増加することが考えられる。そのため、認知症の方の支援、介護者の負担軽減のための新たな介護方式が必要となる。

本研究では、ロボットによる認知症患者を対象とした「忘れ物防止支援」を用いて、二次的要因による認知症の心理行動症状(BPSD: Behavioral and Psychological Symptoms of Dementia)を軽減する方式を検討する。

ロボットを用いた忘れ物支援における従来研究において、我々は BPSD のうちの「物盗られ妄想」に着目し、それを回避するための忘れ物支援システムの方式を提案してきた。谷本らは、RGB-D カメラである Xtion およびロボット Kobuki による「人追従機能」、QR コードや音声での物体認識を用いた「物の預かり機能」を用いた忘れ物支援の提案・改良が行われてきた[3]。

本研究では、ロボットが対象者の部屋の環境地図の作成、人の追従、自己位置推定および自律移動が可能であることを前提として、忘れ物支援の要素の 1 つである「置き忘れの検出」に着目した。「置き忘れ検出」では、2014 年度に南口が人の手と腰の距離の変化から置き忘れを未然に防ぎ、注意の仕方によるユーザのストレス変化を研究している[4]。しかし、手と腰間の距離変化で置き忘れ動作検出を行うと、ユーザが置き忘れていなくても手を伸ばしただけで置き忘れ動作として認識されてしまう問題があった。

本研究では、人動作認識および物体認識技術を用いた置き忘れ検出の方式を提案する。

## 2. 提案手法

置き忘れ検出において、人が「特定のものを取る・置く」ことを認識することが重要な要素であることから、人動作認識および物体認識を用いて置き忘れの検出を行う。

また、人の状態把握の性能を検証するために、置き忘れランキング上位とされる「眼鏡」、「財布」、「携帯・スマホ」、「鍵」、「保険証・鞆」のうち、「眼鏡」、「財布」、「鞆」に対して学習を行い、

- 1) 物体認識
- 2) 人が特定のものを置くこと
- 3) 人が特定のものを特定の場所に置いていないこと
- 4) 人がものを取ること
- 5) 人がものを持ち続けている

の物体認識および 4 つの状態のセンサデータを置き忘れの事例と定義した。以下に、物体認識および置き忘れの事例である 5 つの認識対象の認識の流れを示した。

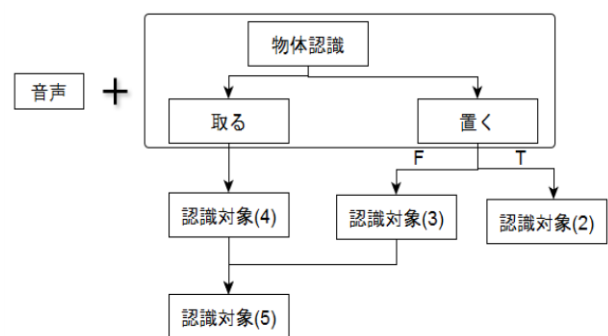


図 1 置き忘れ検出要素の流れ

図 1 では、物体認識結果に基づき「取る」・「置く」のいずれかの人動作を検出した場合、物体認識結果および人動作認識結果から「人がものを取ったかどうか」、「人が特定のものを置いたかどうか」を認識する。この 2 つの認識結果に対して、さらに音声ナビゲーションを行うことで、人がものを持ち続けていたり、人が特定のものを置いていない

ことを認識が可能になると考えられる。なお、ロボットはユーザを常に追従し、ユーザを取り巻く環境情報は既知のものとした。

この置き忘れ検出における5つの認識対象の認識結果およびロボットの会話機能を組み合わせることで、

- 1) 置き忘れ
- 2) 習慣による対象者の特定物の置き場所
- 3) ユーザが特定物を持っているかどうか

の3つの検出が可能となる。

ロボットの会話機能により、人の「置く」動作を検出した際にその動作が「ただ置いた」のか「置き忘れている」のかを判別することが可能であり、ユーザが探し物をしている際に「人が特定のものを置いていない」ことからユーザが特定物を持っているかどうか認識することが可能であると考えられる。

### 3. 物体認識

#### 3.1 物体認識の概要

置き忘れ検出における物体認識には、ロボットが物体の位置を把握するために、3次元点群にて物体検出を行う。検出された物体の重心座標を取得することで、ロボットが物体のキャプチャを行うために能動的に移動することを想定した。置き忘れ検出に用いる物体認識の流れを図2に示した。

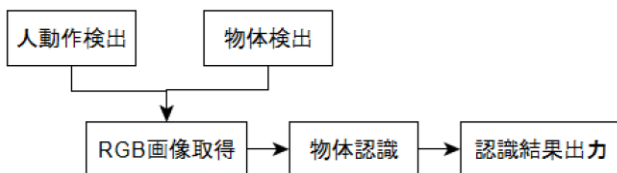


図2 物体認識の流れ

図2では、人動作認識、物体検出、物体認識を異なるモジュールとして用意し、人動作の検出および物体検出が行われた場合にRGB画像としてキャプチャを行い、取得したRGB画像に対して特定物体認識を行う。物体認識を2次元で行う理由として、3次元で物体認識を行う場合、RGB-Dカメラが搭載されたロボットから遠くに位置する物体の3次元点群は、RGB-Dカメラから遠くにあればある程、点群が荒くなり認識が困難であるということ、眼鏡などフレームが細くレンズが透明である物体や光沢物の物体の検出および認識が困難であるため、物体認識には3次元点群ではなく2次元画像を用いた。

#### 3.2 物体検出部

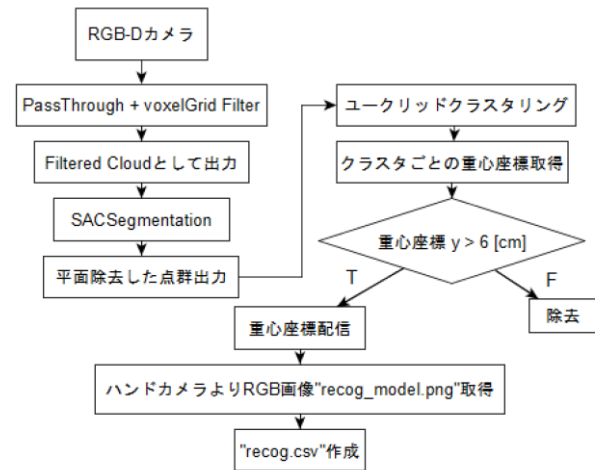


図3 物体検出部の流れ

図3では、RGB-DカメラであるXtionからのセンサデータをリアルタイムに入力し、PCLライブラリを用いて3次元点群として物体検出を行った。まず入力データに対して、PassThroughFilterを用いて入力点群の範囲を指定し、VoxelGridFilterによって点群の間隔を1cmに間引き、FilteredCloudとして出力する。FilteredCloudに対してSACSegmentationによる平面の検出・除去を行う。平面除去された点群に対してユークリッドクラスタリングにより1つの物体を1つのクラスタとして分け、各クラスタの重心座標を計算する。棚の正面部分を物体として検出対象外とするために、計算した重心座標から物体の高さを計算し高さが6cm未満のクラスタは除去する。高さが6cm以上のクラスタは物体として重心座標に基づきRGB画像としてキャプチャを行い、物体認識モジュールのトリガーとなる「recog.csv」ファイルを作成する。取得したRGB画像に対して特定物体認識を行う。なお、物体のキャプチャ画像「recog\_model.png」は物体認識を行うたびに更新する。物体検出の実行例を図4に示した。

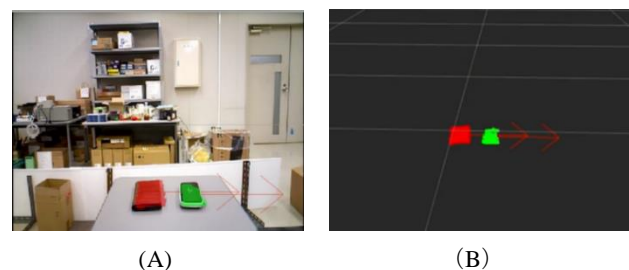


図4 物体検出結果

図4に示したように、画像(A)はRGB-DカメラによるRGB-D映像を入力としたときのRGB画像である。画像(B)では平面に置かれた財布と眼鏡ケースが検出されていることが確認できる。また矢印は原点を検出された各物体の重心として描画しており、検出された財布と眼鏡ケース

それぞれに物体の重心から矢印が描画され、重心がそれぞれ検出されていることが確認できる。以下の図 5 に検出された物体の重心座標を示した。

```
[x y z] = [0] [-0.0262297, 0.219346, 0.804883]
error no
[x y z] = [1] [0.11105, 0.219265, 0.782054]
min : y:0.20481 z:0.72740 max : y:0.24466 z:0.72740
min : y:0.20631 z:0.72636 max : y:0.23562 z:0.72636
```

図 5 重心座標

図 5 では、黄枠の配列の添え字[0]が図 4 の画像(B)における赤色で示された財布の重心座標を表しており、赤枠の添え字[1]の配列が図 4 の画像(B)における緑色で示された眼鏡ケースの重心座標である。min は検出された物体の重心からの物体の底までの長さ[m]であり、max は検出された物体の重心から頂点の長さ[m]である。y は検出された物体の高さ[m]であり、z は RGB-D カメラからの物体の距離である。それぞれ上から添え字[0]および[1]の各値を示している。この座標を基にロボットが物体を撮影可能な位置まで移動するとして、以下の図 6 に平面に置かれた財布に対して取得した RGB 画像を示した。

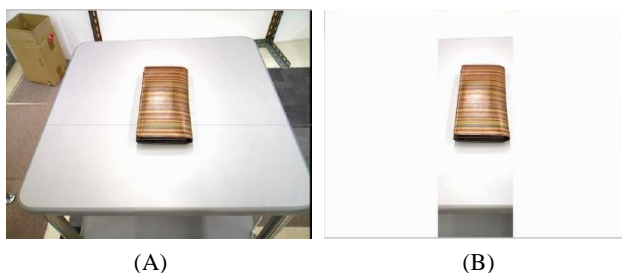


図 6 RGB-D カメラより取得した RGB 画像

図 6 に示したとおり、画像(A)は RGB-D カメラの映像により取得した元画像であり、画像(B)は RGB-D カメラ映像よりキャプチャした物体認識の入力画像である。画像(A)では画像(B)の量幅 200[pix]を白色に塗りつぶした画像である。この取得した RGB 画像に対して物体認識を行う。

### 3.3 物体認識部

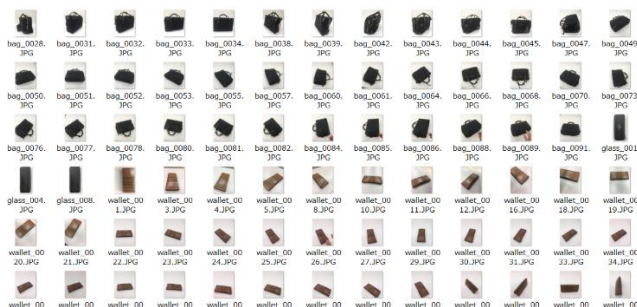


図 7 学習画像群

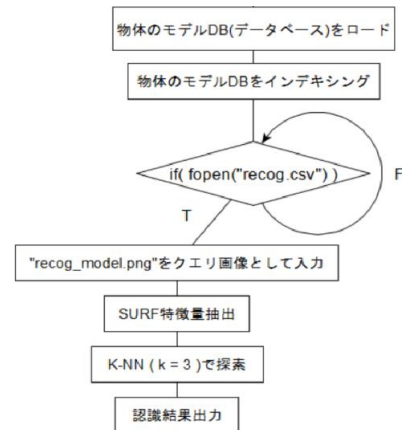


図 8 物体認識部の流れ

図 8 の物体認識部では、図 7 に示したように物体認識のための「鞆」、「財布」、「眼鏡」からなる学習画像群を物体モデルデータベースとして作成しておき、物体認識部のモジュールが実行されると物体モデルデータベースを読み込み、インデキシングを行う。物体が検出され、入力画像「recog\_model.png」を取得し、「recog.csv」が作成されるまで物体認識部を待機する。「recog.csv」が作成されると「recog\_model.png」をクエリ画像として入力し、SURF 特徴によるマッチングを行う。K=3 として K-NN による探索を行い、対象物体と近似する物体を物体モデルデータベースより識別結果として物体モデルデータベース内の画像名および識別結果を表示する。以下に、物体認識を行った実行例を図 9 に示した。

図 6 の画像(B)に示した物体認識に用いた入力画像に対して k-nn 法を用いて k=3 として物体認識を行った。以下の図 7 に物体認識の結果を示した。

```
クエリのキーポイント数: 259
Pass -2
Pass -1
識別結果: wallet_0031.JPG
財布
next query, y or n? :
```

図 9 物体認識結果

図 9 では、入力画像に対する SURF 特徴点数および物体モデルデータベースより、特徴点が近似する画像が識別され図 6 画像(B)の財布の画像に対して、財布と認識できていることが確認できる。

## 4. 人動作認識

### 4.1 人動作認識の概要

人動作認識には、skeleton\_marker による骨格検出および TanvirParhar による neuro\_gesture\_recognition モジュールを用いた。neuro\_gesture\_recognition モジュールとは、入力、中間、出力からなるフィードフォワードニューラルネット

ワークを用いた手動作認識モジュールである。OpenNI および PyBrain を依存ライブラリに持つ。

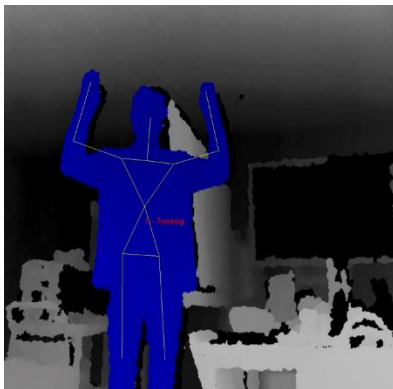


図 10 骨格検出実行例

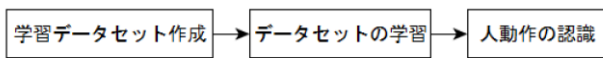


図 11 人動作認識の構成

本研究では図 11 に示したように、人動作認識の構成は 3 つに分けることができる。学習データセットの作成では人動作認識のための図 10 の skeleton\_marker によって検出された骨格情報に基づいて、胴体の重心を原点としたときの右手の x, y, z の座標を入力とし、右手の座標の変化を 1 動作につき 15 回対象の csv ファイルに 3.5 秒間記録する。記録された座標のデータセットに対して、バックプロパゲーションアルゴリズムを用いたフィードフォワードニューラルネットワークにて学習を行う。学習後、学習データに基づき図 10 の骨格情報から右手座標を入力として動作認識を行う。

本研究では、「置く」、「取る」、「四角を描く」、「三角を描く」、「円を描く」の動作に対して学習データセットを用意し、「置く」、「取る」の 2 つ動作に対して認識を行った。以下に、実際の人動作認識実行例を図 12 示した。

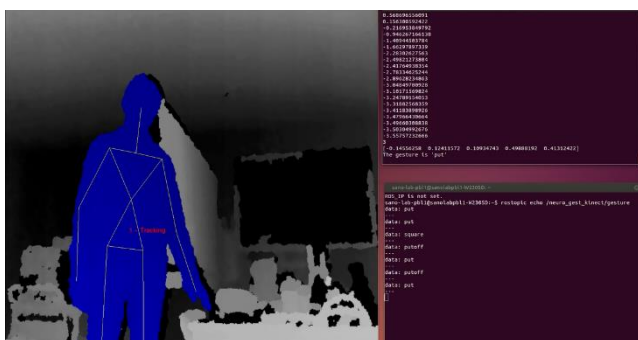


図 12 人動作認識実行例

図 12 では、「置く」の動作に対して人動作認識を行い、右手座標の変化を画像の右上のターミナルに表示し、右下のターミナルにて動作認識結果を表示している。

## 5. 実験・結果と考察

### 5.1 実験内容

本研究では、物体認識、人動作認識、置き忘れ検出のそれぞれに対して認識率を求める。置き忘れにおける認識対象の事例を表 1 として以下に示した。

表 1 置き忘れの事例

認識対象	
認識対象1	人が特定のものを置くこと
認識対象2	人が特定のものを特定の場所に置いていないこと
認識対象3	人が特定のものを取ること
認識対象4	人が特定ものを持ち続けている

表 1 に示したように、本実験で行う置き忘れの認識対象には 4 つの置き忘れの事例がある。この 4 つの置き忘れの事例のうち認識対象 1、認識対象 3 に対して、物体認識および人動作認識を用いて検出が可能かどうかを検証する。

認識対象 1 の実験方法として、ものが置かれていない平面上に人が「鞆」、「財布」、「眼鏡ケース」のいずれかの物体を人が置く。このとき、「置く」の人動作および物体認識より、人動作認識および物体認識が正しく認識できているかを検証する。

認識対象 2 では、ユーザの「鞆」、「財布」、「眼鏡」の 3 つの習慣による置き場所をロボットが既知であるとする。ユーザの習慣の置き場所とは異なる場所にて、「人が特定のものを置いたこと」を検出した場合、人が特定のものを特定の場所に置いていないと定義する。本実験では、対象の環境情報が既知であることから、認識対象 1 の認識が可能であった場合に認識対象 2 の認識も可能であるといえる。本実験では環境情報の認識は行っていないため、認識対象 2 に対する認識率は求めないこととした。

認識対象 3 の実験方法として、ものが置かれていない平面上に「鞆」、「財布」、「眼鏡ケース」の 3 つの物体から 1 つの物体を配置する。配置された物体を人が取る。このとき、認識対象 1 と同様に人動作認識および物体認識が正しく認識されているかを検証する。

認識対象 4 では、認識対象 3 が検出されたと仮定して、一定時間ユーザが「置く」動作を行わず、探し物をしていない場合、音声を用いて探している物をユーザ本人が持っているかどうか質問する。ユーザの応答内容により、人が特定のものを持ち続けているかどうかを決定する。

### 5.2 実験結果と考察

本研究では、物体認識、人動作認識、表 1 に示された認識対象 1 および認識対象 3 のそれぞれに対して認識率を求めた。以下に、「鞆」、「財布」、「眼鏡」の 3 つの物体の認識率を表 2 に示した。

表 2 物体認識率

	試行回数(回)	認識率(%)
靴	10	10
財布	10	90
眼鏡	10	40

表 2 に示したように、「財布」の認識率は 90% の精度で認識が可能であった。「靴」、「眼鏡」の認識率は、それぞれ 10%、40% であった。「財布」の認識率と比較して、「靴」および「眼鏡」の認識精度が低い理由として、物体認識に用いた「靴」と「眼鏡ケース」の色がともに黒色であり、個体差が多いことが考えられた。今後の研究として学習サンプルを増やし改善する必要がある。ここで、人動作認識の認識率を表 3 として以下に示した。

表 3 人動作認識率

	試行回数(回)	認識率(%)
置く	10	80
取る	10	90

人動作認識では、「置く」、「取る」の 2 つ動作に対して認識を行った。表 3 に示したように、「置く」、「取る」の 2 つの動作に対して、それぞれ 80%、90% の精度で認識が可能であった。「置く」と「取る」動作の動作開始位置に対して、明確に差分をとったため、高精度での認識率で実現可能であったと考えられた。ここで、置き忘れ検出における認識率を表 4 として以下に示した。

表 4 置き忘れ検出要素認識率

	試行回数(回)	認識率(%)
認識対象1	10	60
認識対象3	10	40

置き忘れの検出では表 1 に示したように、置き忘れを 4 つの事例に分け定義した。表 4 では表 1 に示された認識対象 1 および認識対象 3 に対する認識率を示した。

認識対象 1 の「人が特定のもの置くこと」は 60%、認識対象 3 の「人が特定のもの取る」は 40 パーセントの精度で認識が可能であった。認識対象 1 および認識対象 2 はともに物体認識および人動作認識に依存しており、物体認識および人動作認識の結果から、物体認識において「財布」以外の物体の認識を改善する必要があると考えられた。

## 6. まとめと今後の課題

本研究では、認知症ケアのための忘れ物支援において「置き忘れの検出」に着目した。表 1 に示したように、4

つの置き忘れの事例を定義し、物体認識および人動作認識を用いた置き忘れ検出の方式を提案し、実験を行った。本研究における「置き忘れ検出」において、物体認識の方式や環境情報の認識、ロボットの会話機能など、さらに改善の必要があることが考えられた。本研究は将来、認知症の方を対象とした忘れ物防止支援のみならず、認知症の方のための支援ロボットとしての実装を考えている。

## 参考文献

- [1] 内閣府：高齢化の現状と将来像，  
[http://www8.cao.go.jp/kourei/whitepaper/w-2016/html/zenbun/s1\\_1\\_1.html](http://www8.cao.go.jp/kourei/whitepaper/w-2016/html/zenbun/s1_1_1.html), (2014).
  - [2] みんなの介護：第 56 回 介護職員の人材不足問題，最新版！2025 年には全国で 38 万人，東京・大阪・愛知…と大都市圏で 2 万人以上が不足する未来は変えられない!?, <http://www.minnanokaigo.com/news/N53135246/> (2015. 8. 7).
  - [3] 谷本鯛介：忘れ物支援ロボットの改良と心理的安心感の評価, (2016. 2. 25)
  - [4] 南口友里佳：認知症患者のための忘れ物支援ロボット, (2015.2.25)
  - [5] 木村孝広：環境の形状情報を付加した人間の動作分類に基づく 3 次元セマンティックマップの生成 (2014. 3. 13)
  - [6] 川崎広一：手の動きと動作状況を用いた行動認識 (2009. 2. 10)
  - [7] 森郷士：RT ミドルウェアに基づくコミュニケーションロボットプラットフォーム構築と強化学習によるコミュニケーションロボットの情動インタラクション制御 (2013)
  - [8] Hatena Blog (id:s0sem0y)：フィードフォワードニューラルネットワークの基本，  
<http://s0sem0y.hatenablog.com/entry/2016/05/22/014656>, (2016. 5. 22)
  - [9] 上間圭祐，平井裕麻，松尾直志，島田伸敬：人つながり支援のための人物位置の実時間マッピングに基づくロボットナビゲーション, (2014. 9. 4)
- 山崎亘：セマンティックオブジェクト獲得のための一人称視点 RGB-D 映像を用いた手動作の識別, (2016. 3. 14)