

人物の位置・姿勢に注目した行動パターンの学習・認識と非日常状態 検出への応用

青木 茂樹[†] 岩井 嘉男^{†*} 大西 正輝^{††} 小島 篤博^{†††}
福永 邦雄[†]

Learning and Recognizing Behavioral Patterns Using Position and Posture of
Human Body and Its Application to Detection of Irregular State

Shigeki AOKI[†], Yoshio IWAI^{†*}, Masaki ONISHI^{††}, Atsuhiko KOJIMA^{†††},
and Kunio FUKUNAGA[†]

あらまし 一般に、人物の行動には規則性や習慣性が存在すると考えられる。本研究では、行動の規則性・習慣性を行動パターンと呼び、行動パターンの学習・認識を目的としている。行動パターンに注目した従来手法では、赤外線センサや電気機器のオン・オフの情報のみを取り扱っていたために、人物が部屋の中でどのような動作を行っているかを詳細に認識することは難しかった。一方、動作を詳細に認識する手法を一般的な環境下で実現するためには、あらかじめ多数のモデルを用意する必要があった。本研究では、学習期間内に行われた動作を自動的に分類して個別モデルを作成することによって、あらかじめ多数のモデルを用意することなく人物の動作を詳細に認識し、更に動作の順序も考慮した人物の行動パターンを認識する手法を提案する。実験では、室内環境における人物の動作と行動パターンを学習・認識し、有効性を確認した。

キーワード 行動パターン, HMM, 人物の位置・姿勢

1. ま え が き

一般に、人物の行動には規則性や習慣性が存在すると考えられる。本研究では行動の規則性・習慣性を行動パターンと呼び、人物の行動パターンを学習・認識することを目的としている。

人物の行動パターンに注目した従来手法として、文献 [1]~[4] では、居間や風呂、トイレなどに設置した赤外線センサから得られる情報や、家庭用電気機器のオン・オフなどの情報をもとに人物の日常的な行動のパターンを学習し、独居高齢者が日常的な行動から逸

脱した場合に、介護者に異常を通知するシステムが提案されている。また、文献 [5], [6] では、電気機器のオン・オフのパターンから人物の行動のパターンを学習し、高齢者の QOL (Quality of Life) の向上を支援する手法が提案されている。これらの研究では、特定の時間に特定のスイッチを操作するなどの情報のみを取り扱っているため、精度良く人物の状態を検出できる反面、人物が部屋の中でどのような動作・行動を行っているかを詳細に知ることが難しいという問題があった。

人物の動作を詳細に認識する手法としては、文献 [7]~[9] などが挙げられる。これらの手法では、あらかじめ用意しておいたモデルと、カメラから得られる入力画像中の人物の動きを比較することによって人物の動作を認識する。一般的な環境下でこれらの手法を用いて詳細に動作を認識するためには、個人差や環境に合わせて多数のモデルを用意しておく必要があるために、記憶容量や計算コストが大きくなるという問題点があった。また、これらの手法では一つひとつの動作の種類を知ることはできるものの、“トイレに行っ

[†] 大阪府立大学大学院工学研究科, 堺市
Graduate School of Engineering, Osaka Prefecture University, 1-1, Gakuen-cho, Sakai-shi, 599-8531 Japan

^{††} 理化学研究所バイオ・ミメティックコントロール研究センター, 名古屋市
Bio-Mimetic Control Research Center, RIKEN, 2271-130 Anagahora, Shimoshidami, Moriyama-ku, Nagoya-shi 463-0003 Japan

^{†††} 大阪府立大学総合情報センター, 堺市
Library and Science Information Center, Osaka Prefecture University, 1-1, Gakuen-cho, Sakai-shi, 599-8531 Japan

* 現在, シャープ株式会社

た後に手を洗う”などといった動作の順序については考慮されていなかった。

これに対して文献[10]では、“部屋に入って、本棚に行き、机に行く”などのような動作の順序をパターンととらえて、このような人物の行動のパターンを、あらかじめモデルを用意することなく学習し、観測されたパターンの日常性や、どのようなパターンで行動を行っているかを認識する手法を提案している。しかし、この手法では人物の位置情報のみを用いているため、“立って歩いた場合”と“這って移動した場合”を正しく判別できないなど、人物の動作を詳細に知ることではできなかった。本論文では、全方位視覚センサから得られる入力画像から、人物の位置及び姿勢の情報を抽出することによって、動作形状をより多くのパラメータでとらえて認識し、更に動作の順序も学習・認識する手法を提案する[11]。本論文で提案する行動パターン学習・認識手法は、多数のモデルをあらかじめ用意することなく人物の状態を詳細に知ることができるため、独居高齢者の見守りシステムなどへの応用が容易であると考えられる。

以下、2. で動作の学習・認識手法について述べ、3. で行動のパターンの学習・認識手法を説明する。その後、4. で実験と考察を行い、5. でまとめる。

2. 動作の学習

2.1 特徴量の抽出

人物の動作をカメラを用いて観察するとき、広範囲の情報を一度に観測できることが望ましい。そこで、周囲 360 度シーンを一度に撮影することができる全方位視覚センサ[12]を用いる。

人物の動作を詳細に知るためには、人物の位置や顔・手の位置関係などが重要な情報であると考えられる。そこで、全方位視覚センサから得られる入力画像中の人物領域と肌色領域を抽出し、これらの領域から人物の動作を表すと考えられる特徴量を抽出する。

2.1.1 人物領域の特徴量

一般的な部屋の中で行われる人物の動作を考えたとき、例えば「ベッド」で「寝る」、「机やテーブルの前」で「座る」など人物が動作を行う位置と動作には関連があると考えられる。また、独居高齢者など要介護者の見守りシステムなどへの応用を考えた場合、同じ“しゃがむ”動作でも、“テーブルの前”で行われた場合と“廊下”で行われた場合では、その動作のもつ意味は異なるものと考えられる。また前述のように、同

じ“ドアからベッドへ行く”動作でも、“立って”移動した場合と“這って”移動した場合とではその意味は大きく異なるものと考えられる。そこで、人物の位置と姿勢を表す特徴量を抽出し、これらの情報を用いて人物の動作を表現する。

まず、入力画像中の動領域を人物領域と仮定して、動領域をロバストに抽出することのできる動エッジ[13]を抽出し、ノイズ除去のため膨張・収縮処理及び孤立点除去を n 回ずつ繰り返す。ここでは、予備実験において処理速度の低下が少なく、効果的にノイズを除去できる回数が 2 回であったため、 $n = 2$ としている。次に、入力画像中には人物が複数いることも考えられるため、得られた動エッジをクラスタリングし、各クラスごとに動エッジの重心 μ_p 、分散共分散行列 Σ_p を次式で求める。

$$\mu_p^{<j>} = \frac{1}{N^{<j>}} \sum_{i=1}^{N^{<j>}} x_i^{<j>} \quad (1)$$

$$\Sigma_p^{<j>} = \frac{1}{N^{<j>} - 1} \sum_{i=1}^{N^{<j>}} (x_i^{<j>} - \mu_p^{<j>})(x_i^{<j>} - \mu_p^{<j>})^T \quad (2)$$

ここで、 $N^{<j>}$ はクラス j に含まれる動エッジの画素数を示しており、 $x_i^{<j>}$ は動エッジの座標を示している。以上で求めた各クラスの重心 $\mu_p^{<j>}$ と分散共分散行列 $\Sigma_p^{<j>}$ を用いて、各画素 x_i のマハラノビス距離を式 (3) によって求め、マハラノビス距離がしきい値未満となる楕円領域を単一の人物領域として抽出する。

$$D_i = (x_i - \mu_p^{<j>})^T \Sigma_p^{<j>^{-1}} (x_i - \mu_p^{<j>}) \quad (3)$$

そして、抽出した人物領域ごとに得られる動エッジの“重心”、“画素数”及び分散と共分散で表される“領域分布”を人物の動作を表す特徴量として保存する。

ここで、動エッジの重心は人物の位置を表し、画素数は動領域の大きさを表すため、人物の動きや移動量の大きさを示す。また、領域分布は手を広げたり、しゃがんでいる場合など人物の形状や姿勢の変化に対応する。

全方位視覚センサから得られる入力画像を図 1 (a) に示す。ここでは画像の左右に 2 人の人物が写っている。この画像に対し、動エッジを抽出した結果を図 1 (b) 中の黒い画素で示している。人物の存在する個所で動エッジが検出されていることが確認できる。

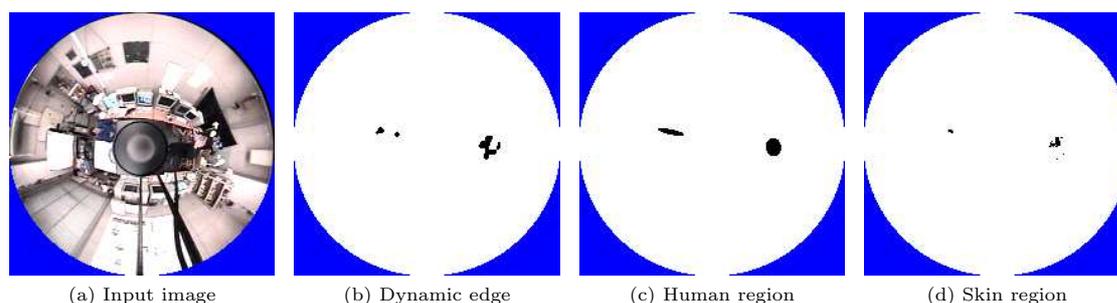


図 1 人物領域と肌色領域の抽出例
Fig. 1 Examples of feature set extraction.

また、動エッジをクラスタリングして抽出した人物領域を図 1 (c) に示す。2 人の人物が抽出されていることが分かる。

2.1.2 肌色領域の特徴量

人物の手や顔の動きを知るために、前節で述べた手法により抽出した人物領域中の肌色画素を抽出する。ここでは人物領域を楕円で表現しているため、手を大きく広げている場合などに、手の先の部分が楕円領域に含まれない可能性が考えられる。そこで、肌色画素を探索する領域（肌色探索領域）に楕円領域だけでなく、その近傍に存在する動エッジも含める。このようにすることで、手を大きく広げている場合などにも手や顔の領域を安定して抽出することができる。

人物の肌色を抽出するための色情報として、比較的明るさの影響を受けにくい CIE1976($L^*a^*b^*$) 表色系を用いる。一般的な環境下で撮影した画像から、人手によって人物の肌色領域の画素を N_s 個（本論文では $N_s = 10000$ 個程度）選び、これら N_s 個の画素の色情報 $C_i = (a_i^*, b_i^*) (i = 1, 2, \dots, N_s)$ に対しその平均値 μ_s 、分散共分散行列 Σ_s から、肌色探索領域中の各画素の色情報が肌色にどのくらい近いかを人物領域を抽出する場合と同様にマハラノビス距離によって求める。

肌色探索領域中のすべての画素においてマハラノビス距離を求め、しきい値処理することによって肌色画素を抽出する。図 1 (d) に肌色画素の検出結果を示す。人物の肌の箇所が検出できていることが確認できる。

以上で求めた肌色画素の“重心”、“画素数”、“領域分布”を各人物領域ごとに求め、これらも人物の動作を表す特徴量として用いる。ここで、重心は手・顔の位置を表し、画素数は人物がカメラ方向を向いている場合には肌色画素の画素数が多く、逆方向を向いてい

表 1 特徴量
Table 1 List of feature set.

特徴量	得られる情報	次元数
動エッジの重心	位置	2
動エッジの画素数	動きの大きさ	1
動エッジの領域分布	人物の形状	4
肌色画素の重心	手・顔の位置	2
肌色画素の画素数	人物の方向	1
肌色画素の領域分布	手・顔の位置関係	4
人物領域の数	人数	1

る場合には少なくなるため、人物の方向を表す。また、領域分布は手・顔の位置関係に対応する。これら肌色領域から抽出した特徴量を用いることで、人物の顔の方向や手を使った動作なども識別できる。

2.2 記号列の生成

人物の動作が観測されたとき、その動作を 2.1 で述べた手法により表 1 に示す 15 次元の特徴量の系列に変換する。次に、得られた特徴量を特徴量の相関関係を保持したまま次元を圧縮することが可能な固有空間法 [14] を用いて次元を圧縮する。更に、作成した固有空間上でクラスタリング処理を行い、得られたクラスにラベル付けを行う。ここで、クラスタリング処理は k -means アルゴリズムですべてのクラスのクラス内分散がしきい値未満となるまでクラス数を 1 から 1 ずつ増加させることによって行っている。その後、人物の動作が観測されたとき、得られた特徴量を固有空間上に投影して最もマハラノビス距離の短いクラスを調べ、そのクラスのラベルを順に出力することによって、人物の動作を記号列で表現する。本研究では、動エッジが抽出されてから抽出されなくなるまでの間を一つの動作と考えて記号列を生成している。ただし、記号列長が 15 未満である短いものはノイズと考えて除去する。

2.3 動作の学習と認識

前節までで述べた手法を用いて人物の動作を記号列で表現したとき、人物が同じ動作を行った場合には類似した記号列が出力され、異なる動作を行った場合には異なる記号列が観測される。しかし、同一の動作を行った場合でも動作の速度が異なる場合などには、長さの異なる記号列が得られる。長さだけが異なる類似した記号列を異なる動作と識別することも考えられるが、動作の種類が増加することが予想される。本研究では、「動作」の識別だけでなく、動作の順序関係を考慮した「行動パターン」を識別することも目的としているため、動作の種類が膨大な場合には、「行動パターン」の種類が増加し、長期にわたる学習期間を設定する必要があると考えられる。そこで、本研究では記号列パターン認識によく用いられる Hidden Markov Model (HMM) を用いることによって人物の類似する動作を一つの動作として学習・認識する。HMM は、時間軸方向の移動や伸縮にロバストなため、本手法のような個人差の大きい記号列の学習に適していると考えられる [7], [10], [11], [15]。

一般に HMM を用いた認識では、学習セットであらかじめ学習させておいたモデルを用意しておき、入力とどのモデルの類似度が高いかを測定する手法が用いられている。しかし、一般的な環境下で詳細な動作を認識する場合には、人物の行う動作の種類が膨大となるためにそれらすべてのモデルをあらかじめ用意しておくことは現実的ではない。そこで、以下に示すアルゴリズムで学習期間内に行われた人物の様々な日常動作を動作ごとに分類して学習する。

[動作の学習アルゴリズム]

- (Step1) 観測されたすべての記号列を学習データとする HMM を作成する。
- (Step2) 学習データのすべての記号列を (Step1) の HMM で出力し、各記号列の出力確率を数直線上にプロットして、ヒストグラムを作成する。
- (Step3) ヒストグラムの度数に対して判別分析法を適用して求めた規準値の近傍で極小値を探索し、その値をしきい値とする。
- (Step4) 出力確率がしきい値以上であればその記号列に “1” のフラグを立て、しきい値未満であれば “0” のフラグを立てる。
- (Step5) 同一のフラグを与えられた記号列の集合ごとに、それらを学習データとする HMM を作成する。

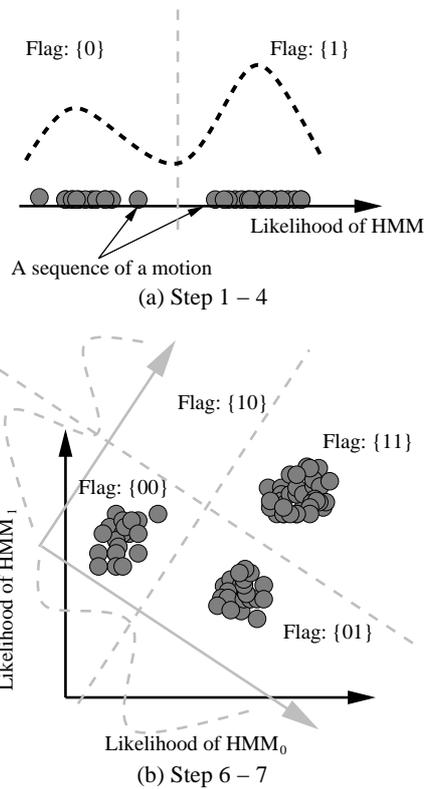


図2 動作の学習アルゴリズム
Fig.2 An algorithm of learning of motions.

(Step6) 作成したすべての HMM で全記号列の出力確率を求め、各 HMM での出力確率を軸とする空間上に全記号列を投影し、主成分分析法で基底ベクトルを求める。

(Step7) 累積寄与率がしきい値以上となるまでの基底ベクトルの軸を用いて、各軸ごとに (Step2, 3) の手順に従ってヒストグラムとしきい値を求め、記号列のフラグを求める。

(Step8) (Step7) で同一のフラグを与えられた記号列の組合せが前回のループで得られた組合せと比べて変わらないようになるまで (Step5~7) を繰り返す。

このアルゴリズムでは、学習期間内に観測されたすべての動作を動作の種類ごとに分類する処理を行っている。動作を分類する指標として、ここでは、HMM で類似する記号列を出力した場合に、HMM 内で同様の状態遷移を行うために出力確率が類似する性質を利用している。

(Step1~4) の処理では、すべての記号列を学習した HMM を用いて算出した、すべての記号列の出力確率を基準として、学習期間内に観測されたすべての動作を “Flag{0}” の動作と “Flag{1}” の動作に分類している。図 2(a) は (Step1~4) の処理の流れを概念的に示したものである。図中、丸印が一つの動作に対応した記号列を示し、HMM の出力確率をもとに数直線上にプロットしている。黒の点線は (Step2) で求めたヒストグラムを表しており、灰色の点線は (Step3) で求めたしきい値を示している。また、上部の Flag:{0}, Flag:{1} は各記号列に付けたフラグであり、動作を二つに分類した結果を表している。

(Step6~7) の 1 回目のループの時の処理の流れを概念的に示したものを図 2 (b) に示す。図 2 (a) と同様に丸印は動作の記号列を表しており、(Step4) で同一のフラグを与えられた記号列ごとに作成した二つの HMM (HMM₀, HMM₁) で求めた出力確率をもとに 2 次元平面上にプロットしている。このとき、(Step4) で Flag:{0} となった記号列を学習した HMM₀ では、Flag:{0} の記号列の出力確率が高くなると考えられるし、同様に Flag:{1} の記号列を学習した HMM₁ では、Flag:{1} の記号列が高くなると考えられる。また、(Step1~4) ではすべての動作を大きく二つに分類しているだけなので、両方の HMM で出力確率の高くなる記号列や、両方の HMM で出力確率が低くなる記号列が存在すると考えられる。これらを効率良く分類するために、(Step6~7) で主成分分析法を利用して基底軸を変換し、基底軸ごとにヒストグラムを作成して、動作を分類する。図 2(b) 中の灰色の矢印が主成分分析法で求めた基底ベクトルを示しており、灰色の点線で示している各軸ごとに求めたしきい値によって各記号列のフラグを決定し、観測されたすべての動作を Flag:{00}, Flag:{01}, Flag:{10}, Flag:{11} の四つに分類している。

以上の処理で動作を四つに分類することができたが、一般に人物の行う動作の種類数は未知なので、いくつに分類すればよいか分からない。そこで、(Step8) の条件を満たすまでの間、(Step5~7) を HMM を増加させながら繰り返すことによって、学習期間内の人物の動作を分類して学習する。

以上で求めた HMM, 基底ベクトル, しきい値及び一定数以上の記号列が含まれるフラグ (図 2(b) の例では {10} 以外の各フラグ) を保存しておく。ここで保存した各フラグが人物の日常的な動作を表しており、

それ以外のフラグ (図 2(b) の例では {10}) が非日常的な動作を表す。

動作の認識は、観測された動作の記号列をすべての HMM で出力し、基底ベクトルで張った空間上に記号列を投影してフラグを求め、保存しておいた動作のフラグと一致する動作を認識結果として出力する。一致するフラグが存在しない場合には、人物が日常的に行わない非日常的な動作を行ったと判断する。

3. 行動パターンの学習

2. では “ベッドからトイレに歩いた”, “トイレから洗面台に歩いた” などという単独の動作の認識手法について述べた。これらの動作のうち短い時間間隔で連続して行われる動作には関連性があると考えられる。例えば前述の例であれば “トイレに行った後、手を洗うために洗面台に行った” と考えることができる。また、短い時間間隔で行われる動作は、規則的あるいは習慣的に行われるものと思われる。ここでは、連続して行われる動作の規則性、習慣性を行動パターンと呼ぶ。以下では、行動パターンの学習・認識手法について述べる。

まず、行動パターンを記号列で表現するために、各動作のフラグに対してラベル付けを行う。次に、観測されたすべての行動パターンに対してラベル系列を求め、2.3 と同様の学習アルゴリズムによって人物の行動パターンを学習する。学習の結果求められた各フラグが人物の一つの行動パターンに対応する。

人物がどの行動パターンで動作を行ったかを認識する場合についても 2.3 と同様に、観測された行動パターンのフラグを求め、求めたフラグと一致する行動パターンで動作を行っているかと判断する。また、一致するフラグがない場合も同様に人物が非日常的な行動パターンで動作を行っているかと判断する。

4. 実験及び考察

4.1 実験

本手法の有効性を確認するために、研究室の中央に全方位視覚センサを設置して動作の学習・認識、行動パターンの学習・認識実験を行った。研究室の部屋の配置図を図 3 に示す。処理は 320 × 240 画素で取り込んだ画像の中央部 210 × 210 画素を用いており、処理速度は毎秒約 3 フレームである。

4.1.1 予備実験

まず、動作の学習に最適な HMM の状態数を調べる

予備実験を行った。実験では、日常的な動作として 8 種類、非日常的な動作として 5 種類を設定し、日常的な動作と非日常的な動作の識別に最適な HMM の状態数を調べた。

日常的な動作には、通常の生活での動作を想定して、図 4 に示す 8 種類を用いた。

動作 1: ドア 2 から入り鏡の前を通って、テーブルの右側からテーブルの周りを 1 周し、ドア 2 から部屋の外へ出る動作

動作 2: 動作 1 と逆の経路を通る動作

動作 3: ドア 1 から入り鏡の前を通って、テーブルの右側からテーブルの周りを 1 周し、ドア 1 から部屋の外へ出る動作

動作 4: ドア 1 から入り鏡の前を通って、ホワイトボードの前まで行き、同じ経路を通って部屋の外へ出る動作

動作 5: ドア 2 から入りテーブルの左側を通って、本棚に沿って部屋の奥まで行き、同じ経路を通って

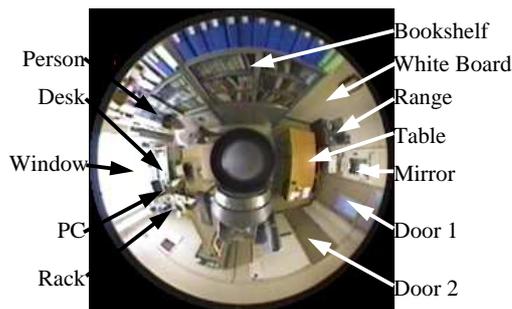


図 3 部屋の配置図
Fig. 3 The layout of the room.

部屋の外へ出る動作

動作 6: ドア 2 から入り鏡の前を通ってホワイトボードまで行き、ホワイトボードに文字を書いて消し、同じ経路を通って部屋の外へ出る動作

動作 7: ドア 2 から入り鏡の前に行って手で髪をかきあげ、同じ経路を通って部屋の外へ出る動作

動作 8: 動作 2 と同じ経路を手を挙げて移動する動作

以上の動作を 30 回ずつ計 240 回行動作を行うたびに記号列を生成し、学習用記号列 (Set A) を用意した。特徴量のクラスタリングに用いた固有空間の次元数は累積寄与率が 83.3% となる 6 次元とし、6 次元の固有空間上でクラスタリング処理した結果得られたクラスタ数は 16 で、記号列の系列長は 30 から 132 だった。

次に、新たに各動作について 20 回ずつの動作を行い、動作を行うたびに記号列を生成して 160 個の認識用記号列 (Set B) を用意した。更に、非日常的な動作として以下の五つの動作を 10 回ずつ行い 50 個の記号列 (Set C) を作成した。ただし、抽出した特徴量とすべてのクラスタ中心の固有空間距離がしきい値以上の場合には、非日常的な位置・姿勢であると考え、学習時には用いていない記号を割り当てている。

動作 a: ドア 2 から入りテーブルの左側を通って机まで行き、椅子に座って立ち上がり、同じ経路を通って部屋の外へ出る動作

動作 b: ドア 2 から入りテーブルの左側を通って、棚に沿ってラックのところまで行き、ホワイトボードの前を通り鏡の前を通ってドア 2 から部屋の外へ出る動作

動作 c: ホワイトボードと机の間を往復する動作

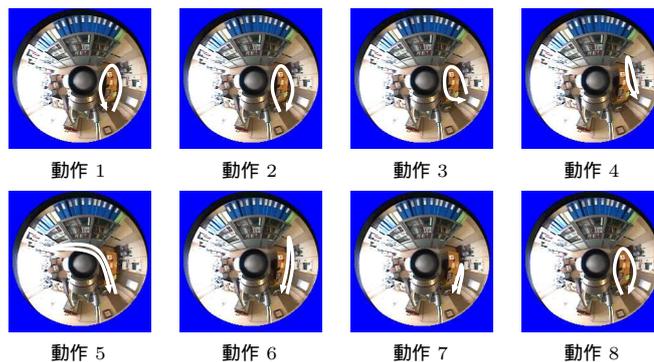


図 4 動作の軌跡の例
Fig. 4 Examples of motions.

動作 d: 机の前を周回する動作

動作 e: ドア 1 から入りドア 2 から部屋の外へ出る動作

HMM の状態数と認識率の関係を調べるために、3 から 70 まで 1 刻みで状態数を変化させて作成した HMM で学習用記号列 (Set A) を学習した。また、学習量と認識率の関係を調べるために、記号列の学習量を模擬的に 4 倍、30 倍、50 倍と変化させて学習した。次に、学習した HMM を用いて Set B と Set C の計 210 個の認識用記号列を出力させ、日常的な動作と非日常的な動作を識別した。HMM の出力シンボル数は 17、学習には Baum-Welch アルゴリズムを用い、記号列の出力確率は Viterbi アルゴリズムで求めたシンボル出力確率を系列長で正規化した値を用いた。また、日常的な動作と非日常的な動作を識別するためのしきい値は、独居高齢者の見守りシステムなどへの応用を考えた場合、非日常状態を確実に検出する必要があるため、非日常的な記号列である Set C をすべて非日常的と識別する最小のしきい値を探索して、その値を用いた。そして、そのしきい値を用いたときに Set B の記号列を日常的と正しく識別できる割合を調べた。実験の結果を図 5 に示す。グラフには、横軸に HMM の状態数を示し、縦軸に Set B を日常的な動作と正しく識別できている割合を示している。また、それぞれのグラフは上から順に記号列の学習量を模擬的に 4 倍、30 倍、50 倍としたときの結果である。

図より、HMM の状態数が少ない場合には十分に識別できていないが、HMM の状態数が 20 以上であれば、学習量に関係なく安定して高い評価が得られており、HMM の性質として一般的に知られている知見と同様の結果が得られた。

4.1.2 動作の学習・認識実験

本手法で人物の動作を動作の種類ごとに学習し、識別する実験を行った。実験は、前節の予備実験と同様の実験条件で行い、HMM の状態数は予備実験の結果を考慮して 30、動作の学習アルゴリズム (Step7) におけるしきい値はあらかじめ行った予備実験の結果を考慮して 70% として、学習用系列 240 個を学習した。実験の結果、8 種類の動作計 240 個を 30 種類の動作に分類して学習した。また、各動作を表すフラグは 14 次元だった。結果の例を表 2 に示す。表は横方向に学習に用いた動作を示し、縦方向に学習の結果分類した動作とそのフラグを示しており、表中の数字は学習に用いられた系列の数を示している。動作 A_1 に注目す

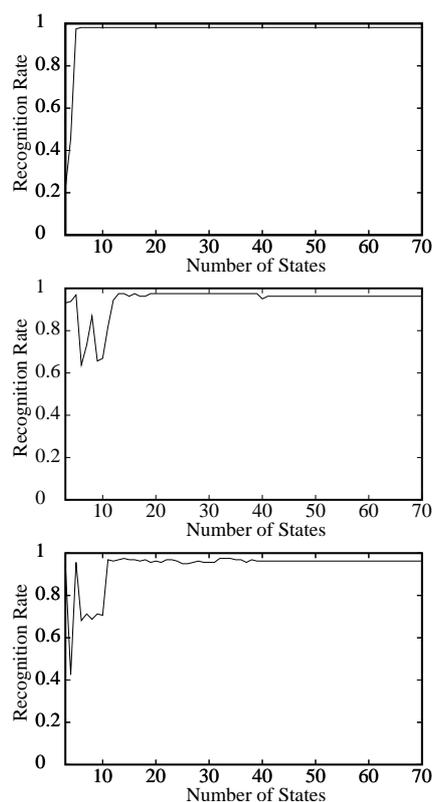


図 5 HMM の状態数を変化させたときの認識率
Fig. 5 Recognition rate with changing number of states.

ると、動作 1 の学習系列 24 個で学習されていることが分かる。動作 A_{26} では、動作 2 の学習系列 3 個と動作 8 の記号列 12 個で学習されていることが確認できる。

次に、8 種類の動作をそれぞれ 20 回ずつ計 160 回行って、動作を行うたびに記号列を生成してフラグを求め、人物がどの動作を行ったのかを認識した。表 3 に結果の例を示す。どの動作にも当てはまらず非日常的な動作であると判断した結果を太線の下に示している。動作 3 に注目すると、動作 3 の観測系列 20 個中のすべてのフラグが動作 A_{10} のものと同じであったことが確認できる。実験の結果、全 160 個の記号列中、異なる種類の動作と誤認識した動作 1 回、どの動作にも当てはまらず非日常動作であると判断された 9 回 (動作 1 の 2 回、動作 2 の 4 回、動作 4 の 1 回、動作 7 の 2 回) を除く 150 回 (93.5%) の動作を正しく認識できた。誤認識した動作を表 3 中に太字で示している。

表 2 動作の学習結果例

Table 2 Results of learning of motions.

分類結果	フラグ	学習させた動作				
		1	2	3	...	8
A ₁	{00100011111011}	24				
A ₂	{00110011111011}	3				
A ₃	{00100011111111}	1				
A ₄	{00100011111101}	1				
A ₅	{10100011111011}	1				
A ₆	{10100101101011}		24			
A ₇	{10100001101011}		1			
A ₈	{10100001101111}		1			
A ₉	{00100111111001}		1			
A ₁₀	{00000011100001}			29		
A ₁₁	{00000011101001}			1		
⋮						
A ₂₆	{10110101101011}		3			12
A ₂₇	{10110101100011}					15
A ₂₈	{00110101101011}					1
A ₂₉	{10100101101111}					1
A ₃₀	{10110101100010}					1

表 3 動作の認識結果例

Table 3 Results of recognition of motions.

分類結果	認識させた動作				
	1	2	3	...	8
A ₁	15				
A ₂	1				
A ₃	2				
A ₄					
A ₅					
A ₆		15			1
A ₇					
A ₈					
A ₉					
A ₁₀			20		
A ₁₁					
⋮					
A ₂₆		1			9
A ₂₇					9
A ₂₈					1
A ₂₉					
A ₃₀					
非日常動作	(2)	(4)		(3)	

更に、人物が通常行わないことを行った場合を想定した5種類の動作 a~e 各10回についても動作を行うたびに観測系列のフラグを求め、動作の非日常性が検出できることを確かめた。実験の結果、すべての動作で日常動作と異なるフラグが得られ、人物が日常的に行わない非日常的な動作を行ったことを正しく認識できた。

4.1.3 行動パターンの学習・認識実験

本手法を用いて行動パターンの学習・認識が行えることを確認する実験を行った。実験では、人物が

表 4 行動パターンの学習・認識結果

Table 4 Results of learning and recognizing of behavioral patterns.

分類結果	フラグ	学習させたパターン (認識させたパターン)		
		1	2	3
B ₁	{010110}	6(2)		
B ₂	{110100}	4(4)		
B ₃	{111001}		2	
B ₄	{101001}		4(3)	
B ₅	{110001}		2	
B ₆	{100001}		1(1)	
B ₇	{101011}		1(1)	
B ₈	{110101}			10(6)
非日常パターン			(1)	

4.1.1 で設定した八つの種類の動作を7回連続して行うパターンを学習させた。また、各回に人物の行った動作は4.1.2の学習結果を利用して認識した。行動パターンの学習に用いたHMMの出力シンボル数は、学習したフラグの総数である30と非日常動作と認識された場合のシンボルを合わせて31、状態数は4とし、学習アルゴリズムの累積寄与率のしきい値は70%とした。

行動パターンの学習は、以下の3種類のパターンを設定して行った。

パターン1: 動作1, 動作3, 動作5, 動作7, 動作2, 動作5, 動作3の順序で連続して動作を行うパターン

パターン2: 動作7, 動作3, 動作2, 動作1, 動作5, 動作7, 動作1の順序で連続して動作を行うパターン

パターン3: 動作2, 動作4, 動作6, 動作8, 動作4, 動作8, 動作6の順序で連続して動作を行うパターン

3種類の行動パターンで動作を10回ずつ行い、学習用系列として30個の記号列を使用して行動パターンを学習した。実験の結果、3種類の行動パターンを八つに分類して学習した。実験の結果を表4に示す。表には、表2と同様に、横軸に学習させた行動パターン、縦軸に学習の結果、分類した行動パターンとそのフラグを示している。表よりパターン1が二つ、パターン2が五つに分割されて学習されていることが確認できる。

次に、新たに各パターンについて6回ずつ計18回、行動パターンに沿って動作を行い、7回の動作を完了するたびに得られた観測系列のフラグを調べてどの行

動パターンで動作を行ったのかを認識させた。実験結果を表 4 中の括弧の中に示している。また、どの行動パターンにも当てはまらなかったものを太線の下に示している。表より、非日常的なパターンと認識された 1 回を除く、18 回中 17 回の 94.4% で正しく認識できていることが確認できる。

更に、人物が通常行わないパターンで動作を行ったときに、非日常的であると正しく認識できることを確認する実験を行った。実験に用いた行動パターンは以下の 2 種類である。

パターン a: 日常的な動作をランダムな順序で 7 回連続して行うパターン

パターン b: 非日常的な動作を 7 回連続して行うパターン

以上二つのパターンでそれぞれ 6 回ずつ動作を行い、計 12 回の行動パターンの日常性を確認した。その結果、すべてのパターンにおいて非日常的なパターンであることを正しく認識できた。

4.2 考察

本論文では、人物の位置及び姿勢に注目して人物の動作を自動的に分類・学習し、更に動作の順序性を考慮した、行動パターンの認識手法を提案した。以下、本手法の特徴について考察する。

(1) 動作の記号化

人物の動作を表す特徴量として人物領域と肌色領域から位置情報や姿勢に関する情報など、15 次元の特徴量を抽出し、特徴量を投影した固有空間内でクラスタリング処理を行うことにより人物の動作を記号列で表現した。これまでに提案されている動作認識手法の多くでは、人物が同じ動作を行った場合には、どこにいても同じ動作として識別することを目的としていた。しかし、本手法では動作と動作の行われる位置の関連性に注目するために、特徴量に人物の位置情報を加えている。また、抽出した特徴量の重要性などは、直接特徴量に重みづけするなどの方法で考慮するのではなく、固有空間を作成する時に、寄与率が低く特徴を表現するのに適さない軸を用いないことで考慮している。個別に設定するパラメータが少ないために、独居高齢者など要介護者の見守りシステムなどへの応用が容易であると考えられる。

実験では、動作 2 と動作 8 で人物の軌跡が同一で姿勢だけが異なる二つの動作の識別を行ったが、学習で 3 回、認識で 1 回誤認識した以外は正しく認識できている。非日常動作と誤認識した他の 9 系列も含めて、

全方位センサからの距離が大きくなったときに特徴量が安定して抽出できなかったために誤認識したものと考えられる。これに対しては、入力画像を高解像度化するなどの方法で解決できると考えている。

(2) 動作の学習と認識

あらかじめ用意した学習用セットを用いて学習させたモデルを用意するのではなく、学習期間中に人物が行った日常的な動作のモデルを自動的に作成して動作を分類・学習する手法を明らかにした。本手法では、類似する記号列の HMM での出力確率が類似することを利用して、記号列の分類を行って動作を学習する。動作を分類するときに主成分分析法を利用して動作を分類するのに相応しい基底軸を利用していることや、分類するときのしきい値を判別分析法で求めた規準値の近傍に存在する極小値としていることから、動作をロバストに分類できると考えられる。また、類似する動作が複数存在する場合でも、それら類似する動作を学習した HMM が作成されれば、その HMM における出力確率が動作ごとに変化するために、安定して動作を分類、識別できると考えられる。

実験では、8 種類 240 回の動作を 14 次元のフラグで 30 種類の異なる動作として学習した。1 種類の動作を複数に分割して学習しているが、これは、動作を行ったときの被験者の位置や姿勢が少し変化したためだと考えられる。認識実験では、8 種類の動作を 160 回行い 93.8% の認識率が得られた。誤認識した 10 系列は、前述したように安定して特徴量を抽出できなかったことが要因であると考えられる。入力画像の高解像度化に加えて、動作のフラグのハミング距離などを考慮することによって解決できると考えられる。

(3) 行動パターンの学習と認識

人物の行う単一の動作を認識するだけでなく、個々の動作の順序性を考慮した行動パターンの学習・認識手法を提案した。本手法を用いることによって、これまでの手法では考慮されなかった、一つひとつの動作の出現順序を考慮した一連の動作を認識することができる。

実験では、8 種類の動作を 7 回連続して行ったパターンの学習・認識を行い、94.4% の認識率を得ることができた。本論文では、行動パターンに含まれる動作の個数を既知として実験したが、実環境では、行動パターンにいくつの動作が含まれるかは未知である。また、行動パターンのなかにも動作の順序が異なるシーンも存在するものと考えられる。このような場合にも、自

動的に学習できることが望ましいが、関連性のある動作のみを選択することは容易ではない。観測された動作から関連性の高いものを選択する手法の検討は今後の課題とする。

(4) HMM の状態数

本論文では、人物の動作、行動パターンを学習するときの HMM の状態数を予備実験の結果をもとに、それぞれ 30, 4 と固定して実験し、動作で 93.8%, 行動パターンで 94.4% と高い認識率を得ることができた。しかし、人物が一般的に行う行動には複雑な動作や行動パターンが含まれると考えられ、それらを学習・認識する場合には、単一の状態数では正しく識別できない可能性も考えられる。このような問題は、動作・行動パターンの種類ごとに状態数の異なる HMM を作成することで解決できる可能性もあるが、どの程度の状態数の HMM が最適かを自動的に決定することは容易ではない。観測された動作と HMM の状態数の関係を考慮して、動作や行動パターンごとに異なる状態数の HMM を作成する手法の検討は今後の課題とする。

(5) 非日常状態検出への応用

学習期間内に観測された動作・行動パターンを分類・学習し、学習期間内に行われなかった動作・行動パターンを非日常状態として検出する手法を提案した。独居高齢者など要介護者の見守りシステムなどへの応用を考えた場合、実時間で非日常状態を検出できることが必要となるが、本手法を用いた実験では、320×240 画素の入力画像を用いて毎秒 3 フレームで処理を行うことができた。動作の認識処理では、動作特徴を抽出する処理の計算負荷が大きく、記号列の認識処理などの占める計算負荷は小さかった。識別する動作の種類が増加した場合には、HMM の数やフラグの次元などが増加することが予想されるが、八つの動作を識別する実験において実時間で行っていたことと、認識処理では学習時に得られたすべての HMM で記号列を出力して動作を認識するために、処理時間は HMM の数に比例して増加することなどを考慮すると、動作の種類数の認識処理時間への影響は少ないと考えられる。学習処理に関しては、学習する記号列によって処理時間が異なると考えられるため、動作の種類数と処理時間の関係を一概に述べることはできないが、学習処理を人物が動作を行わない時間帯などに行うことなどによって、動作の種類が増加した場合にも本手法を適用できると考えている。また、動作が行われるたびにその動作の日常性が識別でき、更に動作は日常的で行動

パターンが非日常的である場合でも、行動パターンの記号列を抽出した直後に非日常状態を検出することができる。

実験では、非日常的な動作・行動パターンを 100% 識別することができたものの、1 個所にとどまって行う動作を識別する必要があるシーンに適用する場合などでは、320×240 画素、毎秒 3 フレームの映像では、特徴を十分に抽出することができず、識別に失敗する可能性も考えられる。このような場合にも、入力画像を高解像度化し、PC クラスタなどを利用して処理速度を向上させることによって解決できると考えている。

5. む す び

本論文では、あらかじめモデルを用意しておく必要のない、動作及び行動パターンの認識手法を提案した。実験では、8 種類の動作を認識する実験及び 3 種類の行動パターンを認識する実験を行い、有効性を確認した。今後の課題としては、1 日や 1 週間といった長い時間間隔での行動パターンの学習手法や非日常状態が検出されたときの通知手法の検討などが挙げられる。

謝辞 本研究の遂行にあたり有益なご助言を頂きました、元大阪府立大学大学院工学研究科助手の猪飼武夫先生に感謝致します。

文 献

- [1] 太田 茂, “高齢者のためのモニタリングシステム,” 情報処理, vol.41, no.6, pp.639-643, June 2000.
- [2] 上山直浩, 山口晃史, 山越憲一, 菅原康博, 八木沢博史, 金戸正人, “人感センサによる行動モニタリングシステム,” 第 38 回日本 ME 学会大会要項集, p.376, April 1999.
- [3] 山口晃史, 上山直浩, 山越憲一, 菅原康博, 八木沢博史, 金戸正人, “高齢者の生活支援情報計測・運用システムの開発,” 第 38 回日本 ME 学会大会要項集, p.134, April 1999.
- [4] 菅原康博, “介護支援情報システム,” 計測制御, vol.40, no.5, pp.379-383, May 2001.
- [5] 松本 勉, 嶋田泰幸, 柴里毅生, 大塚弘文, 川保茂保, “人の在宅行動モデルと非日常性行動判断について,” 第 19 回計測自動制御学会九州支部学術講演会論文集, pp.231-234, Oct. 2000.
- [6] T. Matsumoto, Y. Shimada, K. Shibasaki, H. Ohtsuka, and S. Kawaji, “Creating behavior model of a senior solitary life and detecting an urgency,” Proc. International ICSC Congress on Intelligent Systems & Applications (ISA2000), vol.2, pp.234-238, Dec. 2000.
- [7] J. Yamamoto, J. Ohya, and K. Ishii, “Recognition human action in time-sequential images using hidden markov model,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.379-385, June 1992.

- [8] A. F. Bobick, and J. W. Davis, "The recognition of human movement using temporal templates," IEEE Trans. Pattern Anal. & Mach. Intell., vol.23, no.3, pp.257-267, March 2001.
- [9] 西村拓一, 十河卓司, 小木しのぶ, 岡 隆一, 石黒 浩, "動き変化に基づく View-based Aspect Model による動作認識," 信学論 (D-II), vol.J84-D-II, no.10, pp.2212-2223, Oct. 2001.
- [10] 青木茂樹, 大西正輝, 小島篤博, 福永邦雄, "HMM による行動パターンの認識," 信学論 (D-II), vol.J85-D-II, no.7, pp.1265-1270, July 2002 .
- [11] 青木茂樹, 岩井嘉男, 大西正輝, 小島篤博, 福永邦雄, "人物の位置・姿勢に注目した行動パターンの認識," 信学技報, PRMU2002-82,WIT2002-25, Sept. 2002 .
- [12] 山澤一誠, 八木康史, 谷内田正彦, "移動ロボットナビゲーションのための全方位視覚センサ," 信学論 (D-II), vol.J79-D-II, no.5, pp.698-707, May 1997.
- [13] 大西正輝, 泉 正夫, 福永邦雄, "講義映像における板書領域のブロック分割とその応用," 信学論 (D-I), vol.J83-D-I, no.11, pp.1187-1195, Nov. 2000 .
- [14] 村瀬 洋, シュリーナイヤー, "2 次元照合による 3 次元物体認識 -パラメトリック固有空間法," 信学論 (D-II), vol.J77-D-II, no.11, pp.2179-2187, Nov. 1994 .
- [15] A.D. Wilson, and A.F. Bobick "Parametric hidden Markov models for gesture recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol.21, no.9, pp.884-900, Sept. 1999.

(平成 15 年 3 月 27 日受付, 8 月 27 日再受付)



大西 正輝 (正員)

1997 阪府大・工・情報卒。2002 同大学院博士後期課程了。同年理化学研究所バイオ・ミメティックコントロール研究センターフロンティア研究員, 現在に至る。ロボットの認知・知識処理に関する研究に従事。電気学会, 日本ロボット学会, 映像情報メディア学会各会員。工博。



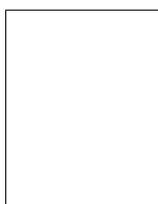
小島 篤博 (正員)

1990 阪府大・工・電気卒。1992 同大学院博士前期課程了。同年富士通(株)入社。1996 年 9 月より同大総合情報センター助手。画像情報から自然言語への変換に関する研究に従事。電気学会, 情報処理学会, 人工知能学会各会員。工博。



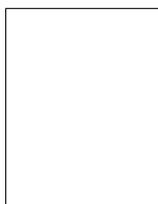
福永 邦雄 (正員)

1967 阪府大・工・電気卒。1969 同大学院修士課程了。同年同大・工・電気助手。現在同大・工・情報教授。コンピュータビジョン, グラフ理論とその応用などの研究に従事。情報処理学会, 人工知能学会, システム制御情報学会, IEEE 各会員。工博。



青木 茂樹 (正員)

1998 阪府大・総合科学部卒。2004 同大学院工学研究科博士後期課程了。同年熊本電波高専電子制御工学科助手, 現在に至る。コンピュータビジョン, パターン認識に関する研究に従事。日本ロボット学会学生会員。工博。



岩井 嘉男

2002 阪府大・工・情報卒。同年シャープ(株)入社。在学中, コンピュータビジョン, パターン認識に関する研究に従事。