

全方位画像を用いた歩行軌跡による確率的人物行動認識システム

小 阪 勇 気^{†,††} 平 山 高 嗣^{†††}
岩 井 儀 雄^{†††} 谷 内 田 正 彦^{†††}

本研究では人物の行動を認識し、未知の行動を行う人物を不審人物として切り出すような監視システムへの応用を目的として、全方位視覚センサにより撮影された画像を用いて人物追跡を行い、その行動を認識するシステムの構築を目指す。従来、部屋の中で人物が入室してから退室するまでを監視し、人物を追跡することにより得られた歩行軌跡から、人物の行動モデルを作成して、その行動を認識するシステムが提案されている。このシステムは、監視システムとして実用する場合に、人物追跡と行動認識ともに処理時間が膨大であることと、認識率が低いという欠点が挙げられる。よって本研究では、従来の MCMC 法により人物行動認識を行うシステムの、人物追跡の高速化と、行動認識の高効率化・高精度化を目的とする。

People Tracking and Action Recognition System Using Omnidirectional Images

YUKI KOSAKA,^{†,††} TAKATSUGU HIRAYAMA,^{†††} YOSHIO IWAI^{†††}
and MASAHICO YACHIDA^{†††}

We propose a system for tracking people and recognizing their actions through indoor scenes captured with an omnidirectional image sensor. The system is composed of two modules : tracking module and recognition module. The system has two disadvantages. First, the process of tracking people is off-line. Second, the recognition rate is low. In this work, we solve these problems. We apply the real-time tracking system to the tracking module. We also improve the performance of action pattern modeling important for action recognition.

1. はじめに

近年、身の回りの危険を自動的に検知し、回避を促すことで、人々に安心して安全な空間を提供するセキュリティシステムの開発が求められている。このシステムには、周りの環境を把握、理解することによって、異常事態の発生を自動で感知し、管理者に通知するといった機能が必要である。そこで、本研究では人物の行動を認識し、未知の行動を行う人物を不審人物として切り出すような監視システムへの応用を目的とし、全方位視覚センサ¹⁾を用いて、入力画像から人物の発見と追跡を行い、行動を認識するシステムの構築を目指す。

近年、人物の発見、追跡について様々な手法が提案されており、代表的なものとして、カラー情報により

肌色モデルと背景モデルを作成し、人物領域の抽出を行った後、線形ガウス型状態空間で対象の状態を推定することにより、カルマンフィルタを用いて予測追跡を行うもの²⁾がある。また、予め背景画像を取得しておき、入力画像との背景差分を取ることで人物の抽出を行い、この人物抽出を時系列間で繰り返す。そして、前フレームの人物領域と現在のフレームの人物領域の重なりから人物追跡を行うもの³⁾がある。この手法は、市販の画像処理ボードを用いて画素間演算の高速化を図り、画像中から複数人物を同時に発見して実時間追跡を実現している。

また、行動認識手法についても様々な研究が行われているが、本システムのように人物の歩行軌跡を用いた行動認識は大きく二つに分類できる。一つは、行動のある離散的な状態に変換し、その遷移を調べることによって行動を認識するもの⁴⁾で、もう一つは、連続的な動きをそのまま扱い、その動きを認識するもの⁵⁾である。前者のように、人物の行動状態を離散的に表現する手法の代表的な例として、隠れマルコフモデル (HMM)⁴⁾ や有限状態オートマトンを用いたもの⁶⁾が挙げられる。HMM はジェスチャーや音声認識でもよく用いられる手法である。一方、後者の連続的な動きを認識するために、近年ベイズ推定を用いた確率的

† 大阪大学大学院基礎工学部
School of Engineering Science, Osaka University

†† 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology,
the University of Tokyo

††† 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka University

な手法⁵⁾が提案されている。ベイズ推定は周辺分布を求める計算が非常に困難であるが、この手法では、Markov Chain Monte Carlo 法 (MCMC 法)^{7)~9)}により、周辺分布を近似計算することで解決している。MCMC 法は、HMM と比較して各状態のパラメータが明確であるという利点がある。

本研究が目指す監視システムとは、部屋の中で人物が入室してから退室するまでを監視し、人物の行動を認識することにより、未知の行動を行う人物を不審人物として実時間で切り出しを行うシステムである。従来、これを目指した人物行動認識システム⁵⁾が提案されている。このシステムは、人物が入室してから退室するまで監視して人物追跡を行い、歩行軌跡をモデル化する。そして、MCMC 法を用いて、人物が入室から退室までのモデルに沿った動きを行ったのかを認識するシステムである。しかし、このシステムは人物追跡と行動認識の処理時間が膨大であることや、認識率が低いという欠点があり、本研究が目標とする監視システムとして実用的でない。

そこで、本研究では人物の発見・追跡処理を高速化するために、従来提案された全方位視覚センサを用いた複数人物の実時間追跡システム³⁾を応用する。また行動認識処理には、MCMC 法を用いたシステム⁵⁾を高効率・高精度なものへと改良し、適用する。

本研究では人物の歩行軌跡を得るために、足下位置を特定しなければならない。しかし、前者の実時間人物追跡システム³⁾は人物領域を抽出することができるが、足下位置特定はできない。そこで、本研究では足下位置特定を行う処理を付加する。

また、後者の人物行動認識システム⁵⁾は、人物の入室から退室までの歩行軌跡を行動モデルとし、これを細かい部分的な軌跡(これを部分行動モデルとする)に分割してモデル化を行い、行動を認識している。行動モデルが正しく歩行軌跡を表現できなければ、どの行動モデルに沿って動いているのかを認識することが困難である。そこで認識率を向上させるために、最も重要な部分である行動モデル生成を改良する。本研究では、従来の行動モデル生成の欠点を挙げ、それを解決する行動モデル生成手法を提案する。

2. システム概要

本システムの概要を図 1 に示す。本システムは、カメラセンサからの入力画像より人物の検出と追跡を行う追跡モジュールと、人物の行動パターンを学習し、その行動を認識する認識モジュールで構成される。

2.1 全方位視覚センサ

本研究では、周囲 360 度の撮影が一度に可能な全方位視覚センサを用いて撮影を行う。標準画角のカメラでは、視野角が狭いために人物がすぐに視野から外れてしまうが、全方位視覚センサであれば、視野角が広

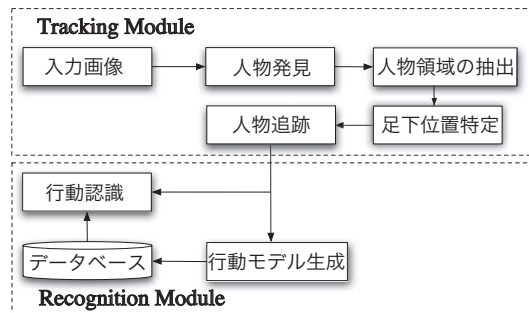


図 1 システム概要



図 2 全方位画像

いため同時に複数の人物を視野から外れることなく撮影することが可能である。追跡モジュールの入力として、この全方位視覚センサでの撮影画像を読み込む。

2.2 追跡モジュール

システムの利用環境は照明変化の比較的小さい屋内を想定し、直立して歩行する人物の入室してから退室するまでを追跡して、人物の足下位置により歩行軌跡を生成する。初めに、撮影された画像から、背景差分と人物の動きの連続性を用いて人物追跡をする。その後、図 2 のように全方位視覚センサは、カメラ中心に最も近い部分が人物の足下位置となるという特性があり、それを利用して人物の足下位置を特定する。最後に足下が床面に接しているという仮定から、人物の足下位置をカメラ座標から世界座標に変換して、世界座標における人物位置を特定する¹⁾。人物追跡を高速に行うために、画像処理ボードを使用する。

2.3 認識モジュール

認識モジュールでは、追跡により得られた軌跡を用いて人物の歩行パターンを学習し、その行動を認識する。本研究では入室から退室までを表す行動モデルと、部分的な軌跡を表す部分行動モデルという 2 つのモデルを用いる。行動モデルとは、人物が右のドアから入室し、コンピュータの前で立ち止まり、また同じドアから退室したというような、一連の行動の流れを表す。部分行動モデルは、右ドアから入室したという動きや、コンピュータの前で立ち止まったなどの一つ一つの動作を表す。このモジュールは予め、サンプル軌跡から

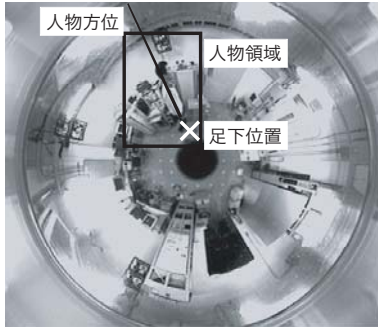


図 3 足下位置特定の様子

行動モデルと部分行動モデルを自動的に生成し、認識処理に必要な部分行動モデル間の遷移確率なども学習しておく。そして、生成された部分行動モデルの時系列データと遷移確率などをデータベースに保存する。行動認識は、入力されたテスト軌跡とデータベースの行動モデルを比較し、確率的な手法により行う。最後に、入力されたテスト軌跡はどの行動モデルに沿っていたかを、認識結果として出力する。

3. 人物追跡と足下位置特定

3.1 人物の発見と追跡

本システムでは人物の発見・追跡処理を高速化するために、従来提案された全方位視覚センサを用いた複数人物の実時間追跡システム³⁾を応用する。室内を想定し、人物が入室し、退室するまでの追跡を行う。人物領域の発見処理においては、背景差分法を用いる。まず、人物領域発見の前段階として、予め背景画像を取得する。そして、全方位視覚センサからの入力画像との背景差分を取り、動的閾値で2値化処理を行うことで、人物候補領域の抽出を行う。次にノイズ除去を行う。このノイズ除去処理は、8近傍により2値画像中の独立した1画素の除去を行う。そして、人物候補領域の膨張と収縮処理を行う。膨張・収縮処理により、抽出した人物候補領域の不連続の部位を結合させ、人物の形状を安定に抽出できる。次に、全方位画像中では、直立している人物は画像中心から放射状に現れるという性質を利用し、抽出された人物候補領域に対して、角度方向の加算投影処理を行う。これによって、人物候補領域が画像上のどの方位にあるのかがわかり、全方位視覚センサに対する方位ごとの差分頻度ヒストグラムを作成する。そして、このヒストグラムに対し、ある一定の閾値 R_{th} 以上の差分頻度を持つ領域を人物領域と定義する。また、その領域の中で最も差分頻度の大きな方位をその人物の人物方位と定義する。

人物の追跡は、人物領域の推定処理を時系列間で繰り返し、前フレームの人物領域と現在のフレームの人物領域の重なりから、各人物の追跡を行う。

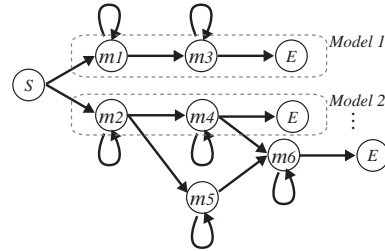


図 4 行動モデル

3.2 足下位置特定

本システムは、追跡と同時に人物の足下位置の特定を行う。人物追跡処理によって、人物領域の方位情報から画像中のどの範囲に人物領域があるのかわかる。全方位視覚センサの特性から、入力画像上でカメラ中心に最も近い部分を人物の足下位置と特定する。図3に人物領域から足下位置特定の様子を示す。次に、人物の足下が床面に接しているという仮定から、人物の足下位置をカメラ座標から世界座標に変換する¹⁾。

4. 人物の行動認識

4.1 人物の行動モデル

本研究では、追跡モジュールによって得られた人物の歩行軌跡を用いて、行動モデルを生成する。歩行軌跡は部分的に同じ軌跡を多く含むため、軌跡をいくつかの部分軌跡に分割し、データベースに保存することで、保持するデータ量の削減を図る。以下、人物が入室してから退室するまでの動きを行動モデル $M = \{M^\eta | \eta = 1, \dots, \mathcal{M}\}$ 、分割された部分的な動きを部分行動モデル $\mathbf{m} = \{\mathbf{m}^\mu | \mu = 1, \dots, m_N\}$ と呼ぶ。図4のように行動モデルは、部分行動モデルの遷移によってモデル化する。ここで、人物が入室した状態を S 、退室した状態を E とする。部分行動モデル \mathbf{m}^μ は、 τ_{max}^μ フレームに渡る人物位置の連続的な点列 (X_τ^μ, Y_τ^μ) の情報を持つ。

4.2 行動モデル生成

4.2.1 従来手法の問題点

初めに、入力された人物の歩行軌跡から部分行動モデルを自動的に生成する従来手法のアルゴリズムについて述べる。このアルゴリズムは新宅の手法¹⁰⁾を応用している。入力軌跡を I_x 、 I_x 上のフレーム s からフレーム e 間の軌跡を I_{se} 、データベースに保持してある部分行動モデルを $\mathbf{m} = \{\mathbf{m}^\mu | \mu = 1, \dots, m_N\}$ 、 I_{se} と \mathbf{m}^μ の類似度を $Dse(I_{se}, \mathbf{m}^\mu)$ とする。入力軌跡 I_x のフレーム t での位置を (X_t, Y_t) とし、部分行動モデル \mathbf{m}^μ のフレーム τ での位置を (X_τ^μ, Y_τ^μ) とする。類似度 $Dse(I_{se}, \mathbf{m}^\mu)$ は、

$$Dse(Ise, m^\mu) = \frac{1}{\tau_{max}^\mu} \sum_{\tau=1}^{\tau_{max}^\mu} \exp \left\{ -\frac{1}{2} \left(\frac{(X_{t-s+\tau} - X_t^\mu)^2}{\sigma_X^{\mu,2}} + \frac{(Y_{t-s+\tau} - Y_t^\mu)^2}{\sigma_Y^{\mu,2}} \right) \right\} \quad (1)$$

と求める．ここで、 σ_X^2 、 σ_Y^2 はそれぞれ部分行動モデル m^μ に対する位置 X と Y の分散である．従来の部分行動モデル生成のアルゴリズムは5つのステップから構成される．

Step 1 入力された軌跡 I_x の中に、データベースに保持されている部分行動モデルが含まれていないか探索する．入力軌跡 I_x 中のある区間の軌跡 Ise と、部分行動モデル m^μ との類似度 $Dse(Ise, m^\mu)$ を算出し、最大の類似度を持つ Ise と m^i を求める．

Step 2 類似度 $Dse(Ise, m^i)$ が閾値 T_d 以上であれば、入力軌跡 I_x から、 Ise を除いた前後の軌跡をそれぞれ I_x と考え、再び含まれる部分行動モデルを探索するため Step 1 を繰り返す．ここで、 I_x の長さ（フレーム数）が閾値 T_L の場合は棄却し、探索を行わない．

Step 3 類似度 $Dse(I_x, m^i)$ が閾値 T_d より小さければ、入力軌跡中に含まれる部分行動モデルはデータベースに存在しないとして、この入力軌跡を新たな部分行動モデルとして、データベースに保存する．これを m^{new1} とする．

Step 4 新しく生成された部分行動モデル m^{new1} が、データベースに保持してある部分行動モデルに含まれていないか探索を行う．部分行動モデル m^μ の軌跡を入力軌跡 I_x とし、その中のある区間の軌跡 Ise とモデル m^{new1} との類似度 $Dse(Ise, m^{new1})$ を求める．部分行動モデル m^j を入力軌跡 I_x とした時の類似度 $Dse(Ise, m^{new1})$ が、閾値 T_d 以上となれば、部分行動モデル m^j に新しく生成された部分行動モデル m^{new1} が含まれるとする．その場合、部分行動モデル m^j の軌跡から Ise を除いた前後の軌跡を、それぞれ新たに部分行動モデル m^{new2} 、 m^{new3} としてデータベースに保存し、部分行動モデル m^j をデータベース内から削除する．従って、この時点でデータベース中の部分行動モデルは $m = \{m^\mu | \mu = 1, \dots, j-1, j+1, \dots, m_N, new1, new2, new3\}$ となる．ただし、 m^{new2} 、 m^{new3} の長さが閾値 T_L 以下であれば棄却する．

Step 5 分割された部分行動モデル m^i 中の軌跡 m^{new2} と m^{new3} がデータベース中の部分行動モデルに含まれていないか探索を行う．

以上の操作を新たな部分行動モデルが生成されなくなるまで繰り返す．このアルゴリズムにより入力軌跡が、

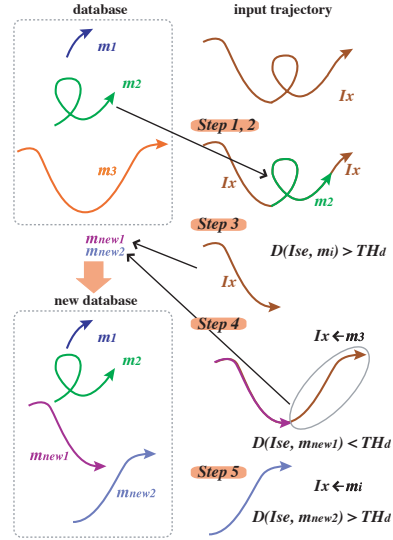


図5 部分行動モデル生成手法

部分行動モデルの遷移によって表され、この遷移は1つの行動モデルとなる．部分行動モデルの遷移が同じ行動モデルは、同一の行動モデルとする．

次に、上で述べた従来手法の問題点を挙げる．入力軌跡と部分行動モデルの類似度を計算する際に、図5に示すように入力軌跡に既存の部分行動モデルが含まれていないか調べる場合、部分行動モデルの長さに固定した探索を行っている．つまり、 $e-s = \tau_{max}^\mu$ である．よって、この従来手法は歩行軌跡の時間伸縮に未対応である．時間伸縮とは、歩行軌跡が同じ時に歩行速度が異なる場合のことを指す．歩行速度が異なる場合に、歩行軌跡が同一のモデルを違うモデルとして考える場合、保持するデータ量が膨大となり、本研究の目的に反する．そこで、時間伸縮に対応した部分行動モデル生成のアルゴリズムを次節で提案し、保持するデータ量を削減する．ただし、不審行動の認識に歩行速度を考慮することが重要となる可能性がある．これに関しては、次節で説明する行動認識アルゴリズムでの対応を考える．

4.2.2 DP マッチングを用いた部分行動モデル生成手法

4.2.1 において、従来の行動モデル生成手法は、部分行動モデルの長さに固定した探索を行っており、時間伸縮に対応できないという問題点を挙げた．そこで、時間伸縮に対応するために、DP マッチング¹¹⁾を用いた部分行動モデル生成手法を提案する．DP マッチングは、時間伸縮に対応した照合手法であるため、DP マッチングのコストを類似度に置き換えることで、入力軌跡と部分行動モデル間の類似度を可変長で算出することができる．入力軌跡 I_x 上のフレーム s からフレーム e 間の軌跡を $Ise = \{i_1, i_2, \dots, i_p, \dots, i_m\}$ とする．また、部分行動モデル

ル \mathbf{m}^μ は, $\mathbf{m}^\mu = \{m_1, m_2, \dots, m_q, \dots, m_n\}$ とする。ここで, I_{se} の長さは τ_{max}^μ に固定されない。入力軌跡と部分行動モデルの類似度 f を式 (2) ~ 式 (5) で求める。

$$f(0, 0) = 0, \quad k = 0 \quad (2)$$

$$f(p, q) = \max \left\{ \begin{array}{l} f(p-1, q) + d(i_{p-1}, m_q), \\ f(p, q-1) + d(i_p, m_{q-1}), \\ f(p-1, q-1) + d(i_{p-1}, m_{q-1}) \end{array} \right\} \quad (3)$$

$$d(i_p, m_q) = \exp \left\{ -\frac{1}{2} \left(\frac{(X_{i_p} - X_{m_q})^2}{\sigma_X^2} + \frac{(Y_{i_p} - Y_{m_q})^2}{\sigma_Y^2} \right) \right\}, \quad k = k+1 \quad (4)$$

$$Dse(I_{se}, \mathbf{m}^\mu) = \frac{f(m, n)}{k} \quad (5)$$

(X_{i_p}, Y_{i_p}) は i_p の世界座標, (X_{m_q}, Y_{m_q}) は m_q の世界座標を表す。部分行動モデル生成のアルゴリズムは, 従来法と同じ5つのステップで構成する。また, 部分行動モデルの遷移確率 $p(\mathbf{m}^i | \mathbf{m}^j)$ と, 各行動モデルに対する部分行動モデルの初期生起確率 $p(\mathbf{m}^j | M_{t=0})$, 部分行動モデルに対する行動モデルの生起確率 $p(M_t | \mathbf{m}^i)$ をモデルの生成と同時に学習する。

4.2.3 確率モデルによる行動認識⁵⁾

生成された行動モデルを用いて, 行動を認識する方法を説明する。時刻 t までの状態と観測値をそれぞれ $\mathcal{X}_t = \{\mathbf{x}_1, \dots, \mathbf{x}_t\}$, $\mathcal{Z}_t = \{\mathbf{z}_1, \dots, \mathbf{z}_t\}$ とする。ただし, 状態の遷移はマルコフ性を持ち, 次時刻の状態は現時刻での状態のみ依存し, それ以前の状態には依存しないとす。

$$p(\mathbf{x}_t | \mathcal{X}_{t-1}) = p(\mathbf{x}_t | \mathbf{x}_{t-1}) \quad (6)$$

行動モデルの認識は, 観測値に対して最大事後確率を持つ行動モデルを最適な行動モデル M_s として認識する。

$$M_s = \arg \max_{M_t} P(M_t | \mathbf{z}_t) \quad (7)$$

ここで, $P(M_t | \mathbf{z}_t)$ は, 行動の状態 \mathbf{x}_t とベイズの定理とを用いて,

$$P(M_t | \mathbf{z}_t) = \int p(M_t, \mathbf{x}_t | \mathbf{z}_t) d\mathbf{x}_t \propto \int p(\mathbf{z}_t | \mathbf{x}_t, M_t) p(\mathbf{x}_t | M_t) d\mathbf{x}_t \quad (8)$$

となる。 $p(\mathbf{x}_t | M_t)$ は行動モデル M_t が起こったときの状態 \mathbf{x}_t の生起確率で, 部分行動モデルを用いて以下のように変形することができる。

$$p(\mathbf{x}_t | M_t) = \sum p(\mathbf{x}_t | \mathbf{m}^i) p(\mathbf{m}^i | M_t) d\mathbf{m}^i \quad (9)$$

さらに, 部分行動モデルの遷移確率を用いると,

$$p(\mathbf{m}^i | M_t) \propto p(M_t | \mathbf{m}^i) \sum p(\mathbf{m}^i | \mathbf{m}^j) p(\mathbf{m}^j | M_{t-1}) d\mathbf{m}^j \quad (10)$$

となる。4.2節により, 得られた部分行動モデルの遷移確率 $p(\mathbf{m}^i | \mathbf{m}^j)$ と, 初期生起確率 $p(\mathbf{m}^j | M_{t=0})$, 生起確率 $p(M_t | \mathbf{m}^i)$ を用いて式 (10) は再帰的に求めることができる。また, $p(\mathbf{x}_t | \mathbf{m}^i)$ は一様分布とする。

式 (8) の確率を求める確率分布の積分計算は, 分布の形状が未知であるため非常に困難である。そこで本研究では, MCMC 法¹²⁾ を利用して分布を近似的に計算する。状態 $\mathbf{x}_t \sim p(\mathbf{x}_t | M_t)$ で独立にサンプルできるとすると, サンプル $\{\mathbf{s}_t^{(n)}, n = 1, \dots, N\}$ を用いて, 式 (8) は以下のように近似できる。

$$P(M_t | \mathbf{z}_t) \propto \sum_{n=1}^N p(\mathbf{z}_t | \mathbf{x}_t = \mathbf{s}_t^{(n)}, M_t) p(\mathbf{x}_t = \mathbf{s}_t^{(n)} | M_t) \quad (11)$$

5. システム性能評価

被験者が部屋の中を動き回る様子を撮影し, 本システムの性能評価実験を行った。

人物追跡処理には, SONY の 3CCD カメラ XC-009 を使用し, 双曲面ミラーは山田光学工業製を用いた。そして, 日立製画像処理ボード IP5000 を使用した。また, これらを制御し画像を記録するために hp 社製ワークステーション xw4000 を使用した。CPU は Pentium 4 2.40B GHz で, メモリは 1GB, OS は Red Hat Linux 7.3 である。本実験では, 全方位画像を 512 × 440 画素で各点 8bit の濃淡画像により保持している。

行動認識処理は, SGI 社製ワークステーションの Origin300 で行った。撮影した部屋の大きさは約 7m × 7m で, カメラを部屋の中央に設置した。カメラの高さは 180.0cm である。図 6 に部屋の見取り図を示す。部屋の大きさは約 7m × 7m であるが, コンピュータや棚が置かれているために実際に人物が歩行できる範囲は約 6m × 6m であるため, 歩行軌跡はその範囲で取得した。

5.1 人物追跡実験

追跡モジュールにより人物を追跡する実験を行った。実験の目的は, 人物追跡と足下位置特定の評価を行うことである。

表 1 に, 画像取り込み, 人物追跡・方位特定, 足下位置特定, メモリ書き込みの各平均処理時間を示す。画像処理ボードを用いた処理により, 画像取り込みから足下位置特定, メモリ書き込みまでの処理を約 0.3[sec] で行うことができた。画像取り込みと人物追跡・方位特定には, 画像処理ボードによる処理が多いことから特に高速処理が可能である。しかし, 足下位置特定とメモリ書き込みは CPU による処理であるために低速

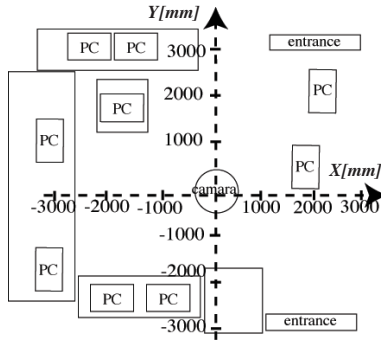


図 6 部屋の見取り図

表 1 人物追跡と足下位置特定の各平均処理時間

処理内容	平均処理時間 [ms]
画像取り込み	1.05
人物追跡・方位特定	92.1
足下位置特定	137
メモリ書き込み	46.3
合計	276

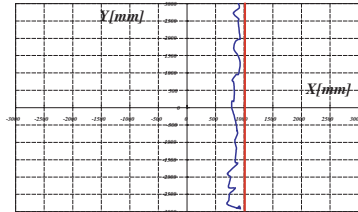


図 7 人物追跡・足下位置特定の誤差 1

で、全体として実時間での処理が不可能となった。よって、より高速に処理を行うために改良する必要がある。ただし、人間の歩行速度は約 $11.1[m/s]$ であることから、本システムでは、最小で約 $0.370[m]$ 間隔で足下位置を特定していることになる。この間隔で十分に行動認識を行うことができると考えている。

次に、実際に被験者に歩行してもらい、どのくらいの誤差があるのかを検証した。被験者には、図 7、図 8 に示す直線に沿って室内を歩行するように指示した。ぎざぎざな線が、被験者の足下を追跡した結果である。この実験を被験者 1 人に対し、10 通りの軌跡を 2 回ずつ行った。その結果、最大 30cm の誤差が確認され、平均で約 15cm の誤差が生じた。誤差の原因は、被験者に「直線上を歩行するように」という指示のみで実験を行ったために、線を踏んだり、線の内側や外側に足の着地点が入ってしまったことが考えられる。誤差が平均 15cm 生じたが、人物の歩行軌跡を十分に表現した足下位置特定となっていることは確認できた。ただし、この誤差は今後の認識モジュールで補うことができる。

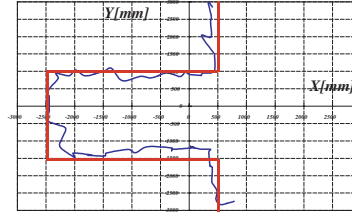


図 8 人物追跡・足下位置特定の誤差 2

5.2 行動モデル生成の実験

部分行動モデルと行動モデルの自動生成実験を行った。本実験は、従来の行動モデル生成手法と提案した DP マッチングによる行動モデル生成手法の比較を目的とする。以下の二種類の実験を行った。一つ目は、被験者 1 人に指定した軌跡 12 通りを 2 回ずつ歩いてもらい、24 個の歩行軌跡から生成される行動モデルの結果を比較する実験である。二つ目は、被験者 9 人に室内を自由に歩いてもらい、合計で 50 個の軌跡を得て、生成される行動モデルの結果を比較する実験である。両手法ともに、部分行動モデルの類似度の閾値 T_d, T_L の設定は同一であり、それぞれ 0.15, 10 とした。また、軌跡を入力する順番は従来手法と提案手法で同じとした。

一つ目の実験で指定した 12 通りの軌跡には、他の軌跡の一部分と似ている部位を意図的に含ませた。また、12 通りの軌跡は全体の動きとして似ている軌跡は無く、理想とする実験結果としては、24 個の歩行軌跡から行動モデル数が 12 個生成される。図 9 に 24 個の歩行軌跡を示す。表 2 に従来手法と提案手法により、24 個の歩行軌跡から生成した行動モデルと部分行動モデルの数を示す。提案手法は、生成した行動モデル数が 12 個になったことから、正確にモデル化できたと言える。しかし、従来手法は生成した行動モデルが 18 個で、その中に似ている軌跡を表す行動モデルが複数個あり、正確にモデル化できていない。12 通りの軌跡を 2 回ずつ行っていることから、従来手法の結果のように行動モデル数が 12 個を超えてしまうことは、モデル生成手法の性能が悪いことを意味する。また、提案手法は従来手法に比べて、行動モデル数と部分行動モデル数ともに少なく、効率よく部分行動モデルを生成していることを実証した。図 10 と図 11 に 24 個の入力軌跡の中から、入力軌跡 1 と入力軌跡 2 を示す。図 12 に、提案手法により生成した入力軌跡 1 を表す行動モデル 1 ($M1$) を示す。行動モデル 1 は、4 つの部分行動モデルの遷移によって成り立つ。図 13 に、提案手法により生成した入力軌跡 2 を表す行動モデル 2 ($M2$) を示す。行動モデル 2 は、4 つの部分行動モデルの遷移によって成り立つ。図 12 と図 13 の部分行動モデル 0 と 3 が、行動モデル 1 と 2 の両方の表現に利用されており、2 つの入力軌跡で、類似した

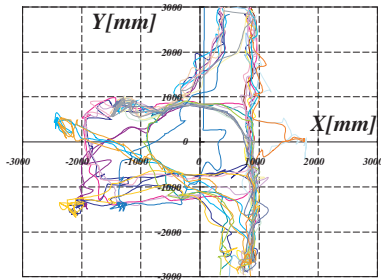


図 9 24 個の入力軌跡

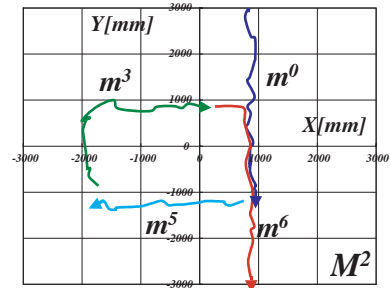


図 13 行動モデル 2

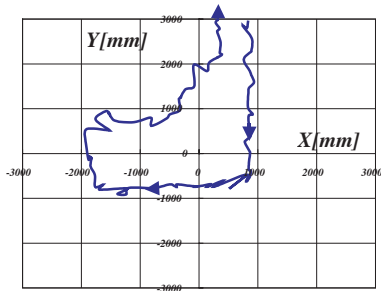


図 10 入力軌跡 1

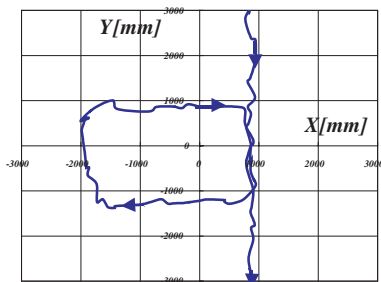


図 11 入力軌跡 2

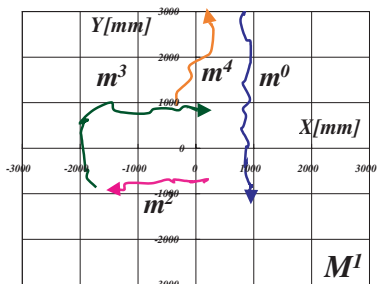


図 12 行動モデル 1

部分を同一の部分行動モデルとしていることから、正しく部分行動モデルを生成できたことがわかる。

二つ目の実験は、被験者 9 人に室内を自由に歩くように指示し、合計で 50 個の軌跡を得た。そして、その歩行軌跡を 10 個ずつ入力して、入力した歩行軌跡の数に対する行動モデルと部分行動モデルの生成の結果を比較する。

表 2 行動モデル・部分行動モデル生成結果

手法	行動モデル数 [個]	部分行動モデル数 [個]
従来手法	18	22
提案手法	12	19

表 3 行動モデル・部分行動モデル生成結果

入力軌跡数	行動モデル数		部分行動モデル数	
	従来手法	提案手法	従来手法	提案手法
10	10	10	12	17
20	18	19	35	33
30	28	26	49	41
40	36	33	71	50
50	46	42	77	59

実験結果を表 3 に示す。入力軌跡が 10 個の結果は、従来手法が提案手法よりも部分行動モデルの数が少ない。入力軌跡が 20 個の結果は、従来手法が提案手法よりも行動モデルの数が少ない。しかし、提案手法は部分行動モデルの数が従来手法よりも少ないことから、少ない部分行動モデルで多くの行動モデルを表していることがわかる。入力軌跡が 30 個、40 個、50 個の結果は、提案手法が従来手法よりも行動モデル数、部分行動モデル数ともに少ない。特に入力軌跡が 50 個の結果では、従来手法と提案手法の行動モデルの数の差が、4 個であるのに対して、部分行動モデルの数の差は 18 個であることから、提案手法は行動モデルの数が増えるほど、少ない部分行動モデルで多くの行動モデルを表すことができる可能性がある。また、被験者に室内を自由に歩くように指示したので、歩行速度が被験者によりばらつきがあり、提案手法が時間伸縮に対応していることから、提案手法の行動モデル数や部分行動モデル数が従来手法に比べて少なくなったと考えられる。これらの実験結果から、従来手法に比べて提案手法の性能が良いことを実証できた。本手法では、行動モデルと部分行動モデルはオンライン学習によって生成されるため、入力する軌跡の順番に大きく依存することも考えられるが、入力する軌跡の順番を入れ替えても、結果に大きな差がないことも確認した。

6. 結 論

本研究は、人物の行動を認識し、未知の行動を行う人物を不審人物として切り出すような監視システムへの応用を目的として、全方位視覚センサにより撮影された画像を用いて人物の追跡を行い、その行動を認識するシステムの構築を目指している。

従来のシステムは、人物追跡と行動認識の処理時間が膨大であることや、認識率が低いという欠点が挙げられる。よって本研究では、この2つの欠点を改良することを目的とし、人物追跡の高速化と行動認識の効率化・高精度化を行った。人物追跡の高速化のために、従来提案されている実時間人物追跡システムを改良して、人物追跡と足下位置特定を約0.3[sec]で行うことができた。そして行動認識について、従来の行動モデル生成手法には時間伸縮に未対応という欠点があることから、時間伸縮に対応した照合手法であるDPマッチングを用いて、行動モデルを生成する手法を提案した。従来手法と提案手法の比較実験を行った結果、提案手法が従来手法に比べて、少ない部分行動モデルでより多くの行動モデルを表すことができ、提案手法が優れていることを実証した。行動モデルが正しく軌跡を表現できなければ、どの行動モデルに沿っているのかを認識することが困難であることから提案手法で行動モデルを生成し、行動認識を行うことで高い認識率が期待できる。

今後の課題は、提案した行動モデル生成手法によって生成された行動モデルと部分行動モデルをMCMC法を用いた行動認識に適用し、従来システムよりも認識率が向上することを実証することである。そして、不審人物を切り出すような監視システムへの応用を実現することである。

なお、本研究は大阪大学基礎工学部で行った。

参 考 文 献

- 1) 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,” 電子情報通信学会論文誌 (D-II), vol.J79-D-II, no.5, pp.698-707, 1997.
- 2) A. Matsumura, Y. Iwai, and M. Yachida, “Tracking Algorithm for Detecting Trajectories of Human Movement from Omnidirectional Images,” Proc. of the 1st International Symposium on System & Human Science(SSR2003), pp.251-256, 2003.
- 3) T. Mituyosi, Y. Yagi, and M. Yachida, “Real-time Human Feature Acquisition and Human Tracking by Omnidirectional Image Sensor,” Proc. of IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems, pp.258-263,

2003.

- 4) 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,” 電子情報通信学会論文誌 (D-II), vol.J85-D-II, no.7, pp.1265-1270, 2002.
- 5) A. Matsumura, Y. Iwai, and M. Yachida, “Stochastic Action Recognition from Omnidirectional Images,” Proc. of Asian Conf. on Computer Vision, Vol. 1, pp.120-125, 2004.
- 6) T. Wada, and T. Matsuyama, “Multiobject behavior recognition by event driven selective attention method,” IEEE Trans. on PAMI, vol.22, no.8, pp.873-887, 2000.
- 7) M.J. Black, and A.D. Jepson, “A probabilistic framework for matching temporal trajectories: CONDENSATION-based recognition of gestures and expressions,” European Conf. on Computer Vision, vol.1, pp.909-924, 1998.
- 8) H. Sidenbladh, M.J. Black, and L. Sigal, “Implicit probabilistic models of human motion for synthesis and tracking,” European Conf. on Computer Vision, vol.1, pp.784-800, 2002.
- 9) M. Isard, and A. Blake, “CONDENSATION - conditional density propagation for visual tracking,” International Journal of Computer Vision, vol.29, no.1, pp.5-28, 1998.
- 10) 新宅聡子, “部分ジェスチャ抽出の試み,” 特別研究報告, 大阪大学基礎工学部, 2002.
- 11) 長尾真, パターン情報処理, コロナ社, 1983.
- 12) W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, MARKOV CHAIN MONTE CARLO IN PRACTICE, Chapman & Hall, London, 1996.
- 13) 近藤邦宏, 加藤直樹, 渡邊俊典, “データ圧縮を利用したオンライン・スケッチ認識手法 OSR,” 情報処理学会論文誌, vol. 38, no.12, pp.2468-2478, 1997.