

視覚移動ロボットと環境カメラの協調による複数人物追跡

Multi-person tracking by cooperation of a mobile robot and environmental cameras

1091011



研究代表者 豊橋技術科学大学 工学部 情報・知能工学系 助 手 佐 竹 純 二

共同研究者 豊橋技術科学大学 工学部 情報・知能工学系 教 授 三 浦 純

〔研究の目的〕

本研究では、移動ロボットにカメラを取り付け、撮影した動画画像から複数の人物を検出・追跡することを目的とする。ショッピングセンターのような広い空間の中で、人やロボットが様々な方向へ自由に移動する状況を想定し、安定な追跡方法を検討する。カメラで撮影した動画画像から人物を追跡する技術は、行動認識や監視システムなどに利用でき、非常に重要な技術の一つである。

移動ロボットに取り付けたカメラで撮影した動画画像を用いる場合、歩行者が他の歩行者に完全に隠されてしまうような複雑なオクルージョンが頻繁に発生する。そこで、本研究ではステレオカメラを用い、距離情報を利用して複数人物の分離や追跡を行う。予測位置をもとに複数人物の重なりあったシルエット形状を認識することで、複数人物を正確かつ安定に追跡する方法を提案する。

本研究では、まず、移動カメラを用いた人物追跡の単体性能を向上させるため、ステレオカメラから得られる距離情報を利用した追跡方法を開発する。それから、静止障害物などによるオクルージョンに対応するため、ロボットカメラと周囲環境に設置されたカメラを併用して人物を追跡する方法について検討する。

〔研究の内容、成果〕

1. ステレオカメラを用いた複数人物追跡

1.1 距離情報を利用した人物検出

移動カメラで撮影した動画画像を用いて安定に人物を追跡するため、本研究では距離画像テンプレート(図1)を用いる。このテンプレートはステレオカメラから得られた距離画像をもとに背景領域と人物領域を2値とするシルエット画像であり、人物がカメラからの距離2[m]の時の距離画像から作成している。画像探索時にはカメラから対象人物までの距離に応じてテンプレートの大きさを変更することで、画像上でのスケールの変化に対応する。また、テンプレート中の人物領域の画素値はカメラから人物までの距離とする。背景領域の画素値については、カメラが移動するため、各フレームの入力画像から背景の距離を推定する(距離のヒストグラムを求め、累積90%となる距離を背景の距離とした)。

サイズ $H \times W$ の距離画像テンプレートを

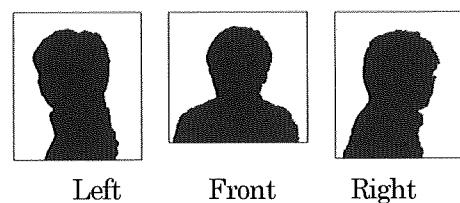


図1 距離画像テンプレート

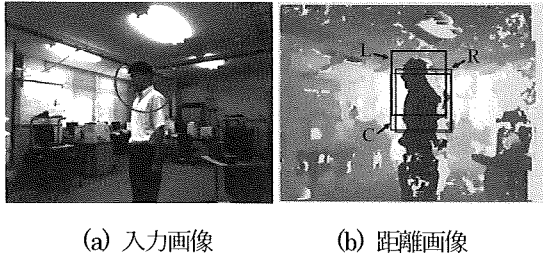


図2 距離画像テンプレートを用いた検出例

$T(x, y)$, 入力された距離画像を $I_D(x, y)$ とした時, テンプレートマッチングの相違度 d は次のように求められる。

$$d = \frac{1}{HW} \sqrt{\sum_p \sum_q [T(x, y) - I_D(x+p, y+q)]^2}$$

なお, 体の向きの変化に対応するため, 図1のような4方向3種類のテンプレートを併用し, 最も相違度が低いものをマッチングの結果として使用する。

距離画像テンプレートを用いた人物検出の例を図2に示す。3種類のテンプレートの探索結果の内, 最も評価値の良いもの(相違度の低いもの)を太線で表示しており, 複数のテンプレートを用いることで体の向きの変化に対応できていることが確認できた。

1.2 パーティクルフィルタを用いた3次元位置推定

使用したロボットやカメラと, 各座標系の関係を図3に示す。時刻 t における人物の3次元位置を (X_t, Y_t, Z_t) とし, 状態変数 x_t を次の

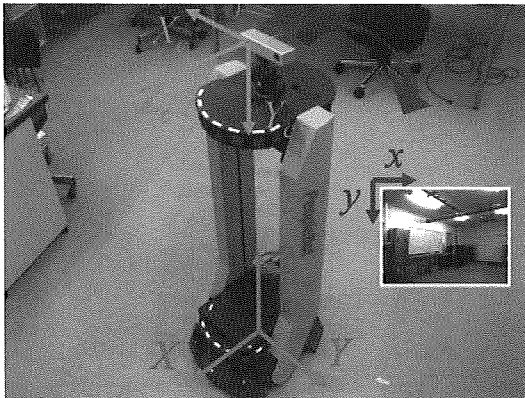


図3 各座標系の定義

ように定義する。

$$x_t = [X_t \ Y_t \ Z_t \ \dot{X}_t \ \dot{Y}_t]^T$$

ただし, $\dot{X}_t, \dot{Y}_t$ は XY 平面上の移動速度である。

自由に移動する人物を安定に追跡するため, 本研究ではパーティクルフィルタを用いて各人物の3次元位置を推定する。各パーティクルの尤度 L は相違度 d を用いて次のように計算する。

$$L = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{d^2}{2\sigma^2}\right)$$

1.3 シルエットの重なりを考慮した複数人物追跡

接近した複数人物を正確かつ安定に追跡するため, シルエットの重なりを考慮した複数人テンプレートを作成する。まず, 追跡対象人物が他の人物と離れている場合には, N 個のパーティクルを用いて独立に追跡する。2人の人物A, Bが接近し, 領域の重なりが生じた時, それぞれの人物の状態変数 x_t^A と x_t^B を結合し, 新たな状態変数 x_t^{AB} を次のように作成する。

$$x_t^{AB} = \begin{bmatrix} x_t^A \\ x_t^B \end{bmatrix}$$

ここで, 各人物のパーティクル数が N 個の時, 組合せの総数は $N \times N$ となる。しかし, 全ての組合せについて評価(マッチング)を行うのは計算コストが大きいため, 本研究ではそれぞれの人物のパーティクルの中で尤度の高いもののみを結合に用いる。結合したパーティクルの初期の尤度 L^{AB} は結合に用いたパーティクルの尤度 L^A, L^B の積 $L^{AB} = L^A L^B$ とする。

複数人テンプレートを用いた追跡の手順を図4に示す。まず, 結合したパーティクルそれぞれについて, 2人のシルエットが重なりあったテンプレートを作成する。2人の3次元位置から画像上での位置関係を計算し, カメラから遠い人物のテンプレートの上に, 近い人物の人物領域を上書きする。なお, ここでは計算コストを抑えるため, 図1の3方向のテンプレートの内, 各人物の移動方向に対応するテンプレート

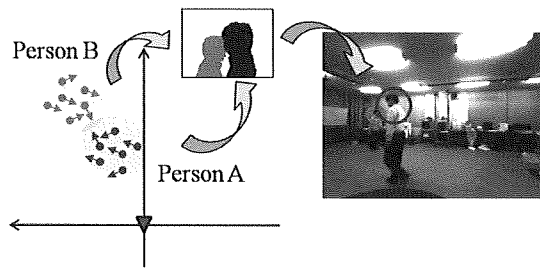


図4 複数人テンプレートを用いた追跡の手順

1つのみを用いる。そして、個別テンプレートの場合と同様に複数人テンプレートと距離画像とのマッチングを行い、尤度 L^{AB} を計算する。

接近していた人物 A, B が離れ、画像上のテンプレートが重ならない距離になった時、状態変数 x_t^{AB} を x_t^A と x_t^B に分離する。分離時には、結合された N^{AB} 個のパーティクルの内、尤度の高い N 個のパーティクルを用いる。

複数人テンプレートを用いた場合と、それぞれ個別のテンプレートを用いた場合の追跡結果の比較を図5に示す。複数人テンプレートを用いた追跡結果を白色、個別テンプレートを用いた追跡結果を赤色と青色の円で描画している。個別のテンプレートを用いた追跡では奥の人物(赤色)の推定位置がずれており、これはシルエットの一部が手前の人物によって隠されてしまったためと考えられる(図5左)。また、奥の人物が完全に隠されてしまった場合には、対象人物を見失い、誤って別の人物を追跡してしまった(図5右)。一方、複数人テンプレートを用いた場合にはそれぞれの人物を正しく追跡することができており、シルエットの重なりを考慮することで他の人物に隠された場合でも安

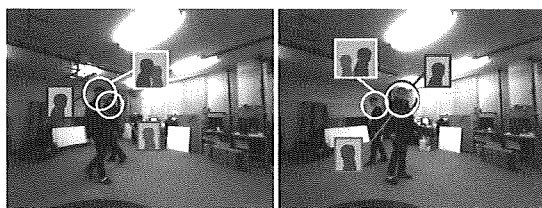


図5 複数人テンプレート(白枠)と個別テンプレート(赤枠・青枠)の追跡結果の比較

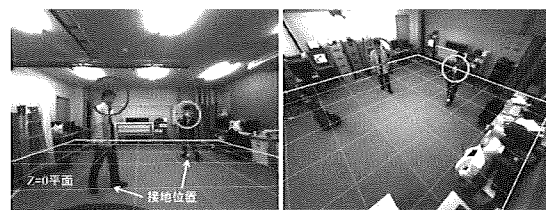
定な追跡が実現できていることが確認できた。

2. 複数台カメラへの拡張方法の検討

静止障害物が多く存在し、ロボットカメラだけで人物追跡を行うことが困難な状況では、周囲環境に設置されたカメラ(以下、環境カメラと呼ぶ)を併用する必要がある。複数のカメラやレーザレンジセンサを用いた研究も数多く行われているが、本研究では特に移動カメラによる人物追跡を主とし、ここでは提案手法を複数台カメラで併用する場合の拡張方法について検討する。ロボットカメラ1台と環境カメラ2台を併用した場合の例を図6に示す。このシステムは各カメラ画像を処理する追跡モジュールとそれらの結果を統合する統合モジュールで構成され、カメラ1台は移動ロボットを想定しているため、各追跡モジュールは別PCで実行するものとする。以下では、図7に示した2種類の統合方法について述べる。

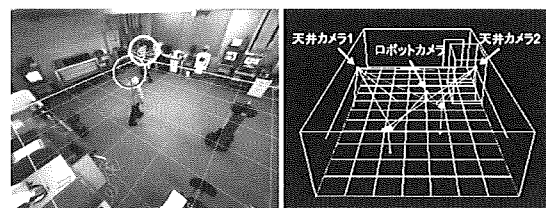
2.1 各カメラで個別に追跡する方法

まず、各カメラで個別に追跡を行う方法を述べる(図7(a))。この方法では、それぞれの追跡モジュールは各人物について N 個のパーティクルを保持し、それぞれ追跡を行う。なお、図6(c)のようにカメラ上で複数の人物が接近した場合には、提案手法を用いて人物シルエット



(a) ロボットカメラ

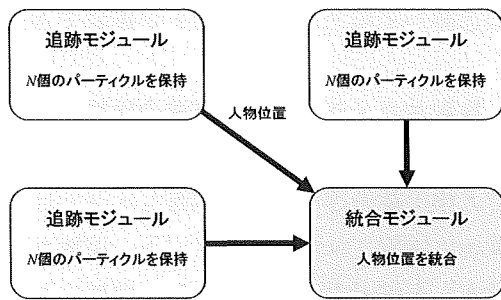
(b) 天井カメラ1



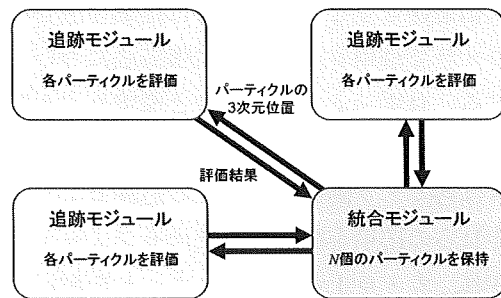
(c) 天井カメラ2

(d) 統合結果

図6 複数台カメラを用いた人物追跡



(a) 各カメラで個別に追跡する方法



(b) 複数カメラを統合して追跡する方法

図7 複数台カメラを用いた追跡の統合方法

トの重なりを考慮した追跡を行う。そして、推定した人物の3次元位置を統合モジュールへ送信する。統合モジュールでは、各追跡モジュールで推定された人物位置の情報を統合する。この時、各カメラ間の人物の対応付けを行い、あるカメラでの追跡に誤りがある場合には、その情報を該当の追跡モジュールにフィードバックする。

2.2 複数カメラを統合して追跡する方法

次に、複数カメラを統合して追跡を行う方法を述べる（図7(b)）。この方法では、統合モジュールで各人物について個のパーティクルを保持する。各パーティクルの3次元位置をそれぞれ全ての追跡モジュールに送信し、追跡モジュールは各パーティクルの評価のみを行う。そして、その評価結果を統合モジュールに送信し、統合モジュールで統合する。この時、カメラの視界の外に居る人物や、評価値が低く、他の人物や静止物体に隠れていると判断される人物については、そのカメラの評価結果を破棄する。このように複数台のカメラを用いて多方向

からの評価を同時に行うことで、より安定な人物追跡が実現できると考えられる。ただし、接近した人物の結合をどのように実装するかは今後の課題である。

なお、単に移動ロボットの見えていないところの情報を環境カメラから取得するだけでなく、オクルージョンの状況に応じてロボットをうまく移動させることで、固定カメラだけでは不十分な情報を補うことができると考えられる。このように、視覚移動ロボットと環境カメラとの相互協調によって安定な人物追跡を実現する方法について今後検討する。

3. 実験結果

3.1 ロボットカメラ1台を用いた人物追跡

図3に示した車輪移動ロボットとステレオカメラを用い、複数人物の追跡実験を行った。ロボットはMobileRobots社製のPeopleBotを用い、その上部にPointGreyResearch社製のステレオカメラBumblebee 2を設置した。1台のノートPCにより、ステレオ処理、人物検出・追跡、ロボット制御の全ての処理を行っている。処理に用いた画像サイズは 512×384 である。

ロボットの動き情報の誤差（滑りや遅延）の影響を排除するため、まずはロボットを移動させず、10 [fps] で撮影したオフライン画像を用いて追跡を行った結果を図8に示す。パーティクル数は $N = 100$ 、 $N^{AB} = 25 \times 25$ とした。白縁の付いた人物はパーティクルを結合していることを表している。#160～#202の人物Bや#310～#324の人物Aのように、他の人物に隠された状態で移動方向が変化した場合でも正しく追跡が行われている。これは生成されたシルエットの重なり方の仮説が、十分な多様性を持っていたためだと考えられる。

60回のオクルージョン（12ケースの画像セットを各5回テスト）について追跡の比較を行った結果を表1に示す。それぞれの画像セットは2人の人物が近づき、交差して離れるまで



図8 オフライン画像を用いた人物追跡の結果

表1 パーティクル数を変化させた場合の追跡結果の比較

(a) 個別テンプレートを用了場合

number of particles	success rate	positional error	processing time
$N = 100$	73.3 [%]	9.56 [pixel]	248 [ms]
$N = 200$	75.0 [%]	9.95 [pixel]	383 [ms]

(b) 複数人テンプレートを用了場合

number of particles	success rate	positional error	processing time
$N^{AB} = 15 \times 15$	86.7 [%]	7.09 [pixel]	226 [ms]
$N^{AB} = 20 \times 20$	90.0 [%]	6.74 [pixel]	314 [ms]
$N^{AB} = 25 \times 25$	93.3 [%]	6.64 [pixel]	436 [ms]
$N^{AB} = 30 \times 30$	90.0 [%]	6.53 [pixel]	568 [ms]
$N^{AB} = 100 \times 100$	91.7 [%]	6.43 [pixel]	4993 [ms]

のオフライン動画であり、各人物の移動方向や移動速度が様々に変化している。追跡の成功率は、のべ60回の交差の内、最後までそれぞれの人物を正しく追跡できたものの割合とした。また、2次元画像上の追跡誤差と1フレーム当たりの平均処理時間の計算には追跡に成功した画像セットのみを用いた。なお、各フレームにおけるそれぞれの人物の正解位置は手動で与えた。以上の結果から、シルエットの重なりを考慮した複数人テンプレートを用いた追跡は、個別のテンプレートを用いた場合よりも正確な人

物位置を推定できていることが確認できた。

最後に、実際にロボットを制御し、特定人物を追跡した結果を図9に示す。この実験では $N = 100$, $N^{AB} = 20 \times 20$ とし、ロボットは最初に検出した人物Aを追跡対象とするようにした。別の人物Bがロボットと人物Aの間を横切り、オクルージョンが発生した場合でも正しく追跡が行われていることが確認できた。

3.2 ロボットカメラと環境カメラを用いた人物追跡

図6に示したロボットカメラ1台と環境カメラ2台を併用して人物追跡実験を行った。3つの追跡モジュールと統合モジュールはそれぞれ別のPCで実行し、4台のPC間でTCP/IPを用いて情報を通信した。図7(b)の統合方法を用いて追跡を行った結果を図10に示す。ロボットカメラの画像中で人物位置を示す円が黒くなっているものは、全パーティクルの評価値が小さく、オクルージョンが発生していると判断されたことを示している。これらの結果から、他の人物や静止障害物によってオクルージョンが発生した場合でも、正しく追跡できていることが確認できた。

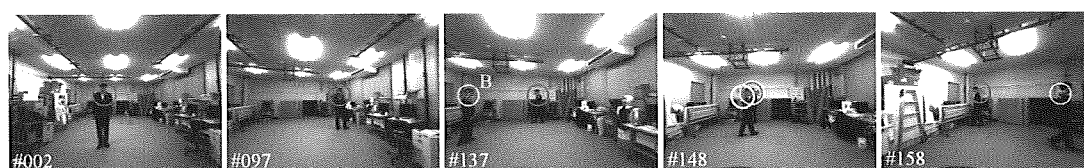


図9 特定人物追跡の結果

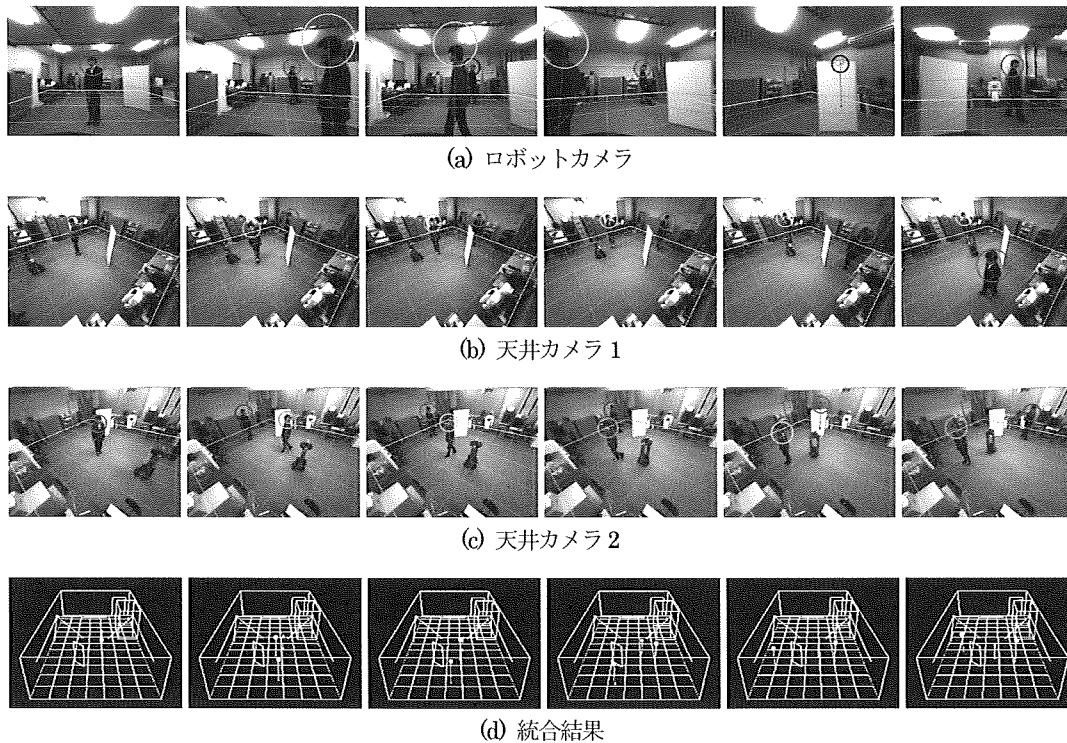


図 10 複数台カメラを用いた人物追跡結果

4. まとめ

本研究では、まず、移動カメラを用いた人物追跡の性能を向上させるため、ステレオカメラから得られた距離情報を利用して複数人物を追跡する方法を提案した。シルエットの重なりを考慮した距離画像テンプレートを用いることで、正確かつ安定な追跡を実現した。そして、実際に移動ロボットによる特定人物の追従実験を行い、提案手法の有効性を確認した。また、複数台のカメラへ拡張する方法を検討した。1台のロボットカメラと2台の環境カメラを併用した追跡実験を行い、静止障害物によってオクルージョンが発生した場合でも正しく追跡できていることを確認した。

[今後の研究の方向、課題]

今後の課題としては、まず、人物の素早い動きに対応するために処理速度を向上させる必要

がある。また、複数人テンプレートを人物の検出に拡張し、既知の人物2人が接近した場合だけでなく、追跡中の人物の後ろに重なっている未知の人物を検出できるようにする。そして、3人以上の人物が接近した場合の追跡方法について検討し、より複雑な状況での複数人物追跡の実現を目指す。

[成果の発表、論文等]

- [1] J. Satake and J. Miura: "Stereo-Based Multi-Person Tracking using Overlapping Silhouette Templates", Int. Conf. Pattern Recognition (ICPR 2010), Aug. (2010)
- [2] 佐竹純二, 三浦純: "人物シルエットの重なりを考慮したテンプレートを用いたステレオビジョン複数人物追跡", 画像の認識・理解シンポジウム (MIRU 2010), Jul. (2010)
- [3] 佐竹純二, 三浦純: "移動ロボット制御のための人物シルエットの重なりを考慮した複数人物追跡", 情報処理学会研究報告, 2010-CVIM-172, May (2010)