

テンプレートの可変分割と統合による人物の動作追跡

齊藤 慎也 佐治 斉

静岡大学大学院 理工学研究科

概要 人物全体をテンプレートとしたマッチングによる人物の動作追跡では、テンプレート内に含まれる背景領域や人物の姿勢変化などにより誤追跡となりやすい。そこで本研究では、人物全体を囲む領域をテンプレートとして抽出した後、人物の形状に合わせてテンプレート領域を分割・統合してマッチングを行う動作追跡システムを提案する。まず、背景差分を利用して人物領域を抽出し、その領域を囲むようなテンプレートを作成する。そして、テンプレート内に背景領域と人物領域が混在している場合は、テンプレートを4つのブロックに分割し、逆に背景領域または人物領域のみの場合は分割しない。この処理を繰り返す。次に、分割されたブロックのうち、背景のみの領域を含むものを削除し、人物の中心に近いものは統合する。実画像上で抽出されたテンプレートに対してマッチングを行った結果、正しい追跡結果が得られた。

Variable Division and Integration of the Template for Human Motion Tracking

Shinya Saito and Hitoshi Saji

Graduate School of Science and Engineering, Shizuoka University

Abstract Human motion is often tracked by matching a template whose area surrounds the human region. However, this method causes mismatching because of the background region and the human form change. We propose a new tracking method using the template matching. In our method, a template is divided and integrated fitting to the human form. First, we extract a template around a human region by computing the difference between the background image and the current image frame. Next, if the template area includes both the background region and the human region, we divide the template into four blocks. If the template area includes only the background region or only the human region, we do not divide the template. We repeat this process. Finally, we delete the blocks including the background area only and integrate the blocks near the human center. We obtain the correct matching result by using the template extracted on the real image.

1 はじめに

近年、画像処理技術の進歩に伴い、動画画像処理に関する研究が多くなされている。特に、人物の動作追跡に関しては、人物の位置や動きを把握する人物の動作追跡の研究 [1-6]、顔や手足といった人体の一部を追跡するジェスチャ認識の研究 [7][8]、人体の複雑な動作を追跡する3次元人体追跡の研究 [9] などが行われ、コンピュータビジョンやCG、スポーツ科学、セキュリティなど様々な分野で期待されている。このような人物の動作追跡に関する研究では、動作追跡の際に使用される手法として、テンプレートマッチングがある。

テンプレートマッチングは、入力画像とテンプレート画像とを重ね合わせることにより比較照合し、両者が一致しているかどうかを判定する処理であり、代表的な手法として、差分法や相関法などがある [10][11]。しかし、人物全体をテンプレートとして扱った場合、テンプレートマッチングによる人物の動作追跡では、次のような問題がある。

まず、人物の形状の変化による問題がある。例えば、人物の手足の動きにより姿勢や向きが変化するため、画像内の人物の形状も変化する。しかし、従来のテンプレートマッチングでは、テンプレートの大きさは変化しないため、テンプレート内の背景領域が大きくなったとき、誤追跡となる。これに対処するため、人物領域の抽出処理に使用したアルゴリズムを、マッチング結果を覆った狭い領域のみに適用することで、マッチングの結果ごとにテンプレートの大きさを更新する手法を考案した [1]。しかし、複数人物が重なるときには、誤追跡となった。

次に、複数人物の交差・交錯による問題がある。これは、複数人物の動作追跡において、画像内の複数の人物が、交差あるいは交錯している状態にあるとき、誤追跡となる問題である。従来の研究では、複数の人物が重なっているとき、カメラ側から見て隠れている人物の追跡を行っていない。そして、人物が離れたとき、再び人物領域を抽出し、追跡を行っている [4][6]。しか

し、複数の人物が重なっているかどうかの判断そのものが困難である。

このような問題に対して、本研究では、テンプレート領域を小さなブロック領域に分割し、分割したブロック領域ごとにマッチングを行う。しかし、テンプレート領域内には、背景領域と人物領域が混在している。そして、背景と人物との境界領域では、ブロックを小さく分割するほど、マッチング結果は正確になる。また、背景のみあるいは人物のみの領域では、ブロックを分割しない方が、マッチング結果は正確になる。従って、テンプレート領域を可変分割してマッチング結果の精度を上げる。また、分割されたブロックの内、マッチングを行う必要のないブロックは削除し、統合できるブロック同士は統合させる。

以上より本研究では、人物全体をテンプレートとしたマッチングによる動作追跡を行うことを目的とし、テンプレート領域を可変分割・統合したブロックマッチングによる人物の動作追跡システムを提案する。

2 処理の流れ

本手法では、前処理、テンプレート領域の可変分割・統合処理、そして、追跡処理の3つの処理で構成される。まず、前処理では、背景画像と入力画像との差分を求め、人物領域を抽出する。そして、人物領域に対してラベリングを行い、人物全体をテンプレート領域として抽出する。次に、テンプレート領域の可変分割・統合処理では、まず、抽出したテンプレート領域を複数のブロック領域に可変分割する。次に、ブロックの削除と統合を行い、マッチングを行うブロックを選択する。そして、追跡処理では、分割・統合されたブロックに対して次の入力画像とのマッチングを行う。そして、ブロックごとのマッチング結果を統合して人物全体での追跡結果を求める。

3 処理の内容

3.1 前処理



図 1: 前処理の流れ

前処理の流れを図 1 に示す。前処理では、背景画像と入力画像に対して、背景差分と二値化処理、雑音除去を行い、移動物体を検出する。そして、検出された領域に対してラベリング処理を行い、人物全体を覆うような領域をテンプレート領域として抽出する。図 2 は実画像における前処理の結果の例である。

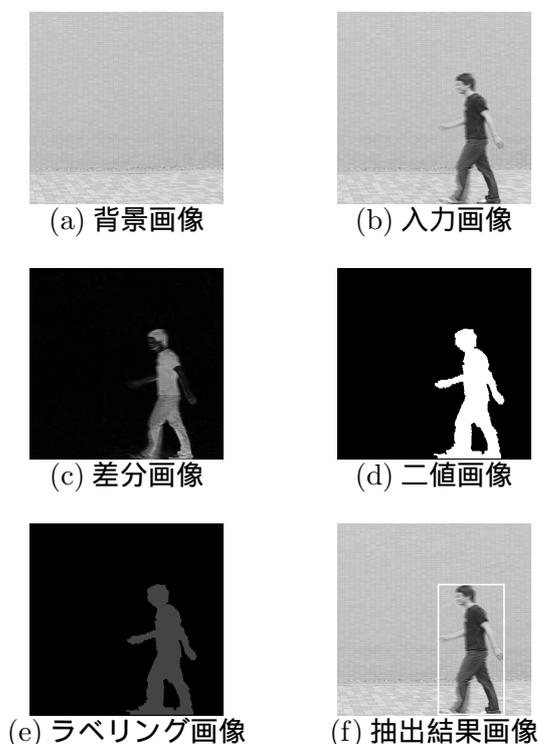


図 2: 人物領域の抽出結果

3.2 テンプレートの可変分割と統合処理

テンプレート領域の可変分割

まず、テンプレート領域を、人物の形状に合わせてテンプレート領域を小さなブロック領域に分割する。本研究では、人物の形状に対して、背景領域を黒画素そして人物領域を白画素とした二値画像を使用して分割を行う。そして、図 6(a) に示すように、テンプレート領域内に背景領域と人物領域の双方が混在しているときは、そのテンプレート領域を 4 分割する。逆に、図 6(b) に示すように、背景領域のみまたは人物領域のみのときは分割しない。これを繰り返すことでテンプレート領域の可変分割を行う。図 4 は実画像からにおけるテンプレート領域の可変分割結果の例である。

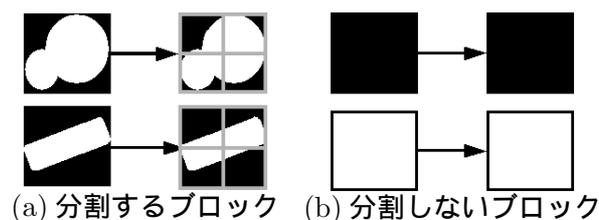


図 3: テンプレート領域の分割基準

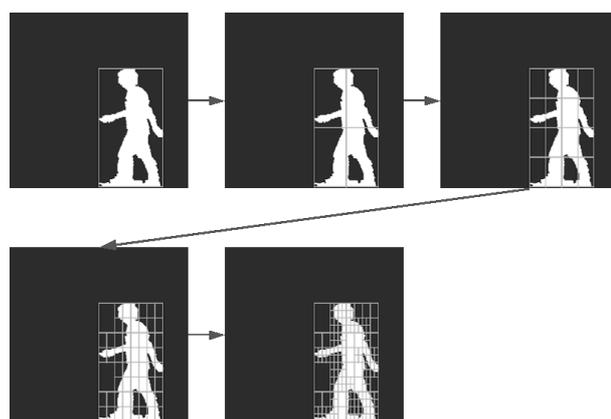


図 4: テンプレートの可変分割の結果

ブロック領域の統合

次に、ブロックの削除及び統合を行う。テンプレート領域を小さなブロック領域に分割したことで、テンプレート内には、背景領域のみを含む図 5(a) に示すようなブロック、背景領域と人物領域が混在している図 5(b) に示すようなブ

ロック，人物領域のみを含む図 5(c) に示すようなブロックが存在する．



図 5: ブロック領域の種類

まず，背景領域のみを含むブロックでは，人物領域がブロック内に存在しないため，マッチングを行う必要は無い．従って，このようなブロックは削除する．次に，背景領域と人物領域が混在しているブロックでは，ブロック内の人物領域の割合が極端に少ない場合，マッチング結果が背景領域の影響を受けやすくなる．従って，そのようなブロックは削除する．次に，人物領域のみで表されるブロックでは，胴体に位置するブロックを中心にブロックの統合を行う．人物領域内におけるブロック領域のマッチングでは，背景領域の影響を受けない分だけ，精度の高いマッチング結果を求めることができる．そこで，本研究では，人体の中心に近い，人物領域のみのブロックの内，隣接する同じ大きさのブロックを統合することで，マッチングの精度を上げる．

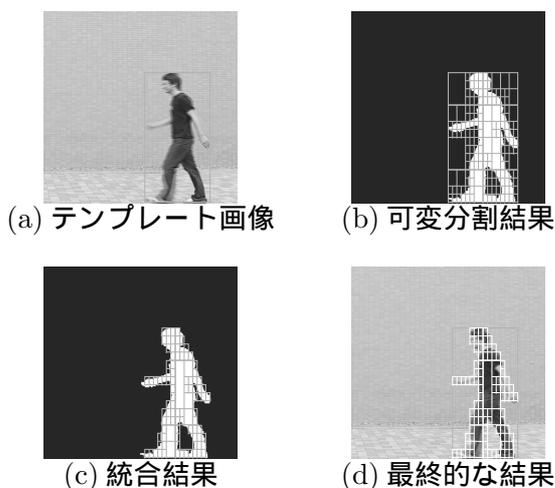


図 6: 可変分割・統合の結果

3.3 追跡処理

マッチングによる動作追跡

追跡処理では，現フレームの画像において，可変分割・統合処理によって作成されたテンプレート領域内の各ブロック領域に対して，次フレームの画像とマッチングを行う．マッチングでは，正規化相関法により算出した類似度を用いる．そして，ブロックの照合位置関係から，人物全体での枠を決定し，動作追跡結果とする．このとき，あるブロックにおけるマッチングにおいて，走査した結果得られた最大類似度が低い場合，そのブロックは削除する．

複数人物の交差・交錯

複数人物における動作追跡において，人物同士が交差・交錯しているかどうかの判断は，ブロック領域におけるマッチングの類似度やお互いの位置関係から決定する．まず，お互いの位置関係や境界領域のブロックの類似度から一方が遮へいしていると判断できる場合，人物全体での枠を更新しない．そして，最大類似度が低いブロックを削除する．そして，追跡処理を続け，胴体部分のブロックにおいて，一方の最大類似度が低くなった場合，その人物は完全に隠れていると判断して追跡処理を止める．この後，手前の人物の周辺において，前処理と同様の処理を行い，新たな人物が検出されたら，隠れていた人物をその人物に割り当てる．

4 実験

4.1 実験条件

本研究では，フレームレートが 30 フレーム/秒の汎用ビデオカメラを使用し，人物が歩いている風景に対して，動作追跡を行い，その精度について検証を行った．

実験では，CPU1.5GHz(Pentium4)，メモリ 384MB のコンピュータを使用した．また，1 フレーム辺りの画像情報は，サイズ 720×480，256 階調のフルカラー画像である．

以上の条件において，人物の動作追跡を行った．

4.2 実験結果

まず，人物全体における動作追跡の結果を図7に示す．この結果において，人体を覆った枠が背景の影響を受けずに正しく追跡できていることが分かる．

分かる．この原因として，手や足が，その振りの速さによってぼやけてしまうことや，振りによる角度の違いが生じていることが考えられる．

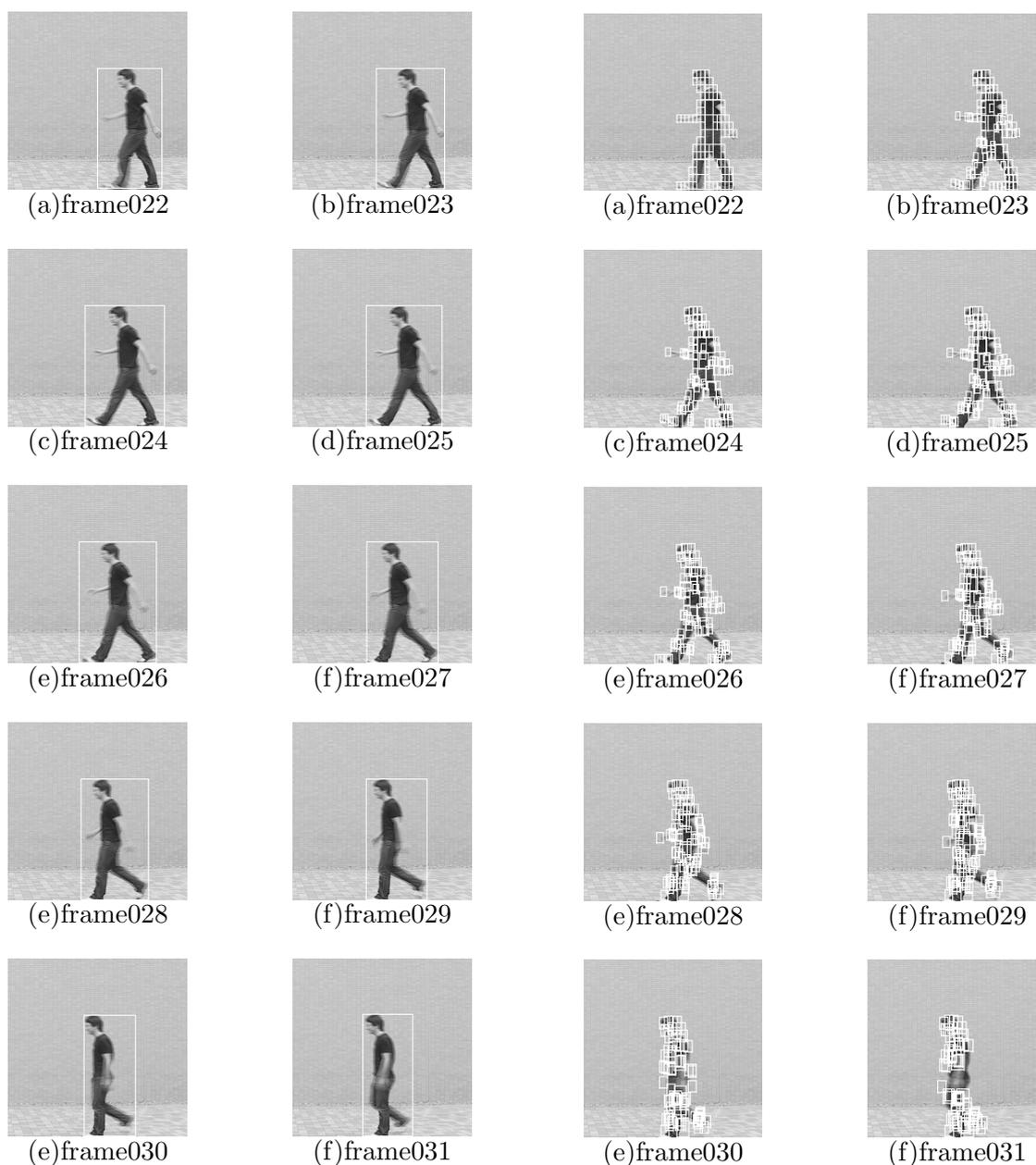


図 7: 実験結果 1

図 8: 実験結果 2

次に，ブロックごとのマッチング結果を図8に示す．この結果において，手や足の動きに対して，ブロックが正しく追跡できていないことが

5 まとめ

5.1 結論

本研究では、テンプレートの可変分割・統合による人物の動作追跡を行った。実験結果から、テンプレートを可変分割・統合することで、人物の形状の変化に対応できることを示した。

しかし、分割されたブロック一つ一つにおけるマッチングの結果を見ると、正しく追跡できていないブロックがいくつかあった。特に手や足の動作によってぼやけたり角度が変化したりする場合や、カメラ側から見て手前にある手の動作によって胴体の一部分が隠れる場合などにおいて、マッチング結果にずれが生じていた。

このような問題に対して、ブロックごとの動きベクトルやブロックの相互位置関係などを利用した位置予測や、一部のブロックにおける角度計算を取り入れることを検討している。

5.2 今後の課題

以下に今後の課題を列挙する。

- 複雑な条件下での実験

本研究では、一人の人物が画像内で交差・交錯するという内容の画像を用いて実験を行った。実利用を目的とした場合、人が屈む動作や複数人物の交差・交錯など、より複雑で実環境に近い条件に対して実験を行う必要がある。

- ブロックの位置予測

手や足がぼやけていたり、胴体と重なっていたりするような場合、ブロックの位置予測を行うことは、マッチングの探索時間を減らすことや余計なマッチングを省くことができ、精度の高いマッチングが行えると考えられる。

- 背景と人物の境界領域における角度計算

背景と人物の境界領域に作成されたブロックのマッチングでは、手足の動作や姿勢などにより境界線の角度が変化するため、境

界線の角度計算を行うことが必要である。

以上に示した課題を解決するために、多数の画像に対して実験を行い、検証し、解決していきたいと考えている。

参考文献

- [1] S.Saito, H.Saji, "Tracking of Team Sports Players in Image Sequences Using a Portable Video Camera," Inter-Academia 2003, Vol.2, pp.355-360 (2003)
- [2] 片山 壽二, 杉山 岳弘, 阿部 圭一, "1次元オプティカルフローに基づく移動物体の方向の推定," コンピュータビジョンとイメージメディア (CVIM), Vol.133, No.4, pp.25-32 (2002)
- [3] 岩瀬 幸子, 斎藤 英雄, "多視点画像を用いたサッカー選手の追跡," コンピュータビジョンとイメージメディア (CVIM), Vol.133, No.16, pp.119-124 (2002)
- [4] 佐竹順二, 尺長健, "階層的注視点制御による動画像上での複数人物追跡," 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.8, pp.1212-1221 (2003)
- [5] 岩元 浩太, 麻田 知嗣, 小松 尚久, "信頼度パラメータを用いた広域動画像における人物追跡," 電子情報通信学会技術研究報告 (TM), Vol.101, No.560, pp.13-18 (2002)
- [6] 大野 義典, 三浦 純, 白井 良明, "サッカーゲームにおける選手とボールの追跡," コンピュータビジョンとイメージメディア (CVIM), Vol.114, No.7, pp.49-56 (1999)
- [7] 小渡悟, 星野聖, "オプティカルフローと色情報に基づく掌の検出と追跡によるジェスチャ認識," CVIM, Vol.125, No.6, pp.47-54 (2001)

- [8] 高橋正貴, 小倉信彦, 渡辺澄夫, ”隠れマルコフモデルによる混合動作の認識”, PRMU, Vol.2000, No.151, pp.31-38 (2000)
- [9] 佐藤明知, 川田聡, 大崎喜彦, 山本正信, ”多視点動画像からの人間動作の追跡と再構成”, 電子情報通信学会論文誌 D-II, Vol.J80-D-II, No.6, pp.1581-1589 (1997)
- [10] 尾上守夫, 前田紀彦, 斎藤優, ”残差逐次検定法による画像の重ね合わせ”, 情報処理学会論文誌, Vol.17, No.7, pp.634-640 (1976)
- [11] 斉藤文彦, ”ブロック照合投票処理を用いた遮へいに強い画像マッチング”, 電子情報通信学会論文誌 D-II, Vol.J84-D-II, No.10, pp.2270-2279 (2001)