

モバイルエージェント技術による 人物追跡システムの構築のための色特徴量について

Color Futures for Human Tracking System

Based on Mobile Agent Technologies

中野 翔太[†] 四元 辰平^{††} 谷川 浩三^{††} 高橋 健一[†] 川村 尚生[†] 菅原 一孔[†]

Shota Nakano[†] Tappe Yotsumoto^{††} Kozo Tanigawa^{††} Kenichi Takahashi[†] Takao Kawamura[†] Kazunori Sugahara[†]

[†] 鳥取大学大学院 工学研究科 情報エレクトロニクス専攻 ^{††} メルコ・パワー・システムズ株式会社

1 はじめに

近年、セキュリティシステムの導入が様々な企業で盛んに行われている。セキュリティシステムのひとつに監視カメラを利用したものがある。これは監視カメラに映った不審人物を監視員が監視・追跡するシステムである。このようなシステムは監視・追跡する対象が複数になると監視員の負担が大きくなる。我々はこの問題を解決するためにモバイルエージェント技術を用いた人物追跡システムの開発をしている [1]。提案するシステムではエージェントが自動で追跡を行う。これにより監視員は追跡を行う必要がなくなるため、負担が軽減される。また、監視員が1人の追跡対象にくぎ付けにならないため、複数の監視カメラの映像を監視することも容易になる。

エージェントが自動で追跡をするには、監視カメラに映っている人物が追跡対象の人物であるかをエージェントに判断させる必要がある。本システムでは、追跡対象を登録する際にエージェントに特徴量を持たせる。特徴量とは人物の特徴を数値に表したもので、監視カメラの映像から計算している。エージェントは自身が持っている特徴量と監視カメラの映像から計算した特徴量を比較することで人物の判定を行っている。

従来は SIFT 特徴量を利用している。しかし、SIFT 特徴量は計算コストが高いためリアルタイムでの処理が必要な人物追跡システムには向いていない。また、SIFT 特徴量は複数の角度から撮影した人物の判定では精度が大きく落ちるため、人物追跡の精度が低くなる原因になっている。そこで、人物の衣服や肌の色を利用した特徴量を提案する。本手法では人物の映像を縦方向に走査し、色の変化の大きな部分を境界として検出する。この境界をエッジと呼ぶ。人物の特徴としてエッジによる分割数と分割された領域の色の平均値を利用する。多くの場合、衣服の色はどの方向から見ても同じ色であるため、複数の角度から撮影した映像であっても判定が可能な手法である。

2 特徴量抽出手法

2.1 累積背景差分法

監視カメラの映像から人物の特徴量を計算するためには、映像から人物の領域を抽出する必要がある。映像

から人物領域を取り出すために、映像から背景を除いた前景領域を取り出す。前景領域を取り出す手段として背景差分法を利用する。しかし、監視カメラが設置されている場所の照明環境が時間経過と共に変化することが考えられる。その場合、背景画像をあらかじめ用意しておく背景差分法では背景の変化に対応できないため、正しく前景領域を取り出せない。この問題を解決する手法として累積背景差分法を利用する [2]。累積背景差分法では時間の経過に合わせて背景画像を変化させるため、背景が変化した場合にも正しく前景領域を取り出すことができる。図1に背景差分法によって取り出した前景領域を示す。



図1: 背景差分による前景領域の抽出

2.2 ラベリング手法

背景差分によって得られた前景領域は人物個人を識別したものではない。監視カメラに複数の人物が映っている場合は前景領域の各画素が、それぞれどの人物を示す画素であるかを識別する必要がある。そこで、ラベリング処理によって画素のグループ分けを行う。ラベリング処理では画像をラスタ走査し、前景領域の画素を見つけるとその画素に隣接した画素にラベルが割り当てられているかを調べる。ラベルが割り当てられていなければ同じラベルを割り当てる。ラベルが割り当てられていなければ新たにラベルを付与する。同一のラベルが割り当てられている画素をすべて含む最小の矩形を人物領域として抽出する。図2にラベリングによって取り出した人物領域を示す。

2.3 人物領域の分割

人物領域の上端の一点と下端の一点を結ぶ一直線上で色の変化が大きな部分をエッジとして抽出する。抽出したエッジによって人物領域を分割する。エッジを



図 2: 人物領域の抽出

検出するための直線は人物の上を通過する必要がある。そこで、人物の上を通過する直線の両端の点を背景差分で得られた前景領域から得られた。人物領域の上端で前景領域の画素は頭部の画素であると考えられる。また、人物領域の下端で前景領域の画素は足の画素であると考えられる。それぞれの画素を両端とした直線は頭と足を結ぶ直線であるため、人物領域の上を通過する直線であると考えられる。ただし、上端の画素と下端の画素を求めたとき複数の画素が検出されることが考えられる。そこで、上端の点が複数検出された場合はその点の中央値を採用する。下端の点は左右の足について条件に一致する画素が検出されることが考えられる。歩行の際に踏み出した足の画素を採用した場合は直線は斜めになるため、人物の上を通過しないことが考えられる。そこで、軸足になっている足の画素を下端の点として採用する。人物領域の下端の画素であり、上端の点に最も近い点を下端の点として採用する。分割された人物の画像を図 3 に示す。



図 3: 分割された画像

2.4 影を考慮したエッジ検出

衣服にはしわが原因で影ができる場合があるが、エッジを検出する際に影による色の変化をエッジとして検出するべきでない。影による色の変化によってエッジが検出されることを防ぐために画像の色を HSV 形式で扱う。HSV 形式は色を色合い・彩度・明度によって表す方式である。影による色の変化では色合いは変わらないため、色合いが変化した部分をエッジとして扱うことで影によって間違ったエッジが検出されることを防ぐことができる。

2.5 エッジの結合

エッジを検出する際に、上記のように定めた上端と下端の点を結ぶ直線が人物の上を通過しない場合がある。直線が人物以外の場所を通過している場合は背景のエッジが検出される。背景のエッジが含まれると本

来エッジでない部分がエッジとして検出されるため、特徴量の精度が落ちる。この問題を解決するために、分割された領域の色の平均値を計算する。隣り合った領域間で色の平均値の変化が小さい場合は誤って検出されたエッジであると考えられる。そのため、そのエッジを消去して領域を結合させる。

2.6 特徴量の抽出

エッジによって区切られた領域の色の平均値と分割の割合を特徴量とする。特徴量の計算の際には、区切られた領域内の前景領域の画素のみを利用する。特徴量は毎フレームごとに計算したものを蓄積しておき、最終的には蓄積した特徴量の平均値を人物の特徴量とする。また、エッジによって分割の割合を特徴量とする。

2.7 特徴量の照合手法

入力された画像から人物の領域を取り出し、照合する対象の人物の特徴量と比較する。最初に照合する対象の人物と同様の割合で入力された人物の画像を分割する。次に分割された領域ごとに色の平均値を計算する。最後に分割されたそれぞれの領域の色の平均値同士の差を取り、各領域の差の平均値を照合の基準とする。

3 実験

実験には複数の場所に設置した監視カメラの映像を録画したものを利用した。監視カメラは部屋・廊下・階段にそれぞれ一台ずつ設置した。映像はそれぞれの場所を一人の人間が歩いているものを三つ、二人で歩いているものを一つ用意した。監視カメラの映像を図 4 に示す。3箇所それぞれで特徴量を取得し、特徴量を



図 4: 監視カメラの映像

取得した以外の場所で人物の判定を行った。結果を表 1, 2, 3, 4 に示す。表は縦が特徴量を取得した場所で横が人物の判定を行った場所である。正しく判定できた場合は○, 判定できなかった場合は×, 特徴量の取得に失敗した場合は△で示す。

表 1, 3 から各場所で取得した特徴量から人物の判定ができていくわかる。表??から部屋で取得した特徴量を用いた照合で判定ができていないことがわかる。また、廊下で取得した特徴量も部屋での照合ができていないことがわかる。表 4 から二人で歩いた場合には特徴量の取得ができていないことがわかる。

表 1: 一人だけ歩いた映像の判定結果 1

	部屋	階段	廊下
部屋	-	○	○
階段	○	-	○
廊下	○	○	-

表 2: 一人だけ歩いた映像の判定結果 2

	部屋	階段	廊下
部屋	-	×	×
階段	○	-	○
廊下	×	○	-

表 3: 一人だけ歩いた映像の判定結果 3

	部屋	階段	廊下
部屋	-	○	○
階段	○	-	○
廊下	○	○	-

表 4: 二人が歩いた映像の判定結果

	部屋	階段	廊下
部屋	-	△	△
階段	△	-	△
廊下	△	△	-

4 評価

実験結果から複数の人物が映った場合に人物の特徴量が取得できていないことがわかる。これは複数の人物が重なり合う場合に特徴量の蓄積が正しく行えていないことが原因であると考えられる。また、本手法はリアルタイムで処理が可能であったため、リアルタイムで処理ができなかった SIFT 特徴量と比較して高速であることがわかる。

5 おわりに

本研究ではモバイルエージェントシステムを用いた人物追跡システムのための特徴量の取得のための画像処理の設計と実装を行った。実験として、監視カメラを複数台用いて人物の特徴量の取得と判定を行った。複数人物が重なりあった場合に特徴量を取得することができなかつたため、今後は人物の重なりの方策が必要となる。SIFT 特徴量と比較して高速に処理できることが分かった。

参考文献

- [1] Hiroto Kakiuchi, Kozo Tanigawa, Takao Kawamura, and Kazunori Sugahara. A construction method for automatic human tracking system

with mobile agent technology. *Recent Developments in Video Surveillance*, pp.21-38, 4 2012.

- [2] 森田真司. 全方位画像センサを用いたネットワーク対応型のテレプレゼンスシステムと遠隔監視システム. 奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻修士論文, 2 2003.