

色ヒストグラムの合成を用いた交差時の人物判定方式
Multiple Human Tracking Method using Mixed Color Histogram for Occlusion小泉 博一† 大網 亮磨† 平田 恭二十
Hirokazu Koizumi Ryoma Oami Kyoji Hirata

1. はじめに

近年、監視カメラの低廉化と共に、防犯目的以外にも様々な用途におけるカメラの導入が進んでいる。例えば、インターネットで情景や店内の様子を提供したり[1]、外出先から家の中を確認したりするモニタリングサービス[2]に利用されている。

これらの従来のサービスに認識技術を組み合わせることで、例えば、店舗モニタリングや独居老人モニタリングなど、得られた人物の行動情報を解析し、売上に役立てる、異常状態を自動的に通知するといった応用が実現できる。

認識技術と組み合わせたモニタリングシステムにおいて、各人物の行動を把握するためには、各人物の時系列データを取得し、データを解析することが不可欠であり、そのために、各人物を追跡する必要がある。店舗など、室内に多数の人物が存在し、人物同士が頻繁に交差する状況では、交差時に各人物を判定し、正確に対応付けなければならない。本稿では、色ヒストグラムを合成して人物と領域に対応付けることにより複数人の交差に対処可能な追跡手法を提案する。

2. 従来手法の問題点

監視分野などではこれまでにも人物同士の交差時の人物追跡手法が研究されている。McKenna ら[3]は、交差前の人物の色ヒストグラムと交差後の領域の色ヒストグラムとを比較して各領域に存在する人物を判定する手法を提案している。この手法では、真横から撮影した画像に対し人物の大きさは一定と仮定し、人物と領域の色ヒストグラムの類似度を比較することにより領域内に存在する人物を判定し、これにより2人が交差した際の各人物の追跡を実現している。

しかし、3人以上が存在し、部屋全体を見渡すために斜め上方にカメラが設置される、通常の室内環境に適用する場合、以下の問題が生じる。

a) 複数人の交差

人物単体の色ヒストグラムと領域の色ヒストグラムとを比較しているため、4人が交差した後に2人ずつの2つのグループに分かれるといった、分離後の領域内に複数人がいる場合には、領域の色ヒストグラムはどの人物の色ヒストグラムとも異なるため、正しく対応付けることができない。

b) カメラアングルによる人物の大きさの変化

カメラアングルによってはカメラからの距離により人物の大きさが変化するため、誤判定する可能性がある。

3. 複数人交差時の人物判定

3.1 概要

上記課題を解決するために、色ヒストグラム合成に基づく人物の判定手法を提案する。提案手法では、人物の色ヒストグラムから複数人が存在している場合の色ヒストグラムを合成し、合成された色ヒストグラムと領域の色ヒストグラムとの距離を算出する。この距離を全領域で合計し、合計値が最小となる人物と領域の対応付けを、可能な全ての組み合わせの中から選択し、人物と領域を対応付ける。色ヒストグラムの合成には、カメラアングルによる人物の大きさの変化に対処するために交差前の人物の面積比を考慮する。

以下、本手法について詳細に述べる。

3.2 色ヒストグラム合成に基づく人物の対応付け

色ヒストグラムは、HSI空間上で算出し、色ヒストグラムの面積が1になるように正規化する。人物同士が交差後に分かれる際、人物を追跡するため、色ヒストグラム同士のマッチングにより各人物と領域とを対応付ける。

具体的には、領域の色ヒストグラムと交差前の人物の色ヒストグラムとを用いて、人物と領域の可能な全ての組み合わせについて距離を計算し、最も距離の小さい対応を選択する。

図1のように、3人の人物AとBとCが交差し、交差後に2つの領域 α と β に分かれた場合、3人の人物A、B、Cと2つの領域 α 、 β を対応付けることで各人物の移動方向を判断するが、このとき領域 (α, β) に対応付ける人物の組み合わせは(A,BC)、(B,AC)、(C,AB)、(AB,C)、(AC,B)、(BC,A)、(ABC, Φ)、(Φ ,ABC)の8通りが考えられる。ここで Φ は物体を示す。各組み合わせについて色ヒストグラム間の距離を算出し、距離が最小となる組み合わせを正解の対応付けとする。領域 (α, β) と人物(A,BC)の組み合わせの距離は式(1)のように、人物と領域との各対応の距離の平均として求める。ここで領域 β と人物のグループBCとの距離は、各人物の色ヒストグラムから、人物の大きさで重

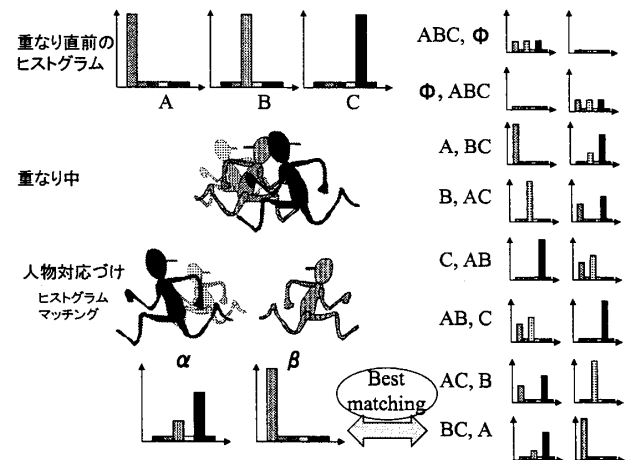


図1. 人物の対応付け手法

† NEC マルチメディア研究所

み付け加算してグループの色ヒストグラムを合成し、合成された色ヒストグラムと領域の色ヒストグラム間距離を求める。また、 (α, β) と (ABC, Φ) の距離は式(2)のように求める。

$$F((\alpha, \beta) = (A, BC)) = \frac{1}{2} \left(\sum (Hr_{\alpha}(i) - Hm_A(i))^2 + \sum \left(Hr_{\beta}(i) - \frac{S_B * Hm_B(i) + S_C * Hm_C(i)}{S_B + S_C} \right)^2 \right) \quad \text{式(1)}$$

$$F((\alpha, \beta) = (ABC, \phi)) = \sum \left(Hr_{\alpha}(i) - \frac{S_A * Hm_A(i) + S_B * Hm_B(i) + S_C * Hm_C(i)}{S_A + S_B + S_C} \right)^2 \quad \text{式(2)}$$

ここで、 Hr_{α}, Hr_{β} は各領域の色ヒストグラム、 Hm_A, Hm_B, Hm_C は各人物の色ヒストグラム、 S_A, S_B, S_C は交差直前における各人物の面積を示す。

3.3 効果

本手法では以下のような特徴を有する。

- a) 色ヒストグラムを合成して人物と領域を対応付け、全ての人物と領域との対応関係を考慮することで、3人以上の交差に対処でき、対応付けの精度を向上できる。
- b) 交差前の人物の面積比を考慮することで、カメラアングルに起因する人物の大きさの変化に対してロバストに判定できる。

4. 評価実験

4.1 実験条件

上記手法を検証するために、従来手法[3]との比較実験を行った。カメラを部屋の隅上方に、部屋全体が見渡せるように設置して撮影した、3~4人が動き回る画像(320x240, 約12fps, 各5~6分程度)8シーンを使用して実験した。人物領域は、人物のいない初期状態であらかじめ撮影された背景画像を用いて、背景差分法により抽出した。色ヒストグラム算出時のHSI空間の量子化数は、H:4, S:4, I:2とした。Iの量子化を粗くしているのは、輝度変動の影響を抑えるためである。

4.2 実験結果

分離発生時の人物判定回数の結果を表1に示す。分離状況別に分離発生回数と、各手法における正解数を求めた。

表1. 分離時の判定回数と判定率

分離状況	発生数	提案手法	従来手法
2人→1人+1人	14	14	7
3人→2人+1人	18	14	11
3人→1人+1人+1人	1	1	0
4人→3人+1人	2	2	1
合計	35	30	19
判定率		89%	54%

この結果、提案手法では89%の判定率を示し、従来手法の54%を大きく上回った。

従来手法では、2人が分かれた場合に多く誤判定した。これは、人物の大きさは一定という仮定により、分離した領域の大きさが異なる場合に、各人物を大きい方の領域に対応付けるためである。また、3人が分かれた場合には、

大きさの変化に加えて、人物単体と領域との色ヒストグラムを比較して対応付けるために誤判定した。

一方、提案手法では、人物の面積比を考慮して合成した色ヒストグラムを用いて距離を計算し、さらに人物と領域の対応づけを全体で最適化した結果、正しく判定した。

図2に複数人の交差時における判定結果の一例を示す。処理画像中の赤、緑、青、黄色の領域はそれぞれ人物1~4の領域を示し、色の混ざった領域は、色に対応する人物が重なっている状態を示す。また、各領域の左上の番号は人物番号を示す。ここでは、人物2と3に人物4が加わって室内左手で打ち合わせをし、室内右手で作業をしていた人物1が室内奥へ移動した際に人物4が呼び止めて室内奥で会話をしている。このとき、人物2~4の領域が、人物2と3のグループの領域(図2中の領域 α)と人物4の領域(図2中の領域 β)とに分かれたために、人物2~4の判定が必要となった。

従来手法では、分離した2つの領域 α と β のうち、領域 α の面積が大きく、各人物の色を持っているために、全ての人物を領域 α に対応づけようとして誤判定した。一方、提案手法では、人物の面積比を考慮して色ヒストグラムを合成するため、正しく判定した。

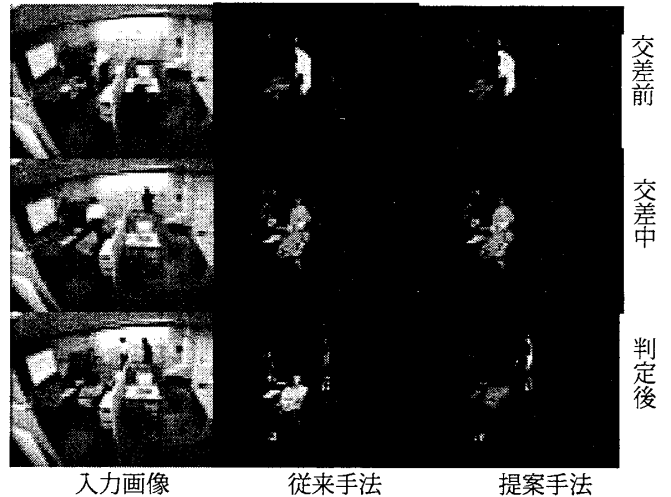


図2. 従来手法と提案手法の比較

5. まとめ

本稿では、複数人が交差した際に各人物を判定する手法を考案し、評価した。本手法では、複数の人物の色ヒストグラムから合成した色ヒストグラムと領域の色ヒストグラムとの距離を算出し、さらに人物と領域の全ての組み合わせから最適な組み合わせを選択することで、交差時に各人物を追跡できる。比較実験では、本方式は正解率が89%であり、従来方式に比べて精度の向上を示した。

参考文献

[1] <http://www.ilive.hokkaido.isp.ntt-east.co.jp/>
 [2] <http://www.ntt-me.co.jp/livinggate/main.html>
 [3] S.J.McKenna, S.Jabri, Z.Duric, H.Wechsler, "Tracking Interacting People", in Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000, pp.348-353, 2000