

分散カメラ間におけるグループ特徴量を用いた人追跡

守口 裕介¹ 岡田 亜沙美¹ 浮田 宗伯^{1,a)} 萩田 紀博^{†1,b)}

概要: 本研究では、視野に重畳のない分散カメラ群によって、広域を移動する人をそれぞれ追跡するために、新たに各人のグループ特徴量を利用する追跡法を提案する。本手法では、カメラ間の移動においてグループ（行動を共にする1人以上の人の集まり）の関係は変化しないという仮説の下、その人が属しているグループに関する特徴量をカメラ間同定に用いるために、まず各カメラにおける追跡結果からグループと非グループそれぞれの軌跡同士の関係を Support Vector Machine (SVM) によって学習し、グループ検出を行う。そして、グループ検出の結果からグループ特徴量を算出し、色情報やカメラ間の移動時間を反映させた既存の特徴量と併用することによって、カメラ間で人を同定する。実験では、まず軌跡同士の関係からグループ検出が平均 83% の精度で行えることを確認した。また、グループ検出結果から得たグループ特徴量によってカメラ間同定の精度が 37% 増すことを確認した。

Grouping People for Tracking across Non-overlapping Cameras

YUSUKE MORIGUCHI¹ ASAMI OKADA¹ NORIMICHI UKITA^{1,a)} NORIHIRO HAGITA^{†1,b)}

Abstract: The groups of people are detected in each camera and then utilized as an additional clue for improving people identification across non-overlapping cameras. It is assumed that if persons belong the same group in a camera view, they may be in the same group also in another camera view. The grouping process is achieved by discriminatively classifying the trajectories of grouped people and non-grouped people with their spatio-temporal features of the trajectories. And then the spatio-temporal features are utilized with other features of each person (e.g. color histogram, transit time between cameras) for people identification across non-overlapping cameras. The experimental results demonstrate the improvement in people grouping and people identification by the proposed methods with a public dataset; the rate of group detection was 83 % on average, and the successful rate of people identification across cameras was improved by 37 %.

1. はじめに

複数カメラでの人追跡において、あるカメラに映った人が別のカメラで映った人と同一人物か判断する人同定が必要である。監視カメラはコストの面から視野に重なりが無い配置になっていることが多い。また、一般的には図1のように広域撮影のため人の撮影解像度は低く、顔認識などの個人同定手法は利用できない。以上の問題点を解決するため、低解像度における個人のアピアランスに着目したカメラ間の人追跡が多数研究されている（サーベイ論文 [1]）。

本研究では、カメラ間同定において個人のアピアランス以外に利用できる特徴量として、「行動を共にする人の集まり」であるグループ（図1内の各枠参照）に着目した特徴量を提案する。カメラ間の移動においてグループの関係が変化しないという仮定の下、その人が属しているグループの人数やグループ内の人のアピアランスを「グループ特徴量」と定義し、人同定に利用する。

提案手法の実現には、同一グループに属する人を検出するグループ検出が必要である。グループ内の人と同じ様な軌跡で移動するため、各カメラにおけるグループ内と非グループの人同士それぞれの軌跡間関係を学習し、グループ検出を可能にする。このグループ検出により得られるグループ特徴量と従来の特徴量を併用し、カメラ間同定の精度を向上させる。

¹ 奈良先端科学技術大学院大学
NAIST, 8916-6 Takayama, Ikoma, Nara 630-0192, Japan

^{†1} 現在、国際電気通信基礎技術研究所
Presently with ATR

^{a)} ukita@is.naist.jp

^{b)} hagita@atr.jp

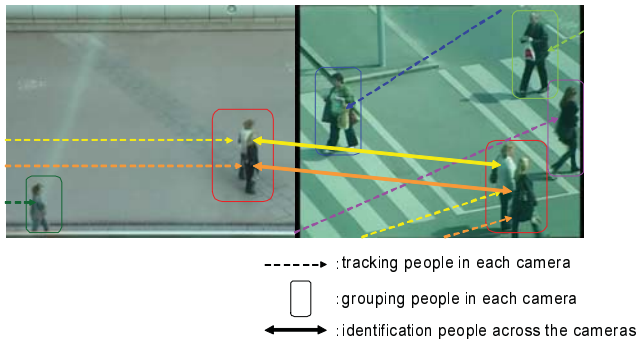


図 1 監視カメラの映像例. 広域対象追跡の実現には, 各カメラの視野内で撮影された人の追跡(各破線)と, 各カメラで撮影された人の中で同一人物判定を行うカメラ間人同定(各実線)が必要. 各枠はそれぞれグループを示しており, 赤枠内の2人は同一グループ.

2. 関連研究

複数台のカメラ間での人同定について, 個人のアピランスに着目することで精度を向上させた研究が多数報告されている. 個人のアピランスの取得方法として, 特徴点群 [4], HOG[5], カラーヒストグラム [6] などの特徴量を使用した手法が提案されている. また, 各カメラの照明条件などによる色の変化を補正する手法を提案した研究 [7], [8] など様々な研究が存在する.

カメラ間人同定において個人のアピランスの特徴量だけでなく, カメラ間移動に必要な時間を表す時空間特徴量を併用した手法も存在する(カメラ間追跡結果からのカメラ間関係モデリング [9], [10], 追跡不要なモデリング [11], [12], 逐次モデリング [13] など).

本研究では, 以上の研究で利用されている特徴量や手法に加えて, 新たに人のグループに着目する. グループに関する情報をカメラ間人同定に用いるためには, グループ検出が必要である. グループ検出に関する研究には静止画を利用した研究 [14] もあるが, 本研究では監視カメラから得られる映像を用いるため, 動画を利用したグループ検出手法に着目する. 動画を利用することで, 各フレームにおける画像の情報や時系列に沿った人の動きをグループ検出に利用できる一方で, 一般的に静止画と比べて監視カメラの動画は人が小さく撮影されるため, 顔認識による年齢の推定などが難しく, 取り出すことのできる特徴が限定される. グループ検出の研究で重要となる考え方に, Social Force Model (SFM)[15] と呼ばれるモデルがある. この SFM を使ってグループ検出を行った研究もある [16], [17]. Yamaguchi ら [18] も, SFM と類似した軌跡間の関係の特徴量として用いてグループ検出を行った.

3. 単眼カメラ視野内での人追跡

単眼カメラ視野内での人追跡を行うためには, 各フレームにおける画像(例: 図 1) から人を検出し, フレーム間で

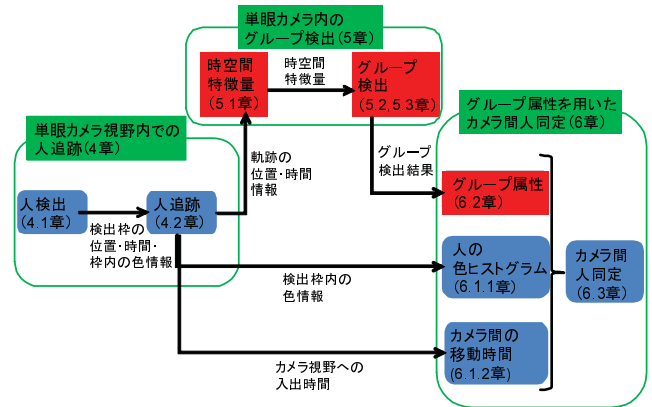


図 2 処理全体の流れ. 赤く表示した部分が新しく提案する処理. 提案手法は, 従来のカメラ間人同定手法に図中の赤く表示されている処理を追加している.

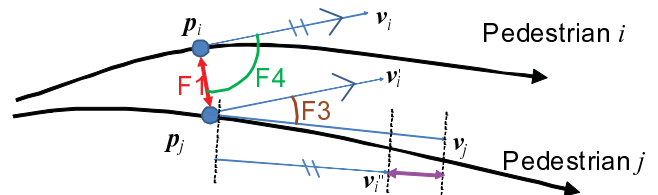


図 3 文献 [18] で提案されている時空間特徴量. v_i と v'_i は並行, v_i と v'_i の長さは等しい.

同一人の検出結果をつなぎあわせるアプローチをとる. 人検出には, 画像中の各対象領域から HOG 特徴量を抽出し, Support Vector Machine (SVM) [19] によって人またはその他の領域に識別することによって行う. この人検出によって, 各フレームの画像中での人の位置とサイズが得られる. この各フレームにおける人の位置と矩形領域内の色情報から, 文献 [3] で提案されている手法を用いて, 各フレームにおける検出結果を大域的最適的に連結することで人追跡を行う.

4. 単眼カメラ内のグループ検出

4.1 軌跡から得られる時空間特徴量

軌跡のグループ内ペア判定をする指標として, 文献 [18] では任意の軌跡ペアに対して, 5 種類の時空間特徴量を提案した. 歩行者 i の平面上での位置と速度をそれぞれ p_i , v_i , 人が観測領域内に存在したフレーム t の集合を $T_i = [t_{i,1}, t_{i,2}, \dots, t_{i,N_i}]$ とした時, 歩行者 i と歩行者 j について 5 つの時空間特徴量 (F1~F5) を以下のように(図 3 参照)定義する.

- (F1) 2 人間の距離: $|p_i - p_j|$
- (F2) 2 人の速度の絶対値の差: $||v_i| - |v_j||$
- (F3) 2 人の速度の向き差: $|\arctan(v_i) - \arctan(v_j)|$
- (F4) 1 人の速度の向きと 2 人の相対位置の向き差: $|\arctan(p_i - p_j) - \arctan(v_i)|$
- (F5) 軌跡ペアの同時観測割合: $|T_i \cap T_j| / |T_i \cup T_j|$

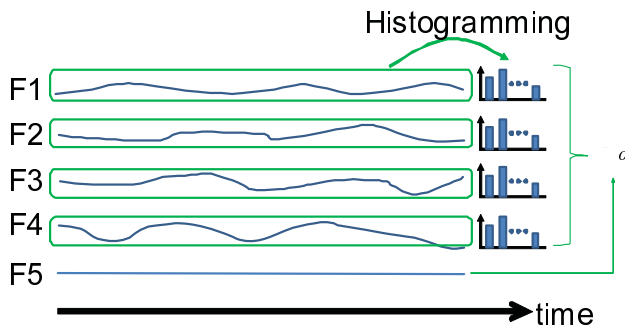


図 4 文献 [18] におけるグループ内ペアの検出法.

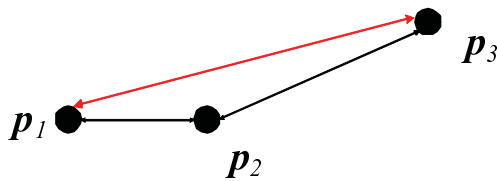


図 5 3人以上のグループにおけるグループ内ペア間の距離.

文献 [18] では、この 5 つの時空間特徴量を各フレームにおいて算出し、全フレームの F1~F4 をそれぞれ m 個のビンに分割されたヒストグラムにし、F1~F4 をそれぞれ m 次元の特徴ベクトルで表す。こうして算出される特徴ベクトルと F5 のスコアを組み合わせた $4m + 1$ 次元の特徴ベクトル f^o を、各軌跡ペアに対して求める (図 4 参照)。学習時には、グループ内ペアとそうでないペアの軌跡に対して f^o を求め、それぞれをポジティブサンプル・ネガティブサンプルとして学習する。そしてグループ内ペア検出時には、その学習結果を元にグループ内ペア検出を行った。しかし、この時空間特徴量には次の問題点があった。

- (1) 各フレームにおける F1~F4 を独立にヒストグラム化してしまうため、各フレームにおける時空間特徴量間の共起性 (例: 一定の距離を保つために、距離が離れると近づこうとする等) が判定結果に反映されない。
- (2) F4 の計算に速度の向きを使用する際に、軌跡ペアのうち一方の速度しか使用しないため、使用する速度によって F4 のスコアが変動する。
- (3) 一般的にグループ内ペアの F1 の値は小さくなるのが期待されるが、3人以上のグループのグループ内ペアの時空間特徴量を計算する際、距離の離れた軌跡ペア (例: 図 5 の赤矢印のペア) の F1 の値が大きくなってしまい、グループ内ペア検出が失敗してしまう。
- (4) 静止している人の速度を求め F3 や F4 を導出する際に、速度の向きが安定しない。

以上の問題点を解決するために、本研究ではそれぞれの問題点を以下の手法を用いて解決した。

- (1) 各フレームにおける時空間特徴量間の共起性を判定結果に反映するために、各フレームにおける時空間特徴量を 1 つの特徴ベクトルとして算出する。図 6 の

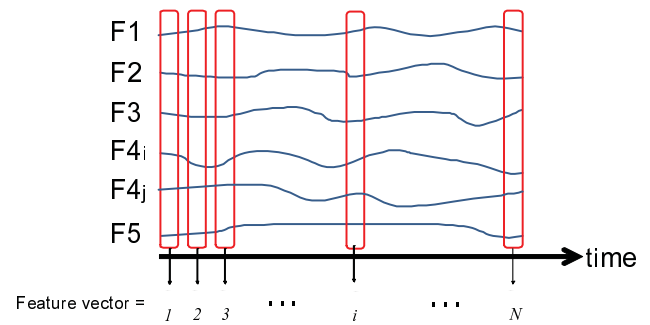


図 6 各フレームにおける時空間特徴量を用いたグループ内ペア判定の概念図.

f_1, \dots, f_N が各フレームにおける特徴量ベクトルである。各フレームにおける時空間特徴量の共起性を表現するために、F1~F4 と同様に各フレームにおける F5 のスコアを求める方法として、 f フレームでの F5 の計算に用いる T_i を前後 T^f フレーム間に限定する方法をとる。つまり $f - T^f$ フレームから $f + T^f$ フレームの間で F5 の計算を行う。

- (2) F4 を 2 つの特徴量に分け、 v_i を計算に用いる $F4_i$ と v_j を計算に用いる $F4_j$ のどちらも計算する。
- (3) 3人以上のグループに対して、学習時は最も近距離にいるグループ内ペアのみを学習させる (例: 図 5 の黒矢印のペアの特徴量のみを学習)。グループ内ペアの検出時には、検出されたグループ内ペアの構成員間に重複があった場合、それらのグループ内ペアの構成員全員を同一グループに属すると判定することで、3人以上のグループの検出も行う (例: 図 5 の黒矢印で示された 2 つの軌跡ペアがグループ内ペアと検出された場合、赤矢印の軌跡ペアもグループ内ペアとする)。
- (4) あるフレームにおける速度が閾値以下となる人については、そのフレームにおけるグループ内ペア学習・判定を行わない。

上記のように定義した 6 つの時空間特徴量 (F1, F2, F3, F4_i, F4_j, F5) を用いてグループ内ペア検出を行う。

4.2 時空間特徴量を用いた各フレームにおけるグループ内ペア判定

4.1 章で定義した時空間特徴量を用いて、軌跡ペアの各フレームにおけるグループ内ペア判定を行う。この判定には、識別器として SVM を用いる。学習時には、グループ内ペアとそうでないペアの軌跡の時空間特徴量の 6 次元ベクトルをそれぞれ正例、負例として SVM で学習する。検出時には、任意の軌跡ペアに対して各フレームにおける時空間特徴量を求め、学習済みの SVM により各フレームのグループ内ペア判定結果を得る。

4.3 各フレームにおけるグループ内ペア判定の結果を用いたグループ内ペア検出

各フレームにおけるグループ内ペア判定結果を用いて、軌跡ペア全体に対してグループ内ペア検出を行う。

4.2章の手法によって各フレームでのグループ内ペア判定した結果、グループ内ペアと判定されたフレーム数が全フレームの内、閾値 T_P 以上となった軌跡ペアをグループ内ペア検出の結果とする。その後、4.1章で述べたように、検出されたグループ内ペアの構成員間の重複を調べ、重複が発見されたグループ内ペアの構成員全員を同一グループに属すると判定することで、3人以上のグループを検出する。

5. グループ特徴量を用いたカメラ間人同定

5.1 カメラ間人同定に用いる従来特徴量

カメラ間人同定の手法として従来利用されている特徴量である、人の色ヒストグラムという人固有の特徴量と、カメラ間の移動時間という観測領域固有の特徴量を求める。

5.1.1 人の色ヒストグラム

文献 [6], [7], [8] 等の多数の研究例のように、人追跡の結果得られる矩形領域内の色ヒストグラム間の類似度をにより人検出結果同士の類似度を求める。 N_B 個の bin を持つ2つの色ヒストグラム間の類似度には、下式の Bhattacharyya 係数を用いた。

$$L = \sum_u^{N_B} \sqrt{P_u^c Q_u^c} \quad (1)$$

Bhattacharyya 係数 L は 0~1 の値をとり、値が高いほどヒストグラム間の類似度が高いということを表す。カメラ間の人同定において、同一人物間の Bhattacharyya 係数は別人間と比べて高くなるため、これを人固有の特徴量のスコア $R^c (= L)$ とする。

各人の色ヒストグラムを取得する際、複数フレームに渡る色情報を統合する。この統合として、本研究では次の3手法をテストする。

- (1) 人が検出されたフレーム全ての色ヒストグラムを平均し正規化する。
- (2) 各カメラにおけるその人が検出されたフレームの前後に検出された人物との間で Bhattacharyya 係数を計算し、Bhattacharyya 係数が最小となるフレームにおける色ヒストグラムを正規化する。
- (3) 各カメラにおけるその人の検出されたフレームの前後に検出された人物との間で Bhattacharyya 係数を計算し、Bhattacharyya 係数の小さいものから n フレーム分だけ取り出し、その n フレームにおける色ヒストグラムを平均し正規化する。

5.1.2 カメラ間の移動時間

人追跡の結果得られるカメラ視野への入出時間を使い、

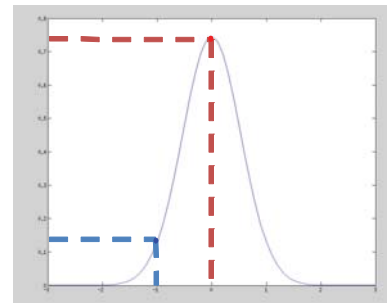


図7 グループ内人数に関するグループ特徴量のスコア導出の概念図。人同士のグループ内人数の差に対するスコアを表す。赤線がグループ内人数が一致した際のスコア、青線がグループ内人数の差が1人の際のスコア。

カメラ間の移動時間カメラ間の移動時間という観測領域固有の特徴量を求める。手法 [12] のように、カメラ間の移動時間の分布はガウス分布で近似した。カメラ間の移動時間の期待値 m^t と分散 δ^{t^2} を求め、確率的にカメラ間の移動時間のガウス分布 $f^t(t; m^t, \delta^{t^2})$ を求めておく。

カメラ間人同定においては、各カメラにおけるカメラ視野への出入りの時間から、カメラ間における任意の人検出結果間の移動時間を求めることができる。ここで、あるカメラ a で検出された歩行者 i の視野からの退出時間を $t_i^{a,out}$ 、別カメラ b で検出された歩行者 j の視野への進入時間を $t_j^{b,in}$ とした時、以下の式のように、カメラ間移動時間 $t_j^{b,in} - t_i^{a,out}$ を算出済みのガウス分布 $f^t(t; m^t, \delta^{t^2})$ に代入することで、観測領域固有の特徴量のスコア R^t を求める。

$$R_{i,j}^t = f^t(t_j^{b,in} - t_i^{a,out}; m^t, \delta^{t^2}) \quad (2)$$

5.2 人のグループ特徴量

カメラ間の移動において、グループ構成は変化しないという仮定の下、以下の2点に着目して、新たな特徴量であるグループ特徴量を求める。

- (1) 同一グループに属する人数
- (2) 同一グループに属する全ての人の色情報

1つ目のグループ特徴量であるグループ内人数を利用し、グループ内人数が等しい人同士のスコアが高くなるようなグループ内人数特徴量のスコア R^{gp} を設定する。具体的には、カメラ間人同定において歩行者 i と歩行者 j のグループ内人数の差 $p_{i,j}^{gp}$ を計算し、図7のように表されるガウス分布 $f^{gp}(p^{gp}; m^{gp}, \delta^{gp^2})$ に代入することで得られた値を、グループ内人数特徴量のスコア $R_{i,j}^{gp}$ とする。このガウス分布は、期待値 $m^{gp} = 0$ 、分散 $\delta^{gp^2} = 0.3$ とした。

$$R^{gp} = f^{gp}(p^{gp}; m^{gp}, \delta^{gp^2}) \quad (3)$$

2つ目のグループ特徴量として、その人の属するグループ内の全員分の色ヒストグラム平均を用いる。異なる2つの人検出結果のグループ内色ヒストグラム P^g, Q^g 間の類似度は、5.1.1章と同様に、Bhattacharyya 係数により求め、

表 1 各カメラにおける正解軌跡を用いたグループ内ペア検出結果. それぞれ, 検出成功 (True-Positive), 誤検出 (False-Positive), 未検出 (False-Negative) のグループ内ペア数

	True-Positive	False-Positive	False-Negative
Camera a	108	14	26
Camera b	130	38	22

これをグループ内色特徴量のスコア R^{gc} とする.

$$R^{gc} = L^g = \sum_{u=1}^{N_B} \sqrt{P_u^g Q_u^g} \quad (4)$$

5.3 グループ特徴量と従来の特徴量を併用したカメラ間人同定

カメラ間人同定に用いるスコア R は, 以下の式で表すように, 5.1 章, 5.2 章で求めた 4 種類のスコア R^c, R^t, R^{gp}, R^{gc} を掛け合わせて算出する.

$$R = R^c R^t R^{gp} R^{gc} \quad (5)$$

この R の値が高いものから, カメラ間人同定の候補として挙げることで, カメラ間人同定を実現する.

6. 実験

Person Re-ID 2011 dataset[20] に含まれる 2 台のカメラの動画を利用した. このデータセットは, 各人の ID 付き追跡結果の正解データも含んでいる. 人 ID は, 2 台のカメラ間で同じ人には同じ ID が与えられるように整合性がとられている. データセットの詳細は以下の通りである.

Camera a

追跡人数 573 人
グループ内ペア総数 134 ペア

Camera b

追跡人数 925 人
グループ内ペア総数 152 ペア

2 つのカメラ間の関係

カメラ間移動時間 平均 65 秒
カメラ間で対応付けされている人数 342 人

各人のグループの正解データは, 手動で与えた.

6.1 単眼カメラ内のグループ検出

グループ検出用の時空間特徴量は, 実世界上でのグループ内ペアの軌跡の特徴に着目して定義されているため, 画像座標を元に時空間特徴量を算出すると, グループ検出の精度が低減してしまう. よって, 人の画像平面上の位置から実世界上の道路面の位置を求めるために, 各カメラに対してホモグラフィ行列を求め, 画像平面上から実世界上への射影を行う. こうして得られた各人の実世界上の位置を用いて, 時空間特徴量を算出する.

4.3 章で述べたグループ内ペア検出時に用いる閾値 T_p

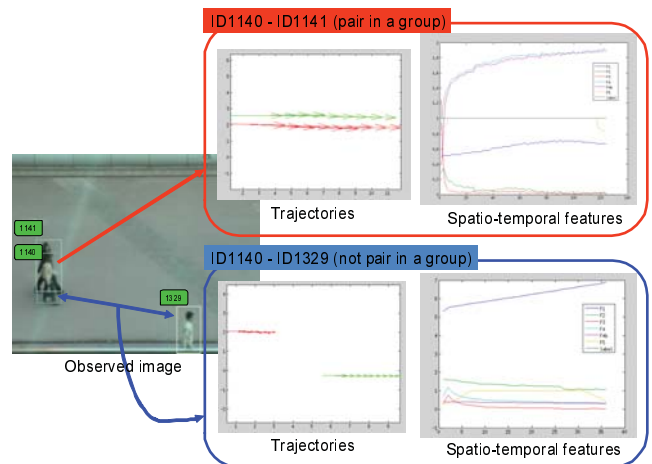


図 8 グループ内ペアの検出成功例 (ID1140 と ID1141) と非グループ内ペア (ID1140 と ID1329) の例. 左のグラフが同時に軌跡ペアが観測されたフレームにおける人の位置 (矢印の始点) と速度ベクトル (矢印) を示す. 右のグラフがその時刻における時空間特徴量とグループ内ペア判定の結果

は, 予備実験の結果から, $T_p = \frac{2}{3}$ とした.

各カメラにおけるグループ内ペア検出結果を表 1 に示す. グループ内ペア全体の 80~86% が検出成功している. グループ内ペアの検出成功例と非グループ内ペアの検出成功例を図 8 に示す. 図 8 に代表される結果から, グループ内ペア判定に F1 のスコア, つまり 2 軌跡の位置関係が強く影響を与えていることが分かる.

グループ内ペアの誤検出例を図 9 に示す. 偶然近距離で移動した軌跡ペアは, F1 が小さくなるため, 誤検出されている. 特に Camera b は, 横断歩道という人通りの多くなる場所を撮影しているため, 局所的な軌跡ペアの接近が多く, その結果 F1 の影響によってグループ内ペアの誤検出が増加していると考えられる.

グループ内ペアの未検出例を図 10 に示す. 各カメラ視野内において部分的に離れて移動していた軌跡ペアは, F1 が大きくなり, 未検出となった. 正解データ作成時には, 一方のカメラ視野のみでグループ内ペアと判定されるペアもグループ内ペアとしているため, そういった軌跡ペアは他方のカメラで未検出とされ易い.

6.2 グループ特徴量を用いたカメラ間人同定

本研究で提案しているグループ特徴量を用いたカメラ間人同定の有用性を確かめた. 双方のカメラで検出されている 342 人におけるカメラ間人同定の精度を確認する. 同定精度は, CMC 曲線 (Cumulative Match Characteristic curve) により示す. CMC 曲線は, 本人同士の類似度が全体の何番目に位置するのかを表している. Camera a の視野内で検出された人 i に対して, Camera b で検出された全人 884 人との間でカメラ間人同定用のスコア $(R_{(i,1)}, R_{(i,2)}, \dots, R_{(i,884)})$ を式 (5) により求め, スコア R の値が高いものから人同

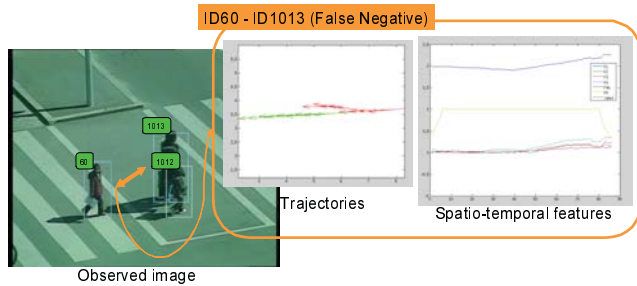


図 9 グループ内ペアの誤検出例 (ID60 と ID1013).

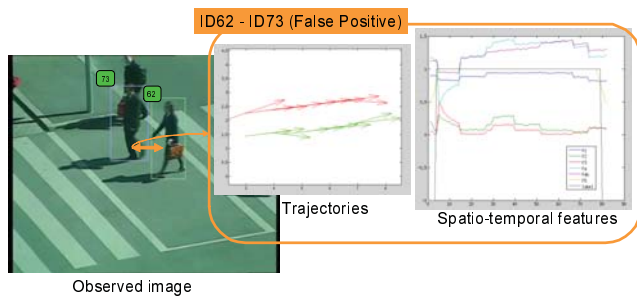


図 10 グループ内ペアの未検出例 (ID62 と ID73).

定の候補として挙げる. この処理を, 双方のカメラ視野内で検出成功している人 (342 人) に対して行い, 全候補 884 人中での本人同士の間で得られたスコア $R_{(i,i)}$ の順位を調べることで, 人同定の精度を計算する.

6.2.1 色ヒストグラム計算時の最良フレーム選択法

まず, 5.1.1 章で述べた色特徴量の取得時のフレーム選択法の中で, 色特徴量のみでカメラ間人同定を行う際に最も精度の良くなるフレーム選択法を調べる.

5.1.1 で述べた 3 手法を用いて人の色特徴量のスコア R^c を計算し, R^c のみを用いてカメラ間人同定を行った結果を図 11 に示す. 図 11 から, 各手法において色特徴量のみを用いた場合の精度に大きな差は見られなかった. このスコアは実験を行う環境によって変化すると考えられる. 本実験の環境においては, 候補の上位において手法 1 の色特徴量を利用した際の結果が最も良いことが分かる.

また, 手法 3 において $N = 5, 10, 20$ と変化させたが, $N = 20$ において手法 1 とほとんど変わらない結果が得られたことから, 人の平均的な色情報は 20 フレーム程度の色情報を用いる事で得ることができると考えられる.

したがって, 以降の実験における色特徴量のスコア R^c は, 本研究の実験環境において, 候補の上位での精度が最良となった手法 1 を用いて算出することとする.

6.2.2 正解軌跡と正解グループを用いたカメラ間人同定

カメラ間人同定におけるグループ特徴量の効果を調べるために, 理想的な環境での実験を行う. そこで, 人同定のスコア R を求める際に, 人のグループ特徴量や軌跡データにはそれぞれの正解データを使用した.

まず, 人の色ヒストグラムのスコア R^c とグループ特徴量を併用した際のカメラ間人同定の結果を図 12 に示す.

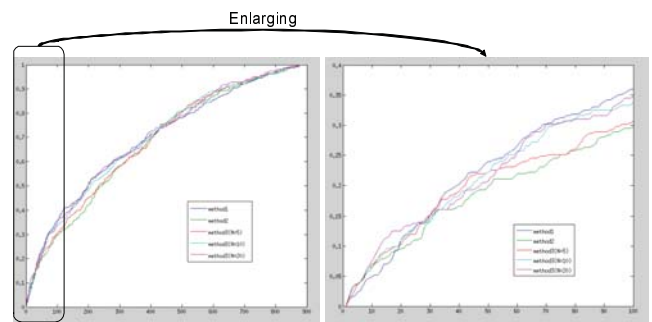


図 11 5.1.1 の 3 つの手法から算出した色特徴量を用いたカメラ間人同定の結果図 (手法 3 におけるフレーム数 $N = 5, 10, 20$). 横軸は本人同士の間で得られた特徴量の全候補 884 人中の順位を示し, 縦軸はその順位以内に入っている 342 人中の人数の割合を示す. 全順位における結果が左側のグラフに, 100 位以内の部分に着目した結果を右側のグラフに示す.

図 12 から, グループ特徴量の中で, 特にグループ内人数特徴量のスコア R^{gp} によって, カメラ間人同定の精度が向上していることが分かる. 一方, グループ内色特徴量のスコア R^{gc} も, R^c と R^{gp} を使用した時の精度と R^c , R^{gp} と R^{gc} を併用時の精度を比較すると, カメラ間人同定の精度向上に貢献していることが分かる. しかし, グループ内人数特徴量による効果と比較するとグループ内色特徴量の効果は小さく, グループ内人数特徴量を併用しなかった場合は有用性が得られなかったため, グループ内色特徴量の取得法については, 新たな取得法を考える必要がある.

色ヒストグラムのスコア R^c と移動時間から求めたスコア R^t にグループ特徴量を併用したカメラ間人同定の結果を図 13 に示す. 図 12 と比較すると, R^t による精度向上が顕著である. しかし, 100 位以内に着目した右側のグラフを見ると, 2 種類のグループ特徴量それぞれによる精度向上の特性は, 図 12 と同様である.

また, グループ特徴量を併用した時と併用しなかった時について, 本人同士の間で得られた特徴量が 10 位以内に入った人数を比較することで, グループ特徴量の効果を定量的に評価する. 結果として, 2 種類の従来特徴量 (色特徴量・時間特徴量) のみを使ったときは 135 人, 従来特徴量と 2 種類のグループ特徴量を併用した時は 195 人となり, グループ特徴量によってカメラ間人同定における精度に 44%の精度向上が見られた.

6.2.3 正解軌跡とグループ検出結果を用いたカメラ間人同定

グループ検出精度が, カメラ間人同定におけるグループ特徴量に与える影響を調べた. そのため, グループ検出の結果として 6.1 章で行ったグループ検出結果を用いる事で, 各人のグループ特徴量を求める. そして, 導出されたグループ特徴量と従来特徴量を併用し, カメラ間人同定の精度を確かめ, 6.2.2 章の実験結果と比較する.

まず, 色特徴量のスコア R^c とグループ特徴量を併用し

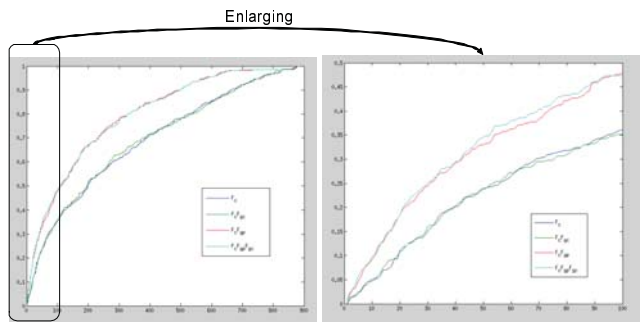


図 12 人の色特徴量とグループ特徴量を併用した際の、カメラ間人同定の結果図 (理想環境中).

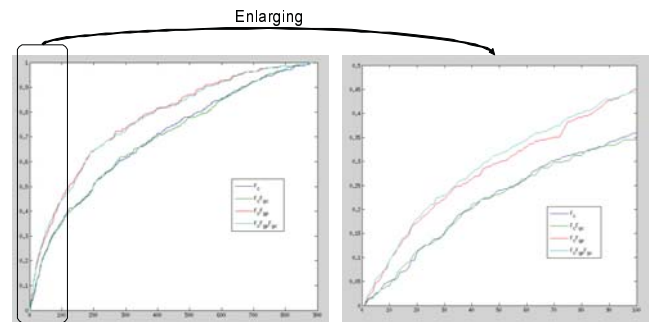


図 14 グループ検出結果によるグループ特徴量と色特徴量を併用した際の、カメラ間人同定の結果図.

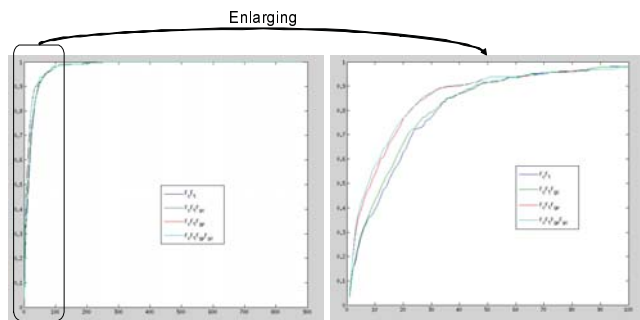


図 13 左側のグラフは人の色特徴量・時間特徴量とグループ特徴量を併用した際の、カメラ間人同定の結果図 (理想環境中).

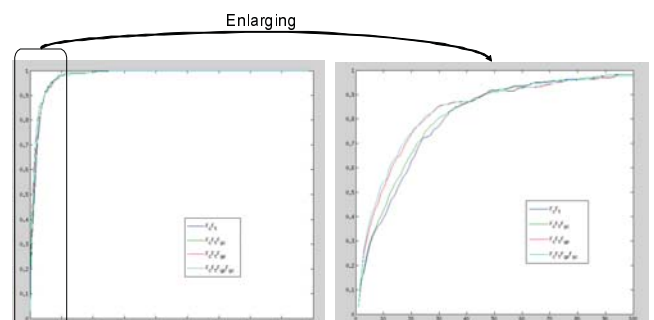


図 15 グループ検出結果によるグループ特徴量と人の色特徴量・時間特徴量を併用した際の、カメラ間人同定の結果図.

た際の結果を図 14 に示す. 図 14 から, グループ検出の結果から得たグループ特徴量を用いてもカメラ間対象同定の精度が向上していることが分かる. 図 14 と図 12 の比較から, グループ内人数特徴量の効果の低減が見られた.

次に, 2つの従来特徴量のスコア R^c , R^t とグループ特徴量を併用した際の, カメラ間人同定の結果を図 15 に示す. 図 13 の右側のグラフの比較から, 時間特徴量を用いない場合と同様に, グループ特徴量によって, カメラ間人同定の精度向上が見られた. また左側のグラフ同士の比較から, これも時間特徴量を用いない場合と同様に, グループ検出の失敗によるグループ特徴量の効果の低減が見られ, 300 位以降においてはグループ特徴量の効果が見られなくなった.

6.2.2 章と同様に, グループ特徴量の効果を定量的に評価する. 結果として, 従来特徴量のみを使ったときは 135 人, 従来特徴量と 2 種類のグループ特徴量を併用した時は 185 人となり, グループ特徴量によってカメラ間人同定における精度に 37%の精度向上が見られた.

6.2.4 推定軌跡とグループ検出結果を用いたカメラ間人同定

実環境中で本手法のシステムを実行した際のカメラ間人同定精度を調べるため, 正解データを完全に除いて, カメラ視野内追跡結果およびグループ検出結果を用いた評価を行った. 双方のカメラで検出されている 342 人中, 双方のカメラ視野内で追跡成功している 272 人について, 各カメラ視野内の追跡結果に基づくカメラ間人同定を行う.

まず, 追跡結果から得たグループ特徴量と色特徴量のスコア R^c を併用した際の, カメラ間人同定の結果を図 16 に示す. 図 16 から, 追跡結果から得たグループ特徴量を用いても, グループ内人数特徴量によってカメラ間人同定の精度が向上していることが分かる.

また, 図 16 と図 14 の比較から, 追跡結果を用いることによるカメラ間人同定の精度の低減が見られた. この原因として, 1) 視野内追跡失敗によるグループ検出の失敗による影響と, 2) 追跡時に得られる矩形が対象を中心としたものになっていないことによる, 色情報に背景情報が含まれてしまった影響, の 2 点が考えられる.

次に, グループ検出の結果から得たグループ特徴量と従来特徴量のスコア R^c , R^t を併用した際の結果を図 17 に示す. 図 17 の右側のグラフの比較から, 時間特徴量を用いない場合と同様に, グループ特徴量による精度向上が見られた. また図 15 との比較から, これも時間特徴量を用いない場合と同様に, 人追跡の失敗によるグループ検出の失敗と色情報への背景色の混入から, 精度低減が見られた.

これまでの実験と同様に, グループ特徴量の効果を定量的に評価する. 結果として, 従来特徴量のみを使ったときは 43 人, 従来特徴量と 2 種類のグループ特徴量を併用した時は 57 人となり, グループ特徴量によってカメラ間人同定における精度に 33%の精度向上が見られた. しかし, 従来特徴量と 2 種類のグループ特徴量の併用時でも, 本人同士の類似度が 10 位以内に含まれるのは全 272 人中 16%で

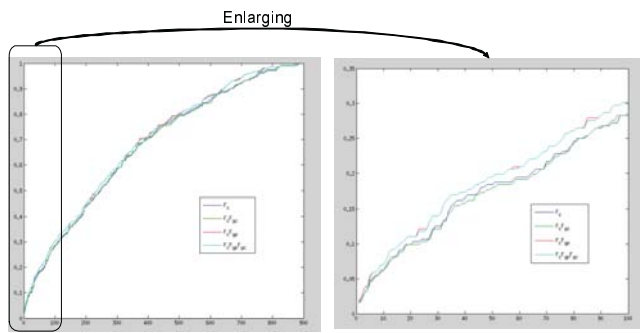


図 16 追跡結果を用いたグループ特徴量と色特徴量を併用した際の、カメラ間人同定の結果図。

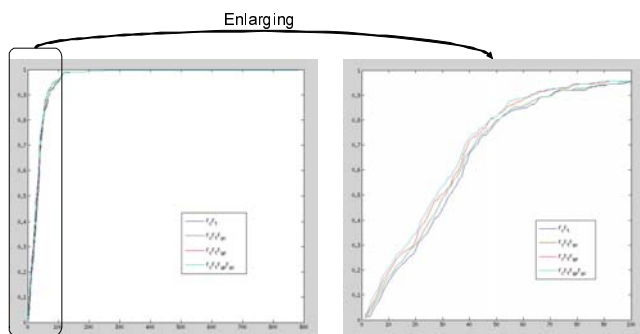


図 17 追跡結果を用いたグループ特徴量と人の色特徴量・時間特徴量を併用した際の、カメラ間人同定の結果図。

ある．ここで、6.2.3 章の正解軌跡を用いた場合の実験結果は全 342 人中 54% であることから、追跡結果を用いる事による精度の低下が著しいことが分かる．本システムを実環境中で利用するには、更なる単眼カメラ視野内での人追跡精度の向上が必要である．

以上より、実環境中で本システムを使用するためには、単眼カメラ視野内における人追跡精度を向上させ、カメラ間人同定の精度を向上させる必要があるといえる．

7. おわりに

本研究では、カメラ間人同定における新たな情報としてグループ特徴量を提案し、グループ特徴量から得たグループ特徴量と従来特徴量を併用することで、カメラ間人同定の精度を向上する手法を提案した．提案手法では、グループ内色特徴量の効果が弱かったため、本研究でのグループ内色特徴量の取得法を検討する必要がある．また、この 2 種のグループ特徴量以外の新たなグループ特徴量を加える事で、カメラ間人同定の精度を向上させる可能性を検討する．また、本研究において、単眼カメラ視野内における人追跡の精度が低かったため、追跡結果を用いたグループ検出・カメラ間人同定の実験において十分な精度を確かめることができなかつた．よって、さらに頑健な単眼カメラ視野内の追跡手法を用いて、追跡結果からグループ検出・カメラ間人同定を行うことで、実環境中で本システムを実行できるように改良をする必要がある．

参考文献

- [1] 井尻喜久, 川西康友, 美濃導彦, 村瀬洋, “サーベイ論文: 視野を共有しない複数カメラ間での人物照合” 電子情報通信学会技術研究報告, 111(317), pp.117–124, 2011.
- [2] Z. Yucel, T. Ikeda, T. Miyashita, and N. Hagita, “Identification of mobile entities based on trajectory and shape information,” *IROS*, 2011.
- [3] H. Pirsivavash, D. Ramanan, and C. C. Fowlkes, “Globally-Optimal Greedy Algorithms for Tracking a Variable Number of Objects,” *CVPR*, 2011
- [4] O. Hamdoun, F. Moutarde, B. Stanculescu, and B. Steux, “Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences,” *ACM/IEEE International Conference on Distributed Smart Cameras*, 2008.
- [5] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, “Shape and Appearance Context Modeling,” *ICCV*, 2007.
- [6] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person Re-Identification by Symmetry-Driven Accumulation of Local Features,” *CVPR*, 2010.
- [7] O. Javed, K. Shafiqe and M. Shah, “Appearance Modeling for Tracking in Multiple Non-Overlapping Cameras,” *CVPR*, 2005.
- [8] A. Gilbert and R. Bowden, “Tracking Objects Across Cameras by Incrementally Learning Inter-camera Colour Calibration and Patterns of Activity,” *ECCV*, 2006.
- [9] V. Kettner and R. Zabih, “Bayesian Multi-Camera Surveillance,” *CVPR*, 1999.
- [10] O. Javed, Z. Rasheed, K. Shafiqe and M. Shah, “Tracking Across Multiple Cameras with Disjoint Views,” *ICCV*, 2003.
- [11] D. Makris, T. Ellis and J. Black, “Bridging the gaps between cameras,” *CVPR*, 2004.
- [12] N. Ukita, “Probabilistic-Topological Calibration of Widely Distributed Cameras,” *Machine Vision and Applications*, Vol.18, No.3, pp.249–260, 2007.
- [13] X. Wang, K. Tieu and W. E. L. Grimson, “Correspondence-Free Activity Analysis and Scene Modeling in Multiple Camera Views,” *PAMI*, Vol.32, No.1, pp. 56–71, 2010.
- [14] A. C. Gallagher and Tsuhan Chen, “Understanding Images of Groups of People,” *CVPR*, 2009.
- [15] D. Helbing and P. Molnar, “Social force model for pedestrian dynamics,” *Physical Review E*, Vol.51, No.5, pp.4282–4286, 1995.
- [16] W. Ge, R. T. Collins, and R. B. Ruback, “Vision-based Analysis of Small Groups in Pedestrian Crowds,” *PAMI*, Vol.34, No.5, pp.1003–1016, 2012.
- [17] S. Pellegrini, A. Ess, and L. van Gool, “Improving Data Association by Joint Modeling of Pedestrian Trajectories and Groupings,” *ECCV*, 2010.
- [18] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, “Who are you with and Where are you going?” *CVPR*, 2011.
- [19] S. Maji, A. C. Berg and J. Malik, “Classification using Intersection Kernel Support Vector Machines is Efficient” *CVPR*, 2008.
- [20] M. Hirzer, C. Belezna, P. M. Roth and H. Bischof, “Person Re-Identification by Descriptive and Discriminative Classification” *Scandinavian Conference on Image Analysis*, 2011.