

分散カメラ間におけるグループ特徴量を用いた人同定

守口 裕介[†] 浮田 宗伯^{†a)} 萩田 紀博[†]

People Identification across Non-overlapping Cameras Using Group Features

Yusuke MORIGUCHI[†], Norimichi UKITA^{†a)}, and Norihiro HAGITA[†]

あらまし 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。この特徴量は、従来広く研究されてきた人の見え以外の情報も表現している。このように見えとは独立な特徴量によって、見た目が類似している人の識別精度を向上させる。まず、各カメラ視野において、撮影されている人々をグループごとに分ける。このグループ検出の結果から各人のグループ特徴量を獲得し、個人の色情報やカメラ間の移動時間を参照し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、軌跡とグループ属性の正解データを与えた場合、従来特徴量のみを用いたときと比較して、カメラ間人同定の精度を18%向上できた。また、グループ検出の実験結果を用いた場合においても、同定精度を17%向上できた。

キーワード 人同定, 追跡, 分散カメラ, グループ検出

1. ま え が き

一台のカメラでは撮影できる範囲は限定されているため、広域人追跡のためには複数台のカメラが連携しなければならない。複数台のカメラにわたる人追跡では、各カメラの視野内での人追跡に加え、あるカメラに映った人が別のカメラで写った人と同一人物か判断する人同定が必要になる。監視カメラはコストの面から視野に重なりのない配置になっていることが多く、継続的な追跡が難しい。また、一般的な監視カメラ画像中では図1のように各人は小さく写っているため、顔認識などの解像度が高くないと利用できない手法を用いるのは困難である。以上の問題点を解決するため、低解像からでも抽出できる個人の見え情報に着目した研究が多数報告されている(サーベイ[1]参照)。しかし、精度の面で頑健なカメラ間人追跡技術は確立されておらず、更なる精度向上が望まれる。

本研究では、カメラ間人同定において個人の見え以外に利用できる特徴量として、「位置関係や軌跡の類似性に着目することで検出される人の集まり」をグルー



図1 監視カメラの映像例。広域人追跡の実現には、各カメラの視野内で撮影された人の追跡(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(実線)が必要。各枠はそれぞれグループを示している。

Fig. 1 Examples of images captured by surveillance cameras. For wide-area human tracking, identification across fields of view (indicated by solid lines) is required in addition to tracking within each field of view (indicated by broken lines). Each rectangle shows a group.

プと定義し、このグループ(図1内の各枠参照)に関する特徴量を提案する。カメラ間をまたいでもグループの関係が変化しにくいという仮定の下、その人が属するグループの人数やグループ内全員の人の見え情報から「グループ特徴量」を得る。この仮定は、通路が限定されるビルなどの屋内や、歩行者間の進路変動が小さいほどにカメラ間の距離が短いような環境において成立する。このグループ特徴量と従来手法の特徴量を併用しカメラ間人同定の精度向上を実現する。

[†] 奈良先端科学技術大学院大学, 生駒市
Nara Institute of Science and Technology, Ikoma-shi, 630-0192 Japan

a) E-mail: ukita@is.naist.jp

2. 関連研究

複数台のカメラ間での人同定について、個人の見え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。個人の見えの特徴量化には、Hamdounら[2]の特徴点ベース、Wangら[3]のHOG[4]ベースなどの手法が提案されている。個人の見えの中でも、人の向きの変化やカメラの視点変化に頑健なものとして色ヒストグラムの取り出し方を工夫する研究がされている。その中には、Farenzenaら[5]のHSVヒストグラムを利用したものなどが存在する。本研究でもこれらの研究と同様に、人の色ヒストグラムを、カメラ間人同定に用いる特徴量の一つとして利用する。個人の見えに着目した研究には、他にも各カメラの照明条件などによる色の变化を補正する手法を提案した研究[6],[7]や、SVMやAdaboostなどの識別手法を工夫した研究[8],[9]など様々な研究も存在する。

カメラ間人同定において、個人の見え特徴量だけでなく、カメラ視野間の移動に費やした時間を表す時空間特徴量を利用する研究も多く存在する(例えば、大半の人がカメラA視野のある座標からカメラB視野のある座標までの移動に60秒費やすシーンでは、移動時間が10秒の同一人物の対応付け成功の可能性は低い)。カメラ視野間の関係をモデル化するためには、

カメラ間追跡結果に基づく方法[10],[11]や、明には追跡を必要としない方法[12],[13]がある。

上述の既存特徴量に加え、本研究では、追跡対象となる人が属するグループの特徴に着目する。グループの検出には、対象の詳細像を撮影できない一般的な監視カメラにも適用できるように、視野内の追跡で得られる移動軌跡のみを参照した手法を用いる。Yamaguchiら[16]は、手法[17],[18]のような動画における対象追跡の軌跡間の関係(位置関係や類似性、同時観測性)を特徴量としてグループ検出を行った。

分散カメラ間の人同定にグループを利用する手法には、手法[19],[20]などが存在する。これらは、同定対象となる人の周辺領域(グループ構成員の写っている領域)の見え方も特徴量とすることにより、識別性の高い見えのマッチングを行っている。こうした方法は、単眼視野内の対象追跡でも利用されていた(例:手法[21])。一方、提案手法は、明にグループ構成員を個別に検出・グループ化した結果から見え以外の新たな特徴量を得ており、見えの特徴量抽出領域をグループ領域全体に拡大するのみの従来法とは異なる。

3. 処理概要

提案手法の概要を図2に示す。図中にて紫枠で示しているように、カメラ間人同定は、単眼カメラ視野内

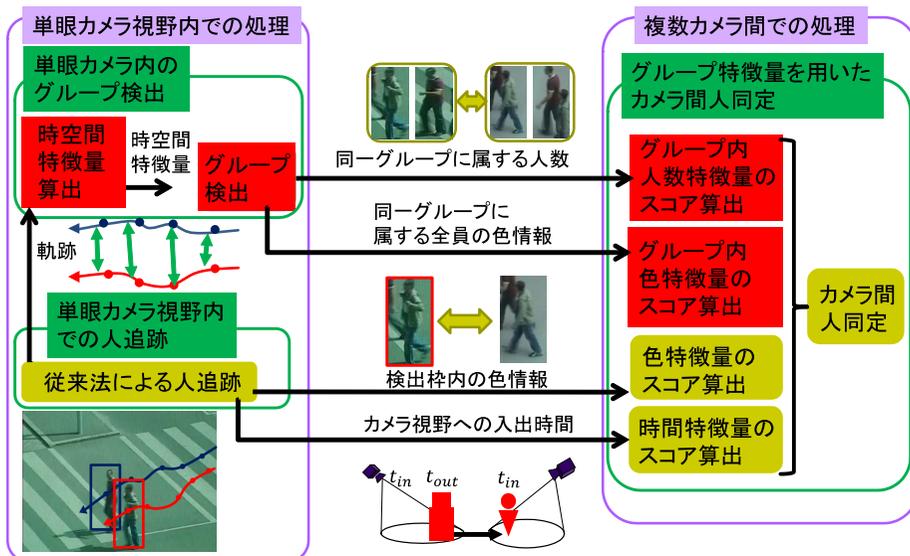


図2 同定処理全体の流れ。赤色の枠で表示されている処理が提案手法にて新たに追加した処理

Fig.2 Abstract of the proposed method. Processes proposed in the paper are colored by red.

で各人の情報を取得する処理と、取得した情報を複数カメラ間で比較し人同定を行う処理に大別される。提案手法は、個人の特徴量のみを参照する手法に図2の赤枠の処理を追加したものとなっている。以下では、全体を「単眼視野内での人追跡」、「単眼視野内のグループ検出」、「カメラ間人同定」の3処理に分けて概要を述べる。

単眼視野内での人追跡： 各視野内での人検出及び人追跡を行う。本論文では、公開データセット中の人領域検出結果（矩形領域）を追跡結果とした。

単眼視野内のグループ検出： 提案するグループ特徴量を求めるため、各視野内での人追跡により得られる軌跡からグループ検出を行う。このグループ検出には、手法[16]を改良した手法[22]を利用した。

カメラ間人同定： まず、従来特徴量である、人の色ヒストグラムと(4.1.1)、カメラ間の移動時間から(4.1.2)、人同定用のスコアを算出する。上記従来特徴量に加え、提案手法では新たにグループ特徴量から人同定用のスコアを求める。そこで、グループ検出結果から、その人の属するグループ内の人数（グループ内人数特徴量）と全員の色情報（グループ内色特徴量）を参照して、人同定用のスコアを求める(4.2)。以上4種類の特徴量（色特徴量、時間特徴量、グループ内人数特徴量、グループ内色特徴量）によりカメラ間人同定を行う(4.3)。

4. カメラ間人同定

単眼カメラ視野内の人領域から特徴量を算出し、異なるカメラで撮影された任意の2人の間で特徴量を比較して、その2人の同一人物らしさを表す人同定用のスコア R を算出し、カメラ間人同定を行う。

4.1 カメラ間人同定に用いる従来特徴量

4.1.1 人の色ヒストグラム

文献[5]~[9]等の多数の研究で行われているように、人追跡の結果得られる矩形領域内から抽出される色ヒストグラム間で類似度を算出する。二つの色ヒストグラム間の類似度には、非剛体対象追跡[23]等で用いられている、Bhattacharyya係数を用いた。色ヒストグラムは、RGBの各チャネルで8階調に分割したのち、正規化して24(=8×3)階調のヒストグラムを得た。任意の2人 i と j の領域に対する色ヒストグラムをそれぞれ C_i と C_j としたとき、 C_i と C_j の類似度を表すBhattacharyya係数 L (0から1の値をとり、値が高いほど類似度が高い)を下式により算出し、これを

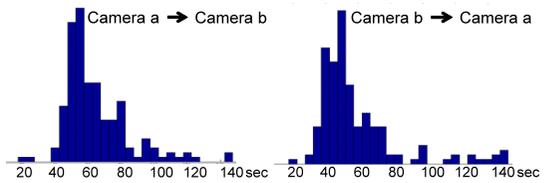


図3 カメラ視野間の移動時間の分布図の例。横軸がカメラ間の移動に費やした秒数、縦軸がカメラ間の移動にその秒数費やした人数を表す。

Fig.3 Distributions of transition times between fields of view. The horizontal and vertical axes show the transition times and the number of people who took each time interval for moving across the fields of view, respectively.

色特徴量のスコア R_C とする。

$$R_C(i, j) = L = \sum_{u=1}^{24} \sqrt{C_i(u)C_j(u)} \quad (1)$$

本手法では、1) 人追跡により得られる検出矩形領域内の背景色の影響を抑制するため、矩形の中心付近のみから色ヒストグラムを抽出し、2) 各対象の全追跡フレームにおける色ヒストグラムの平均を計算し、これを色特徴量とした。

4.1.2 カメラ間の移動時間

文献[10]~[13]等で提案されているように、人追跡の結果得られるカメラ視野への入出時間を利用して、カメラ間の移動時間という観測領域固有の特徴量のスコア R_T を求める。

カメラ間の移動時間の分布の例を、図3に示す。本研究ではカメラ間の経路が一つだけのシーンで実験を行ったため、カメラ間の移動時間は図3のように分布しており、一つのガウス分布で近似できるとみなした。カメラ間の移動時間の期待値 m_T と分散 σ_T^2 を求め、ガウス分布 $g_T(t; m_T, \sigma_T^2)$ を得る。

各カメラにおける視野への入出の時間から、任意の二つの検出結果のカメラ間の移動時間を求める。カメラ a で検出された歩行者 i が視野から退出した時間を $t_{out,a,i}$ 、カメラ b で検出された歩行者 j が視野に進入した時間を $t_{in,b,j}$ としたとき、下式のように、カメラ間移動時間 $t_{in,b,j} - t_{out,a,i}$ を算出済みのガウス分布 $g_T(t; m_T, \sigma_T^2)$ に代入することで、 i と j の同一人物らしさを表す移動時間特徴量のスコア R_T を求める。

$$R_T(i, j) = g_T(t_{in,b,j} - t_{out,a,i}; m_T, \sigma_T^2) \quad (2)$$

4.2 グループ特徴量

グループ検出[22]の結果から、各人が誰とグループ

なのかが分かる．その人が誰ともグループでない（1人で行動している）場合、1人グループが構成されているとする．ここで、カメラ間の移動においてこのグループ構成は変化しにくいという仮定の下、このグループ構成をカメラ間人同定に用いる．提案手法では、以下の二つの情報に注目した．

- (1) グループ内の全ての人の色情報：従来特徴量と同様、人領域検出の結果から比較的頑健に抽出可能
- (2) グループ内の人数：色情報とは独立であるため、色に基づく同定失敗を補う可能性のある特徴

4.2.1 グループ構成員の色特徴量

各人の属するグループ内の全員の領域から色ヒストグラムを一つ抽出し、任意の2グループ間でグループ内色特徴量の類似度のスコア R_{GC} を計算する．各グループの色ヒストグラムは、全構成員のヒストグラムを加算、正規化して得る．まず任意の2人 i と j が属するグループの全構成員の色ヒストグラムを抽出し、各グループの平均色ヒストグラム $C_{G,i}$, $C_{G,j}$ を抽出する．次に、4.1.1と同様に、ヒストグラム間の類似度を表す Bhattacharyya 係数 (4.1.1 参照) をグループ内色特徴量のスコア R_{GC} とする．

$$R_{GC}(i, j) = \sum_{u=1}^{24} \sqrt{C_{G,i}(u)C_{G,j}(u)} \quad (3)$$

4.2.2 グループ構成員の人数特徴量

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．理想的には、歩行者 i と歩行者 j のグループ内人数 N_i と N_j の差 $N_i - N_j$ は0になる．実際には、グループ検出が失敗したり、カメラ間を移動する際にグループ構成員の人数が増減する可能性があるため、差分が0になったときにスコアが高くなり、かつグループ検出の失敗や、グループ内人数の変化にも対応できるように、グループ内人数が異なる場合にはスコアを小さくする．本研究では、このような重みとして平均0、分散 σ_{GP}^2 のガウス分布を採用し、下式をグループ内人数特徴量のスコア $R_{GP}(i, j)$ とする．

$$R_{GP}(i, j) = g_{GP}(N_i - N_j; 0, \sigma_{GP}^2) \quad (4)$$

式(4)は、グループ検出の成功可能性が一定になってしまうように、分散が定数のガウス分布で表現されている．しかし、実際には、グループ検出法における検出の確信度に応じて、グループごとに重みは設定されるべきである．そこで、歩行者 i のスコア計算時に、

i と同時に視野内に観測された他の歩行者 j が、 i と同一のグループ内にいるか、それともグループ外にいるかに応じて以下のように重みを調節する．

- グループ内： i と j が同一グループに属する確信度が低いほど、グループ内人数特徴量の重みを下げる．
グループ外： i と j が同一グループに属さない確信度が低いほど、グループ内人数特徴量の重みを下げる．

この重みは、式(4)の分散の調節により与える．グループ検出の確信度の計算は、グループ検出法に依存する．本手法におけるグループ検出 [22] では、任意の2名が同一グループに属するか否かをサポートベクターマシン (SVM) [24] で識別する．SVMのスコアは、2名が同一グループに属する場合に正値をとり、値が大きいくほど確信度が高く、逆に、属さない確信度が高いほど小さな負値をとる．そこで、下式によりグループ内人数特徴量のスコア $R_{GP}(i, j)$ を得る．

$$R_{GP}(i, j) = g_{GP}(N_i - N_j; 0, e^{-aw_i}) \quad (5)$$

$$w_i = \prod_n^{M_i} \min(1, |s_{i,n}|) \quad (6)$$

ただし、 M_i は i と同時に視野内で観測された人数、 $s_{i,n}$ は i と n が同一グループに属するかどうかを SVM で判定したスコアである．人数が多いほど同一グループ判定の失敗可能性が高くなることに対応するため、0から1の値をとる $\min(1, |s_{i,n}|)$ の積によって人数が増えるほど w を小さくしている． w が小さくなると e^{-aw_i} が大きくなり、その結果、 $N_i - N_j$ が0以外の値をとってもスコア R_{GP} が大きく減少しない．

4.3 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

4.1, 4.2で求めたスコアから、カメラ間人同定を行う．カメラ間人同定用のスコア R は、4種のスコア R_C , R_T , R_{GP} , R_{GC} 全てが高くなった場合に R が高くなるよう、式(1)(2)(3)と式(4)または式(5)より、下式のような4種類のスコアの積で算出する．

$$R(i, j) = R_C(i, j)R_T(i, j)R_{GC}(i, j)R_{GP}(i, j) \quad (7)$$

この $R(i, j)$ が大きいものから、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．

5. 実験

実験に用いた動画は、Person Re-ID 2011

dataset [25] 中の二つのカメラで撮影された動画である。二つのカメラで撮影された画像例は、図 1 に示している。データセットには動画のほかに、各人の ID 付き追跡結果（人の全身を囲む矩形の x, y 座標とサイズ）の正解データが与えられている。異なるカメラ画像でも、同じ人には同一 ID が割り当てられている。

ただし、データセットに添付済みの正解追跡データは、人が遮蔽なく映っているフレームにおいてのみ保存されており、遮蔽の瞬間の正解データが含まれていない。また、誰と誰が同一グループ内ペアかというグループ情報の正解データも含まれていないため、各動画中の位置関係や軌跡の類似度、また親子等の社会的関係などに着目し手動で与えた。この際、片方のカメラでのみグループと判別できるようなペアも、同一グループ内のペアとして設定した。

上記修正を加えた本実験で用いるデータセットの内容は、以下のとおりである。

- カメラ a
フレーム数 92825 フレーム (約 62 分)
追跡人数 573 人
うち 2 名以上のグループ内の人数 103 人
- カメラ b
フレーム数 99997 フレーム (約 67 分)
追跡人数 884 人
うち 2 名以上のグループ内の人数 147 人
- 二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

提案手法の有効性確認のため、以下の実験を行った。

(1) カメラ間人同定におけるグループ特徴量の効果を確認するための、正解データから算出したグループ特徴量を用いたカメラ間人同定実験 (5.1)。

(2) 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るための、グループ検出結果から算出したグループ特徴量を用いたカメラ間人同定実験 (5.2)。

カメラ a の視野内で検出された人 i に対して、カメラ b で検出された 884 人全員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 (7) により求める。この処理を、双方のカメラ視野内で検出成功している 342 人に対して行い、全候補 884 人中での本人同士の間スコア $R_{(i,i)}$ の順位を調べることで、人同定の精度を評価する。この精度評価のために、CMC 曲線 (Cumulative Match Characteristic

curve) を用いる。

グループ内人数特徴量のスコア計算に用いる式 (4) の σ_{GP}^2 と式 (5) の a は、本実験のシーンにおける予備実験から $\sigma_{GP}^2 = 0.222 \dots$, $a = 1$ とした。

上記の同定精度評価を、グループ特徴量を併用した場合と併用しなかった場合、それぞれについて求め、グループ特徴量の有用性を調べる。

5.1 正解グループを用いたカメラ間人同定

まず、カメラ間人同定におけるグループ特徴量の理想的な効果を調べるために、人の所属グループや軌跡に正解データを使用して R を算出する。この実験では、グループの正解データが与えられるため、式 (6) と式 (4) との本質的な差がなくなる。よって、式 (4) によるグループ内人数特徴量のみを評価した。

色特徴量と 2 種類の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 4 に示す。図から、特に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。一方、グループ内色特徴量も、 R_C と R_{GP} を使用したときの精度と R_C , R_{GP} と R_{GC} を併用時の精度を比較すると、多少の精度向上に貢献していることが分かる。

次に、2 種類の従来特徴量 (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を図 5 に示す。図 5 と図 4 の左側のグラフを比較すると、時間特徴量によるカメラ間人同定の精度向上が顕著であることが分かる。また、20 位以内の部分を拡大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ってカメラ間人同定の精度が向上していることが分かり、2 種類のグループ特徴量をどちらも併用した際に最も精度向上が見られた。

また、グループ特徴量を併用したときと併用しなかったときについて、本人同士の間で得られたスコアが 10 位以内^(注1)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。2 種類の従来特徴量のみを使ったときは 140 人、従来特徴量と 2 種類のグループ特徴量を併用したときは 201 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 $(\frac{201}{342} - \frac{140}{342}) \times 100 = 18\%$ の精度向上が見られた。参考までに、本人同士のスコアが 1 位になった人数は、グループ特徴量を用いなかった場合、用いた場合で、それぞれ 39 人、47 人であった。

(注1)：このような監視を主な目的としたシステムでは、最終的には人による確認が行われることが想定される。そこで、スクリーニング目的として、上位 10 位以内という基準で定量的に評価した。

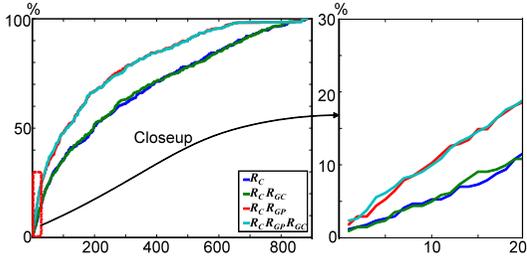


図 4 人追跡やグループ検出について正解データを用いて、色特徴量とグループ特徴量を併用した際の、人同定の CMC 曲線

Fig. 4 CMC curve of identification using color features and two kinds of group features obtained from the ground-truth of people tracking and grouping.

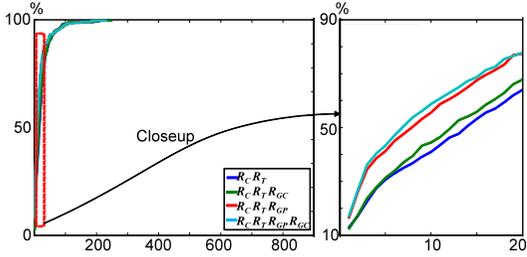


図 5 人追跡やグループ検出について正解データを用いて、2 種類の従来特徴量とグループ特徴量を併用した際の、人同定の CMC 曲線

Fig. 5 CMC curve of identification using the features of color and transition interval and two kinds of group features obtained from the ground-truth of people tracking and grouping.

以上より、グループ特徴量が精度向上に貢献していることが分かり、カメラ間人同定におけるグループ特徴量の有用性が示された。

5.2 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るために実験を行った。手法 [22] による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。この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、5.1 の実験結果と比較する。

色特徴量とグループ検出の結果から得たグループ特徴量を併用した際の、同定結果を図 6 に示す。グループ内人数特徴量として、式 (4) と式 (5) を用いた結果をともに示している。図 6 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。ただし、図 6 と図 4 の比較から、グループ内人数特徴量の効果の低減

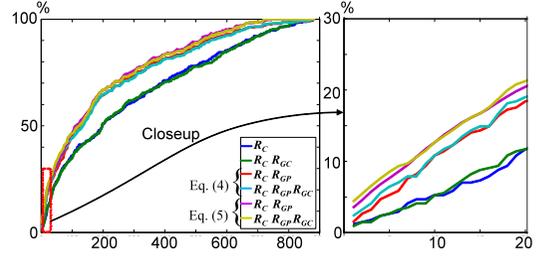


図 6 色特徴量とグループ検出結果によるグループ特徴量を併用した際の、カメラ間人同定の結果図

Fig. 6 CMC curve of identification using color features and two kinds of group features obtained from the results of people grouping.

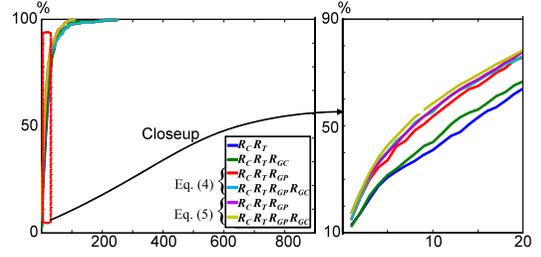


図 7 2 種類の従来特徴量とグループ検出結果によるグループ特徴量を併用した際の、カメラ間人同定の結果図

Fig. 7 CMC curve of identification using the features of color and transition interval and two kinds of group features obtained from the results of people grouping.

が見られた。これは、グループ検出の失敗によるものであり、カメラ間人同定精度への影響が大きいグループ内人数特徴量が特に影響を与えている。式 (4) と式 (5) のグループ特徴量を比較すると、グループ検出の確信度を参照する式 (5) による結果のほうが、特徴量の効果の低減を抑えることができている。

次に、2 種類の従来特徴量 (色特徴量と時間特徴量) とグループ検出の結果から得たグループ特徴量を併用した際の、同定結果を図 7 に示す。図 5 の右側の拡大グラフの比較から、時間特徴量を用いない場合 (図 6) と同様に、グループ特徴量を用いることによりカメラ間人同定の精度向上が見られた。

5.1 と同様に、グループ検出から得られたグループ特徴量の効果を、本人同士の間で得られたスコアが 10 位以内に入った人数で定量的に評価する。結果として、従来特徴量のみを使ったときは 140 人、従来特徴量と 2 種類のグループ特徴量 (グループ内人数特徴量に式 (4) を用いた) を併用したときは 189 人となり、

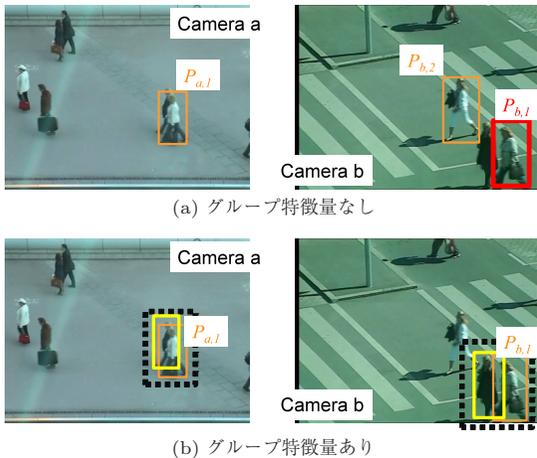


図 8 グループ特徴量の効果の例. 手法により同定された対象領域を橙の矩形で示す.

Fig. 8 Example of the effect of the group features. Each orange rectangle indicates detected region that are identified.

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 $(\frac{189}{342} - \frac{140}{342}) \times 100 = 14\%$ の精度向上が見られた. 更に, グループ内人数特徴量に式 (5) を用いた場合, 従来特徴量と 2 種類のグループ特徴量の併用によって 198 人の対応付けに成功し, 同定精度の向上は 17% であった. 参考までに, 本人同士のスコアが 1 位になった人数は, グループ特徴量を用いなかった場合, 用いた場合で, それぞれ 39 人, 44 人であった.

以上より, グループ検出の正解データを用いたときと比べ, グループ特徴量の導出にグループ検出結果を用いた際には同定精度に多少の低減はあるものの, 従来特徴量のみを用いるよりも同定精度を向上できることが確認できた. このことより, 実際のグループ検出法の誤りを含んでしまう場合でも, グループ特徴量による同定精度向上の効果が得られるといえる.

図 8 に, グループ特徴量の効果の例を示す. グループ特徴量無しの場合, カメラ a で検出された $P_{a,1}$ は, カメラ b の $P_{b,2}$ と同定してしまった. 実際には, $P_{a,1}$ と同一対象は $P_{b,1}$ であるが, 照明条件の変化などによって, $P_{b,1}$ の白い服の輝度が落ちてしまったことなどが同定失敗の原因である. これに対し, グループ特徴量を用いた場合は, 正しく $P_{a,1}$ と $P_{b,1}$ を同定できた. これは, 両カメラでのグループ検出の結果 (図 (b) 中の黒点線の矩形中の 2 人グループ), $P_{a,1}$ と $P_{b,1}$ の間で, 1) グループ内人数特徴量のスコアが高くなり, 2) グループ内色特徴量のスコアも高くなった ($P_{b,2}$

のグループ内色特徴量は主に白のみで構成されるが, $P_{a,1}$ と $P_{b,1}$ のグループ内色特徴量にはもう一方の構成員の服の色である黒が混入する) 結果である.

6. む す び

カメラ間人同定のためのグループ特徴量を提案し, これを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させた. 実験において, 軌跡とグループ属性の正解データを与えた場合, 従来特徴量のみを用いたときと比較して, グループ特徴量を併用することで, カメラ間人同定の精度を 18% 向上させた. また, グループ検出の実験結果を用いた場合においても, カメラ間人同定の精度が最大 17% 向上した.

グループ特徴量の効果は, 観測人数中のグループに含まれる人の割合や, 同時に同人数のグループがどの程度観測されるかによって決まる. 実験では, 2 台のカメラでそれぞれ, $\frac{103}{573} \approx 0.18$, $\frac{147}{884} \approx 0.16$ の割合でグループ中の人間が観測されていた. グループが観測されなかった場合は, グループ特徴量を用いなかった場合によっておおまかに近似できる. 一方, グループが増えた場合, ある程度までは人同定の精度向上が見込まれるが, 各瞬間に同人数のグループ数が増えくると, 効果は薄れてくることが予想される.

今後の課題として, 提案した 2 種のグループ特徴量のうちグループ内色特徴量の効果が弱いため, 更に効果的な色特徴量の取得法を検討する必要がある. 具体的には, 文献 [19], [20] などで提案されている識別性の高い色特徴量の採用などが有効であると考えられる. また, 本論文では, 公開データセットに含まれる人領域の正解データを利用して実験を行ったが, 人領域の検出及び追跡結果の誤りが提案手法に与える影響を調査し, 実応用に向けた更なる開発も重要な課題である.

文 献

- [1] 井尻善久, 川西康友, 美濃導彦, 村瀬 洋, “視野を共有しない複数カメラ間での人物照合,” 信学技報, PRMU2011-119, 2011.
- [2] 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.
- [3] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, “Shape and appearance context modeling,” ICCV, 2007.
- [4] N. Dalal and B. Berg, “Histograms of oriented gradi-

- ents for human detection” CVPR, 2005.
- [5] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person re-identification by symmetry-driven accumulation of local features,” CVPR, 2010.
- [6] O. Javed, K. Shafique, and M. Shah, “Appearance modeling for tracking in multiple non-overlapping cameras,” CVPR, 2005.
- [7] A. Gilbert and R. Bowden, “Tracking objects across cameras by incrementally learning inter-camera colour calibration and patterns of activity,” ECCV, 2006.
- [8] C.H. Kuo, C. Huang, and R. Nevatia, “Inter-camera association of multi-target tracks by on-line learned appearance affinity models,” ECCV, 2010.
- [9] W.S. Zheng, S. Gong, and T. Xiang, “Person reidentification by probabilistic relative distance comparison,” CVPR, 2011.
- [10] V. Kettner and R. Zabih, “Bayesian multi-camera surveillance,” CVPR, 1999.
- [11] O. Javed, Z. Rasheed, K. Shafique, and M. Shah, “Tracking across multiple cameras with disjoint views,” ICCV, 2003.
- [12] D. Makris, T. Ellis, and J. Black, “Bridging the gaps between cameras,” CVPR, 2004.
- [13] N. Ukita, “Probabilistic-topological calibration of widely distributed cameras,” Machine Vision and Applications, vol.18, no.3, pp.249–260, 2007.
- [14] W. Ge, R.T. Collins, and R.B. Ruback, “Vision-based analysis of small groups in pedestrian crowds,” IEEE Trans. Pattern Anal. Mach. Intell., vol.34, no.5, pp.1003–1016, 2012.
- [15] S. Pellegrini, A. Ess, and L. van Gool, “Improving data association by joint modeling of pedestrian trajectories and groupings,” ECCV, 2010.
- [16] K. Yamaguchi, A.C. Berg, L.E. Ortiz, and T.L. Berg, “Who are you with and Where are you going?” CVPR, 2011.
- [17] H.B. Shitrit, J. Berclaz, F. Fleuret, and P. Fua, “Tracking multiple people under global appearance constraints,” ICCV, 2011.
- [18] H. Pirsivash, D. Ramanan, and C.C. Fowlkes, “Globally-optimal greedy algorithms for tracking a variable number of objects,” CVPR, 2011.
- [19] W.-S. Zheng, S. Gong, and T. Xiang, “Associating groups of people,” BMVC, 2009.
- [20] Y. Cai, V. Takala, and M. Pietikäinen, “Matching groups of people by covariance descriptor,” ICPR, 2010.
- [21] M. Yang, Y. Wu, and S. Lao, “Intelligent collaborative tracking by mining auxiliary objects,” CVPR, 2006.
- [22] 守口裕介, 岡田亜沙美, 浮田宗伯, 萩田紀博, “軌跡間の時空間特徴量を用いた人のグループ検出,” 信学論 (D), vol.J96-D, no.11, pp.2776–2783, Nov. 2013.
- [23] V. Ramesh and P. Meer, “Real-time tracking of non-rigid objects using mean shift,” CVPR, 2000.
- [24] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” ACM Trans. Intelligent Systems and Technology, vol.2, Issue.3, pp.27:1–27:27, 2011.
- [25] M. Hirzer, C. Belezni, P.M. Roth, and H. Bischof, “Person re-Identification by descriptive and discriminative classification,” Scandinavian Conference on Image Analysis, 2011.

(平成 25 年 7 月 25 日受付, 11 月 18 日再受付)



守口 裕介

2013 年奈良先端科学技術大学院大学修士課程修了。現在、三菱ビルテクノサービス勤務。在学中、人群解析及び広域対象追跡に関する研究に従事。



浮田 宗伯 (正員：シニア会員)

2001 年、京都大学大学院博士後期課程修了。同年奈良先端科学技術大学院大学情報科学研究科助手。2007 年同准教授。2002 年～2006 年まで科学技術振興機構さきがけ(「情報基盤と利用環境」領域) 研究員兼任。2007 年～2009 年までカーネギーメロン大学客員研究員兼任。2011 年より ATR 客員研究員を兼任。博士(情報学)。コンピュータビジョン, 分散協同視覚, 対象追跡, 人体運動解析・姿勢推定に関する研究に従事。



萩田 紀博 (正員：フェロー)

1978 年慶應義塾大学大学院工学研究科電気工学専攻修士課程修了。同年電電公社(現 NTT)武蔵野電気通信研究所入所。NTT 基礎研究所, ATR メディア情報科学研究所長などを経て、現在 ATR 知能ロボティクス研究所長。この間、文字認識, 画像認識, コミュニケーション科学, コミュニケーションロボットなどの研究に従事。工学博士。IEEE, 電子情報通信学会, 情報処理学会, 日本ロボット学会, 人工知能学会各会員。