

軌跡間の時空間特徴量を用いた人のグループ検出

守口 裕介[†] 岡田亜沙美[†] 浮田 宗伯^{†a)} 萩田 紀博[†]

People Grouping by Spatio-Temporal Features of Trajectories

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

あらまし 複数人の移動軌跡から、人のグループを検出する手法を提案する。このグループ検出は、任意の2人が同一のグループに属しているかどうかを判定することによって行われる。この2人組の軌跡対は、軌跡間の距離や相対速度などの時空間的な関係によって特徴量化される。その特徴量は、識別器によって「グループ内」か「グループ外」に識別される。従来の類似特徴量と比較して、提案する特徴量は、類似した時空間関係の間の識別に優れ、軌跡のノイズに対して頑健である。実験では、歩行者を対象としたレーザーレンジファインダによる計測軌跡の公開データセットを用い、提案手法によるグループ検出の精度向上を示す。

キーワード 人のグループ、歩行軌跡、軌跡間の時空間特徴量

1. ま え が き

人の追跡技術は、様々な応用の基盤。人追跡には、カメラ [1], [2] やレーザーレンジファインダ (LRF) [3], [4] などが主に利用されているが、いずれのセンサにおいても、追跡結果は人の軌跡として出力される。

人追跡の目標は、各人の軌跡をそれぞれ獲得することである。複数人の追跡の場合、相互遮蔽などに対して頑健な追跡が課題となる。こうした各人の追跡結果に加えて、複数人の間の関係性やインタラクションなどの情報も、様々な応用において有用になる。例えば、誰と誰が一緒に移動しているかというグループ情報は、人追跡における有用な手がかりとなる。また、一緒に行動しているグループ内の人々がどのような構成であるかによって、店舗などでの携帯端末を介したナビゲーションや宣伝内容の制御が可能になる。

本研究では、図 1 に示すような複数人の軌跡から、その人々のグループを以下の仮定の下で検出する。

- 観測空間中の複数人は、グループごとにまとまって移動している (通勤ラッシュ時の駅構内のような混雑状況ではない)。図 1 の例のように、店舗などでは家族や友人同士などのグループがまとまって移動

するという仮定は妥当であり、本論文の実験も、実際の店舗 (図 1) における計測データを用いて行った。

- 図 1 (右) に示すような床平面座標における軌跡は、カメラ [1], [2] や LRF [3], [4] で計測済みである。ただし、遮蔽等による計測誤差やミス (例: 一人の軌跡が複数に分離する) は含まれる。実験に用いるデータセット中の軌跡を取得したシステムでは、 $4.6 \pm 2.7 \text{cm}$ [21] の計測誤差が含まれるため、この程度の誤差を許容したグループ検出が、本研究の目的である。

こうした複数人の軌跡間の時空間特徴量を抽出し、この時空間特徴量をグループ内とグループ外に識別す

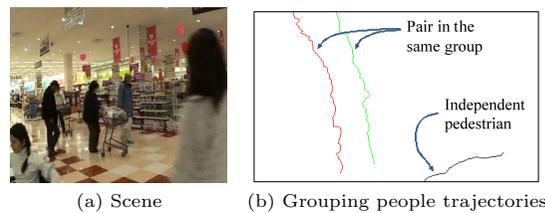


図 1 観測空間中を移動する人々。(a) は観測空間の様子をイメージするためのサンプル画像 (本論文では処理に画像は利用しない)。(b) は人の軌跡の鳥瞰図 (床面に平行な 2D 平面座標)。この人々の軌跡を、それらの時空間的な関係性に基づいて、(b) に示すようにグループごとに分割することを目標とする

Fig.1 Image (a) shows people in a scene. Their trajectories are represented in 2D coordinates (i.e. bird's eye view as shown in (b)) and divided into groups based on spatio-temporal relationships between the trajectories.

[†] 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

ることによって、グループ検出を実現する。

軌跡の時空間特徴量を利用したグループ検出では、軌跡に含まれるノイズや不規則な変動（例：急に立ち止まる）への対処が重要となる。この対処のため、提案する時空間特徴量とそれらの識別法では、従来法と比較して、1) 識別が難しいノイズの大きな特徴量をとのぞき、2) 異なる特徴間の共起性まで評価できる特徴量を加える、ように適切に特徴量を設定することによって、グループ検出の精度を向上させる。

2. 関連研究

カメラや LRF などのセンサにより、観測視野内における人の移動軌跡を得る追跡手法は、これまでに数多く研究されている [7], [8]。

グループには、「共通の志向が分有されている集合」のように社会学などにおいて様々な定義がある。本研究では、Cooley [9] の提唱した第一次集団（直接的接触に基づく親密な関係の社会集団、例：家族、友達等）をグループと呼ぶ。すなわち、家族や友達同士などが、物理的に近接して共に移動・行動している状態をグループと呼ぶ。本研究では、一人で行動する人も一人のグループを構成しているとみなす。

このように物理的な距離に基づいて、グループ検出やグループ構成員の属性推定（例：友人、親子、夫婦）は、静止画像を対象としても研究されている [10], [11]。これに対し、本研究では時系列データである軌跡を対象とする。静止画を対象とした研究と、軌跡を対象とする本研究の差は、以下の 2 点である。

(1) 時系列データでは、時系列属性（速度など）まで利用できる。

(2) 集合写真のような一般的な静止画では、人が高解像度で撮影されているため、年齢や性別などの属性まで推定可能である。本研究が対象にする軌跡は、監視カメラのように人が低解像度撮像されている動画、またはレーザーレンジファインダのように人の見え属性が得られないセンサによって計測される。

グループ検出のように複数人の間の物理的なインタラクションの表現としては、Social Force Model [12] が広く用いられている。このモデルは、人のグループ検出 [1], [5], [6], [15] だけでなく、異常行動検出 [14] のように人を観測対象としたその他の画像認識問題にも利用されている。本研究では、Social Force Model を直接的に利用しないが、類似した時空間特徴量を提案し、その中で従来モデルでは表現できていない複数の

特徴間の共起性まで表現する。こうした特徴間の共起性の有効性は多数研究されており（例：人検出のための濃度勾配の共起性利用 [13]）、本研究でもこの共起性による性能向上を実現する。

特徴量だけでなく、識別器もグループ検出にとって重要な要素である。例えば、階層クラスタリングや条件付確率場を利用して、シーン中の全員を同時にグループ化する手法が、それぞれ文献 [16] と文献 [17] で提案されている。文献 [18] では、シーン中の任意の 2 人の間の物理的な関係を表現した特徴量を、識別的分類器（サポートベクターマシンを利用）を利用して同一グループ内のペアと同一グループに属さないペアに分類することによって、グループを検出している。本研究でも、後者のように、任意の 2 人が同一グループに属しているかどうかを判定するアプローチをとるが、提案する時空間特徴量そのものは全体最適化を行う前者のアプローチにも適用可能である。

3. 軌跡の時空間特徴量によるグループ検出

グループ検出では、同時に観測範囲内に存在する軌跡集合における全ての軌跡ペアにおいて、各ペアの 2 人が同一グループに属するか否かを調べる。この処理を、グループ内ペア判定と呼ぶ。そのため、各軌跡ペアにおいて、軌跡の類似性に着目した時空間特徴量をフレームごとに取得する（ベースとなる従来特徴量を 3.1 で説明し、それを改良した提案特徴量を 3.2 で述べる）。この時空間特徴量を、学習サンプル内のグループ内ペアとそれ以外のペアの軌跡について求め、それらの識別をサポートベクターマシン (SVM) により学習し、未知の軌跡ペアのグループ内ペア判定を可能にする (3.3)。最後に、グループ内ペアの組み合わせから 3 人以上のグループを検出する (3.4)。

3.1 軌跡ペアの時空間特徴量

軌跡のグループ内ペア判定のために、文献 [18] では同一グループに属する人同士が接近して移動することや、軌跡が類似すること、同時に観測領域内に入出することに着目し、五つの時空間特徴量 ($F_1 \sim F_5$) を提案している。歩行者 i と歩行者 j の軌跡間の特徴量は、歩行者 i の位置と速度をそれぞれ \mathbf{p}_i , \mathbf{v}_i , i が観測領域内に存在したフレーム t の集合を $\mathbf{T}_i = [t_{i,1}, t_{i,2}, \dots, t_{i,N_i}]$ (ただし、 N_i は i が観測されたフレーム数) とおくと、 $\mathbf{T}_i \cap \mathbf{T}_j$ の各フレームにおいて以下のように定義される (図 2 参照)。

F_1 2 人の間の距離： $|\mathbf{p}_i - \mathbf{p}_j|$

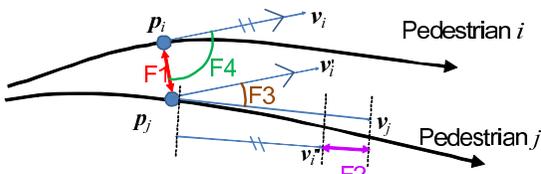


図2 文献[18]で提案されている時空間特徴量。 v_i と v'_i は平行、 v_i と v''_i の長さは等しい
 Fig.2 Spatio-temporal features proposed in [18]. v_i and v'_i are parallel. $\|v_i\|$ and $\|v''_i\|$ are equal.

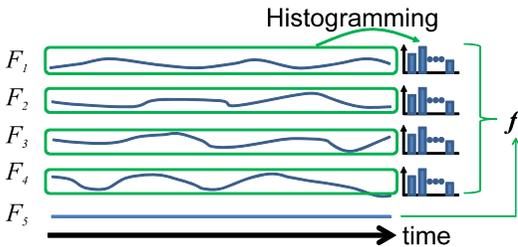


図3 手法[18]のグループ内ペアの検出法
 Fig.3 Pair in the group detection proposed in [18].

F_2 2人の速度の絶対値の差: $\|v_i\| - \|v_j\|$
 F_3 2人の速度の向き差: $|\arctan(v_i) - \arctan(v_j)|$
 F_4 1人の速度の向きと2人の相対位置の向き差: $|\arctan(p_i - p_j) - \arctan(v_i)|$
 F_5 軌跡ペアの同時観測割合: $|T_i \cap T_j| / |T_i \cup T_j|$
 文献[18]では、この五つの時空間特徴量を各フレームにおいて算出し、 $F_1 \sim F_4$ をそれぞれ m ビンに分割されたヒストグラムにして m 次元の特徴ベクトルで表す(図3中のHistogramming)。この特徴ベクトルと F_5 を連結した $4m + 1$ 次元の特徴ベクトル f を、各軌跡ペアに対して求める(図3参照)。

特徴ベクトル f をグループ内ペアとグループ外ペアに分類する識別器の学習時には、それぞれの軌跡ペアの f を正例、負例として2クラス識別器を得る。手法[18]ではSVMを識別器として利用した。

しかし、この時空間特徴量には次の問題点があった。

(1) $F_1 \sim F_4$ をそれぞれ独立にヒストグラム化し、それらを連結するため(図3)、各フレームにおける $F_1 \sim F_4$ 間の共起性(例:グループ内ペアでも、距離 F_1 が小さすぎるフレームでは速度が遅くなり、速度の差 F_2 が更に小さくなりやすい)が表現できない。

(2) 計測期間全体で F_5 を計算するため、遮蔽などによりグループ内ペアの追跡軌跡が途切れてしまった場合、 T_i と T_j の重複が大きく変化してしまう。

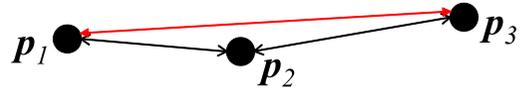


図4 3人以上のグループにおけるグループ内ペア間の距離。赤線で示した p_1 と p_3 間の距離が大きい場合、この2人はグループ外ペアと判定され易い
 Fig.4 Distance between pairs in a group with three or more people. Since p_1 and p_3 (indicated by a red line) are far from each other, they might be regarded as people in different groups.

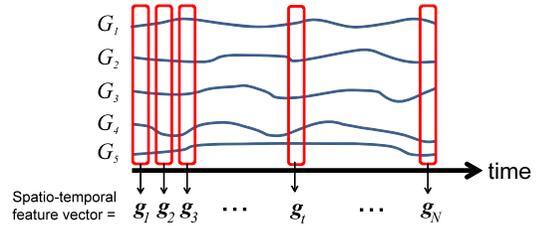


図5 各フレームにおける時空間特徴量の連結による時空間特徴ベクトル g_1, \dots, g_N の獲得
 Fig.5 Spatio-temporal vectors with spatio-temporal features at each frame.

(3) F_4 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、ペアのいずれの速度を使用するかによって F_4 が変動する。

(4) 一般的にグループ内ペアの F_1 は小さくなるのが期待されるが、3人以上のグループにおいて、距離の離れた軌跡ペア(例:図4の赤矢印のペア)の F_1 が大きくなってしまい、グループ外ペアの F_1 との識別が難しくなる。

(5) 静止またはゆっくり進む人の速度方向の計測が安定しないため、 F_3 や F_4 も大きく変動する。

3.2 時空間特徴量の識別性の向上

3.2で述べた手法[18]の問題点のそれぞれを解決するため、本研究では以下の時空間特徴量を提案する。

(1) 各フレームにおける時空間特徴量間の共起性を表現するために、各フレームで独立に時空間特徴量を連結して一つの時空間特徴ベクトルを得る(図5中の g_1, \dots, g_N)。 g_k はフレーム k における特徴ベクトルであり、 N はペアが同時観測されたフレームの総数。各連結ベクトル g_k は、全フレームのヒストグラムを連結する手法[18]の特徴量 f と異なり、 k フレームにおける時空間特徴量間の共起性を表している。

(2) F_5 にも時間的局所性をもたせつつ、他特徴量 $F_1 \sim F_4$ との共起性を評価するため、 k フレームにおける F_5 は、 $k - T_r$ から $k + T_r$ フレームにおける

同時観測割合として、 F_1 から F_5 を全て連結して一つの特徴ベクトル \mathbf{g} とする。

(3) $\bar{\mathbf{v}} = (\mathbf{v}_i + \mathbf{v}_j)/2$ をペアの進行方向とみなし、 $|\arctan(\mathbf{p}_i - \mathbf{p}_j) - \arctan(\bar{\mathbf{v}})|$ を F_4 とする。

(4) 3人以上のグループでは、各人に対して最も近距離にいるペアの特徴量のみを学習させる。図4の例の場合、 $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ の最近傍は、それぞれ $\mathbf{p}_2, \mathbf{p}_1, \mathbf{p}_2$ なので、図中で黒矢印の \mathbf{p}_1 と $\mathbf{p}_2, \mathbf{p}_2$ と \mathbf{p}_3 の2ペアが正例として学習される。グループ内ペア判定時には、全ての任意のペア間の特徴ベクトルに対して識別器を適用し、あるペアがグループ外と判定されても、そのペアの2人が共に他のある人とグループ内ペアと判定されている場合、この3人は同一グループに属すると判定する。図4の例の場合、 \mathbf{p}_1 と \mathbf{p}_3 にいる2人が識別器ではグループ外と判定されても、共に \mathbf{p}_2 にいる人とグループ内ペアと判定された場合、 \mathbf{p}_1 と \mathbf{p}_3 にいる2人もグループ内ペアと判定する。

(5) あるフレームにおける速度がしきい値 T_v 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判定を行わない。

以上を整理すると、提案する時空間特徴量 $G_1 \sim G_5$ 及び時空間特徴ベクトル \mathbf{g} は、 $\mathbf{T}_i \cap \mathbf{T}_j$ 中、 $|\mathbf{v}_i| > T_v$ かつ $|\mathbf{v}_j| > T_v$ の条件を満たす各フレームにおいて、以下のように定義される。

- G_1 2人の間の距離： $|\mathbf{p}_i - \mathbf{p}_j|$
 - G_2 2人の速度の絶対値の差： $||\mathbf{v}_i| - |\mathbf{v}_j||$
 - G_3 2人の速度の向き差： $|\arctan(\mathbf{v}_i) - \arctan(\mathbf{v}_j)|$
 - G_4 2人の速度の向きの平均と2人の相対位置の向きの差： $|\arctan(\mathbf{p}_i - \mathbf{p}_j) - \arctan(\bar{\mathbf{v}})|$
 - G_5 各 k フレームの前後 T_r フレームにおける軌跡ペアの同時観測割合： $|\mathbf{T}'_i \cap \mathbf{T}'_j| / |\mathbf{T}'_i \cup \mathbf{T}'_j|$ 。 $\mathbf{T}'_i = \mathbf{T}_i \cap \mathbf{T}''_k$ 、ただし $\mathbf{T}''_k = [k - T_r, \dots, k + T_r]$ である。
- \mathbf{g} 上記の時空間特徴量 $G_1 \sim G_5$ を連結して、各フレームにおける時空間特徴ベクトルを得る： $\mathbf{g} = [G_1, G_2, G_3, G_4, G_5]$

3.3 時空間特徴量を用いたグループ内ペア検出

前節で定義した時空間特徴ベクトルを用いて、任意の軌跡ペアからグループ内ペアを検出する。各軌跡ペアにおいて、そのペアが同時観測された全フレームの時空間特徴ベクトルを用いてグループ内ペアを検出する手法として、下記2種類の検出法を提案する。

フレーム別判定：図6に示すように、まず各フレームで独立にグループ内ペアの判定を行った後、グループ内ペアと判定されたフレーム数の割合がしきい値以

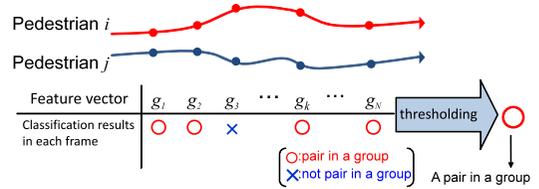


図6 フレーム別判定によるグループ内ペア検出
Fig.6 In-group pair detection by frame-by-frame classification.

上^(注1)となった軌跡ペアをグループ内ペアとして検出する。学習時には、グループ内ペアとそれ以外のペアの各フレームにおける時空間特徴ベクトルを、それぞれ正例と負例としてSVMによって学習しておく。時空間特徴ベクトルの識別をフレームごとで行うため、フレーム数しきい値よりも少ないフレームにおいて、グループ内ペアの軌跡の挙動が一般的なグループ内ペアの挙動から大きく外れる識別不可能な動きでも、そのペアを正しくグループ内ペアとして検出できる。一方、入力軌跡の学習サンプルからの乖離度が小さくとも、すなわち入力特徴ベクトルが識別境界付近に存在する識別困難なフレームでも、このようなフレーム数がフレーム数しきい値以上の場合、識別は難しくなる。Bag-of-Features (BoF) 判定：フレーム別判定のように各フレームにおけるグループ内外の判定を行わず、軌跡全体における時空間特徴ベクトルの分布(ヒストグラム)からグループ検出を行う。学習時には、まず学習データ内の全軌跡ペア、全フレームにおける時空間特徴ベクトルをクラスタリングする。次に、各軌跡ペアについて、同一クラスタに属する特徴ベクトルを一つのビンに投票してヒストグラムを生成する。そのヒストグラムを軌跡ペアの特徴ベクトルとして、SVMでグループ内ペアとグループ外ペアを学習する。検出時も、任意の軌跡ペアの全フレームから時空間特徴ベクトルを抽出し、それらを学習済みのクラスタに分類してヒストグラム化し、特徴ベクトルを得て、SVMで識別する。ただし、各ヒストグラムは、各軌跡ペアにおけるフレーム数で正規化される。この判定法では、ヒストグラム生成時に類似ベクトルは一つのビンに投票されるため、グループ内外の識別が困難、すなわち特徴ベクトルが識別境界付近に存在する類似フレームは同一ビンに投票されやすい。その結果、このような識別困難なフレームにおける微妙な変動は、グループ

(注1)：本論文の実験では、予備実験の結果からこのしきい値を66%とした。

内ペアの判定に大きな影響を与えない。また、SVMが「全フレームの特微量のヒストグラム」に対して適用されるため、正しい判別のためには全フレームに渡る類似フレームの出現頻度がマッチする学習データが必要となる（ヒストグラムなので、類似フレームの観測順序は問わない）。

以上の2手法の性質をまとめると以下ようになる。
 識別不可能なフレーム数： 識別不可能フレームが少なく、フレーム別判定は成功する。BoF判定では、ヒストグラムが変化して判別に悪影響が及ぶ。

識別困難なフレーム数： 識別困難フレームが多いと、フレーム別判定は失敗し、BoF判定には大きく影響しない。

観測期間全体における動き： BoF判定では全フレームに渡り類似した学習データが必要だが、フレーム別判定ではフレーム独立な類似性のみで充分。

3.4 構成員が3人以上のグループ検出

3.3の手法のいずれかで、全フレームの特微量から各ペアのグループ内判定を行った後、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。このペア構成員間の重複チェックにより、SVMによるグループ内ペア検出済みの3人以上のグループをまとめ上げるだけでなく、SVMでは検出できなかったグループ内ペアも検出される（例：図4中の p_1 と p_3 のように空間距離的に離れたグループ内ペア）。

4. LRF計測によるグループ検出実験

LRFで計測された軌跡データによる実験を行った。実験に用いたAPTデータセット[19]は、図1(a)に示したショッピングモールで計測されており、1時間の計測時間に54のペアを含んだ392人の軌跡を記録している。データセットには、全フレームにおける各歩行者の床平面座標での位置と、同一グループに属する全てのペアの歩行者IDが記録されている。ただし、追跡は手法[21]によって自動で行われているため、計測軌跡には歩行者IDの誤った切り替わりや消滅などの誤りも含まれる。一方、ペアの歩行者IDは、アンテーション用に別途撮影されたビデオ（図1(a)）とLRFの座標データとを手で比較しながら与えたため、追跡失敗に影響を受けず正しいペアデータになっている。各歩行者は、様々な方向に進みながら、時に立ち止まって陳列物を眺めたり、雑談をしている。こ

表1 正検出 (TP) と誤検出 (FP) の割合の比較実験結果
 Table 1 Percentages of true-positives (TP) and false-negatives (FP) by different four methods.

	TP	FP
M1 ([5])	79.5	NA
M2 ([18])	90.7	9.3
M3 (Frame-by-frame proposed feature)	94.4	7.4
M4 (BoF proposed feature)	96.2	11.5

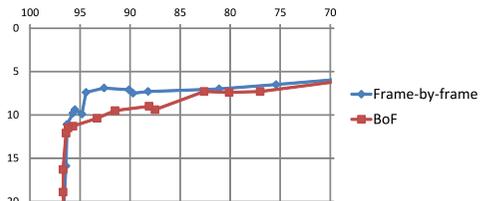


図7 提案手法 (M3 と M4) におけるパラメータ調整に応じた正検出率 (横軸) と誤検出率 (縦軸) の変化
 Fig. 7 ROC curves of two proposed methods depending on two kinds of parameters.

うした歩行者の振る舞いは、その軌跡に基づいた人のグループ検出を難しくしている。データセット中のLRFの時間解像度は、約50fpsである。

比較のため、下記4手法を評価した。

M1: 手法[5]。表1の結果は文献中の数値。

M2: 手法[18]。

M3: 提案した時空間特微量とフレーム別判定。

M4: 提案した時空間特微量とBoF判定。

手法M2, M3, M4の評価では、各試行において全軌跡データからランダムに選ばれた10%を学習に用い、15試行の平均を表1に示した。SVMにはlibsvm[20]を利用した。表中の正検出 (TP) と誤検出 (FP) は、それぞれ「ペアを正しくペアとして検出した割合」と「ペアでない2人組を誤ってペアとして検出した割合 (母数は正しいペアの数)」である。M2には可変パラメータは存在しないが、M3とM4には T_r と T_v の2種類のパラメータがある。 T_r と T_v の幾つかの組み合わせでグループ検出実験を行った結果（図7）の中からTP-FPが最大となる結果を表1に示している。

表1から、提案手法が従来法[5],[18]と比べて良好な結果を得ていることが確認できる。特に、フレーム別判定を採用した手法M3は、正検出、誤検出ともに従来法[18]よりも良い結果を示した。

図8, 9, 10に、それぞれ正検出、誤検出、未検出のペアにおける軌跡（各図中の左の図）、及びフレームごとの時空間特微量とグループ内ペア判定結果（各図中の右のグラフ）を示す。グラフ中で黄色の線で示

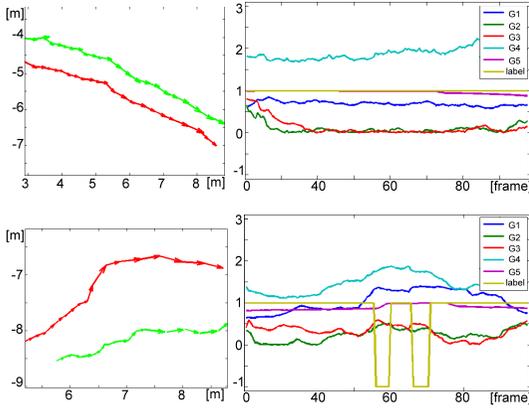


図 8 グループとして正検出したペアの時空間特徴量の例。左図の縦横軸は、床平面座標を示しており、緑と赤の矢印が 2 名それぞれの軌跡を示し、各矢印は各フレームにおける歩行者の位置と移動方向を示す。右のグラフは、 G_1 から G_5 、及びグループ内ペアの判定結果の時系列データを示している。縦軸、横軸がそれぞれ特徴量の値、フレーム ID を示している

Fig. 8 Examples of spatio-temporal features: true-positives. In the left-hand figure, green and red arrows depict the trajectories of two pedestrians. Each arrow depicts the position and velocity of a pedestrian at each frame. The right-hand graph shows the temporal histories of the spatio-temporal features as well as those of classification (denoted by “label”). Vertical and horizontal axes indicate the values of the features and frame IDs, respectively.

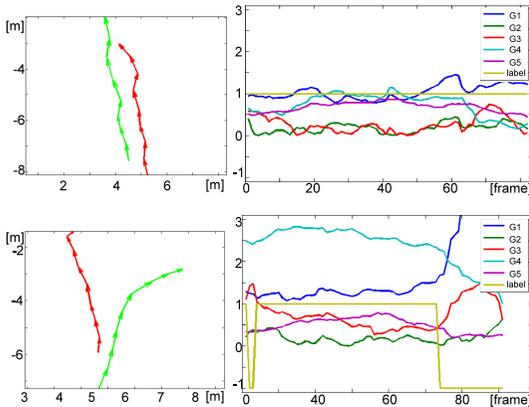


図 9 誤検出されたグループ外ペアの時空間特徴量の例
Fig. 9 Spatio-temporal features of false-positive.

されているグループ内ペア判定の結果は、1 (グループ内と判定) か -1 (グループ外と判定) のいずれかである。いずれの例も、ペアはほぼ同じ期間観測されている。すなわち、 G_5 はペアがグループ内ペアであ

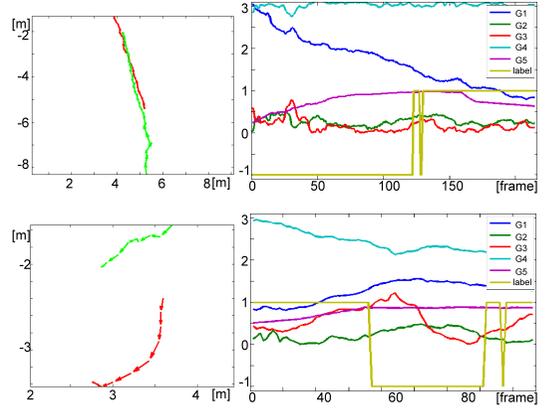


図 10 未検出だったグループ内ペアの時空間特徴量の例
Fig. 10 Spatio-temporal features of false-negative.

ることを示している^(注2)。

図 9 では、上下いずれの例においても、 G_2 と G_3 がともに小さく、 G_1 も比較的小さくなっているフレーム、すなわち偶然近づいて並び歩いたフレームにおいて、グループ内ペアとして判定されてしまったため、最終的に誤検出されてしまった。図 10 では、上の例は、 G_1 が大きくなってしまった、すなわちペアが大きく離れてしまったフレームでグループ外ペアとして判定され、その結果グループ内ペアとして検出できなかった。下の例では、 G_1 が比較的大きく、かつ G_3 が大きい、すなわちペアが異なる方向に進んでいるフレームにおいて、グループ内ペアとして判定できなかった。

図 8, 9, 10 は各特徴量の大小をおおまかに示すための例であり、具体的に各時空間特徴量が、どのような値をとったときにグループ内ペア及びグループ外ペアに判定されるかは、観測シーンの構造 (例: 広い通路か狭い通路か) や属性 (例: 混み合い早足になりがちな駅構内か、ゆったり歩ける広場か) などに依存する。このような場所や時間に応じた時空間特徴量の特性は、本研究では SVM によって学習データから自動的にモデル化される。

図 11 は、3 人以上のグループにおけるグループ化の典型的な例を示している。この例は 3 人グループであるが、その 3 名が横並びに歩いているため、ID1 と ID3 の 2 名間の距離が大きい。そのため、フレーム別判定でも BoF 判定でも、この 2 名はグループ外ペア

(注2) : ほぼ自明だが、ペアが同時に観測されない、すなわち G_5 が小さい場合、確実にグループ外ペアと判定されることは実験的に確認済。

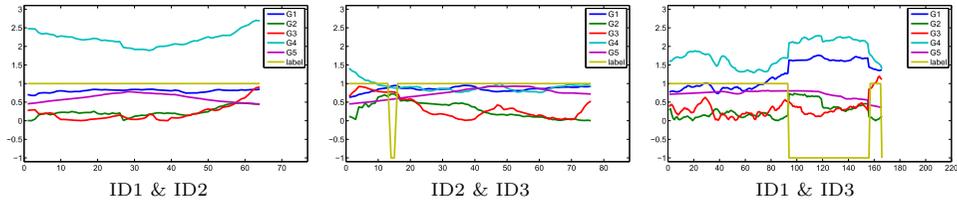


図 11 3人以上のグループにおける時空間特徴量の例. SVMによるグループ内ペア検出により「ID1とID2」「ID2とID3」が検出され、その結果のペア構成員間の重複チェックにより、「ID1とID3」のペア検出に成功している

Fig. 11 Spatio-temporal features in a large group. True-positive detections of “ID1 & ID2” and “ID2 & ID3” were propagated to “ID1 & ID3”.

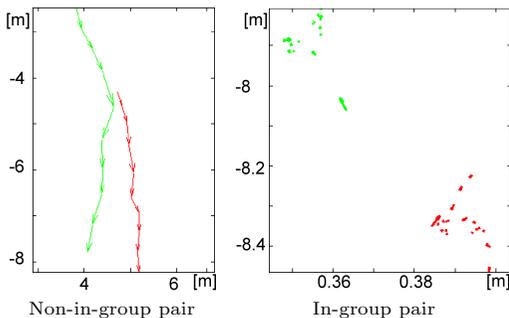


図 12 提案手法による手法 [18] からの検出結果向上の例. 左右の例は、それぞれ手法 [18] ではグループ内ペアとして誤検出及び未検出であった

Fig. 12 Two examples of correctly classified pairs, which are incorrectly classified by the previous method [18].

と判定されてしまった (図 11 はフレーム別判定の結果を示している). しかし、「ID1とID2」及び「ID2とID3」はともにグループ内ペアとして検出されたため、ペア構成員間の重複チェックにより「ID1とID3」の2名もグループ内ペアとして最終的に検出されている.

最後に、提案した時空間特徴量により、従来特徴量 [18] での検出失敗を訂正できた例を、図 12 に示す. 左の例では、無関係の2人が斜め前方後方にずれた状態で同方向に向かって歩いているため、手法 [18] では誤検出が起きていた. 提案手法では各瞬間における G_1, \dots, G_5 の共起性を評価するように特徴ベクトル $g = [G_1, \dots, G_5]$ を判定に用いている. その結果、同方向に同速度で進んでいても (すなわち G_2 と G_3 がグループ内ペアの特徴を示しているも)、それと同時に、2人の距離が多少離れており、また並びが斜めになっている (すなわち G_1 と G_4 がグループ外ペアの特徴を示している) 状態が負例として学習されていることによって、正しく「グループ外ペア」として判定できている. 右の例では、グループ内ペアがほぼ

静止した状態で近接しているため、LRFによる位置計測に失敗し、特に進行方向が乱れてしまった結果、手法 [18] では未検出になってしまっている. これに対し、提案手法では歩行者速度のしきい値処理によりこうした不安定な状態を評価していないため、その他のフレームの評価によって正しくグループ検出できた.

5. む す び

複数人の軌跡データからグループ検出を行う手法を提案した. グループ検出に利用する時空間特徴量は、歩行者間の相対的な位置や速度の関係を表現している. この時空間特徴量の時系列データをグループ内ペアとグループ外ペアに識別することにより、グループ検出を行う. データセットを用いた実験により、類似従来法に対する提案手法の検出精度向上を確認した.

以下に今後の予定を挙げる.

- カメラ画像中の軌跡に対する提案手法の適用.
- 軌跡に加えて画像特徴量 (例: 頭部や胴体の方位推定結果) も併用したより正確なグループ検出.
- 得られたグループ検出結果を利用した応用システム (例: 1. で述べたナビゲーションなど).

文 献

- [1] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, “You’ll never walk alone: Modeling social behavior for multi-target tracking,” ICCV, 2009.
- [2] H.B. Shitrit, J. Berclaz, F. Fleuret, and P. Fua, “Tracking multiple people under global appearance constraints,” ICCV, 2011.
- [3] O.M. Mozos, R. Kurazume, and T. Hasegawa, “Multi-part people detection using 2D range data,” Int. J. Social Robotics, vol.2, no.1, pp.31–40, 2010.
- [4] J. Cui, H. Zhao, and R. Shibasaki, “Fusion of detection and matching based approaches for laser based multiple people tracking,” CVPR, 2006.
- [5] Z. Yucl, T. Ikeda, T. Miyashita, and N. Hagita, “Identification of mobile entities based on trajectory

and shape information,” IROS, 2011.

- [6] Z. Yucel, F. Zanlungo, T. Ikeda, T. Miyashita, and N. Hagita, “Deciphering the crowd: Modeling and identification of pedestrian group motion,” *Sensors*, vol.13, no.1, pp.875–897, 2013.
- [7] W. Hu, T. Tan, L. Wang, and S. Maybank, “A survey on visual surveillance of object motion and behaviors,” *IEEE Trans. Syst. Man Cybern., C*, vol.34, no.3, pp.334–352, 2004.
- [8] A. Yilmaz, O. Javed, and M. Shah, “Object tracking: A survey,” *ACM Comput. Surv.*, vol.38, no.4, 2006.
- [9] C.H. Cooley, *Social Organization: A Study of the Larger Mind*, pp.25–31, Charles Scribner’s Sons, 1909.
- [10] A.C. Gallagher and T. Chen, “Understanding images of groups of people,” *CVPR*, 2009.
- [11] G. Wang, A. Gallagher, J. Luo, and D. Forsyth, “Seeing people in social context: Recognizing people and social relationships,” *ECCV*, 2010.
- [12] D. Helbing and P. Molnar, “Social force model for pedestrian dynamics,” *Phys. Rev. E*, vol.51, no.5, pp.4282–4286, 1995.
- [13] T. Watanabe, S. Ito, and K. Yokoi, “Co-occurrence histograms of oriented gradients for pedestrian detection,” *PSIVT*, 2009.
- [14] R. Mehran, A. Oyama, and M. Shah, “Abnormal crowd behavior detection using social force model,” *CVPR*, 2009.
- [15] P. Scovanner and M.F. Tappen, “Learning pedestrian dynamics from the real world,” *ICCV*, 2009.
- [16] 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.
- [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] APT pedestrian behavior analysis, <http://www.irc.atr.jp/~zeynep/research>
- [20] 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.
- [21] D.F. Glas, T. Miyashita, H. Ishiguro, and N. Hagita, “Laser-based tracking of human position and orientation using parametric shape modeling,” *Advanced Robotics*, vol.23, no.4, pp.405–428, 2009.

(平成 25 年 4 月 17 日受付, 6 月 25 日再受付)



守口 裕介

2013 奈良先端科学技術大学院大学修士課程了。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。



岡田亜沙美 (学生員)

2012 奈良先端科学技術大学院大学修士課程入学。現在、人群解析に関する研究に従事。



浮田 宗伯 (正員：シニア会員)

2001 京都大学大学院博士後期課程了。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 同准教授。2002～2006 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。2007～2009 までカーネギーメロン大学客員研究員兼任。2011 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, 分散協調視覚, 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 慶應義塾大学大学院工学研究科電気工学専攻修士課程了。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。工博。IEEE, 情報処理学会, 日本ロボット学会, 人工知能学会各会員。