

# 歩容特徴と局所的な見えを用いた 特徴点軌跡クラスタリングによる混雑環境下人物追跡

杉村 大輔<sup>†</sup> 木谷クリス 真実<sup>††</sup> 岡部 孝弘<sup>†</sup> 佐藤 洋一<sup>†</sup> 杉本 晃宏<sup>†††</sup>

<sup>†</sup> 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

<sup>††</sup> 電気通信大学大学院情報システム学研究科 〒182-8585 東京都調布市調布ヶ丘 1-5-1

<sup>†††</sup> 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: <sup>†</sup>{sugimura,takahiro,ysato}@iis.u-tokyo.ac.jp, <sup>††</sup>kitani@is.uec.ac.jp, <sup>†††</sup>sugimoto@nii.ac.jp

あらまし 本稿では、特徴点軌跡のクラスタリングに基づいた人物追跡手法を提案する。混雑環境下において頑健に個々の人物を区別するために、本手法では歩容特徴と局所的な見えの時間変動の一貫性という二つの指標を追跡の枠組みへ導入する。周波数空間における歩容特徴は、生体認証の分野において頻繁に利用されている指標であり、個人を識別するための重要な手掛かりであることが知られている。また、局所領域における見えの時間的な変化は、人物の動きが周りと類似する傾向のある混雑環境下において個々の人物を区別するための効果的な指標となる。このような動きと見えの異なる種類の指標を利用することにより、混雑環境下においても安定な追跡を実現することが可能となる。実環境における実験により本手法の有効性を確認した。

キーワード 人物追跡, 特徴点軌跡のクラスタリング, 歩容特徴, 局所的見え, 混雑環境

## Tracking people in crowds based on clustering feature trajectories using gait features and local appearances

Daisuke SUGIMURA<sup>†</sup>, Kris M. KITANI<sup>††</sup>, Takahiro OKABE<sup>†</sup>, Yoichi SATO<sup>†</sup>, and Akihiro SUGIMOTO<sup>†††</sup>

<sup>†</sup> Institute of Industrial Science, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

<sup>††</sup> Graduate School of Information Systems, The University of Electro-Communications 1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585 Japan

<sup>†††</sup> National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

E-mail: <sup>†</sup>{sugimura,takahiro,ysato}@iis.u-tokyo.ac.jp, <sup>††</sup>kitani@is.uec.ac.jp, <sup>†††</sup>sugimoto@nii.ac.jp

**Abstract** We propose a method for tracking individuals in crowds. Our method is based on a trajectory-based clustering approach that groups trajectories of image features that belong to the same person. The key novelty of our method is to make use of a person's individuality, that is, the gait features and the temporal consistency of local appearance to track unique individuals in a crowd. Gait features in the frequency domain have been shown to be an effective biometric cue in discriminating between individuals. In addition, we incorporate the temporal consistency of local patch appearance to differentiate trajectories of multiple people moving in close proximity. Our experiments show that the use of gait features and the temporal consistency of local appearance contribute to significant performance improvement in tracking people in crowded scenes.

**Key words** tracking people, clustering feature trajectories, gait feature, local appearance, crowded scene

### 1. はじめに

近年、朝夕のラッシュ時の駅構内やイベント会場などの混雑環境下(図1)における人物追跡への関心が高まっている。事故や事件の未然防止といったセキュリティ分

野をはじめ、視野内の人数の自動計測、混雑解消のための情報源としての利用など、様々な分野への応用が期待される。

しかしながら、実際の混雑環境において頑健な追跡を実現するためには、頻繁に発生する遮蔽、非常に近接し

ている複数の人物個々の識別といった様々な問題に対処する必要がある。そのため、背景差分に基づく人物の形状モデルを利用した検出手法や見えを手掛かりにした時系列フィルタリング手法は、混雑環境における追跡に不向きであると考えられる。

近年、このような形状や見えを用いた追跡手法とは異なった、画像中の特徴点の動きを利用した手法が提案されている [1-3]。これらの手法は、同じ人物に属している特徴点は類似した動きをするという仮定に基づき、特徴点軌跡群を動きの類似性と位置関係に基づきクラスタリングすることにより人物追跡を実現する。これらの手法は、対象の一部が観測されていれば追跡可能であるため、一般に部分遮蔽に頑健であるといえる。

しかしながら、混雑した状況下では、群集全体により形成される大きな動きの流れにより人物個々の動きは制限されやすい。これにともない、隣接する人物と動きが類似する傾向にあるため、軌跡の動きの類似性と位置関係に基づいた手法では個々の人物を区別することは難しい。

このような問題に対処するために、本稿では、(1) 周波数領域における歩容特徴、(2) 局所領域における見えの時間変動という二つの異なる種類の指標を導入する。これらの特徴点軌跡のクラスタリング手法の考え方をを用いて追跡の枠組みへ組み入れることにより、混雑シーンにおける頑健な人物追跡を実現することが可能となる。

周波数領域における歩容特徴の利用は、群集における個々の人物の識別に非常に有効であると考えられる。互いに近接した複数の人物が同じ方向に同じ速度で歩いている場合においても、各人物の歩幅や歩行周期は異なるため、その周波数特徴には各人物固有の性質が現われる。このような個人特性はその人物に属する特徴点の動きにも現われると考えられることから、本手法では特徴点軌跡の動きの周波数特性を個々の人物の識別に利用する。歩容特徴は個人認証のための重要な手掛かりとしてバイオメトリクスの分野で広く利用されており [4]、これは群集中の人物追跡においても歩容特徴は有用な指標になることを示唆している。

局所領域における見えの時間的な変化を利用することは、動きの周波数特徴さえも類似する場合において効果的に働くことが期待される。本手法では、三つの特徴点の軌跡によって張られる三角パッチを考え、その見えの時間変動に着目する。同じ人物に属する特徴点軌跡により張られる三角パッチ内の見えは時間的にほぼ変化しないことが期待されるが、その一方で他の人物との間に張られる三角パッチは背景や他の人物を含むためその時間変動は大きくなると考えられる。また局所領域のみに注目しているため、追跡対象全体の見えを必要としない。そのため部分的な遮蔽について頑健である。このように、動きと見えの異なる種類の指標を利用することにより、混雑環境においても頑健な人物追跡を実現することが可



図 1 混雑環境の例

能になる。

本稿の構成は以下のとおりである。まず、2 節において本手法に関連する先行研究についてまとめ、本手法の位置づけおよび新規性について整理する。続く 3 節では特徴点軌跡のクラスタリングに基づく本手法の枠組みについて概観する。そして 4 節において歩容特徴、局所領域における見えの変化の二つの指標をどのように追跡の枠組みへ組み入れるのか詳しく述べる。5 節では人工データ、実際の群集の映像を用いた実験を紹介して、本手法の追跡性能、適用範囲について考察する。最後に 6 節で結論と今後の展望について述べる。

## 2. 関連研究

混雑環境下における人物追跡を実現するためにこれまでに様々な手法が提案されている。

一つに、人物の形状をモデル化することにより群集の人物を追跡する方法が提案されている。Zhao ら [5] は人物を、頭部、胴体、両脚の 4 つのパーツに分割し、それぞれ楕円体で表現している。これにより一部が遮蔽された場合においても安定な追跡を実現している。人物のパーツ毎に検出器を用意し、それぞれを学習することにより部分遮蔽に頑健な検出・追跡を実現する枠組みについてもいくつか提案されている [6-8]。これらの手法は実際の混雑した屋外シーンにおいても頑健な人物の検出が可能となることを示している。Dong らは、予め定義した人物の形状モデルを用いることで、重なり合って観測される群集により作られる一つの大きな前景領域から個々の人物の検出を実現している [9]。Khan らはホモグラフィを手掛かりに予め定義した人物モデルを前景領域にあてはめることで人物検出を実現している [10]。しかしながら、図 1 のようなシーンを考慮する場合、前景領域から個々の人物に該当する領域を正しく検出することは難しい。

追跡対象の見えを手掛かりにした、パーティクルフィルタに代表される時系列フィルタを用いた追跡手法もまた数多く提案されている [11-13]。しかしながら、混雑環境下においては頻繁に発生する遮蔽にともない追跡対象が安定に観測されづらい。そのため尤度評価が不安定になり、追跡性能の低下が引き起こされる。オプティカルフローを用いた追跡手法 [14, 15] もいくつか提案され

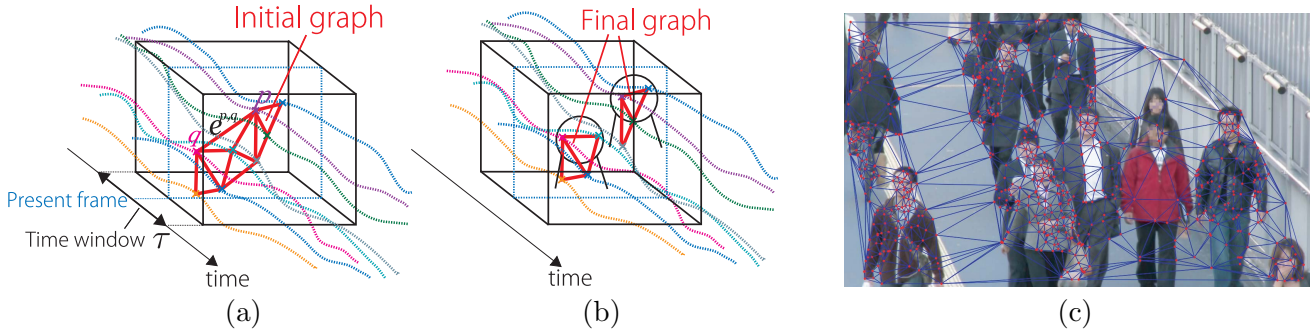


図2 提案手法概要．(a) グラフの構築．時間窓  $\tau$  内における特徴点軌跡群をノードとして用いることにより構成される．(b) クラスタリング結果．得られる特徴点軌跡間相違度に基づきエッジを除去することにより実現される．(c) 実際の映像におけるグラフ構造の例．

ているが，連続するフレーム間での追跡に利用するだけであり，特徴点の動きの軌跡を追跡に利用するアプローチはとられていない．

このような追跡手法とは対照的に，Brostow らはベイズの枠組みのもと画像中の局所特徴点の動きの軌跡をクラスタリングすることにより，混雑環境下において人物追跡を実現している [1]．具体的には，二つの特徴点軌跡のペアの空間関係から事前確率を計算し，その特徴点軌跡のペアの動きの一貫性を計測することにより事後確率が計算される．彼らは混雑した地下鉄の駅において頑健な人物追跡を実現している．Rabaud ら [3] は動きの一貫性の仮定に加え，同じ人物に属する特徴点は同じアフィン運動をするという仮定を導入することにより，屋外の混雑シーンにおいて個々の人物の追跡を実現している．Li らは Brostow らと類似した追跡の枠組みに，動きの一貫性に基づく尤度関数の学習法を組み入れることで追跡の安定化を図っている [2]．しかしながら，混雑した状況においては個々の人物は周りに多数存在する人物により動きを制限され，周りと類似した動き方をする傾向にある．そのため，特徴点軌跡の動きの類似度（動きの一貫性）を考慮するだけでは，個々の人物を識別するためには不十分である．

これに対し，本手法で利用する歩容特徴は個々の人物が持つ固有の特徴であるため，混雑環境においても人物の識別に有効であると考えられる．また局所領域における見えの時間変動（一貫性）を利用することにより，動きが類似する傾向にある混雑状況において個々の人物を区別するための大きな手掛かりとなる．

### 3. 特徴点軌跡のクラスタリングに基づく人物追跡

図2に本手法の概要を示す．本手法は，映像から得られる特徴点の動きの軌跡群を用いてグラフを形成し，それをクラスタリングすることにより人物追跡を実現する．このグラフにおいて各ノードは特徴点軌跡を表し，各ノード間に張られるエッジは特徴点軌跡間の相違度を表す．

最初に，KLT 特徴点追跡法 [16,17] に基づき，画像中の各特徴点の動きの軌跡を獲得する．時間を追うごとに遮蔽により特徴点軌跡の数が減少することが懸念されるため，本手法では [3] において行われていたように毎フレーム新たな特徴点検出を行い，特徴点軌跡の数の減少を防ぐ．信頼度が高い特徴点軌跡を獲得するために，得られた特徴点軌跡群のうち，(1) ほとんど動かないもの，(2) 連続的に変化していない（次フレームにおける特徴点位置が誤って遠い位置に推定される）もの，そして (3) 軌跡の生存時間が短いものを除去する．以上の処理により得られた特徴点軌跡群を用いて，ドロネー三角形分割アルゴリズム [18] により初期グラフを構築する (図2-(a))．実際の映像における初期グラフの例を図2-(c)に示す．

二つのノード（特徴点軌跡）を繋げるエッジは軌跡間の相違度を表す重みを持つ．相違度は，先行研究 ([1-3]) において利用されていた空間関係，動きの一貫性に基づく尺度に加え，(1) 歩容特徴，(2) 局所領域の見えの時間変化の二つの新たな指標を用いることで計算される．それぞれの指標に基づき計算された重み（相違度）をもとに，最終的な二つの軌跡  $p, q$  の間の相違度  $e^{p,q}$  は以下のように計算される．

$$e^{p,q} = e_{freq}^{p,q} \cdot e_{app}^{p,q} \cdot e_{prox}^{p,q} \cdot e_{coh}^{p,q}, \quad (1)$$

ここで (1) 式の右辺はそれぞれ，歩容特徴に基づく重み  $e_{freq}^{p,q}$ ，局所領域の見えの変化に基づく重み  $e_{app}^{p,q}$ ，空間関係に基づく重み  $e_{prox}^{p,q}$ ，そして動きの一貫性に基づく重み  $e_{coh}^{p,q}$  を意味する．これらの指標の詳細は次節で詳しく述べる．もしひとつの指標に基づく重みが0になった場合，他の指標に基づく重みの大きさに拘らず最終的な相違度は常に0になってしまう問題が考えられる．このような状況を避けるために，本手法ではいずれかの重みが0になった場合事前に定めた定数  $e_{min} > 0$  で置き換える．

以上のような指標をもとに，構築された初期グラフはいくつかの部分グラフへクラスタリングされる (図2-(b))．得られたそれぞれの部分グラフは追跡対象それ



それを意味する．クラスタリング処理は，本稿では単純にエッジの重み（相違度）がある閾値  $th_p$  より大きいものを枝打ちをすることで実現する．また本手法ではある有限の時間幅を持つ時間窓  $\tau$  を考え，この窓内における特徴点軌跡群のみを考慮する．時間窓は  $\tau/4$  ずつシフトされる，すなわち  $\tau/4$  時間毎にクラスタリング処理（相違度に基づくエッジの枝打ち処理）が行われる．

#### 4. 軌跡間の相違度の計算

二つの特徴点軌跡（ノード）間のエッジの重みは，歩容特徴，局所的な見えの変化，そして先行研究でも利用されている動きの一貫性，空間関係の4つの指標に基づき計算される．これにより，これらの特徴点軌跡が同じ人物に属しているかどうか判断する．本節ではそれぞれの指標について詳しく説明する．

##### 4.1 周波数領域における歩容特徴

周波数領域における歩容特徴は人物個々が持つ固有の特性であり，混雑環境下において各人物を識別するための非常に有効な指標である．本手法では，人物が歩くときに観測される歩容特徴は人間の身体的な構造から垂直方向の周期的な動きに現われると仮定する．この仮定に基づき，まず特徴点軌跡の動きの垂直方向成分について直線当てはめを行うことにより，垂直方向の周期的な動き  $y_p(t)$  を抽出する．図3に本処理の概略を示す．得られた周期的な動き  $y_p(t)$  について高速フーリエ変換を施すことにより，歩容特徴にあたる振幅スペクトル，位相を獲得する．

歩容特徴に基づく二つの特徴点軌跡の相違度  $e_{freq}^{p,q}$  は，振幅スペクトルと振幅成分が最も含まれている周波数における位相成分の二つの尺度を用いることにより計算される．振幅スペクトル間の相違度は，それぞれの軌跡の振幅スペクトル間のユークリッド距離を計算することにより求められる．この評価尺度は歩容認証に関する研究分野において一般に使われている [4]．また位相に関する相違度については単純に差の絶対値を利用する．これより，歩容特徴に基づく軌跡間の相違度  $e_{freq}^{p,q}$  は次のように計算される．

$$e_{freq}^{p,q} = \sqrt{\sum_{k=0}^{\tau/2} [a_k^p - a_k^q]^2 |\phi^p - \phi^q|}, \quad (2)$$

ここで  $a_k^p$  と  $\phi^p$  は  $p$  番目の特徴点軌跡の  $k$  番目の周波数帯の振幅成分，振幅成分が最も含まれる周波数帯における位相成分をそれぞれ意味する．また周期的な動き  $y_p(t)$  は実数値信号なので，半分の周波数帯のみ考慮すればよい．

##### 4.2 局所領域における見えの時間変動

歩容特徴は混雑状況において個々の人物を特徴づける有効な指標であるが，混雑環境下ではまわりの人物の歩

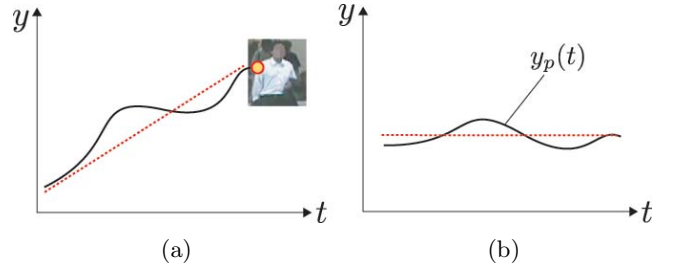


図3 周波数領域における歩容特徴計測のための前処理．(a) 歩容特徴にあたる垂直方向の周期的な動きの抽出．得られる特徴点軌跡に直線当てはめを行うことにより得られる．(b) 抽出された垂直方向の周期的な動き  $y_p(t)$  ．

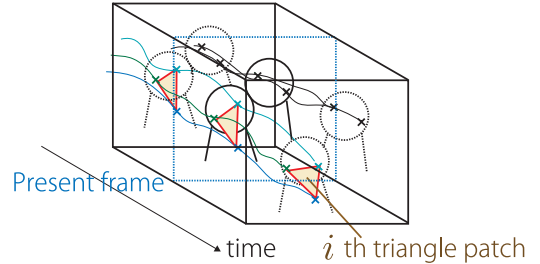


図4 三角パッチの利用．三角パッチは連結されている特徴点軌跡により張られ，パッチ内の見えの時間変動が利用される．

容特徴も類似したものになる傾向にある．このような問題へ対処するためには，特徴点軌跡の動きに基づいたものとは異なる種類の指標が必要となる．そこで本手法では，局所領域における見えの時間的な変動を手掛かりとして利用する．

本手法では，初期グラフにおいて連結された3つの特徴点軌跡によって張られる三角パッチを考え，そのパッチ内の見えの時間変動を利用する．図4に概略図を示す．同じ人物に属する特徴点軌跡により張られる三角パッチ内の見えは時間的にほぼ変化しないことが期待されるが，それに比べて他の人物との間に張られる三角パッチは，背景や他の人物を含むためその時間変動は大きくなると考えられる．これにより動きが類似する傾向にある混雑環境においても個々の人物を識別することが可能となる．

見えに関する尺度は，色ヒストグラムをはじめ，テクスチャや勾配ヒストグラムなど様々なものが考えられるが，本稿では簡単のため HS 色ヒストグラムの時間変動を考える．具体的には，時間窓  $\tau$  内の各時刻の色ヒストグラム  $\mathbf{h}(t)$  と，その平均色ヒストグラム  $\bar{\mathbf{h}}$  の Bhattacharyya 距離の RMS を計算し，これを見えに基づく相違度  $e_{app}^{p,q}$  とする，すなわち，

$$e_{app}^{p,q} = \sqrt{\frac{1}{\tau} \sum_{t=t_s}^{t_s+\tau-1} d_{hist}^2(\mathbf{h}(t), \bar{\mathbf{h}})}. \quad (3)$$

ここで  $t_s$  は時間窓の最初のフレームを意味し， $d_{hist}(\cdot)$  は Bhattacharyya 距離を意味する．得られた相違度  $e_{app}^{p,q}$  は，三角パッチを構成するすべてのエッジに割り当てら

れる．またエッジは隣り合う二つの三角パッチにより共有される可能性があるが，この場合は二つの三角パッチのうち相違度が大きい方を割り当てるものとする．

### 4.3 空間関係と動きの一貫性

先行研究 [1-3] において利用されているように，本手法でも同様に特徴点軌跡間の空間関係，動きの一貫性に基づいた相違度を評価する．もし二つの特徴点軌跡が同じ人物に属しているならば，それらは空間的に近い距離にあり（空間関係），時間窓内において同じ距離関係を保ちやすい（動きの一貫性）と考えられる．そこで，空間関係に基づく相違度  $e_{prox}^{p,q}$  として本手法では時間窓内における二つの軌跡間で最大となる距離を利用する．また時間窓内における二つの軌跡間の距離関係の標準偏差を，動きの一貫性に基づく相違度  $e_{coh}^{p,q}$  として利用する．

## 5. 実験・考察

本手法の有用性を確認するために，人工データと実際の群集の映像を用いてそれぞれ実験を行った．実験に用いた環境は，OS: Windows XP, CPU: Intel Core2 Quad 2.66GHz, RAM: 3.00GB である．効果の比較をするために，歩容特徴と局所の見えの変化に基づく指標を除いたものを基準のシステムとして用いた．これはすなわちエッジ間相違度を空間関係と動きの一貫性のみ ( $e^{p,q} = e_{prox}^{p,q} \cdot e_{coh}^{p,q}$ ) で計算することを意味する．この基準システムは特に明記されない限り以下の全ての実験において同じものが利用される．

### 5.1 人工データを用いた実験

本手法で提案する歩容特徴，局所の見えの時間変化の二つの指標の有効性を確認するために，近接した二つのターゲットが同じ方向に同じ速度で動いている人工データについて実験を行った．また歩容特徴は人物の歩行における垂直方向の周期的な動きに現れるため，ターゲットの動きに垂直方向の周期的な動きを付加した．

#### 歩容特徴の利用の効果

本実験ではまず歩容特徴の効果を確認した．隣接した二つのターゲットは同じ方向，同じ速度で動くが，垂直方向の周期的動き（歩容特徴にあたるもの）の周波数，周期は互いに異なるものとした．また歩容特徴の性能向上への寄与を確認するために，本実験では局所領域における見えの変化の指標は利用しないものとした．

図 5 に比較結果の一部を示す．基準としたシステム（図 5-(a)）では，近接する二つのターゲットを一つの大きなターゲットとみなして誤追跡してしまっていることがわかる．それに対して図 5-(b) では，正しく二つのターゲットを追跡できていることがわかる．これより歩容特徴の利用が追跡性能の向上に寄与していることが見て取れる．



図 5 歩容特徴の利用の効果：(a) 歩容特徴なしの追跡結果．(b) 歩容特徴を組み入れた追跡結果．

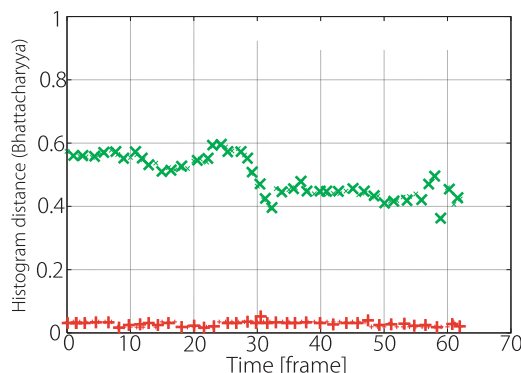


図 6 ヒストグラム距離  $d_{hist}(\mathbf{h}(t), \bar{\mathbf{h}})$  の時間変動結果．赤色の点は一つのターゲット内に張られた三角パッチによるものを意味し，緑色の点は二つのターゲット間に張られた三角パッチによるものを表す．

#### 局所領域における見えの一貫性の利用の効果

混雑環境における個々の人物の識別において最も難しいケース，すなわち動きの周波数特徴についても周りの人物と同調してしまう場合を想定し，二つのターゲットの動きが全く同じケースについて実験した．このようなケースにおいて実験をすることにより，局所領域における見えの変化に基づく指標の効果を確認することができる．そのため本実験において基準とするシステムでは空間関係，動きの一貫性そして歩容特徴の指標に基つきエッジ間相違度 ( $e^{p,q} = e_{freq}^{p,q} \cdot e_{prox}^{p,q} \cdot e_{coh}^{p,q}$ ) を計算した．

実験の結果，基準としたシステムでは前節の実験と同様に一つの大きなターゲットとして誤追跡してしまった．それに対し見えの時間変化を用いた指標を組み入れたものは正しく二つのターゲットを識別することができた．図 6 にある時間窓内における三角パッチ内のヒストグラム距離  $d_{hist}(\mathbf{h}(t), \bar{\mathbf{h}})$  のばらつきを示す．赤色でプロットされた点群は一つのターゲットに属した三角パッチのヒストグラム距離の結果，緑色の点群は二つのターゲット間に張られたもの，すなわち背景を含むパッチの結果を意味する．この結果からわかるように，背景を含む三角パッチに比して一つのターゲットに属した三角パッチ内の色ヒストグラムはほぼ一貫していることがわかる．これより見えの時間変動に基づく指標は二つのターゲットを識別するための効果的な指標であるといえる．

### 5.2 実環境下実験

実際の群集の映像を用いて本手法の有用性を確認した．



表 1 実験で利用した映像のパラメータ.

	解像度	fps	総フレーム数
(A)	1280x720	60	700
(B)	1280x720	60	730
(C)	800x600	30	300
UCSD data [3]	720x480	30	1000

表 2 実環境下実験に利用したパラメータ. 閾値  $th_p$  はエッジ間相違度  $e^{p,q}$  の集合のメディアン値とした. 相違度の集合の獲得のために利用したシーケンスは実験に含まれない.

	$\tau$	提案手法		基準システム	
		$e_{min}$	$th_p$	$e_{min}$	$th_p$
(A)	64	0.1	0.269	0.1	64.759
(B)	64	0.1	1.375	0.1	560.002
(C)	40	0.1	0.675	0.1	66.611
UCSD data [3]	40	0.1	1.102	0.1	207.700

本実験では、それぞれ異なる場所から撮影した 3 つのシーケンス (A), (B), (C) と, UCSD のデータセット [3] を利用した. 表 1 に用いた映像の詳細を示す. また本実験で利用したパラメータを表 2 に示す. エッジの枝打ちのための閾値  $th_p$  は予め獲得したエッジの重み  $e^{p,q}$  の集合から決定した. また時間窓  $\tau$  は、一般的な人物の歩行周期 (およそ 1 秒) を確保するように、データのフレームレートに応じて決定した.

図 7, 8, 11, 12 にそれぞれのシーケンスにおける比較結果を示す. いずれのシーンにおいても、基準としたシステムに対して提案手法の方がより多くの人物を正しく追跡できていることが確認できる.

追跡性能を定量的に評価するために、本稿では再現率と適合率を用いて評価した. また本実験では false positive を、(1) 1 人の人物に複数のクラスタが生成される場合、(2) 複数の人物が 1 人の人物として追跡される場合、の二種類とした. 表 3, 図 9 に各シーケンスにおける追跡結果を示す. いずれのシーケンスにおいても提案手法の方が追跡性能が向上していることがわかる. 主に再現率についてその差が顕著であり、基準としたシステムと本手法にはおよそ 2 倍の性能差が見られる. このことから、歩容特徴と局所領域における見えの時間変動に基づいた指標が追跡性能に大きく寄与しているといえる.

#### 追跡失敗に関する考察

図 10 に頻繁に観測された二種類の追跡失敗例を示す. 図 10-(a) の失敗例は追跡すべき人物が存在しているにも拘らず追跡できていないケースである. このような例は主に追跡対象がテクスチャがほとんどないような服装をしている場合、もしくはカメラに対して後ろ向きである場合に頻繁に観測された. これはクラスタリングすべき特徴点軌跡がほとんど発生していないためである. 図 10-(a) の右側の図は特徴点軌跡を描画したものであるが、明らかに特徴点軌跡の数が不足していることが見て取れ

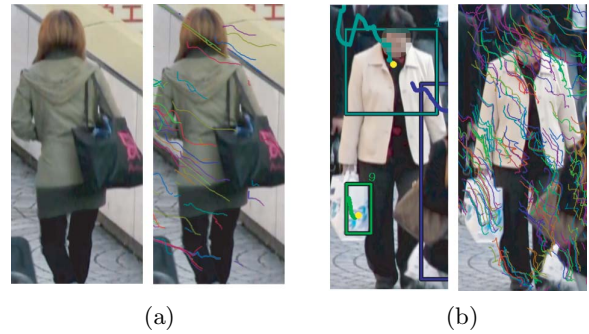


図 10 追跡失敗例. (a) 特徴点軌跡の不足に起因する人物の未検出 (b) 周波数特徴の違いに起因した過検出.

る. またそのような追跡対象に特徴点軌跡が発生しても、上述したような原因からそれが信頼すべきものであるとは言い難いため、結果的に追跡性能が低下することが考えられる.

二つ目の失敗例は、図 10-(b) にあるように一人の人物に複数のクラスタが発生するケースである. このような失敗例において、発生するクラスタは主に人物の体に一つ、そして手、もしくはその人物が持つバッグなどに他のクラスタが現れる. これは、図 10-(b) の右図 (特徴点軌跡の発生の様子) からわかるように、人物の体の動きの周波数特徴とバッグ、もしくは手の動きの周波数特徴が明らかに異なっていることに起因すると考えられる. 本手法は周波数特徴に基づきクラスタリングを行っているため、このような場合、システムは誤って二人の人物が存在するとみなしてしまっていると考えられる.

## 6. おわりに

本稿では、人物の歩容特徴と局所領域における見えの時間変動の一貫性を利用した混雑環境下における人物追跡手法を提案した. 特徴点軌跡のクラスタリングに基づく考え方をを用いて、これらの指標を追跡の枠組みへ組み入れることにより、頑健な人物追跡が可能となる. 人工データ、実際の群衆の映像を用いた実験を通じて、歩容特徴は混雑環境において個々の人物の識別に有効であることを確認した. さらに局所領域の見えの時間変動を計測することにより、動きが類似する傾向にある混雑状況において有効な指標となることを確認した.

実際の群衆の映像を用いた実験において観測された追跡失敗について詳しく考察することにより、現段階における本手法の問題点についてまとめた. 今後は、前述したような人物の各部 (腕や脚など) の動きの特徴のばらつきに起因する追跡失敗について対処する予定である. また、現行のシステムでは同じ人物であるにも拘らず遮蔽前と後で異なる人物とみなして追跡してしまっている. このような人物の対応付けについても今後検討する予定である.

## 謝 辞

本研究の一部は、日本学術振興会特別研究員制度お



図 7 シーケンス (A) における追跡の様子: 左図: 提案手法, 右図: 基準システム

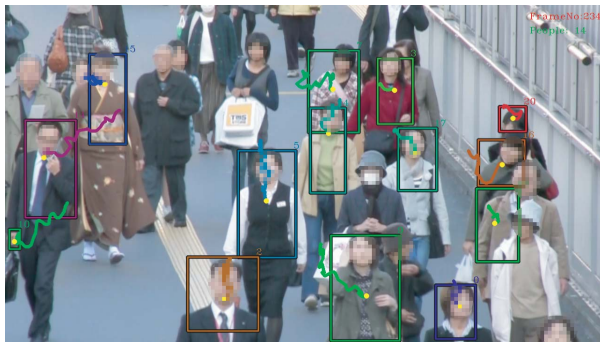


図 8 シーケンス (B) における追跡の様子: 左図: 提案手法, 右図: 基準システム

表 3 追跡結果:  $N$  は視野内における平均人数,  $\mu, \sigma$  はそれぞれ平均, 標準偏差を表す. また  $t_c$  はクラスタリング処理に要する平均計算時間を表す.

	$N$	提案手法					基準システム				
		再現率		適合率		$t_c$ [sec]	再現率		適合率		$t_c$ [sec]
		$\mu$	$\sigma$	$\mu$	$\sigma$		$\mu$	$\sigma$	$\mu$	$\sigma$	
(A)	12.9	0.643	0.132	0.957	0.066	30.45	0.284	0.207	0.746	0.107	4.16
(B)	20.9	0.571	0.117	0.981	0.035	28.75	0.434	0.2	0.924	0.225	4.33
(C)	32.1	0.419	0.145	0.954	0.073	10.09	0.223	0.115	0.862	0.098	2.23
UCSD data [12]	21.3	0.595	0.124	0.888	0.075	6.75	0.363	0.144	0.886	0.103	3.33

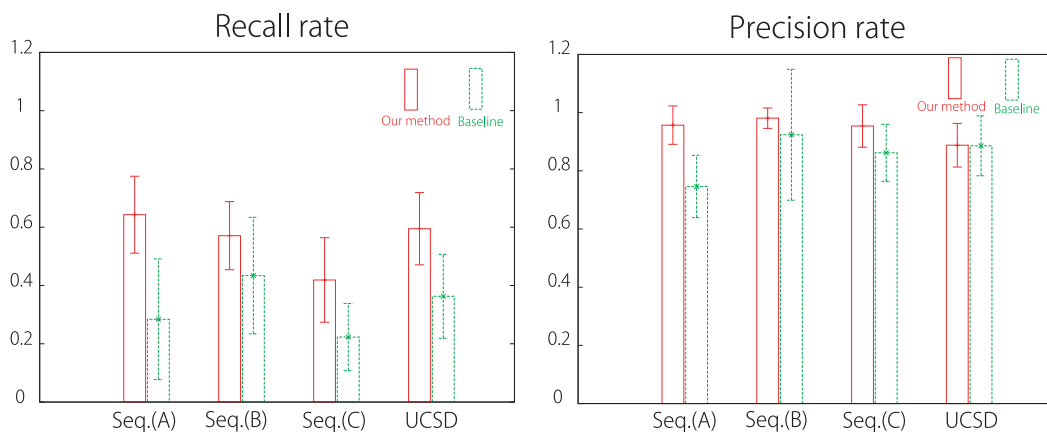


図 9 追跡結果: 左のグラフはそれぞれのシーケンスにおける再現率 (平均, 標準偏差) を表し, 右のグラフは各シーケンスにおける適合率を表す.

よび文部科学省科学研究費補助金基盤 (B) 「混雑状況下における人物追跡にもとづく行動解析」(課題番号: 20300061) の助成により行われた.

文 献

[1] G. Brostow and R. Cipolla: “Unsupervised bayesian

detection of independent motion in crowds”, Proc. of IEEE CS Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 594–601 (2006).

[2] Y. Li and H. Ai: “Fast detection of independent motion in crowds guided by supervised learning”, Proc. of IEEE International Conference on Image Processing, Vol. 3, pp. 341–344 (2007).





図 11 シーケンス (C) における追跡の様子: 左図: 提案手法, 右図: 基準システム

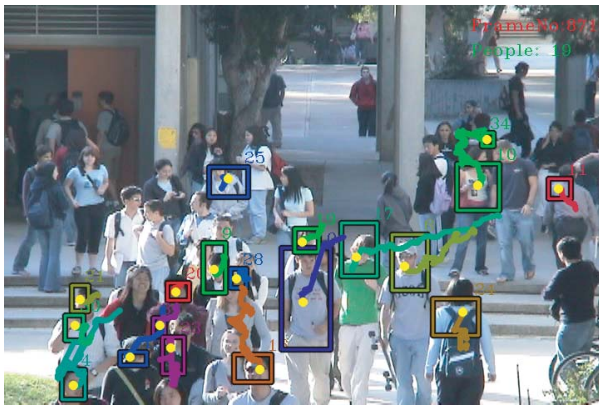


図 12 UDSO のデータ [3] における追跡の様子: 左図: 提案手法, 右図: 基準システム

- [3] V. Rabaud and S. Belongie: “Counting crowded moving objects”, Proc. of IEEE CS Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 705–711 (2006).
- [4] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo and Y. Yagi: “Gait recognition using a view transformation model in the frequency domain”, Proc. of European Conference on Computer Vision, Vol. 3, pp. 151–163 (2006).
- [5] T. Zhao and R. Nevatia: “Tracking multiple humans in complex situations”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **26**, 9, pp. 1208–1221 (2004).
- [6] B. Leibe, E. Seemann and B. Schiele: “Pedestrian detection in crowded scenes”, Proc. of IEEE CS Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 878–885 (2005).
- [7] M. D. Rodriguez and M. Shah: “Detecting and segmenting humans in crowded scenes”, Proc. of ACM International Conference on Multimedia, pp. 353–356 (2007).
- [8] B. Wu and R. Nevatia: “Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors”, International Journal on Computer Vision, **75**, 2, pp. 247–266 (2007).
- [9] L. Dong, V. Parameswaran, V. Ramesh and I. Zogh-lami: “Fast crowd segmentation using shape indexing”, Proc. of IEEE International Conference on Computer Vision, pp. 1–8 (2007).
- [10] S. M. Khan and M. Shah: “A multiview approach to tracking people in crowded scenes using a planar homography constraint”, Proc. of European Conference on Computer Vision, Vol. 4, pp. 133–146 (2006).
- [11] M. Isard and J. MacCormick: “Bramble: A bayesian multiple-blob tracker”, Proc. of IEEE International Conference on Computer Vision, pp. 34–41 (2001).
- [12] Z. Khan, T. R. Balch and F. Dellaert: “An mcmc-based particle filter for tracking multiple interacting targets”, Proc. of European Conference on Computer Vision, Vol. 4, pp. 279–290 (2004).
- [13] K. Okuma, A. Taleghani, N. Freitas, J. Little and D. Lowe: “A boosted particle filter: multitarget detection and tracking”, Proc. of European Conference on Computer Vision, Vol. 1, pp. 28–39 (2004).
- [14] J. R. Casas, A. P. Sitjes and P. P. Folch: “Mutual feedback scheme for face detection and tracking aimed at density estimation in demonstrations”, IEEE Vision, Image and Signal Processing, **152**, 3, pp. 334–346 (2005).
- [15] O. Sidla, Y. Lypetsky, N. Brandle and S. Seer: “Pedestrian detection and tracking for counting applications in crowd situations”, Proc. of IEEE International Conference on Video and Signal Based Surveillance (2006).
- [16] C. Tomasi and T. Kanade: “Detection and tracking of point features”, Technical Report CMU-CS-91-132 (1991).
- [17] C. Tomasi and J. Shi: “Good features to track”, Proc. of IEEE CS Conference on Computer Vision and Pattern Recognition, pp. 593–600 (1994).
- [18] L. P. Chew: “Constrained delaunay triangulations”, Proc. of the Third Annual Symposium on Computational Geometry, pp. 215–222 (1987).