

# 歩容特徴と局所的見えを考慮した特徴点軌跡のクラスタリングによる混雑環境下人物追跡

杉村 大輔<sup>†a)</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>

Incorporating Gait Features and Local Appearances for Tracking People in Crowds Based on Clustering Feature Trajectories

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

あらまし 本論文では、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。周波数領域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

## 1. ま え が き

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、頻繁に発生する遮へい、非常に近接している複数の人物の区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を手掛りに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、一般に部分遮へい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。これに伴い、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、本論文では、(1)

<sup>†</sup> 東京大学生産技術研究所, 東京都

Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

<sup>††</sup> 電気通信大学大学院情報システム学研究科, 調布市

Graduate School of Information Systems, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, 182-8505 Japan

<sup>†††</sup> 国立情報学研究所, 東京都

National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

a) E-mail: sugimura@iis.u-tokyo.ac.jp



図 1 混雑環境の例

Fig. 1 An example of a crowded scene.

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. これらを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, 混雑環境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は, 個人認証のための手掛りとして生体認証の分野で広く利用されていることから分かるように [4], 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固有の性質が現れる. 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歩容特徴とみなし識別に利用する.

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. このような見えの変動を計測することにより個々の人物を識別することが可能となる. 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へい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能となる.

本論文の構成は以下のとおりである. まず, 2. において本手法に関連する先行研究についてまとめ, 本手法の位置付け及び新規性について整理する. 続く 3. 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. 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. では人工データ, 実際の群集の映像を用いた実験結果について述べ, 本手法の追跡性能, 適用範囲について考察する. 最後に 6. 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, 両脚の四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へいされた場合においても安定な追跡が可能となる. 人物のパーツごとに検出器を用意し, それぞれを学習することにより部分遮へいに頑健な検出・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, あらかじ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の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りにあらかじ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りにした時系列フィルタによる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へいに伴い追跡対象が安定に観測されづらい. そのためゆ一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がいくつか提案されているが [14], [15], これらの手法はあくまで連続するフレーム間の動き情報の取得のためにオプティカルフローを利用しているだけであり, いずれも部分遮へいについての効果的な対処はなされていない.

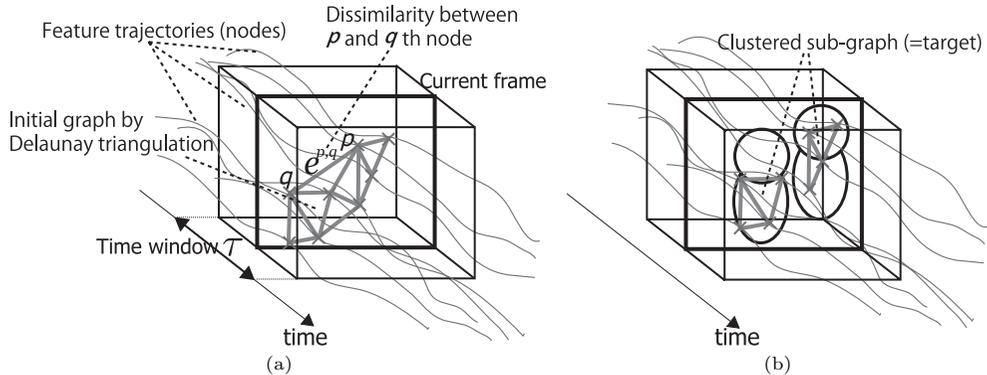


図2 提案手法概要。(a) グラフの構築。時間窓  $\tau$  内における特徴点軌跡群をノードとして用い、ドロネー三角形分割アルゴリズムを適用することにより構成される。特徴点軌跡（ノード）を繋ぐエッジがもつ重みは、軌跡間相違度  $e^{p,q}$  を表す。(b) クラスタリング結果。計算された特徴点軌跡間相違度に基づきエッジを除去することにより実現される。クラスタリングにより得られる部分グラフのうち、あるしきい値  $th_s$  以上のノードをもつものを追跡対象とみなす。

Fig. 2 Graph clustering: (a) The initial graph is constructed via Delaunay triangulation by using the set of feature trajectories in the time window  $\tau$ . A connected edge between  $p$  and  $q$  th trajectory (node) has the dissimilarity  $e^{p,q}$ . (b) The final result by pruning the edges with its edge weight. A clustered sub-graph that has the nodes more than a threshold  $th_s$  is considered as a target that should be tracked.

このような形状モデルや見えを手掛りにした追跡手法とは対照的に、Brostow らは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類似性を計測することにより事後確率を求める。このようにして得られた事後確率をもとに、軌跡群を階層的にクラスタリングすることで人物追跡を実現する。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 らは動きの類似性の仮定に加え、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 [3]。Li らは Brostow らと類似した追跡の枠組みに、動きの類似性に基づくゆがみ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るため、周りとは類似した動き方を示す傾向にある。そのため、特徴点軌跡の動きの類似性を考慮するだけでは個々の人物を識別することは難しい。

これに対し、本手法で利用する歩容特徴は個々の人物がも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。それに加え局所領域における見えの時間変動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も頑健な個々の人物の追跡が可能となる。

### 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、また各ノード間に張られるエッジの重み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] により画像中の各特徴点の動きの軌跡を獲得する。遮へいなどの様々な要因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（背景から抽出された特徴点）、(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、(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、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))。実画像における初期グラフの例を図 8 に示す。

二つのノード (特徴点軌跡) をつなげるエッジは軌跡間の相違度を表す重みをもつ。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、動きの類似性に基づく指標に加え、(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て計算される。本手法では、それぞれの性質は独立に観測できるものと仮定し、最終的な二つの軌跡  $p, q$  間の相違度  $e^{p,q}$  を各指標に基づく相違度を累積することにより計算する。各指標に基づく相違度の対数をとると、相違度  $e^{p,q}$  は次式により与えられる。

$$e^{p,q} = \ln(e_{\text{freq}}^{p,q}) + \ln(e_{\text{app}}^{p,q}) + \ln(e_{\text{prox}}^{p,q}) + \ln(e_{\text{coh}}^{p,q}) . \quad (1)$$

ここで式 (1) の右辺はそれぞれ、歩容特徴に基づく相違度  $e_{\text{freq}}^{p,q}$ 、局所領域の見えの変化に基づく相違度  $e_{\text{app}}^{p,q}$ 、空間関係に基づく相違度  $e_{\text{prox}}^{p,q}$ 、そして動きの類似性に基づく相違度  $e_{\text{coh}}^{p,q}$  を意味する。また、それぞれの指標に基づく相違度は、各指標における相違度の最大値により正規化されており、0 から 1 の値をとる。これらの指標の詳細は次章で詳しく述べる。

もし一つの指標に基づく相違度が 0 になった場合、他の指標に基づく相違度の大きさにかかわ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指標による相違度が 0 になった場合、事前に定めた定数  $e_{\text{min}} > 0$  で置き換える。

以上のような指標を用いて計算された相違度をもとに、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))。本手法におけるクラスタリング処理は、軌跡間相違度があるしきい値  $th_p$  より大きいものを枝打ちすることにより実現される。クラスタリングにより得られる部分グラフのうち、属するノードの数があるしきい値  $th_s$  以下の小部分グラフは除外し、残りの部分グラフが各追跡対象であるとみなす。また本手法では時間窓  $\tau$  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  $\tau/4$  時間ごとにクラスタリング処理が行われる。

#### 4. 軌跡間の相違度の計算

二つのノード (特徴点軌跡) 間のエッジの重み (相違度) 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類似性、空間関係の四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。本章ではそれぞれの指標に基づく相違度の計算方法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も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有効な指標である。本手法では、歩容特徴は人物が歩くときに観測される垂直方向の周期的な動きの周波数特徴に現れると仮定する。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から歩容特徴にあたる周期的な動きを抽出する。具体的には、特徴点軌跡について直線当てはめを行い、推定された線形成分を特徴点軌跡から取り除く。これにより、垂直方向の周期的な動き  $y_p(t)$  が得られる。図 3 に本処理の概略を示す。このようにして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行うことにより、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間の相違度  $e_{\text{freq}}^{p,q}$  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指標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利用されている [4]。また位相に関する相違度については位相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  $e_{\text{freq}}^{p,q}$  は次のように計算される。

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

ここで  $a_k^p$  は  $p$  番目の特徴点軌跡の  $k$  番目の周波数帯の振幅成分を意味し、 $\phi^p$  は振幅成分が最も含まれる周波数帯における位相成分を意味する。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

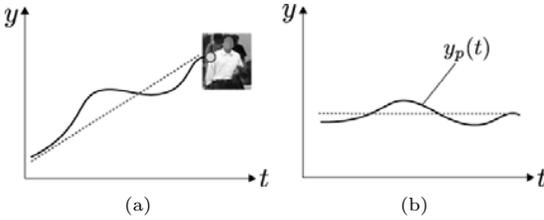


図 3 周波数領域における歩容特徴計測のための前処理．(a) 歩容特徴にあたる垂直方向の周期的な動きの抽出．得られる特徴点軌跡に直線当てはめを行い，その線形成分を除去することにより得られる．(b) 抽出された垂直方向の周期的な動き  $y_p(t)$ ．

Fig. 3 Pre-processing for measuring the gait features in frequency domain. Line fitting for removing linear motion component of feature trajectory (left image). Extracted periodic motion  $y_p(t)$  (right image).

る有効な指標であるが，混雑環境下では周りの人物の歩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新たに利用する．

本手法では局所領域を連結された三つの特徴点軌跡によって張られる三角パッチで表現し，そのパッチ内の見えの時間変動を考える．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，それに対し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．このような三角パッチの見えの変動の違いを計測することにより，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こ配ヒストグラムなど様々なものが考えられるが，本論文では簡単のため HS 色ヒストグラムを利用する．具体的には，時間窓  $\tau$  内の各時刻の色ヒストグラム  $h(t)$  と，その平均色ヒストグラム  $\bar{h}$  の Bhattacharyya 距離の RMS (Root Mean Square) 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，

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

と表すことができる．ここで  $t_s$  は時間窓の最初のフ

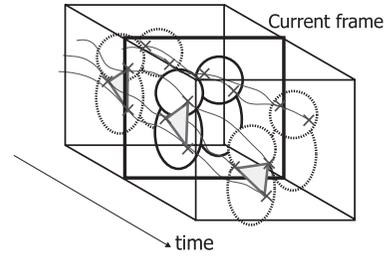


図 4 局所領域としての三角パッチの利用．三角パッチは連結されている特徴点軌跡により張られ，そのパッチ内の見えの時間変動が相違度計算に利用される．

Fig. 4 A sequence of triangular patches is used to measure the variation of appearance over time. Each vertex of a triangular patch is created by the three adjacent trajectories at each time.

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 空間関係と動きの類似性

先行研究 [1] ~ [3] において利用されているように，本手法においても同様に特徴点軌跡間の空間関係，動きの類似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（動きの類似性）と考えられる．そこで本手法では時間窓内において二つの軌跡間が最大となる距離を空間関係に基づく相違度  $e_{prox}^{p,q}$  として利用する．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，動きの類似性に基づく相違度  $e_{coh}^{p,q}$  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 2.66 GHz，RAM: 3.00 GB である．効果の比較をするために，軌跡間の相違度計算において歩容特徴と局所見えの変化に基づく指標を除いたものを基準システムとして用いた．これはすなわち軌跡間相違度を空間関係と動きの類似性の指標のみを用いて計算することを意味する ( $e^{p,q} = \ln(e_{prox}^{p,q}) + \ln(e_{coh}^{p,q})$ )．こ

の基準システムは特に記述がない限り以下のすべての実験において同じものが利用される．後に示す追跡結果の図中に描画されている追跡対象の外接方形は、部分グラフに含まれる特徴点軌跡の現フレームにおける位置を包括する方形領域に対応する．また同様に図中に描画される太線軌跡は、部分グラフの重心位置の軌跡、すなわち追跡対象の動線を意味する．

### 5.1 人工データを用いた追跡実験

本手法で提案する歩容特徴、局所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と仮定しているため、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．

#### 5.1.1 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、垂直方向の周期的な動きの周波数、位相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分かる．それに対し本手法(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分かる．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．

#### 5.1.2 局所領域における見えの一貫性の利用の効果

動きの周波数特徴も周りの人物と同調する個々の人物の識別が最も難しいケースを想定し、二つのターゲットの動きが全く同じシーケンスについて実験した．このようなケースの実験をすることにより、局所領域における見えの利用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は空間関係、動きの類似性そして歩容特徴の指標を用いて軌跡間相違度を計算した ( $e^{p,q} = \ln(e_{\text{freq}}^{p,q}) + \ln(e_{\text{prox}}^{p,q}) + \ln(e_{\text{coh}}^{p,q})$ ) ．

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できることを確認できた．



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

Fig.5 Effect of using gait features. (a) Neighboring targets are mistakenly tracked as one target without the use of gait features (b) two targets are properly tracked by using the gait features.

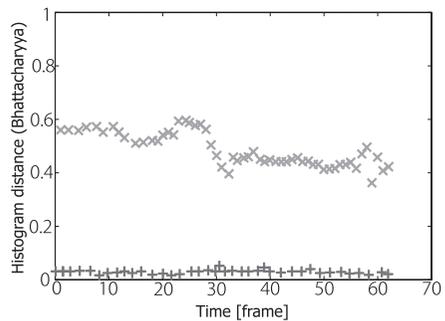


図6 ヒストグラム距離  $d_{\text{hist}}(\mathbf{h}(t), \bar{\mathbf{h}})$  の時間変動結果． 十字点は一つのターゲット内に張られたある一つの三角パッチのヒストグラム距離を意味し、x 点は二つのターゲット間に張られたある一つの三角パッチのヒストグラム距離を表す．

Fig.6 Histogram distance of two local patches over time. The crosses represent a patch located on one target in contrast to the x's which represent a patch which straddles two targets. There is more variation in appearance when the patch straddles two targets.

図6に三角パッチ内のヒストグラム距離  $d_{\text{hist}}(\mathbf{h}(t), \bar{\mathbf{h}})$  の変動結果の例を示す．十字点でプロットされた点群は、一つのターゲットに属したある一つの三角パッチのヒストグラム距離の時間変動結果を意味し、x 点の点群は二つのターゲット間に張られたもの、すなわち背景を含んだある一つの三角パッチの結果を意味する．この結果から分かるように、背景を含む三角パッチに比べて一つのターゲットに属した三角パッチ内のヒストグラム距離は時間的にばらつきが小さいことが見て取れる．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有効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

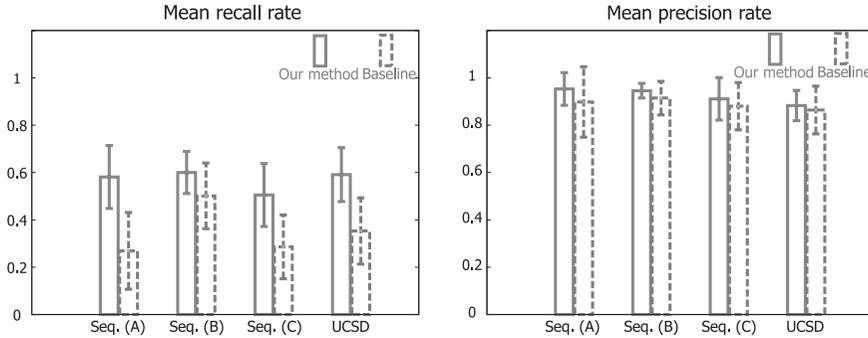


図 7 追跡結果：左のグラフはそれぞれのシーケンスにおける再現率（平均，標準偏差）を表し，右のグラフは各シーケンスにおける適合率を表す。

Fig. 7 Tracking result: the left bar graph shows the recall rate for each sequence, and the right bar graph shows the precision rate for each sequence. The improvement of tracking performance can be observed in all of sequences.

表 3 追跡結果： $N$  は視野内における平均人数， $\mu$ ， $\sigma$  はそれぞれ再現率，適合率の平均，標準偏差を表す．また  $t_c$  はクラスタリング処理に要した平均計算時間を表す。

Table 3 Tracking result:  $N$  shows the average of number of individuals,  $\mu$  and  $\sigma$  show the mean and the standard deviation of recall and precision rate, respectively.  $t_c$  shows the average of computation time for clustering processing.

	$N$	Our system					Baseline				
		Recall rate		Precision rate		$t_c$ [sec]	Recall rate		Precision rate		$t_c$ [sec]
		$\mu$	$\sigma$	$\mu$	$\sigma$		$\mu$	$\sigma$	$\mu$	$\sigma$	
Seq. (A)	14.0	0.581	0.133	0.953	0.069	30.45	0.269	0.162	0.898	0.149	4.16
Seq. (B)	22.0	0.607	0.065	0.943	0.065	28.75	0.514	0.090	0.926	0.069	4.33
Seq. (C)	29.4	0.505	0.133	0.911	0.090	10.09	0.286	0.135	0.880	0.100	2.23
UCSD	21.3	0.591	0.114	0.883	0.064	6.75	0.353	0.140	0.864	0.101	3.33

表 1 実験で利用した映像のパラメータ

Table 1 Characteristics of the videos used in the experiments.

	Resolution	Frame rate	Frames
Seq. (A)	1280 × 720	60	700
Seq. (B)	1280 × 720	60	730
Seq. (C)	800 × 600	30	300
UCSD	720 × 480	30	1000

た．本実験では，それぞれ異なる場所から撮影した三つのシーケンス (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しきい値  $th_p$  は，あらかじめ獲得した相違度  $e^{p,q}$  の集合から決定した．具体的には，対象とするシーケンスからパラメータ設定用の短いシーケンスを用意し，それらについて特徴点軌跡群の生成，グラフ構築，軌跡間相違度の計算を行う．このようにして得られる相違度の集合からメジアン値を計算し，これをクラスタリングのしきい値  $th_p$  として

表 2 実環境下実験に利用したパラメータ．しきい値  $th_p$  はエッジ間相違度  $e^{p,q}$  の集合のメジアン値とした．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．

Table 2 Parameters used for our experiments with real video. We set the threshold  $th_p$  as the median value computed from the set of the edge weight  $e^{p,q}$ . Note that we don't use the sequence used for making the set of weight  $e^{p,q}$  for experimental evaluations.

	$\tau$	$th_s$	Our system		Baseline	
			$e_{min}$	$th_p$	$e_{min}$	$th_p$
Seq. (A)	64	13	0.1	-17.43	0.1	-11.95
Seq. (B)	64	10	0.1	-15.80	0.1	-9.79
Seq. (C)	40	8	0.1	-16.51	0.1	-11.92
UCSD	40	8	0.1	-16.02	0.1	-10.78

利用した．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，映像のフレームレートに応じて決定した．

図 10, 図 11, 図 12, 図 13 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

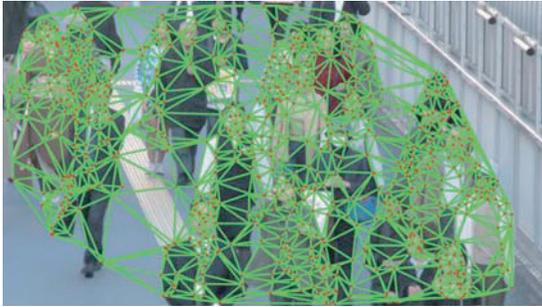


図 8 実際の映像におけるグラフ構造の例  
 Fig. 8 An example of the initial connected graph in a real sequence. Each node represents a feature trajectory and the lines are the set of edges in the graph.

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、本論文では再現率と適合率を用いた。また本実験では false positive を、(1) 一人の人物に複数のクラスタが生成される場合、(2) 複数の人物が一人の人物として追跡される場合、の 2 種類とした。表 3、図 7 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優れていることが見て取れる。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 倍程度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の利用が追跡性能の向上に大きく寄与しているといえる。

### 5.2.1 追跡失敗に関する考察

図 9 に頻繁に観測された二種類の追跡失敗例を示す。図 9(a) の失敗例は追跡すべき人物が存在しているにもかかわ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人物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、若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と考えられる。図 9(a) に生成された特徴点軌跡を描画したものを示す。この図にあるように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る。またこのような場合に生成された特徴点軌跡も上述した理由から信頼できる軌跡であるとはいいがたく、追跡性能の低下が懸念される。

二つ目の失敗例は、図 9(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



図 9 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) 特徴点軌跡の周波数特徴の違いに起因する過検出

Fig. 9 Examples of failure cases. (a) shows that our system does not detect an individual due to the lack of feature trajectories. (b) shows the failure case where multiple clusters are on one individual. This is because there are different frequencies between the individual's body and the baggage.

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つく、そして手、若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 9(b) の右図（特徴点軌跡）からも分かるように、人物の体の動きの周波数特徴とバッグ、若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。本手法は特徴点軌跡の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しているとみなしてしまう。

### 5.2.2 追跡性能に対する枝打ちしきい値の影響

枝打ちしきい値  $th_p$  の選択による追跡性能への影響を確認するために、シーケンス (B) において、しきい値を変化させたときの追跡性能について評価した。具体的には、前述した実験において用いたエッジの枝打ちしきい値 ( $-15.80$ ) を、 $-\ln 10$  倍 ( $-18.10$ )、 $-\ln 5$  倍 ( $-17.41$ )、 $\ln 5$  倍 ( $-14.19$ )、 $\ln 10$  倍 ( $-13.50$ ) と変化させたときのそれぞれの再現率、適合率を評価した。これに加え、しきい値を変化させることにより、クラスタリングされた部分グラフの誤統合、誤分割の発生がどのように変化するかを確認するために、本実験において定義した 2 種類の false positive ((1): 一人の人物に複数のクラスタが生成される、(2): 複数の人物が一人の人物として追跡される) の割合 (検出結果に対する誤検出の割合) の変化についてそれぞれ調べた。

表 4 に各しきい値における追跡結果を示す。適合



図 10 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  
Fig. 10 Tracking results on sequence (A). (left: our system, right: baseline)

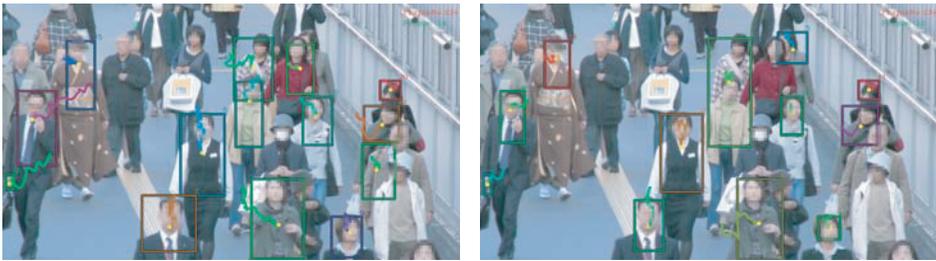


図 11 シーケンス (B) における追跡の様子 : 左図 : 提案手法, 右図 : 基準システム  
Fig. 11 Tracking results on sequence (B). (left: our system, right: baseline)



図 12 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  
Fig. 12 Tracking results on sequence (C). (left: our system, right: baseline)

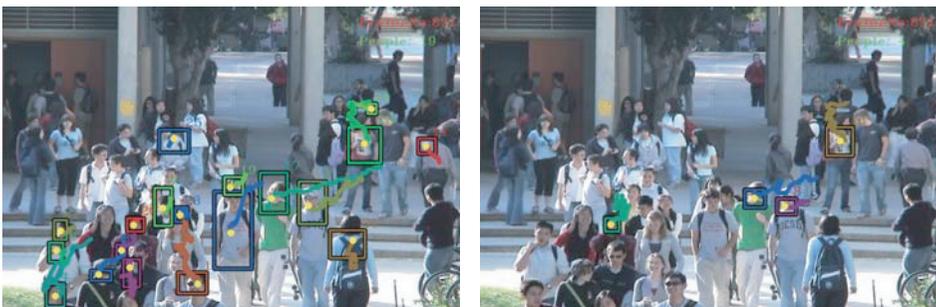


図 13 UCSD のデータ [3] における追跡の様子 : 左図 : 提案手法, 右図 : 基準システム  
Fig. 13 Tracking results on UCSD data [3]. (left: our system, right: baseline)

表 4 枝打ちしきい値  $th_p$  の選択による追跡性能の変動． $\mu, \sigma$  はそれぞれ再現率，適合率の平均，標準偏差を表す．FPR は検出結果に対する誤検出の割合の平均を表す (1): 一人の人物に複数のクラスタが生成されたもの, (2): 複数の人物が一人の人物として追跡されたもの).

Table 4 Changes in tracking performance by selecting a threshold for pruning edges  $th_p$ .  $\mu$  and  $\sigma$  show the mean and the standard deviation of recall and precision rate, respectively. FPR shows the mean of false positive rate ((1): multiple clusters are detected on a single person and (2): multiple people are clustered as one person).

$th_p$	Recall rate		Precision rate		FPR	
	$\mu$	$\sigma$	$\mu$	$\sigma$	(1)	(2)
-18.10	0.441	0.108	0.937	0.073	0.063	0.000
-17.41	0.500	0.091	0.941	0.051	0.053	0.006
-15.80	0.607	0.061	0.943	0.065	0.057	0.000
-14.19	0.618	0.102	0.892	0.082	0.051	0.057
-13.50	0.574	0.101	0.877	0.074	0.025	0.098

率の変化について注目すると，しきい値を大きくした (しきい値を緩めた) 場合に性能の低下が確認される．その一方で，再現率の変化は，しきい値を緩めた場合の変化に対してしきい値を小さくした (しきい値を厳しくした) 場合の変化が比較的大きい．この現象の主な要因は次のように考えられる．枝打ちのしきい値を小さくすることにより，より多くのエッジが除去されるため，部分グラフがもつ特徴点軌跡の数は減少する傾向にある．特徴点軌跡数が少ない部分グラフは追跡対象外として除外するため，結果的に検出率が低下したと考えられる．

次に誤検出率 (FPR) について着目すると，しきい値を大きくすることにより，誤分割の割合 (1) よりも誤統合の割合 (2) の方が大きくなっていることが見て取れる．逆にしきい値を小さくした場合，誤分割の割合が大きくなっていることが分かる．

これらのしきい値変化に伴う追跡性能結果から，エッジの枝打ちのしきい値をメジアン値として選択することは比較的妥当であるといえる．

## 6. むすび

本論文では，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用い，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．更に局所領域の見えの時間変動を計測し利用することにより，局所領域の見えは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かかわらず遮へい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，文部科学省科学研究費補助金 (課題番号: 20300061) 及び科学技術振興機構 CREST 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. 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. IEEE International Conference on Image Processing, vol.3, pp.341-344, 2007.
- [3] V. Rabaud and S. Belongie, "Counting crowded moving objects," Proc. 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. European Conference on Computer Vision, vol.3, pp.151-163, 2006.
- [5] T. Zhao and R. Nevatia, "Tracking multiple humans in complex situations," IEEE Trans. Pattern Anal. Mach. Intell., vol.26, no.9, pp.1208-1221, 2004.
- [6] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," Proc. 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. 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," Int. J.

- Comput. Vis., vol.75, no.2, pp.247-266, 2007.
- [9] L. Dong, V. Parameswaran, V. Ramesh, and I. Zoghalmi, "Fast crowd segmentation using shape indexing," Proc. 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. European Conference on Computer Vision, vol.4, pp.133-146, 2006.
- [11] M. Isard and J. MacCormick, "Bramble: A Bayesian multiple-blob tracker," Proc. 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. 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. 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, vol.152, no.3, pp.334-346, 2005.
- [15] O. Sidla, Y. Lypetsky, N. Brandle, and S. Seer, "Pedestrian detection and tracking for counting applications in crowded situations," Proc. IEEE International Conference on Video and Signal Based Surveillance, pp.70-75, 2006.
- [16] C. Tomasi and T. Kanade, "Detection and tracking of point features," Technical Report, CMU-CS-91-132, pp.1-20, 1991.
- [17] C. Tomasi and J. Shi, "Good features to track," Proc. IEEE CS Conference on Computer Vision and Pattern Recognition, pp.593-600, 1994.
- [18] L.P. Chew, "Constrained delaunay triangulations," Proc. Third Annual Symposium on Computational Geometry, pp.215-222, 1987.
- (平成 21 年 10 月 9 日受付, 22 年 2 月 26 日再受付)



杉村 大輔

2005 阪大・基礎工・システム科学卒。2007 東京大学大学院情報理工学系研究科電子情報学専攻修士課程了。2010 同大学院同研究科同専攻博士課程了。現在、日本学術振興会特別研究員 (PD)。コンピュータビジョンに関する研究に従事。



木谷 クリス真実

1999 南カリフォルニア大・工・電子卒。2000 ケーエルエー・テンコール (株) 入社。2005 東京大学大学院情報理工学系研究科電子情報学専攻修士課程了。2008 同大学院同研究科同専攻博士課程了。現在、電気通信大学大学院情報システム学研究科助教。MIRU2008 学生優秀論文賞受賞。



岡部 孝弘 (正員)

1997 東大・理・物理卒。1999 同大学院理学系研究科物理学専攻修士課程了。2000 同博士課程中退。2001 より東京大学生産技術研究所技官, 助手を経て, 現在同研究所助教。コンピュータビジョン, コンピュータグラフィックス, 画像パターン認識に関する研究に従事。平 17 年度本会論文賞, MIRU 2004 優秀論文賞, 平 16 年度 PRMU 研究奨励賞, 平 19 年度情報処理学会山下記念研究賞などを受賞。情報処理学会, IEEE 各会員。



佐藤 洋一 (正員)

1990 東大・工・機械卒。1997 カーネギーメロン大学計算機科学部ロボティクス学科博士課程了。Ph.D in Robotics。同年より東京大学生産技術研究所研究機関研究員, 講師, 助教授を経て, 現在同大学院情報学環准教授。コンピュータビジョン, ヒューマン・コンピュータ・インタラクション, コンピュータグラフィックスに関する研究に従事。本会論文賞 (平 17 年度, 平 19 年度), 平 11 年情報処理学会山下記念研究賞, 平 11 日本バーチャルリアリティ学会論文賞等を受賞。情報処理学会, 日本バーチャルリアリティ学会, ACM, IEEE 各会員。



杉本 晃宏 (正員)

1987 東大・工・計数卒。1989 同大学院工学系研究科修士課程了 (数理工学専攻)。日立製作所基礎研究所, ATR, 京都大学を経て, 2002 より国立情報学研究所。現在, 同研究所教授。総合研究大学院大学複合科学研究科教授併任。2006 ~ 2007 Paris-Est 大学客員教授。博士 (工学)。視覚情報処理や離散システム・アルゴリズムなどに広く興味をもち, 数理的手法に基づいた手法を確立する研究に従事。2001 年情報処理学会論文賞。IEEE, ACM 等各会員。