

適応的特徴選択を用いた入退室する人物の対応付け

本田 拓也[†] 出口 大輔[†] 高橋 友和^{††,†} 井手 一郎[†] 村瀬 洋[†]

[†] 名古屋大学 大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

^{††} 岐阜聖徳学園大学 経済情報学部 〒500-8288 岐阜県岐阜市中鶯 1-38

E-mail: [†]{thonda,ddeguchi,ttakahashi,ide,murase}@murase.m.is.nagoya-u.ac.jp

あらまし カメラ映像を用いて、部屋に入退室する人物を対応付ける手法を提案する。その際、在室者に応じた適応的な特徴選択を行うことで、高精度な入退室人物間の対応付けを行った。近年、マーケティングを目的とした映像の解析では、美術館やコンビニエンスストアなど、不特定多数の人物が出入りする場の人流解析技術が重要となってきた。例えば、出入口を入退室する人物映像からの平均滞在時間の算出などが挙げられる。こういった情報は、部屋に入室する人物と退室する人物を対応付けることで得られる。しかし、服装の類似した人物が多数存在する場合、単純に色のみでは正しい対応付けを行うことが難しい。そこで本研究では、人物間の局所的な見えの違いに注目し、在室者に応じて対応付けに有効な画像特徴を適応的に選択することにより、対応付け精度の向上を図る。本稿では、適応的特徴選択を利用した特徴の評価実験、および、8日分の入退室映像からの滞在時間や入退室人数の算出実験により、手法の有効性を確認した。

キーワード 人物対応付け、監視カメラ、画像特徴、人流解析、滞在時間

Person Re-identification at a Room Entrance by Adaptive Feature Selection

Takuya HONDA[†], Daisuke DEGUCHI[†], Tomokazu TAKAHASHI^{††,†}, Ichiro IDE[†], and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University, Nagoya, Aichi, Japan

^{††} Faculty of Economics and Information, Gifu Shotoku Gakuen University, Gifu, Gifu, Japan

E-mail: [†]{thonda,ddeguchi,ttakahashi,ide,murase}@murase.m.is.nagoya-u.ac.jp

Abstract This paper proposes a method to re-identify people at a room entrance from a surveillance video. The method allows us to re-identify people in high precision by adaptively selecting features according to the individuals in the room. Recently in the field of a video analysis for marketing, flow analysis of customers is important. It is used in a place such as convenience stores and museums, where an arbitrary number of customers enter and exit; For example, it can be used for measuring the mean time of stay by monitoring the entrance. Such information can be obtained by re-identifying the individuals (entering and exiting the room) using color features. However, in the case that multiple individuals are dressed in similar colors, it is difficult to re-identify them correctly with only the color features. Therefore, focusing on the difference in local features, the person re-identification accuracy is improved by selecting features effective for re-identification per individual. In this paper, we show the experimental results for evaluating adaptive feature selection. We also measured the number of entry/exit events, and the time of stay for each individual from a 8 days video by using the proposed method. From these, the effectiveness of the proposed method was confirmed.

Key words Person Re-identification, Surveillance camera, Image feature, Time of stay

1. はじめに

近年、安全や防犯への意識の高まりから、商店街や店舗内、事務所などへの監視カメラの設置が急速に進んでいる [1]。しかし、多数の監視カメラ映像の全てを常に監視するのは現実的ではなく、自動で映像から情報を得る技術が重要である。一方、交通量・マーケティング調査などを目的として、カメラ映像から人物・車両などの移動物体を検出し、追跡する技術が注目されている。これらの実現には、(a) 連続したフレーム画像から移動物体を検出し追跡する技術、それらを (b) 関連した動作のまとまりであるイベント単位にまとめる技術、(c) 切り出したイベントごとに特徴を抽出し、対応付けを行う技術が必要である。中でも (c) の対応付けに関する技術は重要で、これまでにいくつかの研究グループから手法の提案がなされている。

小林らは、高速道路等に設置されたカメラを用いて、2 地点で収集された車長と車高のデータから同一車両を対応付け、2 地点間の旅行時間を算出した [2]。一方、Hu [4] らは撮影領域が重複し、かつ視野が異なるカメラ映像間で、歩行者の主軸に注目した人物の対応付けを行っている。各カメラ視野において検出された主軸を同一座標系に投影し、主軸間の距離を用いて同一人物の対応付けを行っている。この手法では、集団で移動している場合や、車などで体が隠れた場合についても、正しく歩行者の対応付け、追跡が可能となっている。佐藤ら [5] や由雄ら [6] は、顔や服装の特徴を用いた高速人物検出手法を提案した。これらの手法では、カメラ視野内を移動する人物の追跡を行い、その中で人物検索に最適なフレームを選ぶ。しかし、顔特徴のみでは高速な検索を実現できないため、人物それぞれの服装などの色特徴を利用することで高速かつ高精度な検索を実現している。

以上のように、移動物体の対応付けは、カメラ映像を入力とした統計解析には不可欠な技術である。しかし、類似した服装をした人物が多数存在する状況下では、単純に色特徴を用いただけでは、正しい対応付けは困難である。そのような場合であっても、人物間の部分的な特徴の違いにより、類似した服装の人物を識別できる可能性がある。そこで本稿では、複数の特徴から在室者の状況に対して適応的に有効な特徴を選択し、人物の入退室の対応付けを高精度に行う手法を提案する。

以降 2 節では、適応的特徴選択を用いた人物対応付け手法について述べる。続く 3 節では、提案手法の有効性を評価するための実験に加え、本手法を用いた統計情報の収集実験についても述べる。最後に 4 節でまとめる。

2. 入退室する人物の対応付け手法

本手法は、部屋の出入り口を横方向から撮影したカメラ映像を入力とし、入室および退室した人物の対応付け結果を出力とする。処理の流れを図 1 に示す。処理は、大きく分けて (1) 入退室イベントの検出、(2) 入退室イベントの対応付け、の 2 段階で構成される。(2) の対応付けの処理の際には、入力画像からいくつかの画像特徴を抽出するが、その特徴全てを対応付けに用いるのではなく、その時々々の在室者に応じて対応付けに有

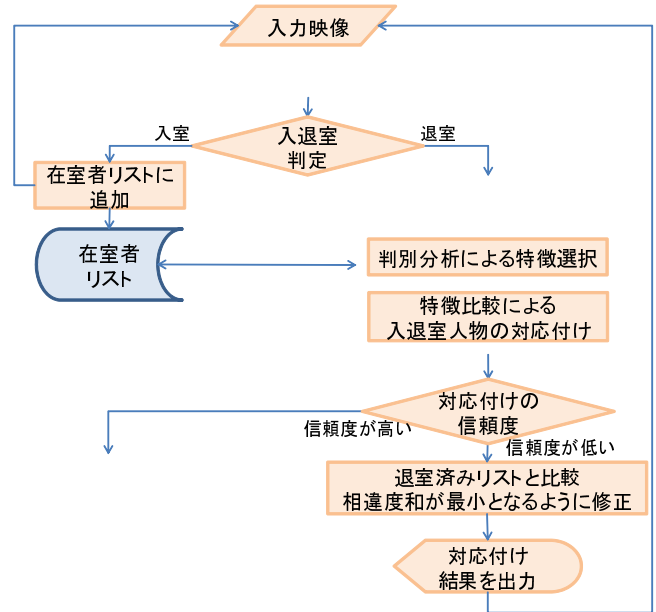


図 1 処理の流れ



図 2 更新型背景モデルを用いた背景差分法の処理画像

効な特徴を選択する。ここでは、1 人の人物が入室する際の一連の動作を入室イベント、退室する際の一連の動作を退室イベントとし、それらを総称して入退室イベントと呼ぶ。

2.1 入退室人物の検出

入退室人物の検出のため、まず人物の検出を行う。固定カメラを利用しているため、撮影領域に変動が少ないという特性を活かして、背景差分法により前景領域を人物領域として抽出し、入退室の判定を行う。背景差分法には様々な改良手法が提案されているが、本環境では急激な照明変動や風などによる揺らぎの影響が無視できるため、更新型背景モデルを用いた背景差分手法を採用する。この手法では、最近の背景のみの 50 フレームから平均背景画像を作成し、入力画像との差分を求める。図 2 に、(a) 入力画像、(b) 背景差分により得られたマスク画像 (白い領域が前景)、(c) 出力画像を示す。

求めた人物の画像をイベントごとにまとめ、人物領域の重心の追跡により人物の移動方向を求める。これにより一連の歩行画像を入退室ごとにまとめるとともに、入室／退室の判定をして、対応付けに用いる。

2.2 入退室人物の対応付け

在室者に応じて有効な特徴を選択し、対応付けに用いる手法について述べる。適応的特徴選択のコンセプトを図 3 に示す。

この手法は、任意の在室者の状況に応じて、在室者を区別し

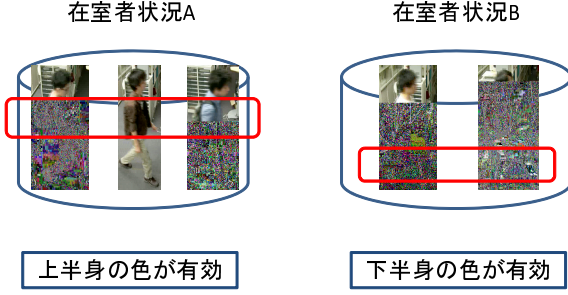


図 3 在室者の状況に応じた特徴選択の例

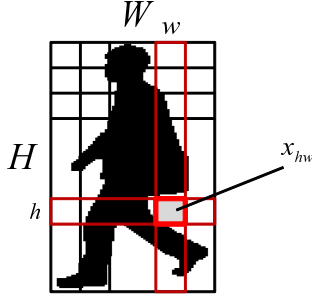


図 4 局所領域の平均色特徴

やすい特徴を見つけ、対応付けるといものである。例えば、図 3 の在室者状況 A に示すように、上着の色が在室者を対応付けるために有効な特徴である場合には、上着の特徴によって退室者と在室者の対応付けを行う。また、在室者状況 B のように靴やズボンの色が有効な特徴である場合には、これによって対応付けを行う。このように、在室者の状況に応じて人物の対応付けに有効な画像特徴を適応的に選択することにより、入退室する人物の高精度な対応付けを行う。提案手法では、人物の画像特徴として色特徴と形状特徴を用い、在室者の画像特徴の判別分析を行うことによって、人物の対応付けに有効な特徴を適応的に選択し、在室者の状況に適した対応付けを行う。

2.2.1 画像特徴の抽出

利用する特徴は大きく分けて、色に関する特徴と形状に関する特徴の 2 つである。抽出した特徴は、連結して 1 つの特徴ベクトルとして対応付けに用いる。

a) 色 特 徴

図 4 のように、入力画像を横方向に W 、縦方向に H でメッシュ状に分割した局所領域の平均色を特徴として用いる。ある局所領域の平均色を $\mathbf{x}_{hw}$ ($w = 0, \dots, W, h = 0, \dots, H$) とすると、求まる特徴は

$$\begin{aligned} \mathbf{x} &= [\mathbf{x}_{11}, \dots, \mathbf{x}_{1W}, \mathbf{x}_{21}, \dots, \mathbf{x}_{HW}], \\ \mathbf{x}_{hw} &= [x_R, x_G, x_B], \end{aligned} \quad (1)$$

のような $3 \times W \times H$ 次元のベクトルとなる。

b) 形 状 特 徴

矩形状の人物領域の幅と高さ、および HOG (Histogram of Oriented Gradients) 特徴 [7] を形状に関する特徴として用いる。HOG 特徴の計算により得られたヒストグラムと、人物領

域矩形の幅と高さの 2 次元特徴を連結したものを形状に関する特徴ベクトルとする。

2.2.2 適応的特徴選択を用いた人物の対応付け

ある時刻において、部屋の中に P 人が在室している状況を考える。まず、2.2.1 で述べた特徴ベクトルを用いて、人物 p のフレーム f における特徴ベクトル $\mathbf{x}_p(f)$ を求める。

$$\mathbf{x}_p(f) = (x_1, x_2, \dots, x_{3WH})^T \quad (f = 1, 2, \dots, F_p). \quad (2)$$

ここで、ある人物 p の入室映像のフレーム数は F_p である。この特徴ベクトルの人物内 (within) 平均 $\mathbf{E}_p^w$ と、人物 p 内での共分散行列 $\mathbf{S}_p^w$ を次式で求める。

$$\mathbf{E}_p^w = \frac{1}{F_p} \sum_{f=1}^{F_p} \mathbf{x}_p(f), \quad (3)$$

$$\mathbf{S}^w = \frac{1}{F_p} \sum_{f=1}^{F_p} (\mathbf{x}_p(f) - \mathbf{E}_p^w)(\mathbf{x}_p(f) - \mathbf{E}_p^w)^T. \quad (4)$$

次に、式 (5) のように在室者の総フレーム数を F として、人物間の特徴ベクトルの平均 $\mathbf{E}^b$ と、人物間 (between) の共分散行列 $\mathbf{S}^b$ を求める。

$$F = \sum_{p=1}^P F_p, \quad (5)$$

$$\mathbf{E}^b = \frac{1}{F} \sum_{p=1}^P F_p \mathbf{E}_p^w, \quad (6)$$

$$\mathbf{S}^b = \frac{1}{F} \sum_{p=1}^P (\mathbf{E}_p^w - \mathbf{E}^b)(\mathbf{E}_p^w - \mathbf{E}^b)^T. \quad (7)$$

求めた $\mathbf{S}^w$ と $\mathbf{S}^b$ を用いて、次の一般固有値問題を解く。

$$\mathbf{S}^b \mathbf{u}_i = \lambda_i \mathbf{S}^w \mathbf{u}_i, \quad (8)$$

$$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{3WH}), \quad (9)$$

$$\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{3WH}). \quad (10)$$

ここで、 λ_i ($\lambda_1 \geq \dots \geq \lambda_i \geq \dots \geq \lambda_{3WH}$) と $\mathbf{u}_i$ は対応する固有値と固有ベクトルであり、固有値の大きいものから $\mathbf{u}_i$ を K 個選択し、特徴変換行列 $\mathbf{U}'$ を求める。

$$\mathbf{U}' = (\mathbf{u}_1, \dots, \mathbf{u}_K) \quad (11)$$

そして、図 5 のように入退室の対応付けを行う。退室が発生した時に、退室イベントと在室者リスト中の各入室イベントの特徴ベクトル間の距離計算を行う。この際、退室者の各フレームに対して、入室者のフレームの中から最も類似したフレームを探す。その際に、まず式 (11) で求めた特徴変換行列を用い、退室イベント o の m フレーム目の画像 I_{om} と入室イベント i の n フレーム目の画像 I_{in} の距離を式 (12) のように求める。

$$\delta_m = \|(\mathbf{U}'^T \mathbf{x}_o(m) - \mathbf{U}'^T \mathbf{x}_i(n))\|. \quad (12)$$

ここで、 $\|\mathbf{x}\|$ は $\mathbf{x}$ のユークリッド距離を表し、入退室 1 回の映

入室件数(件)

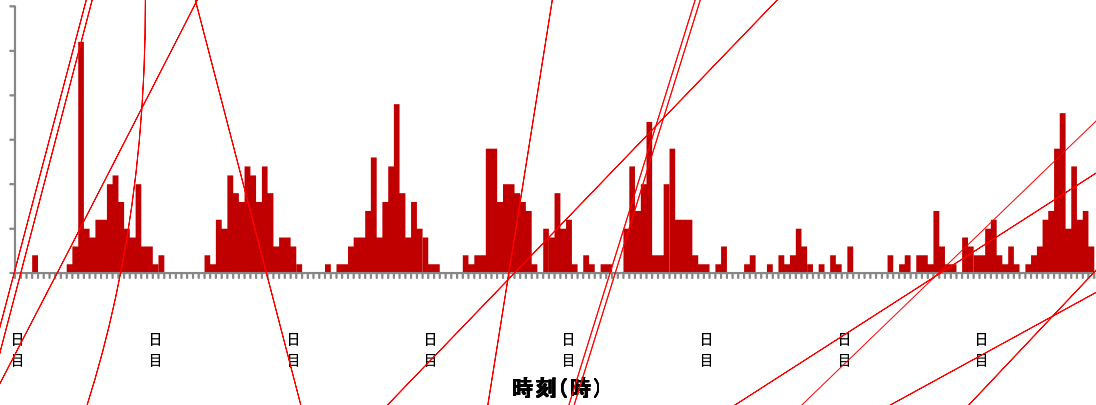


図 7 時間別入室者数分布

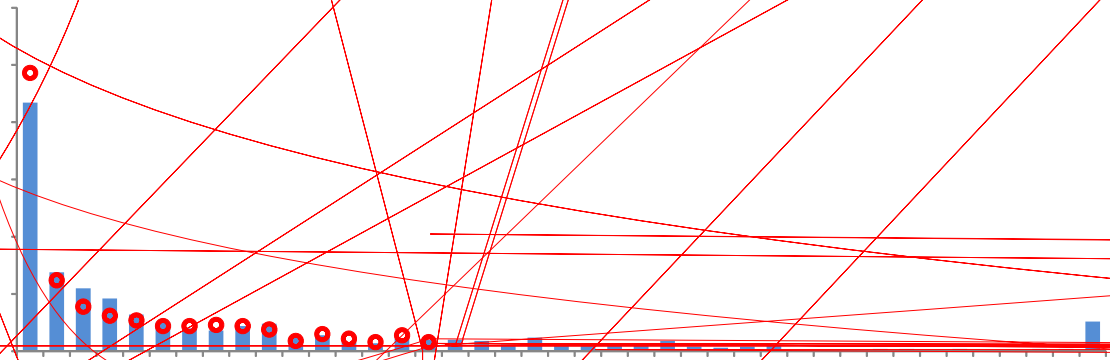
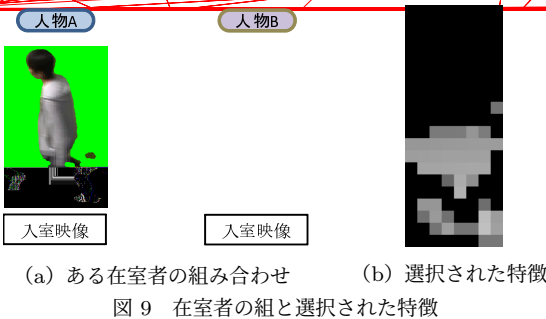
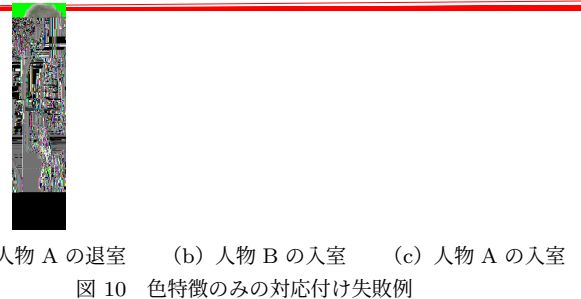


図 8 滞在時間の分布。青い棒グラフは提案手法による分布を表し、赤い丸は実際の分布を表す。



(a) ある在室者の組み合わせ (b) 選択された特徴
図 9 在室者の組と選択された特徴



(a) 人物 A の退室 (b) 人物 B の入室 (c) 人物 A の入室
図 10 色特徴のみの対応付け失敗例

4. 考 察

4.1 適応的特徴選択手法

表 1 のように、判別分析を用いた適応的特徴選択により、良好に対応付けが行えることを確認した。図 9 に特徴選択無しと有りの場合に対応付けられた画像対の例を示す。図 9 の (b) は、どの領域の色特徴が対応付けに用いられたかを示している。(a) の在室者の組み合わせの場合、上半身の色は類似しており、下半身の色が対応付けに有効である。(b) を見ると、下半身を中心とした色特徴が選択されていることを確認した。これにより特徴選択の有効性を確認した。

4.2 利用特徴および対応付け修正手法

まず利用する特徴をみると、色特徴と形状特徴を組み合わせることで対応付け精度の向上が確認できる。図 10 にその一例を示す。色特徴のみを用いた場合は、図 10 の (a) と (b) のよ

うな誤った対応付け結果が得られたが、形状特徴を組み合わせることで、(a) と (c) のような正しい対応付け結果が得られた。これは、色特徴のみでは識別が困難な場合に、形状特徴（人物領域の縦の大きさや体の輪郭など）を用いて対応付けを行った結果であると考えられる。

また、対応付けごとに在室者の組み合わせは異なり、各特徴への依存率も様々である。各在室者状況における色特徴と形状特徴への依存率を図 11 に示す。これは、適応的特徴選択により生成した新たな特徴が、両特徴にどの程度依存しているかを表した図である。特に、一方の特徴への依存率の大きい、在室者状況 6 および在室者状況 9 を、図 12、図 13 に示す。在室者状況 6 の場合は、上下の服装の色の組み合わせが類似した 2 人を含む 3 人が在室しており、そのため形状特徴に大きく依存していると考えられる。また、在室者状況 9 の場合は、在室者 3 人の服装が様々であるため、色特徴への依存率が大きくなっていると思われる。

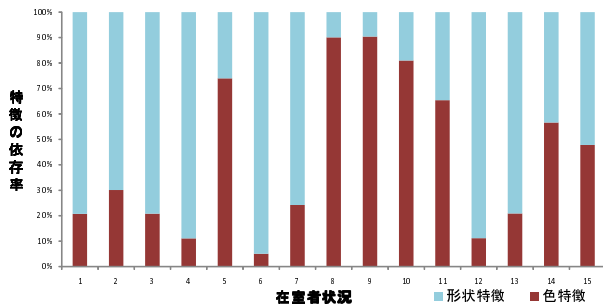


図 11 各在室者状況における特徴への依存率

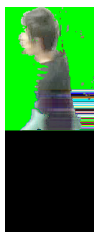


(a) 在室者



(b) 退室人物

図 12 在室者状況 6



(a) 在室者

(b) 退室人物

図 13 在室者状況 9

対応付けの修正に関して、正しく修正できた場合もあれば、逆に誤って修正した場合もあり、全体としては大きな精度向上は見られなかった。この原因として、以下の3つが挙げられる。

- (1) 対応付けの修正が起こらない場合がある
- (2) 修正が誤っている場合がある
- (3) 修正が不完全である

特に原因(3)は、提案する対応付け修正手法が、単に「現在」と「相違度の小さい過去」の2点間の入れ替えを行っていることが影響していると思われる。時系列順に対応付けを行っていった場合、ある時点で対応付けを誤ると、その後の対応付け全てについて精度が低下する可能性がある。そのため、2点間の入れ替えだけでなく、その間の全ての点において再度対応付けを行い、相違度の総和が小さいものを最適な組み合わせとすることが必要である。しかし、計算コストは膨大なものとなり、効率よく対応付けを行うシステムの実現のためには、更なる検討が必要である。

4.3 長期間映像に対する統計情報分析

長期間映像に対して本手法を適用し、得られた滞在時間や時間別在室者数などの統計情報に関する分析を行う。

図7を見ると、基本的に昼間に入室が集中しているのがわかる。特に、昼食前後や夕食前後の時間帯にその日のピークが起きている場合が多い。そして深夜になるに従い、人物の在室者数

が減少している。また、休日である6日目と7日目の入室数が著しく少ない傾向も見取れる。図8では、滞在時間を5分ごとに区切り、提案手法により対応付けた場合と、実際に手動で正解を与えた場合の2種類を比較した結果を示している。提案手法により得られた分布は実際の分布と同様に、5分未満の滞在が最も多く、そこから急激に減り、以降は滑らかに減少している。両分布の一致率は89.9%であり、本手法でも十分な統計情報の取得が可能であることを確認した。しかし、小規模なサンプル映像の結果と比較すると、長期間映像での対応付け精度は低下している。これは、扱う在室者数が多いことが主な原因であり、在室者数がさらに増加するとさらに精度が低下すると予想される。今後は、室内の在室者数に応じた対応付け精度の評価実験が必要だと考えられる。

5. む す び

本稿では、カメラ映像を用いて部屋に入退室する人物に対応付ける手法を提案した。その際に、在室者状況に応じた適応的な特徴選択を行うことで、高精度な入退室人物間の対応付けを実現した。

5名の入退室データを用いた対応付け実験により、特徴選択なしの場合で72.7%、特徴選択ありの場合で77.6%の精度が得られた。また、30件の入退室対応付け実験では、色特徴のみの場合で66.7%、形状特徴のみの場合で40.0%、両特徴を組み合わせることで73.3%と高い精度が得られた。

本稿では、特徴選択に関して詳しく検討したが、姿勢変動などに対する頑健性も求められるため、今後は利用する特徴に関しても検討する必要がある。また、動き特徴の導入と、それに伴う特徴の組み合わせ方法についても検討していく。

謝辞 日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏、その他撮影にご協力いただいた皆様に深く感謝する。本研究の一部は、科学研究費補助金による。また、本研究では画像処理にMISTライブラリ(<http://mist.murase.m.is.nagoya-u.ac.jp/>)を使用した。

文 献

- [1] 小出治, “防犯カメラの効果と利用について,” 予防時報, No.244, pp.30–35, Jan. 2006.
- [2] 小林弘忠, 石関隆幸, 遠藤雅也, 今井浩, 西村茂樹, 下浦弘, 天目健二, “2 地点車両観測情報からの全域的交通流解析システム,” 情報学研報, ITS2000-31, Mar. 2000.
- [3] 赤塚久哉, 高橋友和, 井手一郎, 村瀬洋, “撮影領域に重なりのないカメラ間における人物軌跡の対応付けに関する検討,” 画像の認識・理解のシンポジウム講演予稿集 (MIRU)2006, pp.583–588, Jul. 2006.
- [4] Weiming Hu, Min Hu, Xue Zhou, Tieniu Tan, Jianguang Lou, Steve Maybank, “Principal Axis-Based Correspondence between Multiple Cameras for People Tracking,” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.28, No.4, pp.663–671, Apr. 2006.
- [5] 佐藤正章, 山崎龍次, “服装・顔を用いた高速人物検索,” Panasonic Technical Journal, Vol.52, No.3, pp.67–71, Jan. 2006.
- [6] 由雄宏明, 松川隆行, “顔・着衣特徴による高速人物検索,” Panasonic Technical Journal, Vol.54, No.4, pp.30–32, Jan. 2009.
- [7] Navneet Dalal, Bill Triggs, “Histograms of Oriented Gradients for Human Detection,” Proceedings of Computer Vision and Pattern Recognition, pp. 886–893, Jun. 2005.