

共可視特徴に基づく複数カメラ間の人物同定

伊藤 誠也[†] 金子 俊一^{††}(正会員)[†]株式会社日立製作所 研究開発グループ 制御イノベーションセンタ, ^{††}北海道大学大学院 情報科学研究科

People Re-identification in Multi-Camera based on Co-visible Features

Masaya ITOH[†], Shun-ichi KANEKO^{††} (Member)[†]Center for Technology Innovation - Controls, Research & Development Group, Hitachi, Ltd.,
^{††}Graduate School of Information Science and Technology, Hokkaido University

あらまし セキュリティやマーケティングのための監視カメラシステムの分野において、人物検出や追跡のための画像認識技術に対する期待が高い。本論文では、視野の共有しない複数のカメラネットワークから取得した映像から、同一人物を追跡する手法について提案する。多くの監視システムにおいて、複数のカメラの環境から個人を特定することが課題である。しかしながら、視野を共有しない環境において、視点や姿勢の不明な状況下では人物の特徴が安定しないことにより、人物を照合することは大変困難である。そこで、本論文では人物間の観察方向の共有領域に着目した、共可視特徴に基づく複数カメラ間での人物同定手法を提案する。本手法を公開データセット (VIPeR) を用いた累積照合特性により検証した結果、従来手法と比較して観察方向の違いに対して安定に人物同定が可能となり、提案手法の有効性を確認した。

キーワード：人物同定, 人物検出, 監視カメラシステム

<Summary> Computer vision technologies for people detection and tracking are widely required in various fields such as surveillance system applications for security system and marketing services. In this paper, we propose a method for tracking a person using several video frames in multi-camera network with non-overlapping views. In many surveillance camera systems, it is desirable to determine if a given individual has been previously observed multi-camera environment. However, since the human features are unstable under the change of viewpoint and posture, this is likely due to the difficulty in identifying individuals with non-overlapping cameras. Therefore, this paper presents a method of people re-identification in multi-camera based on co-visible features using observed crossover region between people. The method is evaluated using a public viewpoint invariant pedestrian recognition dataset (VIPeR) and the results are shown to be superior to the previous benchmarks.

Keywords: person identification, human tracking, surveillance camera systems

1. はじめに

近年、全世界的なテロの増加や犯罪の多様化を受けて社会的不安が高まっている。また、業務のIT化が進む中で、内部統制や個人情報保護の観点などからセキュリティへの意識が向上している。そのため、映像監視システムやアクセスコントロールシステムなどの防犯設備への需要が高く、特に防犯カメラの設置台数が急激に伸びている。映像監視システムは、ビル・オフィス、流通小売業、金融機関、公共機関などへ導入され、犯罪の抑止効果や有事の際の証拠としての使い方に加えて、マーケティングや社員教育を目的として利用方

法も多様となっている。例えば、犯罪発生時に不審者を特定する場合や、人物の挙動を観察するためには、取得した膨大な画像データを分析する必要があり、監視員の負担低減が最も重要な課題であった。これに対して、映像認識アプリケーションによる自動監視や検索支援システムが提案されてきた。なかでも、特定の人物を追跡し、移動軌跡を取得する技術は必要性の高い技術の一つである。これにより、特定の人物がどのように施設内を移動したかなどの移動経路を取得することがきるため、窃盗や万引きの際に、容疑者の逃走経路を明らかにすることができる。また、マーケティングでは、特定顧客の来店時の行動を捉えることで、購買に至るまでの行動

履歴を知ることができ、店舗の改善施策などに活用することが期待できる。

このようなアプリケーションを実現するためには、単一カメラ内での行動のみならず、複数のカメラが設置された広範な対象エリア内で特定人物を追跡することが必要である。言い換えれば、複数のカメラ間を横断して移動する人物を対応付けて追従する技術を要する。これは、例えば、カメラ間で視野を共有する部分を頼りに、あるカメラ内での人物の追跡結果をカメラ間で受け渡すなどして、比較的容易に実現することができる。

一方で、対象エリア全体の死角を少なくし、セキュリティレベルを高めたシステムを構築するためには、膨大な数の監視カメラを密に配置することが必要とされる。これは、ユーザのコスト負担や密に設置されたカメラ群の外観などを考慮すると現実的ではない。したがって、多くのセキュリティシステムでは、視野を共有しない複数の監視カメラ間で、特定の人物の対応付けを行いつつ、広範囲に人物を追跡することが求められる。

本論文では、このような視野共有のない複数カメラ間の人物追跡について扱う。この人物追跡については、あるカメラで取得した人物と、異なる視野のカメラで取得した人物の情報を照合する問題として、人物同定と呼ばれる。人物同定の定義は、取得した日時や対象エリアに限らずに、特定の人物に関連づける広義の認証問題としても用いられるが、本論文では、特定の期間を対象とし、屋内施設などの限定的なエリア内での問題として取り扱う。

カメラ間の人物同定を実現するには、顔認証¹⁾、歩容認証²⁾の既存手法を利用することも考えられる。顔認証を用いた方法は、基本的には顔が正面を向いている必要がある。また歩容認識は、人間の歩き方の違いを頼りに人物を特定する方法であるが、動画像に基づく方法であるため、サンプリングレートや大量カメラを用いたときに計算量が増える問題がある。さらに、カメラの配置関係に基づいて人物の対応付けを行う手法では、カメラ間での移動速度などを推定するため、その間での行動が一様であることが前提であったり、通路の分岐がある程度単純にモデル化できる必要がある。また、服装などの色情報に基づいて人物を対応づける手法は広く採用されているが、多様なカメラの設置条件下では設置の制約もあるため、人物の観察される向きや大きさなどの条件が自由に決定できないことを考慮しなければならない。例えば、人物の撮影方向と移動方向の関係により取得される色特徴も異なり、人物同定の精度が低下するなどの課題がある³⁾。

そこで本論文では、色情報に基づく手法における、人物の観測方向の影響を改善するために、共可視特徴に基づく人物同定手法を提案する。共可視は共可視領域とも言い換えることができ、複数のカメラ間で異なる時刻で観測される人物において、それぞれのカメラで共通に観測できる外観の部分領域とする。その領域で取得された画像特徴を共可視特徴とし

て定義する。カメラからの観測方向と人物の移動方向から共可視領域を推定し、そこから取得した共可視特徴を用いて人物を照合する。すなわち、本論文の提案手法は、対象となる人物間の観測方向を考慮した部分領域への分割手法と、共可視に基づく部分領域間の照合方法に着目した手法である。これにより、人物の観測方向の変化による性能の低下の問題に対して、カメラ配置と人物の移動方向を考慮した安定な照合手法を提供し、視野を共有しない複数カメラでの人物追跡を実現する。

本論文の構成は次のとおりである。2章では、マルチカメラの人物追跡における人物同定の定義、関連研究とその問題点について述べる。3章では、共可視の定義と共可視特徴の抽出方法について説明し、共可視特徴を用いた人物照合手法について述べる。次に、4章では、3章にて詳述した提案手法について、評価データによってその実用性について確認し、5章では、本研究の結論と今後の課題について述べる。

2. マルチカメラ人物追跡

2.1 人物同定の定義

本稿では、複数カメラ間の人物を追跡し、人物の移動軌跡を取得する要求に対して、視野を共有しない複数カメラでの人物の特徴を対応づける人物同定の問題を扱う。ここで取り扱う問題の定義について図1に示す。カメラ C_i, C_j に対して、それぞれのカメラから取得する映像の撮像範囲である視野を A_i, A_j とする。ここで、各カメラから取得する視野は $A_i \cap A_j = \phi$ となり共有されない。カメラが設置された環境では複数の人物の往来がある中で、カメラ C_i に出現する m 番目の人物 p_m^i と、カメラ C_j に出現する n 番目の人物 p_n^j が同一人物 ($p_m^i = p_n^j$) であった場合であっても、各カメラで同時刻に出現することはない。それにより、カメラの幾何構造に基づいて各々のカメラで同時刻に取得した人物を対応付ける手法を適用することができない。また、図1における p_m^i と p_n^j とはカメラによる撮影条件が異なるため、カメラから得られる特徴も異なる。その上で、 p_m^i と p_n^j が同一人物であるか否かを判断する必要がある。さらに、1対1の対応ではなく、任意のカメラ C_k で撮影された一定の時間内に取得された複数の人物データ $P^k = \{p_l^k | k = 1, \dots, K; l = 1, \dots, L\}$ から注目する人物を特定することが求められる。

図2に人物同定の概要を示す。図2では、あるカメラで撮影された人物のクエリー画像から、人物画像データベースに登録された別のカメラで撮影された人物とを照合する手続きを示している。クエリー画像から任意の画像特徴を抽出する。人物 p_m^i の画像から取得した複数の特徴ヒストグラムで構成される特徴 f_m^i 、画像データベースから選択した参照データ p_n^j の特徴を f_n^j とする。 f_m^i と f_n^j の一致度について、任意の距離計算や類似度計算によって算出する。すべての参照データについて同様に計算し、データベース内のデータに順位づけを行う。結果の出力の方法として、例えば、類

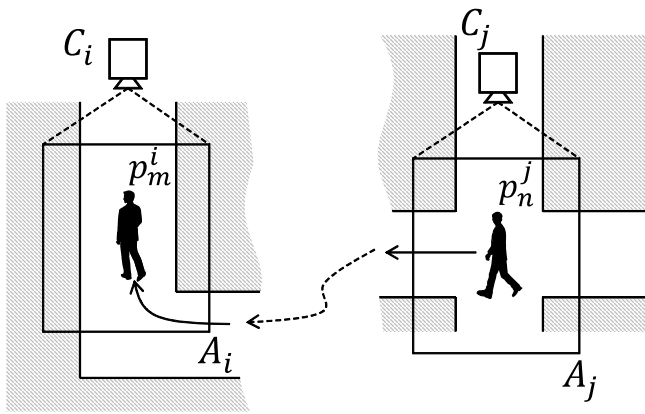


図1 問題の定義
Fig. 1 problem definition

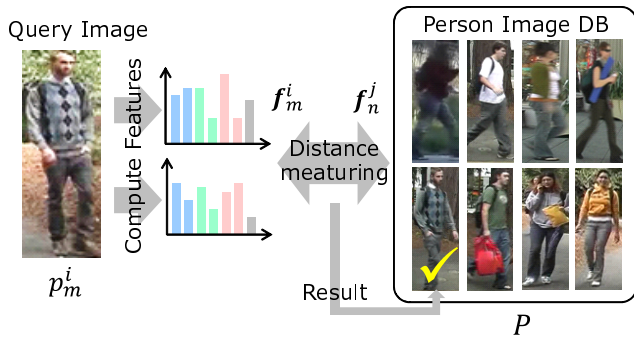


図2 画像特徴に基づく人物照合
Fig. 2 people re-identification based on image features

似度の上位数を候補として提示して、ユーザとのインタラクション作業を介し、候補から対象人物を選定するようなアプリケーションが考えられる⁴⁾。

2.2 画像特徴に基づく人物同定手法

異なるカメラ間での人物同定において画像特徴に基づく手法は多数提案されており、(i) 顔認証^{5)~8)}、(ii) 歩容認証^{9)~13)} (iii) 色やテクスチャなどを用いた方法^{14)~21)}、に大別できる。ここでは、画像特徴に基づく関連研究について順に説明する。

- (i) 顔認証は、外観から個人を識別する手法として認証率が高く広く普及している。しかし、正面から撮影された顔画像を前提としており、防犯カメラなど、俯角のあるカメラ設置では頭頂部しか撮影できない場合も多く、人物の移動方向も様々であるため、人物の理想的な正面顔が取得できる場合が少ない。また、遠方を通過する場合や、解像度や画質が低い場合などは、認証に必要な情報が欠落するために精度が低下することがある。このために、顔認証の利用は特定の状況にのみ有効であって、他の人物同定の方法との組合せが必要とされる。
- (ii) 歩容認証は、歩容 (Gait) と呼ばれる歩き方を違いによって人物の同定を行う方法である。歩容認証は、より容易

に導入できるアピアランスベースの手法が適用されることが多い。これは、人物の移動する動画から人物領域を抽出し、人物シルエットを用いた時系列情報によって、周波数解析などにより人物の歩行の特徴を求める。この手法では、歩き方を示す特徴が取得できれば、低解像の画像であっても比較的安定して照合することができる一方で、人物の抽出の精度が外乱光やシーンによって低下する場合がある。また、時系列データを扱うため、データの取得レートによっては信頼性が低下する。さらに、動画を扱う点からカメラ台数が増えると計算量が膨大になるという点を考慮しなければならない。

- (iii) 色やテクスチャを用いる手法は、複数カメラ間の人物同定を解決する手段として多く提案されている。色特徴は人物の姿勢などの変化に対して大きな変化がなく、比較的特徴量も安定しているため、多くのアプリケーションに適用できる。このように、色ヒストグラムは大域的な特徴が捉えられる一方で、類似の服飾により精度低下をもたらす場合がある。そこで、高さ方向に領域を分割する方法¹⁵⁾や、人物領域を分割して特徴記述をする方法¹⁶⁾、特徴的な色に着目した手法¹⁷⁾が提案されている。また、カメラ間での明るさやカラーバランスの設定によって、同一色であっても画像データ上では異なる問題があり、これに対して、カラーキャリブレーション¹⁸⁾が提案されている。さらに、服飾のテクスチャに注目し照合する手法も多く提案されている。テクスチャ情報として局所特徴を抽出し動画から人物同定を行う手法¹⁹⁾、色特徴とテクスチャ特徴を組み合わせ、領域の分割方法を工夫することで精度向上を図る研究も提案されており²⁰⁾、さらには、色特徴と画像中の空間的な特徴を表す色コリログラムとを組み合わせた手法²¹⁾が提案されている。色特徴の照合問題として、距離指標の学習によってカメラによる影響を受けにくくする手法²²⁾も提案されている。

このように先行研究が多数提案されており、実用化の段階では想定されるシーンにより、これらの手法を組み合わせることでアプリケーションの高度化を試みている^{3);23);24)}。関連研究で分類したアプローチのうち、(iii)の色特徴を用いた手法が最も基本となっている。そのため、色特徴の抽出方法や照合の精度が、アプリケーション全体の精度へ大きく寄与しており、その改善への期待は大きい。そこで本論文においても、最も汎用性の高い色特徴を人物照合に採用する。

なお、本研究の成果が適用されるアプリケーションは、特定施設内での人物の行動を追跡することで、不審者の移動経路を特定することや、マーケティング向けの購買行動観察などを想定している。これにより、人物の追跡を行う対象エリア及び時間については、施設内などの一定の空間内であり、また日をまたがない期間とする。これは、行動範囲の拡大や時間経過によって服装などの変化が発生し、その場合に複数の人物情報を対応づけることは、個人認証問題の領域となり

うるためであり、本論文では取り扱わない。

2.3 問題点

色特徴を用いた人物同定について、その問題点について述べる。色ヒストグラムに基づく人物同定については、撮影条件が異なる画像データに対して照合を行うという問題がある。これは、人物の移動とカメラの撮影方向によって、撮影される人物の向き（観察方向）が異なり、これにより見かけの色特徴に変化が生じることが考えられる。本論文では、カメラにより確認できる部分を観察方向、人物の動きの方向を移動方向として区別して表現する。例えば、あるカメラではカメラに向かって移動する人物に対して、他方では背を向けた状態で映像が撮影されたとする。同一人物であってもカメラに対する移動方向の違いによって観察方向が異なるため、正面の背面の服飾の色特徴の違いが発生し類似度が低下する場合があります。他者を上位に選択することが考えられる。

具体例を図3を用いて説明する。図3は2台のカメラから取得した人物の画像を並べ、観察方向と色特徴について例示している。人物の移動方向について目視で求め、人物を俯瞰した視点で移動方向を矢印にて表現している。この場合、カメラを下から撮影しているものとして示している。図3の(a)はほぼ同じ方向の撮影条件であるが、(b)(c)については、異なる方向となっている。色特徴について、(a)はほぼ同等であり、色ヒストグラム間の距離を任意の指標で算出した場合、その距離が近いことを示す。一方で、(b)(c)についてはヒストグラム間の距離が離れている。これは、人物の観察方向が異なることにより、持ち物や服飾の特定部位が出現するもしくは隠蔽されることによるためである。これは、一般的に自己オクルージョンと呼ばれ、同一物体の一部によりその物体の一部が遮蔽されるという現象である。この問題に対して観察方向の安定性を向上させるべく、カメラ配置などで解決することは現実的ではなく、観測方向が多様であることを考慮しなければならない。実環境での適用を考えた場合、ある人物の服飾や所有物を含めた色特徴が全方位で一様であるという前提に基づく手法では、前述のように精度低下につながる場合がある。

本論文で提案する共可視特徴に基づく人物同定手法は、複数のカメラ間で取得した画像情報に対して、各カメラで共に観測可能な共可視の領域に着目して人物データの類似度を取得する。上述の問題のうち、観測方向が異なることに起因する精度低下を改善することができる。一方で、人物のある向きにおいて観察可能で、識別に適した顕著な特徴を能動的に利用することも可能である。例えば、所持品の情報を頼りとして人物同定の適合率を向上させる効果も期待できる。本手法は、領域分割及び照合方法において共可視と呼ぶ新たな概念を導入するもので、領域内の画像特徴の表現方法や部分領域間の類似性評価については既存の手法と組み合わせることにより精度向上が期待できるものである。次章より提案手法の詳細について述べる。

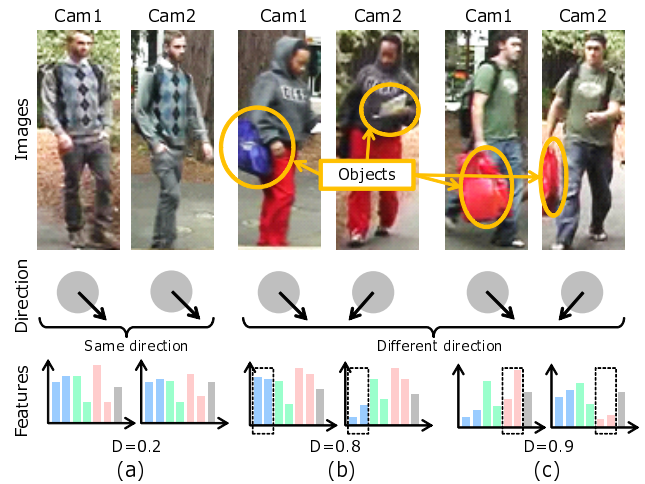


図3 観測方向による人物照合の違い
Fig. 3 Difference in results of people re-identification in each view angle

細について述べる。

3. 共可視特徴による人物同定

本章では、共可視特徴の概念とそれに基づく特徴抽出方法、及び照合方法について順に説明する。

3.1 概要

複数カメラをまたがって移動する人物を追跡するアプリケーションにおいて、本稿の提案手法を用いた手法では、次の手順により人物追跡を実現する。

- (1) 単一のカメラ C_i 内において任意の手法によってある人物 p_i の位置を検出し、時系列フレームに渡って追跡処理を実行する。
- (2) 人物の追跡結果によって、人物 p_i の移動方向 d_i を取得する。
- (3) 任意のフレームにおいて人物 p_i を含む領域の画像 S^i を抽出し、移動方向 d_i とともに人物データベースへ登録する。この場合、背景差分法などの手法によって、人物の前景領域と背景領域を分離する。
- (4) 人物 p_i と同一の人物を、人物データベース内の人物データ P_k から求めるため人物照合処理を実行する。この際に、カメラ C_i 及び C_j より取得された人物 p_i, p_j の2つのデータ間にて、移動方向の差分に基づき共可視領域 \tilde{V} を取得し共可視特徴を求め、それらを照合することで類似度 $R_{i,j}$ を取得する。
- (5) すべての人物データから最も類似度の高い人物ペアを選択し、一定以上の信頼度が得られた場合に、各カメラの追跡結果の結合を行う。

以下、人物領域のセグメンテーション画像及び移動方向は、例えば和島らの手法²⁵⁾などを用いて取得されたものとして、

- (4) 以降の処理について詳細に述べる。

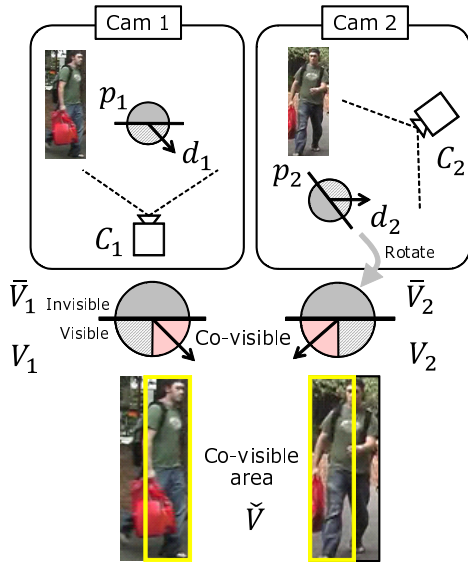


図4 共可視領域
Fig. 4 Co- visible area

3.2 共可視領域の定義

図4を参照しながら、共可視について述べる。2台のカメラによって取得された同一の人物について、カメラ C_1 で観測された人物 p_1 とカメラ C_2 で観測された人物 p_2 について考える。図4左上部は、 p_1 は C_1 に対して移動方向 d_1 で移動しており、実空間の俯瞰視点によって表現したものである。この場合、カメラ C_1 側の観測可能な領域を可視領域 V_1 、自己オクルージョンにより観測ができない領域を不可視領域 \bar{V}_1 とする。カメラと可視領域の関係については、人物の移動方向 d_1 によらずカメラ C_1 と人物 p_1 との位置関係により決定され、カメラ側が可視領域となるのは明らかである。一方で、人物の移動方向 d_1 を考慮すると、人物 p_1 の可視領域は、カメラの撮影方向と移動方向の関係により異なる。すなわち、カメラと人物の移動方向が決定されれば、カメラで撮像される人物の領域が決定される。同様にカメラ C_2 においても、移動方向 d_2 と可視領域の関係が得られる。

ここで、 p_1, p_2 についてカメラの撮影方向が同一となるように回転し可視領域を合わせた上で比較すると、図4に示す人物では正面方向から右半身にかけての部位がカメラによらず共通に観察される領域であることがわかる。これを本稿では共可視領域 (Co-visible area) \check{V} と定義する。カメラによる撮影方向と人物の移動方向とを考慮することで、可視領域における複数カメラとの共可視領域を取得することができ、人物照合のための特徴量取得を安定に行うことができる。

3.3 共可視特徴の抽出

3.2節において述べた、共可視領域の考えに基づく特徴抽出について説明する。まず、複数カメラで取得した人物の一致度を共可視領域に注目し求めるにあたり、共可視領域を分割する。図5に示すように、カメラ C_i で撮影した人物 p_i が d_i 方向に移動する関係を示している。ここで、カメラ C_i で

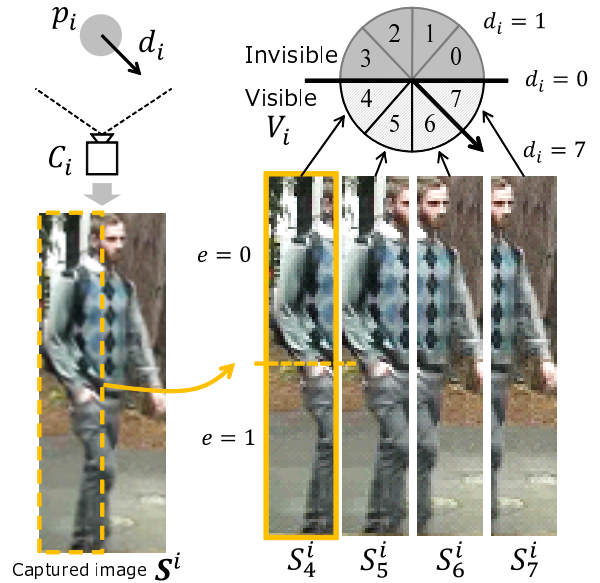


図5 共可視の特徴抽出
Fig. 5 E traction of co- visible features

取得した画像データから任意の手法によって人物領域を抽出したものを S^i とする。この S^i から共可視領域間での照合を行うために、水平方向に分割した部分画像を取得することを考える。

そのため、直立する人物の3次元形状を最も単純な円筒モデルと仮定して定義する。移動方向 $d_i = 0$ を基準として、この円筒モデルの中心軸回りに、分割数 N の量子化幅 $\Delta\theta$ に量子化し、符号 c_i を割り当てる。なお、特定のカメラに着目し比較などを必要とする場合を除き、本稿では単に c と表現する。また、移動方向についても同様に、 $d_i = 0, \dots, N-1$ とする。また、本稿では、実験的に得られた知見から分割数 $N = 8$ とする。

次に、可視領域 V_i における共可視領域 S^i を、円筒モデルを分割した方向ごとに部分画像を取得することで、 $S^i = \{S_c^i | c = 0, 1, \dots, 7\} \in V_i$ のように求める。画像分割の方法を考える場合、取得される画像において、人物の服飾などの色特徴が円筒モデルの円筒面へ投影されたものと仮定し、カメラで撮像した画像平面と円筒モデルとの関係により算出することが可能である。ところが、人物の動きや移動について手足などの動きの自由度や姿勢の揺らぎが大きく、本稿では厳密な記述方式は取らずに部分領域画像を水平方向に均等に分割する。画像の分割については、人物の姿勢や人物領域の抽出位置のずれを考慮して、領域ごとに一定の重複をもって分割する。本稿では画像データの画像幅 w に対して、 $w/2$ を各分割領域の幅として、ウィンドウをスライドさせ取得する。

各共可視領域 S^i における各領域の特徴量 f^i について、特徴抽出の方法を述べる。人物の服飾の色特徴の抽出では、RGBのみならず様々な手法が利用されている。本論文の提案手法では、特徴については限定していないが、評価においてはHSV表色系に基づくHSヒストグラムが従来研究にお

いても比較的高精度とされているためこれを用いる³⁾。また、人物の服飾については上下の組み合わせなど、垂直方向の組み合わせが人物の外観を示す情報であると考えられる。垂直方向での分割と色ヒストグラムの抽出については従来手法で提案されており¹⁵⁾、共可視領域における特徴抽出においてもこの手法を導入する。垂直方向に各領域 $e(e = 0, \dots, E - 1)$ に分割した領域ごとに特徴量をもとめ、これらを連結した $f_c^i = \{f_{c,1}^i, \dots, f_{c,E}^i\}$ を求める。本論文では、経験的に人物の領域を上半身と下半身に2分割 ($E = 2$) する。色相 H 、彩度 S 成分のヒストグラムをそれぞれピン数 b_h, b_s として求め、1次元のベクトル (次元数: $b_h \times b_s \times E$) として並び替える。そして、共可視領域の部分領域全体について、求めた f_c^i について、 $\sum_e f_{c,e}^i = 1$ となるように正規化処理を行う。以降、領域 c ごとの色特徴については、特に注目しない限り、垂直方向へ分割したものを統合したヒストグラムで構成される特徴量 f_c^i として説明する。

なお、本手法の適用が想定されるアプリケーションにおいて、カメラ配置によっては人物画像に傾きや歪みが生じる場合がある。そのため、人物画像の取得もしくは円筒モデルの当てはめについて、本手法を適用するために事前に補正処理を実施する必要があると考える。本論文では上記については今後の課題とし、カメラに対して傾きの少ない正対する人物画像を用いて提案手法の特徴について述べる。

3.4 近似特徴量の算出

人物の観察方向と共可視領域との関係及び特徴抽出の方法について述べる。図6は、カメラ C_i と C_j で観察された人物 p_i, p_j について、それぞれの移動方向 d_i, d_j とした場合の共可視領域の変化について一部を示している。なお、組み合わせごとの表現については、 d_i を基準として表している。共可視領域 \check{V} は、カメラに対する人物の移動方向 d_i と d_j の差分 Δ_c に応じて決定される。例えば、 $\Delta_c = 0$ では、完全一致となるが、 $\Delta_c = 2$ においては可視領域のうち2領域分のみが共可視領域となる。

しかしながら、このように観察方向の関係性において、共可視領域の割合が減少するにしがたい特徴取得が不安定となるため、照合率の低下をもたらす場合がある。そこで、観測可能な共可視領域の特徴を用いて、自己オクルージョンによって観察ができない領域の近似特徴量の取得を行う。

ここでは、人物 p_j の可視領域 V_j における特徴量を基準として、人物 p_i の特徴量の生成方法について考える。 p_j で可視領域 V_j に対応する、 p_i の特徴量 $f^i = \{f_{c_i}^i | c_i \in V_j\}$ のうち $c_i \in \check{V}_i$ となる位置、すなわち C_i からは観察できない領域での近似特徴量 $\hat{f}_{c_i}^i$ を式(1)で求める。

$$f_{c_i}^i = \begin{cases} f_7^i & c_i = 0 & (a) \\ f_{c_i+4}^i & 1 \leq c_i \leq 2 & (b) \\ f_4^i & c_i = 3 & (c) \end{cases} \quad (1)$$

式(1)のそれぞれの条件について述べる。まず (a) 及び (c)

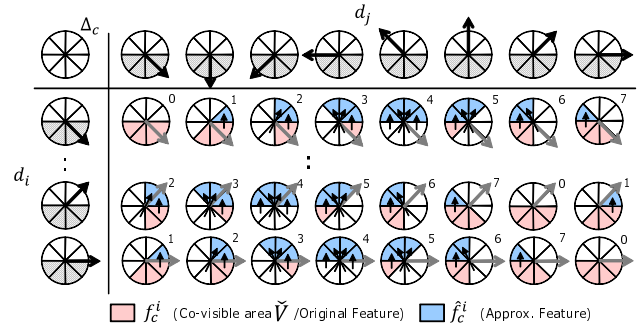


図6 共可視領域の例と近似特徴量
Fig. 6 Example of co-visible area and feature approximation

は、可視領域に隣接する不可視領域での場合である。この場合、人物の服飾特徴や所持品の特徴が隣接する可視領域にも出現している可能性が高く、可視領域の特徴を近似的に利用することができる。一方で、(b) の場合は、完全に不可視である領域である。この場合、人物の服飾の色特徴などが一定の確率で対称性を持つとの仮定から、可視領域における対角に位置する領域の特徴を遷移させる。図6に共可視特徴量の補完の関係について示している。前述の近似特徴量の抽出は、対象となる人物間の移動方向の差分 Δ_c によって決定されることがわかる。つまり、 Δ_c が同じであれば、それに伴う特徴量の生成の手続きについても同様となる。

3.5 照合

人物間の照合については、前節までに述べた共可視特徴により得られた特徴量を用いて、任意の距離指標によって照合する。図7に概要を図示する。ここでは、 C_i, C_j により抽出された人物領域 S^i, S^j から人物を照合する場合を考える。それぞれ、 $d_i = 7, d_j = 5$ の方向に移動している。共可視領域を部分領域に分割し各領域から色特徴 f_c^i, f_c^j を抽出する。図7では、カメラ C_j の可視領域に対して、カメラ C_i の共可視特徴及び近似特徴量を算出し、 \hat{f}^i を求める手続きについて示している。

人物 p_j における $c_j = 4, 5$ の領域は共可視領域であるので、カメラ C_i で取得した特徴量は $f_{4+\Delta_c}^i = f_6^i, f_{5+\Delta_c}^i = f_7^i$ のように、取得した特徴量を照合に直接利用する。それに対して、 $c_j = 6, 7$ の場合は、対応する人物 p_i の特徴量 $f_{6+\Delta_c}^i, f_{7+\Delta_c}^i$ は不可視領域に存在する。そのため、 C_i で取得可能な可視領域における特徴を用いて近似特徴量を生成する。それぞれ、式(1)を用いて、 $f_{6+\Delta_c}^i = f_7^i, f_{7+\Delta_c}^i = f_{7+\Delta_c+4}^i = f_5^i$ のように求める。上述の説明では、ある特定の方向の組み合わせを例として述べたが、その他の組み合わせにおいても同様の考え方により特徴抽出を行う。

次に、各方向符号ごとに対となる特徴量間における色特徴の照合について述べる。色特徴の類似度 r_c を式(2)より求める。

$$r_c = w_\delta \cdot D(f_{c+\Delta_c}^i, f_c^j) \quad (2)$$

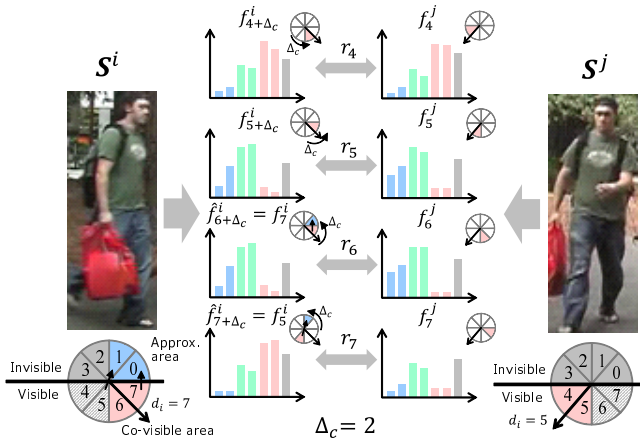


図7 共可視特徴による照合方法

Fig. 7 roposed matching process using co- isible features

式 (2) はカメラ C_j を基準として、カメラ C_i の特徴を生成した場合の記述である。ここで、 w_δ は、観察方向の差分 Δ_c と方向符号 c によって決定される特徴量の信頼度を考慮した重み係数である。前節にて述べた近似特徴量は、近接する領域と、完全に隠蔽される領域の特徴量を取得する場合との2つの条件にて信頼度が異なる。例えば前者の近接した特徴を参照する式 (1) の (a)(c) の場合には、取得した特徴量はある程度の信頼度が確保できると考えられる。しかし、(1) の (b) のように、もう一方の完全に隠蔽された領域では、人物の服飾などが前後左右である程度対象性を持つという仮定を基に、対角に位置する領域から推定している。そのため、前者と比較して信頼度を低く設定する。つまり、 w_δ は、共可視領域の場合と近似特徴量の算出方法による2種類の合計3種類の値を持つ。 $D(\bullet, \bullet)$ は、ヒストグラムの一致度を示す距離計算及び類似度計算を示す。ヒストグラムの距離計算では、ヒストグラム交差値や χ^2 統計量、など様々な方法が考えられる。本稿では、関連研究で広く利用される式 (3) に示す Bhattacharyya 距離指標⁽²⁶⁾を用いる。

$$D(a, b) = \sum_i \sqrt{a_i b_i} \quad (3)$$

式 (2) で算出した r_c を用いて、 S^i と S^j の人物の一致度を式 (4) により求める。

$$R_{i,j} = \sum_{c \in V_j} r_c = \sum_{c \in V_j} w_\delta \cdot D(f_{c+\Delta_c}^i, f_c^j) \quad (4)$$

$R_{i,j}$ は、カメラ C_j を基準とした人物照合結果であるが、参照側を入れ替え $R_{j,i}$ についても同様に求める。カメラ間での人物間の最終的な類似度は、 $R_{i,j}$ は、 $R_{i,j} = R_{i,j} + R_{j,i}$ とする。例えば、水平方向への分割数が $N = 8$ の場合、 $R_{i,j}$ は8組みの共可視領域を含む部分領域間の照合結果を統合したものとなる。



図8 I eR データセット

Fig. 8 E amples of I eR dataset

表1 I eR データセットにおける観察方向の分布
Table 1 The distribution of viewpoint angles in the I eR dataset

| d_i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | total |
|-------|---|----|-----|----|-----|----|-----|----|-------|
| camA | 3 | 0 | 1 | 5 | 49 | 78 | 456 | 40 | 632 |
| cam | 0 | 38 | 242 | 52 | 268 | 30 | 1 | 1 | 632 |

4. 実 験

4.1 評価データ及び実験方法

提案手法である共可視に基づく領域分割及び特徴照合の有効性を確認するため、人物照合の実験を行う。利用するデータセットは、図8に示す Gray らによる VIPeR (Viewpoint Invariant PEdestrian Recognition) DB⁽²⁷⁾を用いる。このデータは、人物同定問題に対する研究では最も良く利用されるもので、2つの異なるカメラから撮像された人物データのペアが632枚あり、様々な方向から撮像されている。なお、このデータの撮影場所などは明示されておらず、視野を共有せず時間的制約条件などを活用できないことから、本論文の問題設定に合致するものである。このデータは照明条件がカメラごとに異なるため、輝度伝播関数⁽²⁸⁾などの色補正が難しく、視野を共有しない人物同定問題としては現実的なデータセットの一つである。

本実験では、このデータセットの人物について、人物の移動方向を手動で45度ごとに分類した。その結果を表1に示す。実アプリケーションにおいては人物の移動方向はカメラ内追跡によってある程度自動的に導出できるが、誤差も一定以上発生する。手動で分類する場合であっても作業者の認識の違いによって誤差が発生し得ることも考慮すると、両者の相違は大きく発生しないと考える。また、同一人物間での移動方向の差分 Δ_c ごとに計数した結果を表2に示す。移動方

表2 人物ペア間における観察方向差分 Δ_c の分布

Table 2 The distribution of difference of viewpoint angles Δ_c in people pairs

| | | | | | | | | |
|--------------|----|-----|----|-----|----|----|---|---|
| Δ_c | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| um. of pairs | 31 | 272 | 92 | 165 | 59 | 11 | 2 | 0 |

向の差分については、比較的少ないデータが多いが、極端な偏りはないデータセットと考える。

評価データにおける背景情報は、人物照合の精度検証において影響を与えるため、人物領域のセグメンテーションについても手動で実施している。これは実際に人物位置を検出する際においても、背景差分法や色セグメンテーションなどによって人物領域を抽出することができるため、実用を考慮しても問題はないと考える。ただし、その精度については環境によっては低下することがあるため、今回の実験方法については比較的良好な状況における実験と考えられる。なお、評価データのうち、本論文の検証に影響を与えられとされる232人分のデータについて、ノイズが多いもの(約9)や、色校正が大きく離れているもの(約22)、回転方向ではなく傾きを多く含むもの(約5)を除外し、合計400人分のデータにより評価データを構成した。

評価指標としては、バイオメトリクス認証の評価などで最も一般的である累積照合特性(CMC Cumulative Match Characteristics)を用いた^{29),30)}。CMCは、1対 n の照合問題において、本人ペアの順位に基づく照合率によって評価する。評価データベース内のそれぞれのペアに対して照合結果を求めて、本人ペアの順位がある順位以内となる確率が得られる。CMCは曲線グラフとして表現され、横軸には注目する順位 X で、これは評価データベースに登録された人数まで表すことができる。縦軸は X 位以内の本人ペアの照合率となる。評価では特定の順位に着目して、1位照合率(CMC rank=1)、5位以内照合率(CMC rank=5)、10位以内照合率(CMC rank=10)について比較した。

特徴抽出については、HSV表色系の色相 H 及び彩度 S ごとに特徴量を求め、それぞれを正規化して結合した色特徴ヒストグラムを利用する。この場合のパラメータ設定について、色特徴ヒストグラムのビン数は予備実験において良好であった $b_h = 8, b_s = 8$ と設定した。また、垂直方向での分割数は $E = 2$ として上半身と下半身での分割とした。共可視特徴の水平方向での分割数については、 $N = 8$ としている。さらに、共可視特徴における重み係数 w_δ については、共可視領域が隣接する場合、対角に位置する場合について、評価データからランダムに抽出した25枚を用いて、照合率が高くなるように $w_\delta = \{1.0, 1.1, 1.2\}$ と設定した。共可視の隣接性や対角に位置する場合の重み係数の決定は、撮影環境の整備された様々な観察方向のデータセットが必要であり、この設計については今後の課題としたい。

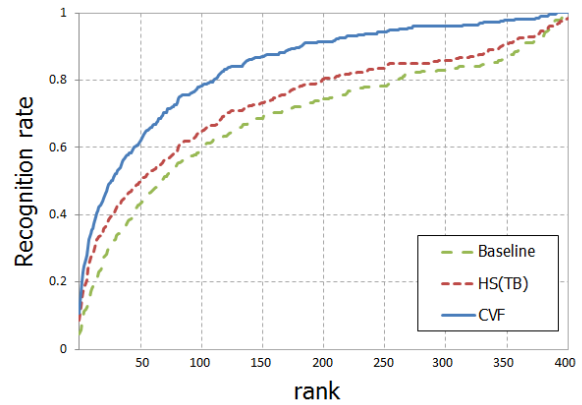


図9 CMC曲線による比較
Fig. 9 Comparison result of CMC curves

表3 累積照合特性の実験結果

Table 3 Experimental result of cumulative match characteristics

| CMC rank | 1 | 5 | 10 |
|---------------------------------------|--------------|--------------|--------------|
| Baseline | 0.045 | 0.113 | 0.170 |
| HS(TB)($E = 2$) | 0.088 | 0.183 | 0.273 |
| CVF($E = 2, N = 8$) | 0.108 | 0.248 | 0.343 |

4.2 結果及び考察

図9に評価結果のCMC曲線を示す。比較対象として、画面全体についてHSV表色系のHSヒストグラム取得し、Bhattacharyya距離指標によって照合したものを基準(図9のBaseline)とした。さらに、同じくHSヒストグラムを用い、画像を垂直方向に上下分割($E = 2$)し特徴抽出したものをHS(TB Top, Bottom)とし、図9ではHS(TB)と示す。これらと提案手法である共可視特徴(図9のCVF Co-Visible Features)による結果を比較した。この結果によると、照合順位の全体にわたって、従来手法と比較して提案手法が精度が高いことがわかる。

各手法ごとの1位照合率(CMC rank=1)、5位以内照合率(CMC rank=5)、10位以内照合率(CMC rank=10)の結果を表3に示す。手法の略記については、図9と同様である。それぞれのCMC rankにおいても、共可視特徴(CVF)の結果が上回っており提案手法の有効性を表しており、1位照合率は10.8%、10位照合率では34.3%の結果を得た。以上の結果から、従来の共可視を考慮しない手法と比較して、提案手法の領域分割及び照合方法が有効であることが確認できた。

なお、既存手法のなかでも最も精度の高いといわれる手法²⁰⁾であっても、VIPeRデータセットにおいては、1位照合率は20%程度にとどまる。この手法は、Maximally Stable Color Region (MSCR)と呼ぶ顕著性を考慮した領域内の特徴抽出と、Recurrent High-Structured Patches (RHSP)を用いた特徴ベクトル間の照合方法を実施している。事前処理の領域分割においては上下左右の分割を実施しているのみである。つまり、本論文の共可視特徴抽出及び照合方法と組み合わせることにより、大幅な精度向上が期待できるものと考

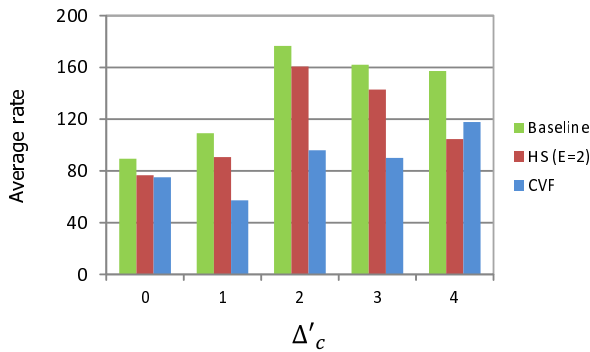


図 10 共可視領域別の順位平均値
Fig. 10 Average rate in each co-visible area

える。

次に、共可視特徴を用いた場合の効果について、観察方向の違いによる性能の変化について考察する。図 10 にその結果を示す。図 10 では、移動方向の差分 Δ_c について、 $\Delta'_c = \Delta_c - 4$ (if $\Delta_c \geq 5$) とし、 Δ'_c の値、すなわち可視領域の重複度合ごとに比較手法との違いをまとめた。 $\Delta'_c = 0$ については同じ観察方向の人物間の照合であるため大きな差は発生しない。 $\Delta'_c = 1$ の場合も向きが近い場合 CMC ran の平均値も比較的低い値を保っている。一方で、 $\Delta'_c = 2, 3$ の場合には、観察方向が異なるため、CMC ran 平均値も低下しているなかで、提案手法では従来手法と比較して改善されていることがわかる。共可視領域に注目した照合によって、観察可能な領域を切り所として照合精度の低下を抑えたものとする。ところが、 $\Delta'_c = 4$ においては逆に精度低下する結果となった。これは、観察方向が完全に逆となることで、共可視特徴の効果を得られなかったためである。分割された可視領域のすべてが近似特徴量となるため、重み係数 w_δ によって類似度が低く抑えられてしまう影響がある。本論文では、重み係数 w_δ を近似特徴量を取得する参照領域との位置関係により決定している。しかしながら、 Δ'_c や人物の撮影方向 (撮影される部位) によって、照合精度への影響度が異なるため、重み係数 w_δ を含めた照合方法を変更する必要があると考える。共可視領域や観察方向による照合精度への影響については、撮影環境を整えたデータベースを構築した上で今後検証を試みたい。

以上の結果から、提案手法の共可視特徴に基づく人物照合においては、観察方向と共可視に着目した照合方法により、人物の移動やカメラによる撮像方向による人物照合の性能低下を低減させることができ、視野を共有しない複数カメラでの人物照合に対する有効性を確認することができた。

5. む す び

本稿では、視野を共有しないカメラ間での人物の追跡を目的として、共可視領域に基づき色特徴を抽出する共可視特徴を用いた人物同定の手法について提案した。これは、カメラによる観察方向の違いによって照合精度が低下する問題に対

して、人物の移動方向から共可視領域を定義する方法と、カメラから観察されない領域に対して近似特徴量を人物の簡易モデルに基づいて算出する方法によって改善を行った。評価データによる累積照合特性の検証の結果、従来手法に対して照合率を高く人物同定が行えることを確認した。本稿では、画像からの領域分割と不可視領域への特徴量遷移について示した。実用に向けてアプリケーションへ適用するには、色特徴の取得方法やカメラ内での人物追跡結果を利用することによる共可視特徴の取得方法の改善などが挙げられる。これらについては、今後の課題としたい。

参 考 文 献

- 1) K. Hirasawa, H. Kage, S. Miwa, K. Hashima, M. Hashimoto, M. Jones, J. Thornton: "Development of Face Image Capturing System for Best Appearance: The Improvement of the Face Detection Algorithm", Technical Report of IEICE. HIP, Vol. 104, pp. 61-66 (2004).
- 2) Wang, G. Doretto, T. Sebastian, J. Rittscher, P. Tu: "Shape and Appearance Context Modeling", Proc. of IEEE 11th International Conference on Computer Vision, pp. 1-8 (2007).
- 3) 井尻善久, 川西康友, 美濃導彦, 村瀬洋: "サーベイ論文: 視野を共有しない複数カメラ間での人物照合", 電子情報通信学会技術報告, PRMU2011-21, pp.117-124 (2011) .
- 4) S. Ali, O. Javed, N. Haering, T. Kanade: "Interactive Retrieval of Targets for Wide Area Surveillance", Proceedings of the International Conference on Multimedia, pp.895-898 (2010).
- 5) P. Viola, M. J. Jones: "Robust Real-Time Face Detection", International Journal of Computer Vision, Vol. 57, pp. 137-154 (2004).
- 6) Chang, S. S.: "A Survey of Recent Advances in Face Detection", Microsoft Research Technical Report, MSR-TR-2010-66 (2010).
- 7) M. Nishiyama, M. Yuasa, T. Shibata, T. Wakasugi, T. Kawahara, O. Yamaguchi: "Recognizing Faces of Moving People by Hierarchical Image-Set Matching", Proc. of 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp.1-8 (2007).
- 8) A. M. Bronstein, M. M. Bronstein, R. Kimmel: "Three-Dimensional Face Recognition", International Journal of Computer Vision, Vol. 64, pp. 5-30 (2005).
- 9) R. Urtasun, P. Fua: "3D Tracking for Gait Characterization and Recognition", Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition, pp.17-22 (2004).
- 10) C. Yam, M. S. Nixon, J. N. Carter: "Automated Person Recognition by Walking and Running via Model-Based Approaches", Pattern Recognition, Vol. 37, pp. 1057-1072 (2004).
- 11) R. Kawai, Y. Makihara, C. Hua, H. Iwama, Y. Yagi: "Person Re-Identification Using View-Dependent Score-Level Fusion of Gait and Color Features", Proc. of 21st International Conference on Pattern Recognition, pp. 2694-2697 (2012).
- 12) A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, R. Chellappa: "Gait Analysis for Human Identification", Proc. of 3rd International conference on Audio and Video Based Person Authentication, pp. 706-714 (2003).
- 13) S. Sarkar, P. J. Phillips, Liu, I. R. Vega, P. Grother, K. W. Bowyer: "The Humanid Gait Challenge Problem: Data Sets, Performance, and Analysis", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 27, pp. 162-177 (2005).
- 14) C. Nakajima, M. Pontil, B. Heisele, T. Poggio: "Fullbody Person Recognition System", Pattern Recognition, Vol. 36, pp.

- 1997–2006 (2003).
- 15) N. D. Bird, O. Masoud, N. P. Papanikolopoulos, A. Isaacs: “Detection of Loitering Individuals in Public Transportation Areas”, IEEE Trans. on Intelligent Transportation Systems, Vol. 6, pp. 167–177 (2005).
 - 16) N. Gheissari, T. B. Sebastian, R. Hartley: “Person Re-Identification Using Spatiotemporal Appearance”, Proc. of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 1528–1535 (2006).
 - 17) Y. Cai, M. Pietikäinen: “Person Re-Identification Based on Global Color Context”, Proc. of Asian Conference on Computer Vision (ACCV 2010), pp. 205–215 (2011).
 - 18) F. Porikli: “Inter-Camera Color Calibration by Correlation Model Function”, Proc. of 2003 International Conference on Image Processing, Vol. 2, pp. II-133 (2003).
 - 19) O. Hamdoun, F. Moutarde, B. Stanciulescu, B. Steux: “Person Re-Identification in Multi-Camera System by Signature Based on Interest Point Descriptors Collected on Short Video Sequences”, Proc. of 2nd ACM/IEEE International Conference on Distributed Smart Cameras, pp. 1–6 (2008).
 - 20) M. Farenzena, L. Bazzani, A. Perina, V. Murino, M. Cristani: “Person Re-Identification by Symmetry-Driven Accumulation of Local Features”, Proc. of 2010 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2360–2367 (2010).
 - 21) 上村和広, 池亀幸久, 下山功, 玉木徹, 山本正信: “ネットワーク上の複数カメラを用いた実時間人物照合システム”, 電子情報通信学会技術報告, PRMU2003-242 (2004).
 - 22) 井尻善久, 芳世, 村瀬洋: “非線形距離指標学習によるカメラ間人物照合”, 電子情報通信学会技術研究報告, パターン認識・メディア理解, Vol.111, pp.139–146 (2011).
 - 23) 小林貴訓, 佐藤洋一, 杉本晃宏: “視野を共有しないカメラ群を用いた人物行動軌跡の推定”, 情報処理学会研究報告, コンピュータビジョンとイメージメディア, Vol.2005, pp.169–176 (2005).
 - 24) U. Park, A. K. Jain, I. Kitahara, K. Kogure, N. Hagita: “Vise: Visual Search Engine Using Multiple Networked Cameras”, Proc. of 18th International Conference on Pattern Recognition, Vol. 3, pp. 1204–1207 (2006).
 - 25) 和島直哉, 高橋悟, 伊藤誠也, 佐藤雄隆, 金子俊一: “放射リーチフィルタに基づく背景変動を考慮した移動体追跡”, 電気学会論文誌 D (産業応用部門誌), Vol.131, pp.137–143 (2011).
 - 26) A. Bhattacharyya: “On a Measure of Divergence Between Two Multinomial Populations”, Sankhyā: The Indian Journal of Statistics, pp. 401–406 (1946).
 - 27) D. Gray, H. Tao: “Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features”, Proc. of the 10th European Conference on Computer Vision (ECCV 2008), pp. 262–275 (2008).
 - 28) O. Javed, K. Shafique, M. Shah: “Appearance Modeling for Tracking in Multiple Non-Overlapping Cameras”, Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 26–33 (2005).
 - 29) J. L. Wayman, A. K. Jain, D. Maltoni, D. Maio, Biometric Systems: Technology, Design and Performance Evaluation, Springer Science & Business Media (2005).
 - 30) D. Gray, S. Brennan, H. Tao: “Evaluating Appearance Models for Recognition, Reacquisition, and Tracking”, Proc. of IEEE International Workshop on Performance Evaluation for Tracking and Surveillance, Vol. 3 (2007).

(2015年6月24日受付)

(2016年5月10日再受付)



伊藤 誠也

2002年 北海道大学工学部卒業。2004年 同大学院工学研究科システム情報工学専攻修了。同年株式会社日立製作所日立研究所入社。現在、同研究開発グループ制御イノベーションセンター所属。2012年 カーネギーメロン大学ロボティクス研究所客員研究員。2014年 北海道大学大学院情報科学研究科博士後期課程単位取得退学。画像理解、パターン認識の研究に従事。博士(情報科学)。



金子 俊一 (正会員)

1978年 北海道大学工学部卒業。1980年 同大学院工学研究科修了。東京農工大学工学部数理情報工学科助手として画像計測の研究に従事。1990年 東京大学より工学博士。1991年より同助教授。1996年 北海道大学工学部助教授、2004年より同大学院情報科学研究科教授としてロボティクス、ロバスト画像処理の研究に従事。