

# 顔が判別できない低解像度映像を用いた複数カメラ間の人物同定

鄭明燮<sup>†1</sup> 辻健太郎<sup>†1</sup> 宮崎信浩<sup>†1</sup> 松田裕司<sup>†1</sup>  
馬場孝之<sup>†1</sup> 瀬川英吾<sup>†1</sup> 上原祐介<sup>†1</sup>

あらまし プライバシーの問題に配慮しながら人の移動や流れの傾向を認識するために、顔が判別できない低解像度映像を用いて、複数カメラ間で移動する同一人物を対応付ける人物同定手法を提案する。低解像度映像では、人物特定ができないと同時に、人物の輪郭や服の模様のような細かな特徴が得られないため、カメラ間での人物同定が困難になる。提案手法では、低解像度映像で最も基本的かつ重要な特徴量となる服の色情報を画素単位で抽出するのではなく、人の頭部、上着など上下関係に基づいて広い範囲での色特徴量を抽出すると同時に、カメラ間で変化が少ない共通する色特徴量を選別することで、人物同定を実現する。頭部の大きさが 4~8pixel の顔が判別できない低解像度映像を用いた評価実験では同定正解率が 72%であり、低解像度映像を用いた目視と同等の同定精度を確認した。

キーワード 監視カメラ, 低解像度映像, 人物検出, 人物同定

## 1. はじめに

近年、人物の移動経路を表す人流情報が、混雑緩和やマーケティングなどの顧客サービスに活用されている[1]。例えば、商業施設では、購買客の人流情報を利用して、店舗の配置や店舗内の品ぞろえ、人員の配置などの最適化が行われている[2]。また、街中では、イベント時の混雑解消や交通機関の運行計画策定などへの活用が検討されている。

人流情報を獲得するために、様々な手法が提案されている。例えば、ビーコンやスマートホンの Wi-Fi の機能で位置を逐次検出する手法や、レーザレーダを利用して人物を検出し、追跡する手法などである[3]。しかし、これらの手法は、センサの設置コストの問題や、デバイスやタグを携帯していない不特定多数の通行者を捕捉できないなどの課題がある。

そこで、防犯用途として既に設置されている監視カメラの映像から人流情報を獲得する手法が注目されている。この手法は、スマートホンなどのデバイスを携帯していない人物の情報も取得することができ、さらに既存の機材を利用するため、設置コストがかからない利点がある。

しかし、街中や商業施設内に設置されたカメラを活用する際の大きな課題として、被撮影者のプライバシー保護がある。最近の監視カメラは高解像度化が進み、映像中の人物の顔を視認できることが多い。そのため、撮影された人物の行動に関する情報を第三者に知られる懸念があり、個人のプライバシー保護に関する問題が生じる可能性がある。映像から個人の特特定を防ぐことが必要である。

個人の特特定を防ぐために、映像中の顔部分や動きのある部分にスクランブルを施す手法が提案されている[4][5]。しかし、いずれの手法も、検出漏れや誤検出が避けられないため、個人を特定できる情報が残る可能性がある。プライバシーの問題に対処するには、個人が特定できるような映

像を入力することなく、人流情報が獲得できる手法を実現すべきである。その手法として、すべての映像に対して、被写体が誰であるか特定できないぐらいまで解像度を低下させた映像（以降、低解像度映像と呼ぶ）を用いる手法が考えられる。例えば、人物の顔が数画素程度で表される解像度にすれば、個人は特定できない。

我々は、低解像度映像から人流情報を獲得可能なシステムの開発を進めている。特に、商業施設や街中のような広範囲にわたる人流情報を獲得するため、複数のカメラで撮影した各映像から人物を検出し、検出した人物同士の特徴がカメラ間で一致する組み合わせを見つけて同一人物として同定する手法を開発している。これまで我々は、低解像度映像を構成する各時刻の低解像度画像から人物を検出する手法について、既に提案している[6][7]。

本稿では、低解像度映像から検出された人物を対象に、複数カメラにおける人物同定手法について述べる。

## 2. 関連研究

カメラ間の人物同定は、それぞれの映像から検出された人物に関する全身の見え方の特徴を抽出し照合することで実現する。異なる場所に設置された監視カメラで撮影された多数の映像の中から、全身の見え方が手ごかりに同一人物を同定する場合、照明や向きや姿勢が変動すること、および類似した服装の人物が存在することが問題を困難にする。この解決に向けた先行研究として、特徴量ベースで人物同定する手法と、機械学習を用いて同定する手法が提案されている[8][9]。

特徴量ベースの手法では、服装の色や模様を取り出すための特徴抽出手法が提案されている[10][11]。これらの手法では、人の見え方が変動する中で精度を維持するため、複数の特徴量を冗長に抽出する。例えば、Farenzena らは、Symmetry-Driven Accumulation of Local Features(SDALF)という色及び形状特徴を抽出する手法を提案している[11]。この手法は、服の色特徴量として、重み付き HSV ヒストグラムと Maximally Stable Color Regions(MSCR)、模様特徴量

<sup>†1</sup> 株式会社富士通研究所  
FUJITSU LABORATORIES LTD.

として、Recurrent High-Structured Patches(RHSP)の、3種類の特徴量を用い、照合結果を統合することで、高い人物同定精度を実現している。この手法で、重み付き HSV ヒストグラムは単純な距離指標で照合されるが、カメラ間の照明変化がある場合は、精度低下に繋がると考えられる。

色特徴量は、低解像度の場合に限らず、一般的に照明条件の変化に影響されやすい。監視カメラは、時間によって変化する環境に対応するため、自動設定で撮影できるようになっているのが一般的である。特に、ホワイトバランスの自動設定によって、昼夜や朝夕などでも見やすい映像が撮影できるようになっている。このような設定では、例えば、色のある壁などからの反射や、色のついたアーケードの日光の透過によって、監視カメラ間で異なる照明であるとみなされ、同じ物体であってもカメラ間で異なる色として検出される。このような照明条件の変化に対応するために、Brightness Transfer Function(BTF)によってカメラ間の色ヒストグラムの変化の仕方を機械学習する手法が提案されているが、学習のために訓練サンプルを必要とする[12][13]。

また、カメラ間で撮影環境の変動や撮影対象の見え方の変化を機械学習によって適応する手法が数多く提案されている[14][15]。しかし、実環境での映像に適用するためには、大量の人物データセットを集める手間がかかる点などが課題である。

これまでに、通常サイズの画像を用いた人物同定手法は提案されてきたが、低解像度画像を用いて人物同定を目指した研究はほとんどない。

一方、人物の顔が判別できないように低解像度化した映像から取り出した画像(以降、低解像度画像と呼ぶ)では、服の模様のような細かな特徴が得られないため、服の色情報が最も基本的で重要な特徴量となる(図1)。この色特徴量により正しく人物同定を行うためには、異なるカメラ間において人物領域内の相対的な位置が同じであれば色も同じになることが前提となる。しかし、低解像度画像の場合、以下のような課題がある。



図1 VIPeR DB 画像の低解像度化による人物特徴。

1) 服が複色色を含む場合、低解像度化の際に、画像上で色の境界部分は平均化により周囲の色が混合された色になる。境界の位置に応じて周囲の混ざる割合が変わるため、混合色も変わる。そのため、同じ服でも同一色とならない。

2) 低解像度化により輪郭などがぼけるため、画像上に映る人物の位置の違いにより、同じ姿勢でも頭部や胴体の輪郭が欠けるなどパターンが異なる画像となる。そのため、異なるカメラの画像間では、人物領域内で相対的に同じ位置の画素であっても同一色とならない(図2)。

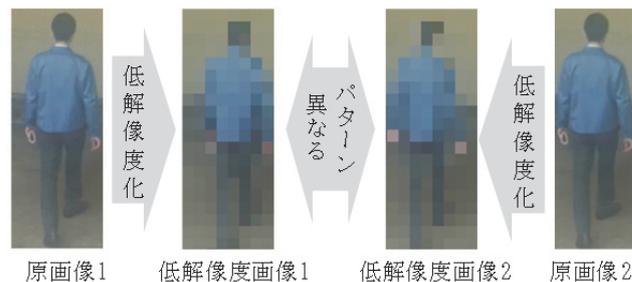


図2 低解像度画像で異なる特徴量。

### 3. 低解像度画像を用いた人物同定手法の提案

#### 3.1 概要

本稿では、機械学習を必要としない特徴量ベースでの人物同定手法に焦点を当てて、服の色特徴量と位置情報を用いたカメラ間の人物同定手法を提案する。

具体的には、2章で述べた低解像度画像の課題を解決するために、以下を提案する。課題1)については、カメラ間で同一人物の色特徴量を合わせるために、カメラ間で異なる混合色を前処理で除去する。課題2)については、画素単位での色特徴が変わる場合でも、人の頭部や上着などの一定の面積を有する領域での色は変動が少なく、その上下の位置も大きく変わらないことに着目し、画素毎に色特徴量を抽出するのではなく、人物の大きな領域に分割し、領域毎に色特徴量を抽出する。特徴量の照合においては、大きな領域毎に抽出した色特徴量を用いて類似度を求める。次節以降で詳細を述べる。

#### 3.2 特徴抽出

特徴抽出処理では、前処理として、カメラ間で異なる混合色を除去したのち、低解像度画像の場合に限らず、照明変化の一般的な問題に対応するために、道路など背景の同一反射特性を利用して照明変化で異なるカメラ間の服装の色を補正する。続いて、体の位置ずれや向きの変化に対応するために、人物領域に対して、位置ずれに頑健な大局的な色特徴量を抽出するとともに、体の向きを考慮した局所的な色特徴量を抽出する。

##### 3.2.1 混合色の除去

人物領域において、混合色を除外し、服装のもともと

の色のみを特徴量として利用する。混合色か否かは、各画素において、隣接する画素の色との関係を調べることにより判定できることに着目した。ここでは、入力画像から背景差分処理によって背景部分を取り除いた人物の領域画像を人物領域としている。この人物領域上で、着目する画素の色成分が、画像の2次元座標上隣接する画素の色成分で合成可能な色かどうかで混合色を判定する。

画像の2次元座標上ある注目する混合色の画素値は、画像の縮小方法によって、X軸方向の左右またはY軸方向の上下に隣接する画素の画素値との線形和で表すことができる。ここで、隣接する第1の画素の画素値を $V_0$ 、隣接する第2の画素の画素値を $V_i$ とした場合、判定対象画素の画素値 $V_c$ は式(1)で表される。

$$V_c = \alpha V_0 + (1 - \alpha)V_i \quad (1)$$

ここで、混色率 $\alpha$ は、式(2)で算出される。

$$\alpha = (V_c - V_i)/(V_0 - V_i) \quad (2)$$

尚、画素値 $V_c$ 、 $V_i$ 、 $V_0$ がそれぞれRGBの3成分で表される場合、画素値 $V_c = (V_{cR}, V_{cG}, V_{cB})$ 、 $V_i = (V_{iR}, V_{iG}, V_{iB})$ 、 $V_0 = (V_{oR}, V_{oG}, V_{oB})$ とすると、式(2)における画素値の差分は、式(3)と式(4)のようにユークリッド距離で算出される。

$$V_c - V_i = ((V_{cR} - V_{iR})^2 + (V_{cG} - V_{iG})^2 + (V_{cB} - V_{iB})^2)^{\frac{1}{2}} \quad (3)$$

$$V_0 - V_i = ((V_{oR} - V_{iR})^2 + (V_{oG} - V_{iG})^2 + (V_{oB} - V_{iB})^2)^{\frac{1}{2}} \quad (4)$$

ここで、混色率 $\alpha$ が閾値以上である場合、注目画素は混合色であると判定し、混色率 $\alpha$ が閾値未満である場合、混合色でないと判定する。

### 3.2.2 カメラ間の色補正

カメラ間共通の背景領域の色を利用してカメラ間画像の色補正を行う。ここでは、複数の監視カメラ間で、以下の二つの仮定が成り立つものとする。一つは、一般的に近くの間所同士の道路では同じ材質で整備されることが多いため、カメラ映像間で、同じ反射特性の道路を撮影していることと仮定する。もう一つは、異なるカメラ間で色が変化する主要な原因がオートホワイトバランスによるものとする。このとき、観測される道路の色の違いは、オートホワイトバランスによるものであり、道路の色が同一になるように、ホワイトバランスを補正すれば、道路以外の同一物体においても同じ色になる。そこで、二つのカメラ間で道路の色を一致させるように、一方の監視カメラのホワイトバランスを修正し、画面全体の色を補正する。

ここでは、撮影時間帯によって変化する道路面の色が常に一定(定めた基準値)となるようにRGBの割合を補正する。しかし、カメラの設置場所によって、一つの同じ光源(太陽光)以外の照明や反射などで道路面の色が変化する場合もあり、オートホワイトバランス以外の要素で部分的に色が変化することがある。一つの色とみなした道路に一つの同じ光源があると想定した場合、同じ画像内では、太陽光以外の照明や反射がある道路部分はそれ以外の道路部分に比べて明るい。道路面からこの明るい領域を除くことで、オートホワイトバランスによる変化領域を絞ることができる。そこで、画像から検出された道路領域の中で、明るい領域を除いてオートホワイトバランスによって変化の関連性が高い領域の色を抽出し、道路の床面の色として補正する。具体的に以下の手順に従う。図3に、色補正処理の画像例を示す。

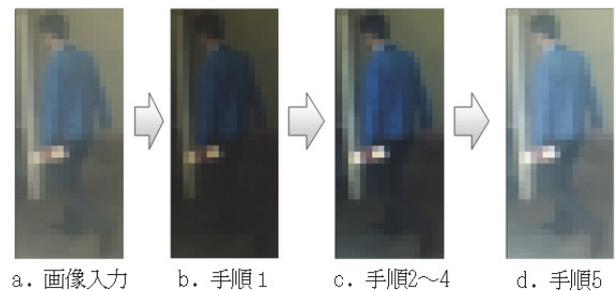


図3 色補正処理の画像例。

手順1) 画像上の任意の座標(i, j)のsRGB色空間(r, g, b)を、リニアな色空間(r', g', b')に変換する(以後、デガンマと呼ぶ)。例えば、rの場合には、式(5)を用いて、デガンマ補正でr'に変換する。

$$r' = 255 \times (r/255)^{\gamma} \quad (5)$$

ここで、色空間は0~255で表現している。また、ガンマ値は一般的に用いられている $\gamma=2.2$ と仮定する。図3bに、デガンマ処理の例を示す。

手順2) 予め指定した道路面の領域の中から、検出された人物領域を除いた部分を道路面の色抽出範囲とする。道路面の色抽出範囲に対して、色(R, G, B)の各チャンネルのヒストグラムを算出する。ヒストグラムからピークを検出し、ピークが複数ある場合は、画素値が高いピークを除いて画素値が低いピークを用いる。これによって、1色としてみなしている道路領域のヒストグラムの中で、太陽光以外の照明や反射で明るくなった部分を除外する。

手順3) 道路面の色空間(r', g', b')を予め定めた基準値(TR, TG, TB)になるように、RGBチャンネルそれぞれの変換係数を求める。例えば、基準値をTRとした場合、検出された道路面の色r'から、式(6)によって変換係数Kr

を求める。

$$K_r = TR/r' \quad (6)$$

手順 4) 変換係数を用いて、画像全体の色空間 ( $r'$ ,  $g'$ ,  $b'$ ) の色補正を行い、補正後の色空間 ( $Tr$ ,  $Tg$ ,  $Tb$ ) を得る。図 3.c に、変換係数による色補正例を示す。

手順 5) 色補正を行った色 ( $Tr$ ,  $Tg$ ,  $Tb$ ) を、ガンマ補正によって元の sRGB 色空間に戻す。図 3.d に、ガンマ補正の例を示す。

### 3.2.3 大局的な色特徴量の抽出

画素単位での色特徴が変わる場合でも、色が類似する部分領域の代表的な色特徴は変わらないことに着目した。そこで、画素毎に色特徴を抽出するのではなく、人物の大まかな部位毎に代表的な色特徴量を抽出する。ここで、異なるカメラの画像においても、画像中での上下関係は変わらないことから、人物を上下に 3 分割した頭部、上半身、下半身を、色特徴を抽出する領域とする。分割比は、カメラが設置された高さで傾きの条件に基づいて、予め人の映り方を確認して決めることとする。図 4 のように、類似した色の分布特性から、人物を上下に頭部(赤枠)、上半身(緑枠)、下半身(青枠)の、3 つの領域として切出して、切出した領域から HSV 色空間の各成分のヒストグラムを色特徴量として抽出する。

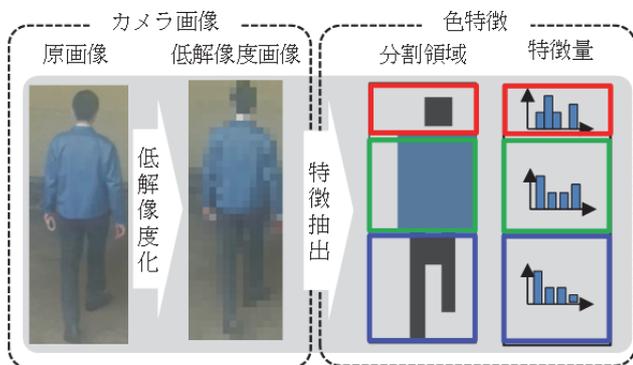


図 4 分割領域の特徴抽出。

### 3.2.4 局所的な色特徴量の抽出

検出された人物領域に対して、領域分割処理で近傍の類似する色を小領域として分割した局所的な色特徴量を抽出する。ここでは、MSCR 特徴量を抽出する従来手法[16]を適用し、人物領域内の類似した色ごとにクラスタリングを行う。そして、クラスタリングで求めた複数のクラスタの、座標位置と色平均値で人物領域を表現する。ここでは、各クラスタは、HSV 色空間の各成分として H (色相), S (彩度), V (明度) の平均値及び分散値を色特徴量として用いる。また、人物領域における各クラスタの中心座標位置 ( $x$ ,  $y$ ) 及び幅と高さ ( $w$ ,  $h$ ) を特徴量として記録する。

### 3.3 特徴量の照合

特徴量照合では、3.2.3 節で述べた大局的な色特徴量を用いた照合処理と、3.2.4 節で述べた局所的な色特徴量を用いた照合処理を行い、それぞれの類似度を算出する。次に、服装の色が類似する対象人物が増えて、同定精度が低下することを防ぐために、カメラ間で移動する人物の位置情報を利用して、照合対象を絞り込むため、人の移動位置の類似度を求める。最後に、それぞれの類似度がすべて高くなった場合にトータル類似度が高くなるように、3 種類の類似度を積算で算出する。

#### 3.3.1 大局的な色特徴量の類似度

分割した部位の色の類似度は、人の頭部、上半身、下半身の 3 つの部位の色のヒストグラムを比較することで求める。色のヒストグラムは HSV 色空間の各成分のヒストグラムを用いて、それぞれの対象人物に対して明暗の変化を吸収するために正規化相関で類似度を求める。そして、3 つの類似度がすべて高くなった場合に類似度が高くなるように積で色特徴量の類似度を算出する。

#### 3.3.2 局所的な色特徴量の類似度

複数カメラで検出された人の体の向きが異なる場合は、抽出された色特徴量が異なるため、人物領域内の同一座標の画素であっても、カメラ間で同一色とならない。一方、カメラの光軸に対して人の移動方向が分かれば、前向いて歩く人の移動方向毎に映る範囲は特定可能である。移動方向が異なってもカメラに映る人の体領域は重畳する共通の部分領域が観測できる。移動方向は、連続して検出される人物領域の中心座標を用いて、現在の座標位置と 1 時刻前の座標位置から推定できる。なお、正面向きと後向きの正反対方向は、移動方向を比較し、対象外に設定している。

具体的には、人の体領域を円柱と例えた場合(図 5)、円柱の横表面を一定の刻みで部分領域に分割すると、部分領域は人の部分領域として見なせることができる。図 6 に、移動方向の違いによって異なる 2 つカメラでの見えの領域と、2 つの見えの領域を重畳したときの共通領域を示す。なお、一つのカメラで観測可能な範囲は予め設定する。この共通領域に含まれるクラスタの色特徴量を用いて類似度を算出する。これによって、異なる領域の特徴が混じることが少なく正しい照合が期待できる。

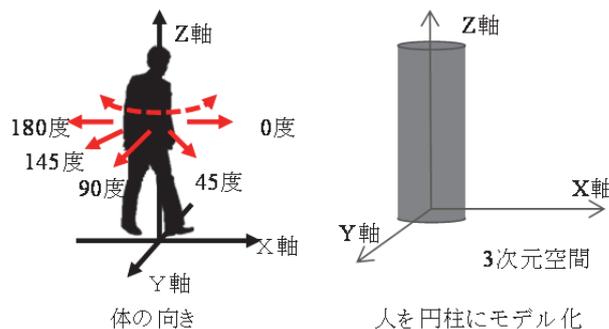


図 5 体領域のモデル化。

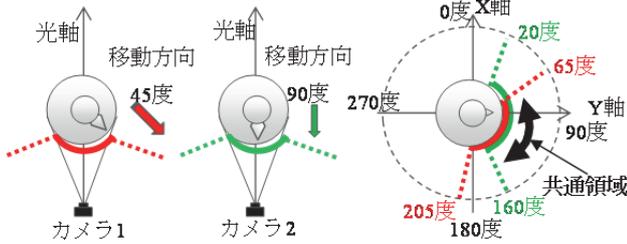


図 6 カメラ間で重畳する共通領域.

### 3.3.3 移動位置の類似度

移動位置の類似度は、異なるカメラで撮影された人物それぞれの検出時刻、移動速度、及び検出位置の整合性を利用して求める。具体的に、異なるカメラで検出された人物を、それぞれ人物 A、人物 B とすると、移動速度が一定で、2 人が同一人物である場合、移動した人が検出された場所の距離が求められる。そこで、あらかじめカメラ間の距離を定義しておき、それぞれの人物がカメラ視野から出た時刻と出現した時刻の差から整合性を求める。カメラ間の移動時間は、カメラ間距離と移動速度で求まる。移動速度は、移動する人物によって異なるため、予め人物毎の平均移動速度を調べておき、カメラ間の移動可能時間の最小値と最大値を  $a_1$  と  $a_2$  とする。さらに、同じ人物に対して、移動速度は映像から推定するが、その移動速度がカメラ間の移動中に変化する可能性があるため、許容できる速度変化を考慮し、移動可能な時間の範囲  $b_1$  と  $b_2$  を設ける。移動時間が  $b_1$  以下または  $b_2$  以上は同一人物の可能性が低いとして、スコアを 0 に決めておく。例えば、画像から推定した人の移動速度を  $S_v$ 、予め定めたカメラ間の距離を  $D_s$  とし、許容できる速度変化の割合  $rate\_a$  ( $>1$ ) と、これ以上はあり得ない変化の割合  $rate\_b$  ( $>rate\_a$ ) を決めておき、以下の式に従って設定範囲を求めることができる。

$$a_1 = \frac{D_s}{S_v} \cdot rate_a$$

$$a_2 = \frac{D_s}{S_v} \cdot \frac{1}{rate_a}$$

$$b_1 = \frac{D_s}{S_v} \cdot rate_b$$

$$b_2 = \frac{D_s}{S_v} \cdot \frac{1}{rate_b}$$

移動位置の類似度は、以下の式(6)で求める。

$$cor\_time(t) = \begin{cases} 1 & \text{if } a_1 \leq t \leq a_2 \\ \frac{t-b_1}{a_1-b_1} & \text{if } b_1 \leq t < a_1 \\ \frac{t-b_2}{a_2-b_2} & \text{if } a_2 < t \leq b_2 \\ 0 & \text{if } b_1 \leq t, b_2 > t \end{cases} \quad (6)$$

式(6)では、予め調べた移動速度で移動した場合の経過時間と実際の経過時間が等しいときに 1 を返し、離れるほど

0 に近づく値になる。

## 4. 実験

### 4.1 実験概要

評価用の映像として、屋外に設置された監視カメラ映像を利用した。このカメラは、高さ約 3.5m から斜め下向きに設置されており、撮影された映像は 352×240 画素、10fps のフレームレートである。今回の実験では、この映像の中から延べ 7000 フレームの連続画像を取り出し、顔が判別できない 88×60 画素の低解像度画像を作成した。また、頭部と上半身が撮影画像中に映っている人物のみを対象とした。具体的には、頭部の大きさ 4~8pixel の人物を対象として、2 台のカメラの両方に検出された 100 人の評価画像セットを用いた。評価はカメラ 1 から検出された人物毎に、カメラ 2 で最も一致度が高い人を 1 人抽出し、これが正解として定義した同一人物であれば人物同定成功、他人であれば人物同定失敗とした。なお、正解は、352×240 画素の原画像を用いて目視確認によって作成した。

また、低解像度画像による目視の正解率を確認するために、88×60 画素に低解像度化した評価データを用いて、初めて評価データを見る実験者の目視確認による人物同定を実施した。その結果、正解率 70%を確認した。ここで、正解率は、対象人物 100 人の中、正しく同定できた人数の割合としている。

### 4.2 人物同定の評価結果

提案手法の有効性を示すために、3章で述べたすべての処理を実施した同定手法2と、混合色の除去、色補正処理及び局所特徴量の共通領域選択を実施していない同定手法1との正解率を比較した。表1に、人物同定正解率の比較結果を示す。

表 1 人物同定の正解率.

	同定手法 1	同定手法 2
正解率	56%(56/100)	72%(72/100)

表 1 より、同定手法 2 の人物同定正解率は、同定手法 1 の 56%から 72%に向上でき、低解像度画像を用いた目視と同等の同定精度が確認できた。

### 4.3 色特徴量の比較結果

次に、混合色の除去及び色補正処理を行う前処理の効果を確認するために、前処理の実施有無について、色特徴量の距離変化を調べた。服装の色が異なる20名の人物画像を選択し、任意に 1 人10か所抽出し、異なるカメラで検出された同一人物の同じ部位の色を比較した。ここで、カメラ間の色の絶対値の変化を確認するために、色はRGB値 (0~255で表現) を用いて二乗誤差を算出し、二つのカメラ間で人物毎に10か所の二乗誤差の平均値を比較している。二乗誤差が小さいほど、二つのカメラ間で色の差異が小さく

なる。表2に、20名の人物の平均二乗誤差を示す。表2の結果より、前処理ありの提案手法は、前処理なしに比べて、カメラ間で色特徴量の差異が小さくなっており、提案手法の有効性を確認した。

表 2 二乗誤差の比較.

	前処理なし	前処理あり
二乗誤差	16.1	12.4

また、図7には、20名の人物毎の二乗誤差を示している。

図7から、カメラ間で人の姿勢の大きな変化で二乗誤差が多少大きくなった人物ID07とID19は除いて、二乗誤差が大きかった人物においては、提案手法が、カメラ間での色特徴量の差異が小さくなっていることが確認できる。

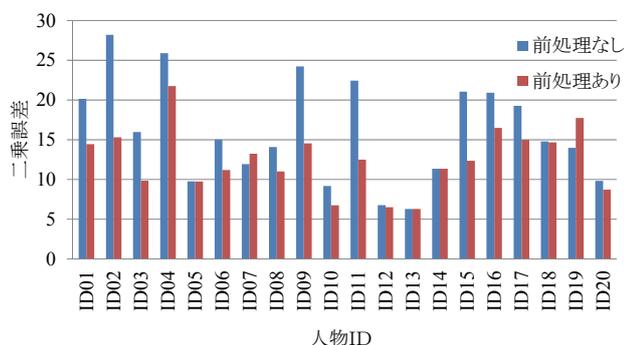


図 7 カメラ間の同一人物の二乗誤差の比較結果.

#### 4.4 考察

提案手法では、顔が判別できない低解像度映像に対し、正解率72%の人物同定精度を実現し、その有効性が確認できた。一方、人物同定失敗の主な原因は服装の色特徴量が非常に類似していたためであることが分かった。この問題については、複数フレーム間での人の動き特徴を求めるなど色特徴以外の情報を利用することで、さらなる性能向上が期待できる。また、今回人物同定が失敗した原因として、人の姿勢変化でカメラ間の映り方が異なり同じ部位の同定ができなかったケースがあった。現状、同定の照合方式は人物の検出領域を複数の部分に分割し、同じ位置の色の類似度を比較しているため、姿勢変化などで位置ずれがある場合は正しい照合ができない。この課題については、時系列の画像を観測し、同じ座標位置ではなく、色特徴量の上下または左右の相対位置関係に基づいた照合を実現することで、改善できると考える。

#### 5. おわりに

本稿では、プライバシーの問題に配慮した人流情報の獲得のため、顔が判別できない低解像度映像を用いてカメラ間の同一人物を同定する手法を提案した。低解像度映像では服装の色などが有力な特徴量となるが、色の境界部分に量子化誤差で色特徴が変化するため、同一人物の同定が困

難になる。提案手法では、服の隣接画素の色成分を調べ、混合色は除外した代表色の特徴量を抽出し、照合することで混合色の問題を解決した。評価実験では人物同定の正解率が72%であり、低解像度映像を用いた目視と同等の同定精度を確認した。今後、低解像度映像の時間変化の情報を利用し、様々な撮影環境や対象に対する性能評価を実施し、さらなる高精度化を実現する。

#### 参考文献

- 1) 島崎 康信, 関本 義秀, 柴崎 亮介, 秋山 祐樹, “人の流れによる時間帯別人口と店舗数との相関関係についての研究,” 都市計画学会都市計画論文集, vol. 44, no. 3, pp.781-786, 2009.
- 2) 山下 倫央, 大西 正輝, “オリンピックにおける人の流れの解析,” 情報処理, vol. 55, no.11, pp.1189-1195, 2014.
- 3) 山下 倫央, 副田 俊介, 野田 五十樹, “人流計測による避難誘導効果の実証的検証,” 情報研報, vol. 2009-UBI-24, no. 25, pp. 1-8, 2009.
- 4) H. Sohn, W. D. Neve, and Y. M. Ro, “Privacy Protection in Video Surveillance Systems: Analysis of Subband-Adaptive Scrambling in JPEG XR,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 21, no. 2, pp. 170-177, 2011.
- 5) S.-C.S. Cheung, M.V. Venkatesh, J.K. Paruchuri, J. Zhao, and T. Nguyen, “Protecting Privacy in Video Surveillance,” in Protecting and Managing Privacy Information in Video Surveillance Systems, Springer, pp. 11-13, 2009.
- 6) 辻 健太郎, 鄭 明燮, 中島 望夢季, 松田 裕司, 宮崎 信浩, 皆川 明洋, “既設カメラなどの低解像度映像による人物検出システム,” 設情処研報, vol. 2015-CVIM-196, no. 21, pp. 1-6, 2015.
- 7) N. Miyazaki, K. Tsuji, M. Zheng, M. Nakashima, Y. Matsuda, and E. Segawa, “Privacy-conscious Human Detection using Low-resolution Video,” 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 326-330, 2015.
- 8) 井尻義久, 川西康友, 美濃導彦, 村瀬洋, 義視野を共有しない複数カメラ間での人物照合, 野 信学技報, pp.117-124, 2011.
- 9) S. Gong, M. Cristani, S. Yan, and C.C. Loy, “Person Re-Identification,” Advances in Computer Vision and Pattern Recognition, XVIII, 445 p. 163 illus., 154 illus, in color, 2014.
- 10) D. Gray and H. Tao, “Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features”, Proc.of ECCV, pp. 262-275, 2008.
- 11) M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person re-identification by symmetry-driven accumulation of local features,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 2360-2367, 2010.
- 12) B. Prosser, S. Gong, and T. Xiang, “Multi-camera Matching using Bi-Directional Cumulative Brightness Transfer Functions,” Proceedings of the British Machine Vision Conference, pp. 64.1-64.10, 2008.
- 13) O. Javed, K. Shafique, Z. Rasheed, and M. Shah, “Modeling Inter-camera Space-time and Appearance Relationships for Tracking Across Non-overlapping Views,” CVIU, 109, 2, pp. 146-162, 2008.
- 14) N. McLaughlin, J. Martinez del Rincon, P. Miller, “Recurrent Convolutional Network for Video-based Person Re-identification,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 1325 - 1334, 2016.
- 15) J. You, A. Wu, X. Li, and W.-S. Zheng, “Top-push Video-based Person Re-identification,” in IEEE Conference on Computer Vision and Pattern Recognition, pp. 1345-1353, 2016.
- 16) P.-E. Forss'en, “Maximally Stable Colour Regions for Recognition and Matching,” in IEEE Conference on Computer Vision and Pattern Recognition, pp.1-8, 2007.