

頭上方向から撮影されたカメラ画像を利用した人物識別

中谷 良太[†] 香野 大地[†] 嶋田 和孝[†] 遠藤 勉[†]

[†] 九州工業大学大学院情報工学府

〒 820-8502 福岡県飯塚市川津 680-4

E-mail: [†]{r_nakatani,d_kouno,shimada,endo}@pluto.ai.kyutech.ac.jp

あらまし 画像を用いた人物識別では、人物の顔やシルエットから特徴量を抽出し用いる手法が提案されている。しかし顔やシルエットを用いる場合、人物の重なりにより隠れが生じ正しく特徴量を抽出できないことがある。この問題に対処するため、本論文では頭上方向に設置したカメラにより撮影を行う人物識別の手法を提案する。手法の流れとしては、まず背景差分法を適用した画像からノイズを除去することで人物領域を特定する。続いて人物領域から体型、髪型、つむじ、髪の色の特徴量として抽出する。これらの特徴量を学習アルゴリズムである AdaBoost に適用することで人物識別を行う。実験により、少人数での出席管理のようなシステムとして提案手法が有効であることを示した。

キーワード 頭上カメラ、人物識別、体型、髪型、つむじ、髪の色

1. はじめに

人物識別はセキュリティや利便性、商用など多くの場面に応用可能である。人物識別の手段には静脈や虹彩、筆跡、画像を用いるものなどさまざまなものがあるが、画像を用いた人物識別には機器に非接触で実現できるため対象者にかかる負担が小さいというメリットがある。多くの研究者が画像を用いた人物識別の手法を提案しており、その中でも顔画像を用いた人物識別の手法が多く提案されている。Kanade [7] は鼻や目などを用いた手法を提案している。CLAFIC 法 [1] や EigenFace 法 [9] も顔から抽出した特徴量を用いて人物識別を行う。これらの手法のように画像を用いた人物識別では顔画像を使用するものが多いが、混雑している場所などでは人物の重なりによって隠れが生じ、顔画像を正しく取得することが困難な場合がある。

そこで我々は人物の重なりによる隠れが生じるという問題を解決するために、頭上方向から撮影されたカメラ画像に着目した手法を提案する。図 1 は人物の重なりによる隠れが生じる問題を解決する方法を表している。正面から人物を撮影した場合では隠れが生じ、正しく顔画像を取得できないような場面においても、頭上方向から撮影したカメラ画像を用いることで隠れが生じることなく人物識別を行うことができる。図 2 は実際に頭上方向から人物を撮影した画像例である。この手法には人物の重なりによる隠れの問題を解消できること以外にもメリットがある。1 つ目は、対象者の視野に入らない位置にカメラを設置していても画像を取得することができるため、撮影されているという心理的負担を軽減することができるということである。2 つ目は、顔を撮影する必要がないためカメラの設置場所に関する制限が軽減される

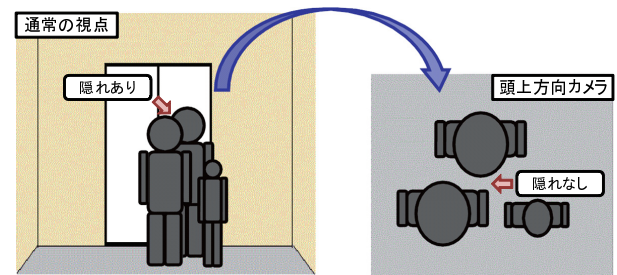


図 1 隠れが生じた際の解決法



図 2 頭上方向から人物を撮影した画像

ということである。

本論文では、頭上方向に設置したカメラを用いて撮影した画像により人物識別を行う手法を提案する。人物識別はセキュリティやロボットとの対話、出席管理など様々な場面で応用が可能である。本論文で提案する手法は主にオフィスや研究室のように少人数の環境での出席管理としての使用を想定している。図 3 は本手法におけるカメラの設置場所を表している。対象者がドアの前で一度立ち止まる場面を想定しており、カメラの設置場所はドアの頭上方向であり容易に設置することができる。このように設置が容易であることは、出席管理や人物検索、行動認識を行うような多くのシステムにとって重要なことである。

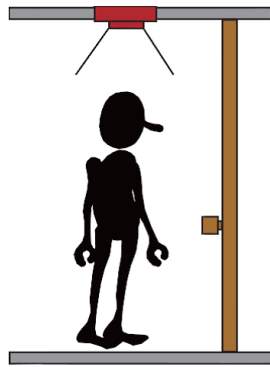


図 3 カメラの設置場所

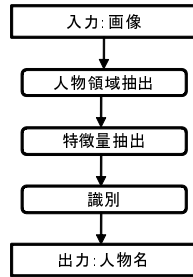


図 4 提案手法の概要

本手法では、まず初めに人物領域を特定する。人物領域を特定するためには、背景差分法を用いることで大まかな人物の位置を把握し、その画像からノイズを除去することで人物領域を特定する。続いて人物領域から“体型”、“髪型”、“つむじ”、“髪の色”の4つを特徴量として数値化して抽出する。“体型”は人物領域の縦幅と横幅を使用する。“髪型”と“つむじ”は頭部領域の輝度勾配情報を、“髪の色”は頭部領域の輝度情報を使用する。最後に4つの特徴量を学習アルゴリズムである AdaBoost に適用することで人物識別を行う。

2. 提案手法

この節では、提案手法について述べる。図4は提案手法の大まかな流れを表しており、大きく(1)人物領域抽出、(2)特徴量抽出、(3)識別、により構成される。

2.1 人物領域抽出

人物を識別するためには、撮影した画像から人物領域を抽出する必要がある。本手法では、背景差分法を適用することにより大まかな人物領域を特定し、その画像からノイズを除去することにより人物領域を抽出する。

(1) 人物領域候補の抽出

人物領域を抽出するため、まず初めに背景差分法を実行することで人物領域候補を抽出する。背景差分法は一般的に輝度変化などの環境変動に影響を受けやすい。この問題を解決するために、あらかじめ用意された複数の背景候補画像から最適な背景画像を選択する。具体的には、識別対象となる画像とそれぞれの背景候補画像につ

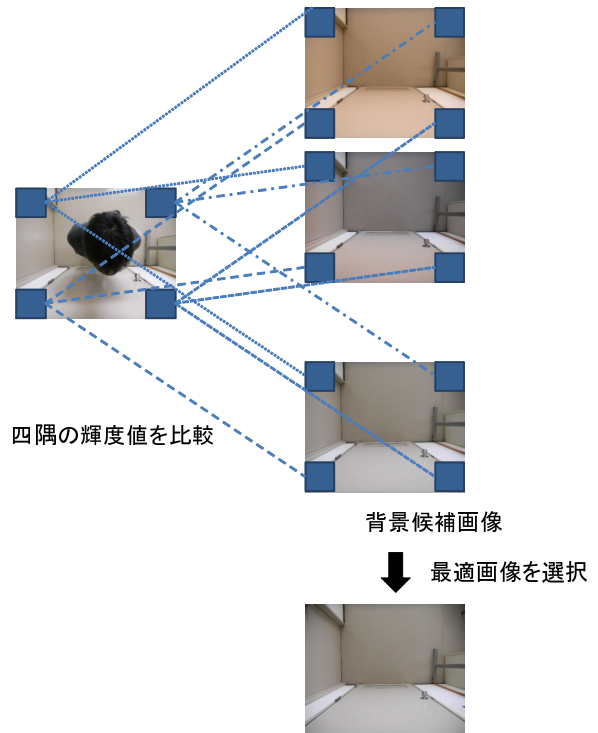


図 5 背景画像の選択

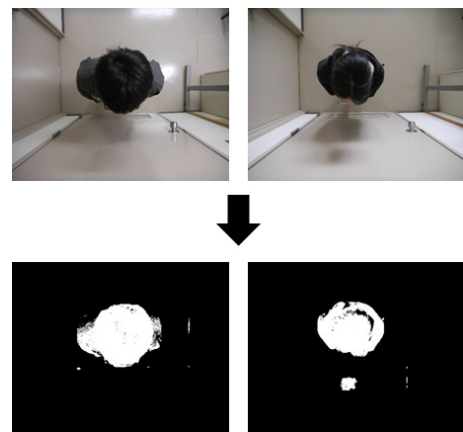


図 6 人物領域候補抽出

いて、それぞれの背景候補画像との四隅の輝度差を計算する。識別対象となる画像との輝度差が最も小さくなる背景候補画像を背景画像とする。図5は背景画像の選択を表している。選択された背景画像を使用して背景差分法を適用することにより、人物領域候補を抽出する。図6は人物領域候補の抽出を表している。

(2) 小さなノイズの除去

次に、各ピクセルの周囲の情報を参考にすることで小さなノイズの除去を行う。背景差分法により人物領域候補だとみなされた領域の中には輝度変化により発生したノイズも含まれる。このようなノイズを除去するため、人物領域候補の中に孤立した小さな領域がないかを確認する。人物領域候補とみなされた中に孤立した小さな

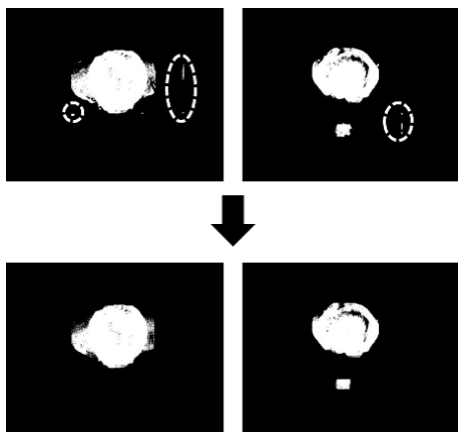


図 7 小さなノイズの除去

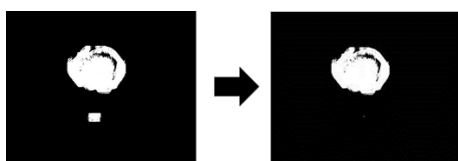


図 8 大きなノイズの除去

領域があった場合は、その小さな領域は背景が輝度変化により人物領域候補であると判断されたノイズだとする。図 7 は小さなノイズの除去を表している。

- (3) 大きなノイズの除去のための人物中心座標推定
小さなノイズの除去を行った画像から大きなノイズの除去をするため、仮の人物中心座標を推定する。背景差分画像には小さなノイズだけでなく大きなノイズを含む場合がある。大きなノイズの主な例として影がある。大きなノイズを除去するためには、 x 軸の輝度値が最大になる座標 x と y 軸の中心となる座標 y から成る座標 (x, y) を仮の人物中心座標とする。図 8 は大きなノイズの除去を表している。
- (4) 大きなノイズの除去
人物領域候補は多くの場合、小さなノイズだけでなく大きなノイズも含む場合がある。正確な人物領域の抽出を行うために大きなノイズの除去が必要となる。(3) で算出した仮の人物中心座標を用いて、人物領域はその座標から限られた距離にあるとすることで大きなノイズを除去する。
- (5) 頭部領域を決定するための人物中心座標推定
大領域のノイズを除去した画像に対して (3) と同じ処理を行うことで、再度人物中心座標の推定を行う。

2.2 特徴量抽出

この節では人物識別に用いる特徴量について述べる。

本手法では 4 つの特徴量を用いる。1 つは“体型”を数値化することにより特徴量として識別に用い、他の 3 つは髪に関するものである“髪型”、“髪の色”、“つむじ”を数値化することにより特徴量として識別に用いる。髪に関する特徴量を算出するためには頭部領域を決定する必要がある。頭部領域を決定するために 2.1 節で算出した人物中心座標を用いる。本論文では、頭部中心座標が人物中心座標と同じ座標となる 50×50 ピクセルの領域を頭部領域とする。

- 体型
体型は人物識別を行う上で重要な特徴量の 1 つとなる。本手法では両肩間の距離と頭部の前後間の距離を特徴量として用いる。これらの値は人物領域画像のヒストグラムから抽出する。
- 髪型
髪型は個人により大きな違いがある。頭部領域における輝度勾配情報を用いて髪型の特徴量を抽出する。
- 髪の色
髪の色は個人間での違いが見られることがある。本手法では頭部領域の輝度値の和を髪の色の特徴量として抽出する。
- つむじ
髪型や髪の色は同一人物でも時間的に変化する。それらの特徴量に比べ、つむじの位置や形は時間経過に頑健な特徴である。そのため、つむじは個人を識別する特徴として大いに期待できる。本システムでは頭部領域を平滑化する。その結果、髪型の輝度勾配情報が削減され、残った輝度勾配情報をつむじの特徴量として用いることができる。

図 9 は特徴量の抽出例を表している。

続いて、髪型とつむじを特徴量として数値化する方法について詳しく述べる。これらの特徴量は輝度勾配情報に基づいている。そのため、輝度勾配情報を特徴ベクトルに変換する必要がある。本手法では髪型とつむじを特徴量化するために、Dadal と Triggs [2] により提案された Histograms of Oriented Gradients (HOG) を用いる。HOG 特徴量はの人物検出や車両検出のように輝度勾配に着目する際に有効とされている特徴の一つであり、局所領域の輝度勾配に着目する輝度勾配強度で重み付けされた輝度勾配方向ヒストグラムのことである。HOG 特徴量を算出するためには、まず画像をセル (5×5 ピクセル) とブロック (3×3 セル) という領域に分割する。続いて、勾配強度 $m(x, y)$ と勾配方向 $\theta(x, y)$ を各ピクセル (x, y) について計算する。

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2}$$

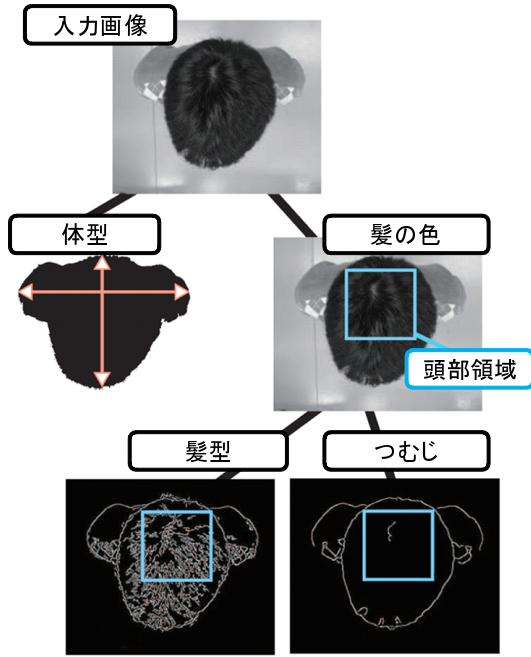


図 9 特徴量の抽出例

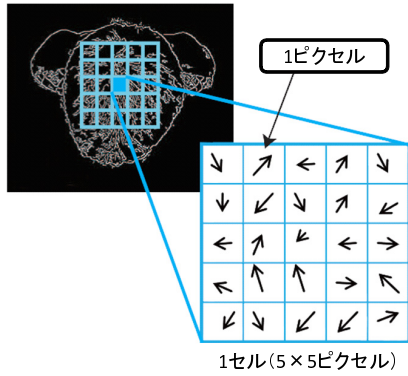


図 10 Histogram of oriented gradients

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)}$$

$$f_x(x, y) = I(x+1, y) - I(x-1, y)$$

$$f_y(x, y) = I(x, y+1) - I(x, y-1)$$

各セルにおけるヒストグラムを作成し、セルのヒストグラムから各ブロックのヒストグラムを作成する。

$$L2 - norm = \frac{v}{\sqrt{\|\mathbf{V}\|^2 + \epsilon^2}}$$

$\mathbf{V}$ はブロックの特徴ベクトルである。 ϵ はブロックを正規化するための要素であり、 $\epsilon = 1$ である。 図 10 は HOG 特徴量の算出方法を表している。

2.3 分類器による識別

最後に抽出した 4 つの特徴量に基づく人物識別を行う。 本論文では、機械学習アルゴリズムとして AdaBoost [3]

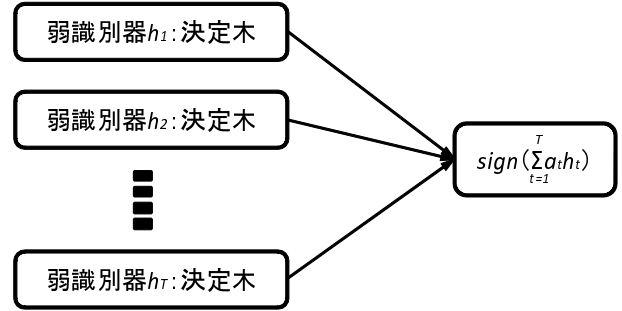


図 11 Adaboost

を用いる。 AdaBoost は機械学習アルゴリズムの中で多く用いられているものの 1 つであり、多数の弱識別器から 1 つの強識別器が構成されている 2 クラス分類器である。

弱識別器としては C4.5 アルゴリズム [8] を用いる^(注1)。 C4.5 は決定木により分類を行うアルゴリズムであり、これもまた機械学習アルゴリズムとして有名なものの 1 つである。 図 11 は AdaBoost アルゴリズムの概要を表している。

本論文では、実装にオープンソースソフトウェアである Weka^(注2)を使用する。 AdaBoost は 2 クラス分類器であるため、One-versus-Rest 法を用いることにより AdaBoost による多クラス分類を行う。

3. 実験

本実験では 8 名 (内男性 6 名、女性 2 名) の被験者に対して人物識別を行った。 画像サイズが 400×500 ピクセルのものを各被験者に 30 枚、計 240 枚の画像を使用し、10 分割交差検証による識別精度の確認をした。

3.1 実験結果

本実験では 4 つの特徴量の組み合わせ全てについて評価実験を行った。 表 1 は実験結果の一部である。 “+” は特徴量の組み合わせを表している。 例えば、“体型 + 髪の色” は“体型”と“髪の色”を特徴量として用いたときの結果である。“all” は 4 つの特徴量を全て用いたときの結果である。 精度は

$$\frac{\text{正解枚数}}{\text{全画像枚数}} \times 100[\%]$$

とした。 表 1 からわかるように、“体型 + 髪の色 + 髪型”の特徴量を組み合わせた際に精度が最も高かった (92.5%)。

3.2 考察

この節では実験結果と本手法の有効性について考察する。 本実験では、前処理である人物領域抽出の段階で

(注1)：本論文で使用するフリーソフト Weka では“J48”。

(注2)：<http://www.cs.waikato.ac.nz/ml/weka/>

表 2 “体型”, “髪の色”, “髪型” の組の詳細な結果

被験者	A	B	C	D	E	F	G	H	身長
A (男性)	28	0	0	0	2	0	0	0	175
B (女性)	0	28	0	0	0	0	2	0	155
C (男性)	0	0	27	0	1	1	1	0	171
D (男性)	0	0	0	30	0	0	0	0	181
E (男性)	0	0	2	0	28	0	0	0	176
F (男性)	0	0	0	0	0	25	0	5	168
G (女性)	0	0	0	0	0	0	30	0	148
H (男性)	0	0	0	1	0	3	0	26	175

表 1 実験結果

特徴	精度 [%]
体型	88.8
髪の色	28.3
髪型	63.3
つむじ	64.6
体型 + 髪の色	89.2
体型 + 髪型	91.7
体型 + つむじ	88.8
髪型 + 髪の色	62.9
髪型 + つむじ	66.7
体型 + 髪型 + 髪の色	92.5
体型 + 髪型 + つむじ	88.3
体型 + 髪の色 + つむじ	88.8
髪型 + 髪の色 + つむじ	66.3
all	88.8

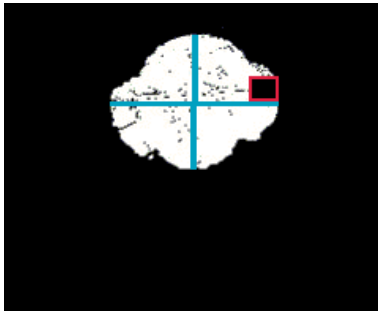


図 12 肩の一部が背景とみなされた画像

240 枚の画像中 8 枚の画像で誤抽出があった。これは被験者が床の色に近い色の服を着ていたことが原因となり、背景差分法が失敗していたためである。人物領域抽出の失敗により人物識別の精度が低下していたため、今後は背景差分の精度を向上させるか、深度情報を用いるなど色情報に頼らない方法で人物領域抽出を行う必要がある。

本実験では“体型 + 髪の色 + 髪型”の組が最も高い精度となった(表 1 では 92.5%)。表 2 はそれらの特徴量の組を用いたときの詳細な結果であり、“A”から“H”は各被験者を意味する。本実験では“C”と“G”を比較した場合のように異性間での誤識別はほとんどなく、特に女性が男性として識別されてしまうことはなかった。これは、男性と女性では体格に大きな差があることに起因している。“F”と“H”を比較すると、身長(すなわち本質

的には体格)が違うにも関わらず 8 個の誤識別があった。この問題は服装が原因となっていた。被験者“F”はよくダウンジャケットのような厚い服を着ていた。その結果、本システムは被験者“F”の体型特徴量を実際の体型よりも多く見積もってしまい、その被験者よりも体格のよい被験者“H”との誤識別が多く発生した。このように本手法では服装が体型特徴量の一部として反映され、誤識別の原因となってしまうことがある。しかし、服装の情報は人物識別においてコンテキスト情報として有効な特徴量の一つとして使用することが可能である。Gallagher と Chen [5] は服装の情報が画像中の人物を識別する際の有効性を示している。山口ら [11] は部分的に顔に隠れが生じた人物識別に対して服装をコンテキスト情報として適用している。本手法においては、服装の情報をコンテキスト情報として追加することにより精度向上することが今後の課題となる。

体型特徴量は単独の特徴としては最も高い精度を示している(表 1 では 88.8%)。体型特徴量は先に述べた背景差分の影響を受けやすい。しかし本論文では体型特徴量として縦幅と横幅を用いており、これによって人物識別の精度の低下を抑えることができた。具体的な例を図 12 に示す。図 12 は肩の一部が背景とみなされた画像例である。この場合、肩の一部は欠損しているが、縦幅と横幅は概ね正確に推定が可能であり、その結果人物識別への影響を抑えることができた。このような場合、例えば体型特徴量として身体の面積を用いる手法では実際と推定値が大きく異なる可能性がある。これは体型特徴に縦幅と横幅を用いることの妥当性を示している。

髪の色は本実験において単独の特徴量としては最も識別精度が低かった(表 1 では 28.3[%])。その主な原因は 2 つある。1 つ目は輝度情報の特徴量として用いているため照明変化に脆弱であるということである。2 つ目の理由は、日本人は多くの人が黒い髪であるからである。これら 2 つの理由により、ほとんど色に差がない被験者の髪の色では照明変動の影響もあり人物を識別することができないという結果になった。その一方で、この特徴量は髪を染めるなどしていることにより対象者の中で少数派に属する人物に対しては有効な特徴量となることが期待できる。

つむじは単独で 64.6%というあまり高くない精度であ

り、さらに組み合わせによっても必ずしも有効ではなかった。本手法ではつむじ特微量抽出の際に頭部領域を平滑化することで髪型の情報を削減したが、平滑化画像にも髪型による輝度勾配情報をノイズとして含むことがあった。そのため、つむじを特微量として正しく抽出できないことがあり、つむじが精度の向上につながらなかった。しかし、つむじは時間的な変化に頑健であるというメリットがある。本実験の期間内では髪型が大きく変化した被験者がいなかったため、つむじ特微量が必ずしも有効には機能しなかったが実験の対象がより長期間になれば、髪型に比べてつむじは変化が小さく精度の向上に期待できる。つむじ特微量による識別の精度を向上させるためには、頭部領域の輝度値をより強く平滑化することで髪型がノイズとして含まれないようにするか、つむじの特微量化の方法として他の手法を考える必要がある。

本手法で用いた特微量の抽出方法は人物が立ち止まっていることを前提としている。この前提条件はドアの前など日常生活において見かけることのある場面であるが、より自由度の高いシステムを構築するために歩いている人や動いている人も対象とすることも今後の課題の1つである。また本手法の人物領域抽出手法は、画像内の人物が一人であることを前提としているが、人数計測を行った後に人物領域を抽出するなどして複数人が写っている画像に対応することで、より自由度の高いシステムを構築することが可能である。

提案手法は識別結果の出力として部外者を含んでいない。この理由は、本手法がオフィスや研究室のような少人数の環境での出席管理や人物検索システムを対象としているためである。より高度なセキュリティシステムとして用いるためには提案手法に棄却処理などの導入が必要であり、これは今後の課題の1つである。

3.3 関連研究

内田ら[10]はエレベータホールにおける頭上方向カメラを用いた人物検出手法を提案している。彼らの研究の目的は人物と車椅子の計測であった。我々の研究の目的は画像中の人物の識別である。IwashitaとStoica[6]は頭上方向カメラを用いた歩行認識手法を提案している。彼女らは歩行認識を行うため影に着目している。我々は体型や髪型などの4つの特微量を用いている。福添ら[4]は人物の輪郭を用いた人物識別の手法を提案している。しかし彼らは頭上方向カメラを用いていないため、システムの有効性については画像中の隠れが問題となっている。

4. ま と め

本論文では人物識別における新しい手法を提案した。画像を用いた人物識別では顔特徴を用いる手法が主流であるが、混雑している場所などでは顔画像を正しく取得できない場合があるという問題がある。その問題を解決

するため、本論文では頭上方向からのカメラ画像を用いた人物識別の手法を提案した。本手法はオフィスや研究室のような少人数の環境を対象としており、出席管理や人物検索、行動認識など多くの場面に応用可能である。

本手法は背景差分を行うことで人物領域抽出を行い、その画像から4つの特微量を抽出することで人物識別を行う。学習アルゴリズムとしてはAdaBoostとC4.5を適用している。“体型”、“髪の色”、“髪型”の特微量を用いることで、92.5%の識別精度となり、提案手法の有効性を確認した。本実験では、“つむじ”は人物識別の精度向上には貢献しなかったが、髪型などと比較し、時間変動に頑健であり、本質的には有効な特徴であると考えられる。実際に“つむじ”が有効に機能しなかったのは、頭部領域からのつむじの抽出精度が十分ではなかったことが大きな要因であり、その特微量抽出処理の改善は識別精度の向上のためには不可欠である。

今後の課題をまとめると、(1) つむじ特微量の抽出手法の改善、(2) 多人数のデータを用いた実験、(3) コンテキスト情報として付属品(鞆など)や服装(種類や色、形など)の情報を使用、(4) 歩いている人物への対応、(5) 複数人が写っている画像への対応、が挙げられる。

文 献

- [1] S. Watanabe and N. Pakvasa. “Subspace method in pattern recognition,” Proc. of 1st Int. J. Conf on Pattern Recognition, pp.2–32, 1973.
- [2] N. Dalal and B. Triggs. “Histograms of Oriented Gradients for Human Detection,” Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886–893, 2005.
- [3] Y. Freund and R. E. Schapier. “Experiments with a new boosting algorithm,” Proc of ICML pp. 148–156, 1996.
- [4] 福添孝明, 伊藤雅人, 水戸大輔, 渡邊 睦, “体型特徴と習慣性特徴の確率的統合認識に基づく非拘束状態下での人物同定,” 電子情報通信学会 2008, Vol. J91-D, No. 5, pp. 1369–1379, 2008.
- [5] A. C. Gallagher and T. Chen. “Using Context to Recognize People in Consumer Images,” IPSJ Transactions on Computer Vision and Applications, Vol.1, pp.115–126, 2009.
- [6] Y. Iwashita and A. Stoica. “Gait Recognition using Shadow Analysis,” Proc. of Symposium on Bio-inspired, Learning, and Intelligent Systems for Security 2009, pp. 26–31, 2009.
- [7] T. Kanade. “Picture processing by computer complex and recognition of human face,” Technical report, Kyoto University, Dept. of Information Science, 1973.
- [8] J. R. Quinlan. “C4.5 Programs for Machine Learning,” Morgan Kaufmann Publishers, (1993).
- [9] M. Turk and A. P. Pentland. “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71–86, 1991.
- [10] 内田 光, 猪田良介, 辻 俊明, 阿部 茂, “実時間画像処理によるエレベータ乗場の人数計測と車椅子の識別,” 電気学会論文誌, 129, No. 6, pp.578–584, 2009.
- [11] 山口 純平, 嶋田 和孝, 榎田 修一, 江島俊朗, 遠藤 勉, “顔特徴とコンテキスト情報に基づく顔の隠れに頑健な人物識別,” 電子情報通信学会技術報告書, 電子通信学会, PRMU2009-237, pp.25–30, 2010.