

# コンテキスト情報を利用した人物識別の改良と機械学習の適用

小松 和朗<sup>†</sup> 嶋田 和孝<sup>††</sup> 遠藤 勉<sup>††</sup>

<sup>†</sup> 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

<sup>††</sup> 九州工業大学大学院情報工学研究院 知能情報工学研究系

〒 820-8502 福岡県飯塚市川津 680-4

E-mail: †{k\_komatsu,shimada,endo}@pluto.ai.kyutech.ac.jp

あらまし 本稿では、隠れに頑健な人物識別手法として、顔特徴とコンテキスト情報を用いた手法について報告する。人物識別においては、衣服や時間情報は重要なコンテキスト情報になりうる。本研究では HOG に基づき得られた衣服特徴と、画像データの撮影時間や曜日などの時間特徴をコンテキスト情報として新たに導入する。顔に隠れが生じた画像に対して、顔特徴のみでの識別結果と本手法を用いた識別結果を比較することで、本手法で取得したコンテキスト情報の有効性を確認する。また、各特徴量に対する重みづけを機械学習によって自動化する。人手による手動での重みづけによる識別結果と機械学習による自動での重みづけによる識別結果を比較することで、機械学習を用いることの有効性を検証する。

キーワード 人物識別, コンテキスト情報, 顔特徴, HOG, 時間特徴, 機械学習

## 1. はじめに

近年、人物識別が様々なタスクにおいて重要な役割を持ってきている。例えば、セキュリティ分野では、駅や空港、ショッピングモールなどの公共の場において、監視カメラの映像に映る不特定多数の人々の中から特定の人物を探し出すためなどに用いられている。また、ロボットの分野でも、生活支援ロボットや受付ロボットなど、人間とコミュニケーションをとるロボットとの対話において、より高度なコミュニケーションを行うために相手の情報を得る手段として用いられている。この人物識別であるが、用いられる特徴には様々なものがあり、指紋、虹彩、音声、顔、筆跡、シルエット、ジェスチャなどがあげられる。これらの特徴の中でも、特に顔特徴を用いた人物識別は対象者に意識させずに非接触で実現できるため様々な分野で用いられている [1]。

しかしながら、顔特徴のみによる手法では、対応することが困難な様々な問題がある。このような問題の解決方法として、コンテキスト情報の利用が挙げられる。コンテキスト情報とは、衣服や時間など顔特徴以外で人物識別に役立つ特徴のことである。Gallagher ら [2] は双子や姉妹のような顔の似た人物の画像のデータベースに対し、コンテキスト情報が顔特徴を用いた人物識別にどの程度影響を与えるかを調べるために実験を行った。実験の結果、顔のみの画像で 58%、衣服まで含む画像で 88% という平均正解識別率を得た。この結果は衣服を使うことで顔の似た人物に対する人物識別を改良できる可能性を示しており、コンテキスト情報は人物識別に有効であると Gallagher らは述べている。また別の問題として、

顔の隠れに対する対応もある。実環境下では、サングラスやマスク、帽子などによって顔の一部が隠れることは十分に起こりうる。この顔特徴の欠損問題にコンテキスト情報を用いて対応した研究として、山口らの人物識別手法 [3] がある。山口らは衣服から取得した特徴をコンテキスト情報として加え、それらを統合することで顔の隠れに頑健な人物識別を凶っている。

本研究では、研究室やオフィスなどにおいて顔の一部、また、顔全体に隠れが生じた状態での人物識別を想定し、山口らが提案した人物識別手法を以下のように改良する。

- HOG に基づく衣服特徴と時間情報に基づく時間特徴の追加
- 機械学習による自動での重みづけの適用

画像処理において有効性がある手法である HOG が山口らの手法では用いられていないこと、時間情報が Gallagher らの提案したコンテキスト情報に当てはまることから、新たな手法で取得した特徴量がコンテキスト情報として有効であるかの検証を行う。また、山口らの手法では顔特徴と衣服特徴の類似度統合時に手動での重みづけを行っているため、データの変更があった場合の頑健性に問題がある。この問題を解決するために特徴量の統合時に機械学習による自動での重みづけを行うことが有効であるかの検証を行う。

本論文は、本章を含めて 5 章から構成され、第 2 章では本研究のベースとなる人物識別として、衣服特徴をコンテキスト情報として用いた先行研究を紹介する。続いて、3 章では本研究の提案手法の概要、4 章では実験とその考察について述べる。最後に 5 章でまとめる。

## 2. 先行研究

本章では、衣服特徴をコンテキスト情報として用い、顔の隠れに頑健な人物識別を行った先行研究を紹介する。

### 2.1 先行研究の概略

山口らの手法による人物識別手法の概略を図 2.1 に示す。まず、USB カメラから人物画像を取得する。次に、顔特徴類似度算出部では図 2.2 のような顔領域を検出し、その検出領域から特徴量を取得する。そして、その特徴量とデータベースの学習データを用いて類似度を求める。衣服特徴類似度算出部でも図 2.3 のような衣服領域を検出し、顔特徴と同様にして類似度を求める。求められたそれぞれの類似度は統合部で統合され、最後に人物の識別結果を出力する。

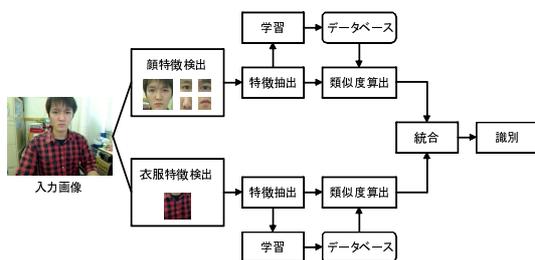


図 2.1 山口らの手法の概略図。



図 2.2 類似度算出に用いる顔領域 (左から正面顔, 右目, 左目, 鼻, 口)。



図 2.3 類似度算出に用いる衣服領域の例。

### 2.2 領域検出と特徴抽出

顔特徴の検出では、Viola ら [4] が提案し、Lienhart ら [5] によって改良された物体検出器に、元吉ら [6] によって追加された処理に基づく検出方法を利用する。まず、入力画像から正面顔、右目、左目、鼻、口の検出を行う。検出された正面顔、右目、左目、鼻、口の各領域から図 2.4 のように特徴量として画素値を特徴ベクトルとして求める。求めた特徴ベクトルに CLAFIC 法 [7] を適用することで各領域の類似度を求める。

また、衣服は顔の下にあるという仮定の下、図 2.5 のように衣服の領域を検出を行って衣服領域を検出する。その際、山口らの手法では衣服領域の特徴量には以下の 4 つを用い、それぞれの特徴量で類似度を求める。

- 色ヒストグラム特徴  
HSV 表色系に基づき得られた値により作成されたヒストグラムから得られる特徴
- モザイク画像の色特徴  
衣服画像のモザイク化により取得できる局所的な色特徴
- パワースペクトル画像特徴  
2次元離散フーリエ変換による周波数分解で得られたパワースペクトルから画像を生成し、その画像の画素値から取得した特徴
- 高次局所自己相関特徴  
着目した輝度値に様々な種類のマスクパターンを適用することで得られる特徴

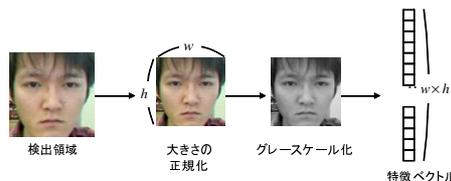


図 2.4 顔領域からの特徴抽出例。

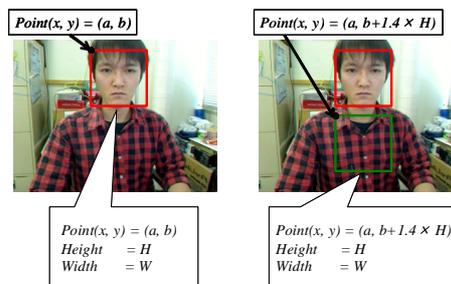


図 2.5 衣服領域の検出方法。

山口らの手法では取得した顔領域の各特徴量と衣服特徴量の統合を式 (2.1)、式 (2.2) のように行っている。

$$S_u = \alpha \times S_f + \beta \times S_c \quad (2.1)$$

$$\begin{cases} S_f = S_{face} + S_{reye} + S_{leye} + S_{nose} + S_{mouth} \\ S_c = S_{hist} + S_{mosaic} + S_{fourier} + S_{mask} \end{cases} \quad (2.2)$$

$S_f$  は顔特徴類似度算出部で求められる類似度、 $S_c$  は衣服特徴類似度算出部で求められる類似度、 $S_u$  は顔特徴類似度と衣服特徴類似度を統合した類似度を表している。式 (2.2) において、 $S_{face}$  は正面顔領域から得られた類似度、 $S_{reye}$  は右目領域から得られた類似度、 $S_{leye}$  は左目領域から得られた類似度、 $S_{nose}$  は鼻領域から得られた類似度、 $S_{mouth}$  は口領域から得られた類似度を表している。また、 $S_{hist}$  は色ヒストグラム特徴による類似度、 $S_{mosaic}$  はモザイク画像の色特徴による類似度、 $S_{fourier}$  はパワースペクトル画像特徴による類似度、 $S_{mask}$  は高次局所自己相関特徴による類似度を表している。顔特徴類似度算出部で求められる類似度と衣服特徴類似度算出部で求められる類似度では、それぞれスケールが異なる。そのため、統合時にスケールが一定になるようにスケール

リングを行う。

また、顔特徴のみでも高い精度が得ることができるとが実験的に分かっているため、統合時に式 (2.1) のように重みづけを行う。山口らは実験的に  $S_f$  の重み  $\alpha$  を 1.0,  $S_c$  の重み  $\beta$  を 0.5 と設定している。

### 3. 提案手法

本章では、本研究で提案する新たなコンテキスト情報と特徴量の統合時に機械学習を適用することについて述べる。まず、山口らの手法での問題点を述べた後、解決策として新たなコンテキスト情報の実装、機械学習の適用について述べる。最後に本研究の提案手法の概要を述べる。

#### 3.1 提案手法概要

山口らの手法ではコンテキスト情報として4つの衣服特徴を用いており、顔に隠れが生じない場合における人物識別は非常に優れた結果を挙げている。しかし、新たなコンテキスト情報を追加することでより識別率の向上が見込まれる。また、顔特徴と衣服特徴の統合時には手動での重みづけを行っており、データの変更があった場合、現在の重みが最適であるとは限らない。そのため、手動で設定した重みの値の頑健性について議論が残る。

##### 3.1.1 本研究で用いる手法

本研究では、先ほど述べた山口らの手法の問題点に対して以下の改良を行っている。3.2 節以降から具体的な説明を行う。

- 新たなコンテキスト情報の実装  
HOG [8], 時間情報
- 自動での重みづけ手法の適用  
Adaboost [9], C4.5 [10]

新たなコンテキスト情報の実装では、画像処理において有効性が確認されている HOG 特徴と時間情報から得られる時間特徴をコンテキスト情報として実装する。重みづけの自動化では、類似度統合時において上記の機械学習による自動での重みづけを行う。なお、実装の際にはデータマイニングツールである Weka [11] を使用する。

##### 3.1.2 提案手法概略図

本研究で提案する手法の概略図を 3.1 に示す。まず、学習データ、テストデータの2種類のデータを用意する。これらのデータに山口らの類似度算出手法を適用することで画像の類似度を算出する。次に、算出された各類似度を用いて機械学習を行い、学習・評価した結果を用いて人物識別を行う。

#### 3.2 HOG

山口らの先行研究では色ヒストグラム特徴、モザイク画像の色特徴、パワースペクトル画像特徴、高次局所自己相関特徴の4つの手法を用いて衣服特徴を取得してい

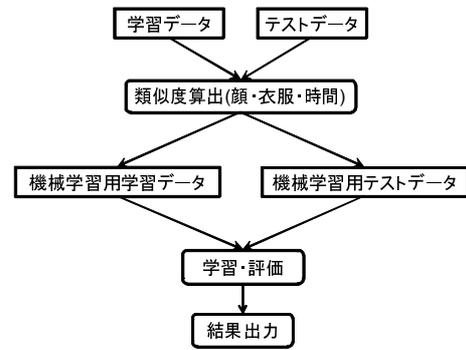


図 3.1 提案手法の概略図。

る。その中で高次局所自己相関特徴という輝度値に基づく特徴に着目する。

近年の画像処理研究において、輝度値に基づく特徴量を用いる手法が注目されている。その1つである HOG (Histograms of Oriented Gradients) は、2005 年に Navneet Dalal と Bill Triggs によって提唱された特徴量 [8] であり、入力画像の輝度勾配を求め、局所領域ごとに勾配方向で区間分割してヒストグラム化することで特徴量を得ることができる。HOG で得られた特徴量は、局所的な幾何学的変化・明度変化に対して不変であるが、回転・スケール変化に対しては不変ではないとされている。しかし、一定領域に対する特徴量を算出することができるため、大まかな物体形状を表現することが可能である。そのため、人検出や車体検出等の一般物体認識等に用いられている [12]。

そこで、本研究では HOG に基づき得られた衣服特徴をコンテキスト情報として適用し、その有効性を検証する。

##### 3.2.1 HOG に基づく衣服特徴抽出

以下に HOG 特徴量の算出法について述べる。

Step1 入力画像を一定サイズに正規化する。

Step2 各ピクセルにおける勾配強度  $m$  と勾配方向  $\theta$  を式 (3.1), 式 (3.2), 式 (3.3) を用いて算出する (図 3.2)。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (3.1)$$

$$\theta(u, v) = \tan^{-1} \frac{f_v(u, v)}{f_u(u, v)} \quad (3.2)$$

$$\begin{cases} f_v(u, v) = I(u+1, v) - I(u-1, v) \\ f_u(u, v) = I(u, v+1) - I(u, v-1) \end{cases} \quad (3.3)$$

Step3 算出された勾配強度  $m$  と勾配方向  $\theta$  を用いて、 $5 \times 5$  ピクセルを 1 セルとした領域における輝度勾配ヒストグラムを作成する。勾配方向は  $0^\circ - 360^\circ$  であるが、勾配方向の向きを考慮する必要がないため  $0^\circ - 180^\circ$  とする。これを  $20^\circ$  ずつに分割するため、9 方向の輝度勾配ヒストグラムが作成される。

Step4 各セルで作成した輝度勾配ヒストグラムを  $3 \times 3$  セルを 1 ブロックとして正規化を行い特徴量  $Fh$  を算

出する。正規化はブロックを1セルずつずらしながら全領域に対して行う。

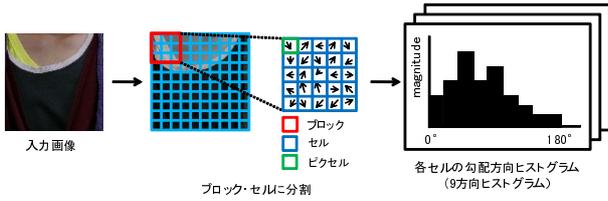


図 3.2 輝度勾配の算出および輝度勾配ヒストグラムの作成。

### 3.2.2 HOG に基づく衣服特徴の類似度算出

HOG に基づく衣服特徴の類似度算出には、衣服特徴の学習データとテストデータのセルごとの特徴量の差分を求め、それを全てのセル分合計する。HOG 特徴の類似度  $S_{hog}$  を算出するにあたり、式 (3.4) を用いる。式 (3.4) において、 $Fh_{test}$  はテストデータのセルの特徴量、 $Fh_{std}$  は学習データのセルの特徴量、 $C$  はセルの総数を表している。

$$S_{hog} = \sqrt{\sum_{i=1}^C \sum_{j=1}^C (Fh_{test_i} - Fh_{std_j})^2} \quad (3.4)$$

なお、HOG 特徴の類似度  $S_{hog}$  は、2.1.1 節で述べた式 (2.2) の衣服特徴類似度算出部  $S_c$  に追加されるため、衣服特徴類似度算出部  $S_c$  は式 (3.5) のようになる。

$$S_c = S_{hist} + S_{mosaic} + S_{fourier} + S_{mask} + S_{hog} \quad (3.5)$$

## 3.3 時間情報

本研究は研究室やオフィスなどの限定された空間における人物識別を想定している。上記のような限定された空間ではある特定の人物による入退室が多く、入退室の時間帯や曜日が習慣的になることが容易に想像できる。このことから、時間情報もコンテキスト情報として有効であると考えられる。

そこで、本研究では撮影時間に基づく時間特徴及び曜日に基づく時間特徴をコンテキスト情報として適用し、その有効性を検証する。以降、撮影時間に基づく時間特徴を撮影時間特徴、曜日に基づく時間特徴を曜日特徴と呼ぶことにする。

### 3.3.1 時間情報からの特徴抽出

時間情報としては、日付、曜日、入退室の時間帯、滞在時間などが挙げられる。図 3.3 のように、画像の撮影時間から先ほど列挙した時間情報を取得して数値化することで特徴抽出を行う。本研究では撮影時間と曜日を時間情報として用いる。

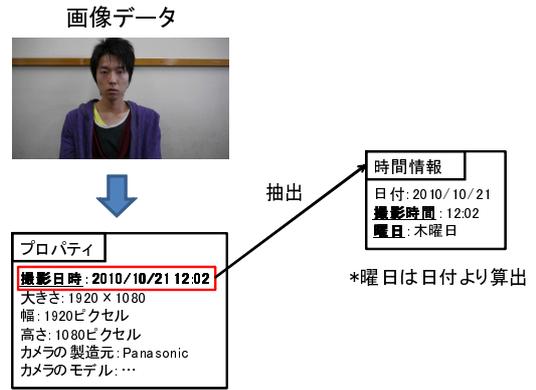


図 3.3 時間特徴の情報抽出例。

### 3.3.2 時間特徴の類似度算出

時間特徴の類似度算出には、学習データの最頻値とテストデータとの差分を用いて算出する。最頻値とは、データの出現率が最大の値のことである。最頻値は平均値などとは異なり、外れ値の影響を受けないため、データの特性が掴みやすいという特徴がある。また、撮影時間は 24 種類の時間帯、曜日は 7 種類の曜日のサイクルで構成されていることに着目する。これは、最頻値とテストデータとの距離を類似度とするためである。例えば撮影時間特徴の類似度算出において、1 時と 23 時との距離を単純に差分で求めると 22 となるが、実際の距離は 2 であり、後者の数値が類似度として妥当であると考えられる。

時間特徴の類似度算出式を式 (3.6) に示す。

$$S_T = S_t + S_d \quad (3.6)$$

式 (3.6) において、 $S_T$  は撮影時間特徴の類似度  $S_t$  と曜日特徴の類似度  $S_d$  を統合した値を表している。以下に時間情報に基づく時間特徴の類似度算出方法について述べる。

Step1 学習データを時間帯別・曜日別にそれぞれ集計した後、カウント数を比較することで最頻値  $Mo$  を求める。  
Step2 最頻値  $Mo$  とテストデータの時間特徴  $F_{test}$  の差分の絶対値を求め、正規化後にその絶対値を 1 から引いた値を時間特徴の類似度とする。

この算出方法を撮影時間情報、曜日情報に適用すると式 (3.7) のようになる。

$$\begin{cases} S_t = 1 - \frac{|Mo_t - F_{t_{test}}|}{SV_t} \\ S_d = 1 - \frac{|Mo_d - F_{d_{test}}|}{SV_d} \end{cases} \quad (3.7)$$

式 (3.7) において、式中の語句は以下を表している。

- $Mo_t, Mo_d$   
学習データの撮影時間特徴・曜日特徴からそれぞれ算出した最頻値
- $F_{t_{test}}, F_{d_{test}}$   
テストデータの撮影時間特徴・曜日特徴からそれぞれ算出した特徴量



- 撮影期間：2010/10/20～2010/12/02

本研究では顔の一部、また、顔全体に隠れが生じた状況下での人物識別を行う。今回は、サングラスの着用により両目が隠れた状態、マスクの着用により鼻と口が隠れた状態、サングラスとマスクの両方の着用により顔が全て隠れた状態の3つの状態で人物識別を行う。なお、山口らの手法で顔特徴類似度と衣服特徴類似度を統合した類似度  $S_u$  を算出する際、式 (3.8) で設定した重みである  $\alpha$ ,  $\beta$ ,  $\gamma$  は、それぞれ適切な値に設定した。

## 4.2 実験データ

実験データとして、被験者7名から学習データ875枚とテストデータ175枚を取得した。この学習データとテストデータについて以下で説明する。

### 4.2.1 学習データ

学習データは、1人の人間が正面を向いて写った画像を約1か月間にわたり撮影した。その画像の一例を図4.1に示す。図4.1のように、被験者には異なる種類の衣服を着てもらっている。この学習データを用いて、顔特徴と衣服特徴の学習を行うが、各被験者の顔特徴は撮影した画像のうち、顔特徴が取得可能な画像を被験者1名につき125枚を選んでいる。これを被験者7名に対して計875枚を用意した。なお、撮影時間情報を被験者1名につき25種類、これを被験者7名分用意した。

本実験では山口らの手法に加え、3章で述べた手法で特徴量を抽出し、学習用データとしてデータベースに登録した。



図 4.1 学習データ例.

### 4.2.2 テストデータ

テストデータは、学習データと同様に1人の人間が正面を向いて写った画像を約1か月間にわたり撮影した。その画像の一例を図4.2に示す。学習データと同様に被験者には異なる種類の衣服を着てもらっている。各被験者の顔特徴は撮影した画像のうち、顔特徴が取得可能な画像を被験者1名につき25枚選んでいる。これを被験者7名に対して計175枚用意した。このテストデータを用いて人物識別を行う。

なお、今回の実験では実際にサングラスやマスクを着用していない。そのため、サングラス着用時は正面顔と鼻と口の特徴量を、マスク着用時は正面顔と左目と右目

の特徴量を、サングラスとマスクの両方着用時は顔特徴全ての特徴量を削除することで、仮想的に着用している状態にしている。



図 4.2 テストデータ例.

### 4.2.3 評価方法

この実験の評価基準には、識別率を採用した。本研究の識別率は入力された画像に対して正しく人物を判断できる割合を表しており、以下の式で求めることができる。

$$\text{識別率 (\%)} = \frac{\text{正しく本人と識別された画像数}}{\text{全テスト画像数}} \quad (4.1)$$

## 4.3 検証実験

### 4.3.1 スコアリング判定による人物識別

3章で述べた評価式 (3.8) に基づくスコアリング判定を適用し人物識別を行った。その結果を表1に示す。表中の語句は以下のことを表している。

- $PREV_{Scoring}$  : 先行研究 [3] の素性のみ
- $+HOG_{Scoring}$  : 先行研究の素性に HOG 特徴量を追加
- $+撮影時間_{Scoring}$  : 先行研究の素性に撮影時間特徴量を追加
- $+曜日_{Scoring}$  : 先行研究の素性に曜日特徴量を追加

なお、「サングラス」はサングラスを着用した状態、「マスク」はマスクを着用した状態、「サングラス+マスク」はサングラスとマスクの両方を着用した状態をそれぞれ表している。

表1より、HOGに基づく衣服特徴はマスク着用時には0.5%、サングラスとマスクの両方着用時には2.3%と識別率の向上がみられた。しかし、サングラス着用時には識別率が1.1%低下しているなど、HOGに基づく衣服特徴を用いることにより識別に失敗している画像もある。全体的にみるとHOGに基づく衣服特徴は識別率の向上に有効であるといえる。一方、撮影時間特徴量・曜日特徴量を追加した場合、識別率が減少するなど、時間情報の有効性はみられなかった。

表1 スコアリング判定による実験結果.

	サングラス	マスク	サングラス+マスク
$PREV_{Scoring}$	91.4 %	98.9 %	81.7 %
$+HOG_{Scoring}$	90.3 %	99.4 %	84.0 %
$+撮影時間_{Scoring}$	85.1 %	96.6 %	71.4 %
$+曜日_{Scoring}$	85.1 %	96.6 %	71.4 %

### 4.3.2 考察

最初に HOG 特徴量について考察する。まず、山口らの設定した素性である高次局所自己相関で得られた衣服特徴と今回提案した HOG で得られた衣服特徴を用いて検証を行う。図 4.3 は高次局所自己相関に基づく衣服特徴では識別に成功し、HOG に基づく衣服特徴では識別に失敗した衣服画像例である。実験データの解析を行ったところ、HOG で識別に失敗した衣服画像というのは衣服模様がシンプルなものが多く、模様情報がほとんど取得できないことがわかった。このことから、模様情報がほとんど取得できない衣服については HOG で取得できる特徴量が少なく、その結果、識別率向上につながらなかったと考えられる。



図 4.3 HOG で識別に失敗した画像例。

また、図 4.4 は高次局所自己相関・HOG の両方とも識別に失敗した衣服画像例である。これは、設定した衣服領域内では衣服の形が似ており、輝度勾配で取得できる特徴量に差異が表れなかったためである。その結果、識別率向上につながらなかったと考えられる。対策としては、現在、衣服領域を顔の高さ×1.4ピクセル分下の領域と設定しているが、上記以外に両肩周辺や衣服全体などを含む領域を新たに衣服領域と設定する。これにより、例えば同じ上着を着用している人物が異なるインナーを着用しているなど、現在設定した衣服領域のみの特徴抽出だけでは特徴量の差異が現れない場合の誤認識を防ぐことができる。また、各領域に最適な手法を適用することで、より質の高い衣服特徴のみを多く取得できるため、衣服特徴によるさらなる識別率向上が期待される。

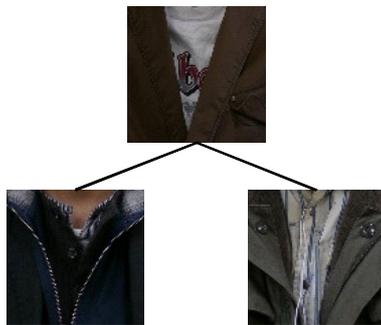


図 4.4 輝度勾配が有効に機能しない画像例。(真中：識別対象画像，左下：高次局所自己相関で選択された画像，右下：HOG で選択された画像)

次に、時間情報について考察する。時間情報の有効性がみられなかった最大の原因は、学習データに設定した時間情報が顔情報や衣服情報に比べて少ないことが挙げ

られる。また、時間情報そのものから取得できる特徴量が少ないことも原因の一つとして挙げられる。今回使用した撮影時間特徴や曜日特徴の他に入室の時間帯・滞在時間などの特徴が撮影時間から算出可能であることから、これらの時間特徴を用いることで時間情報の特徴量増加を図る必要がある。

## 4.4 機械学習の検証実験

### 4.4.1 実験結果

機械学習による自動で適切な重みづけを行い、人物識別を行った。その実験結果を表 2 に示す。また、表中の語句は以下のことを表している。

- $PREV_{ML}$  : 先行研究の素性のみ
- $+HOG_{ML}$  : 先行研究の素性に HOG 特徴量を追加
- $+撮影時間_{ML}$  : 先行研究の素性に撮影時間特徴量を追加
- $+曜日_{ML}$  : 先行研究の素性に曜日特徴量を追加
- $+HOG \& 撮影時間_{ML}$  : 先行研究の素性に HOG 特徴量, 撮影時間特徴量を追加
- $+HOG \& 曜日_{ML}$  : 先行研究の素性に HOG 特徴量, 曜日特徴量を追加
- $+撮影時間 \& 曜日_{ML}$  : 先行研究の素性に撮影時間特徴量, 曜日特徴量を追加
- $+ALL_{ML}$  : 先行研究の素性に HOG 特徴量, 撮影時間特徴量, 曜日特徴量を追加
- $Best_{Scoring}$  : スコアリング判定における識別率の最大値 (表 1 のベスト)

表 2 から、スコアリング判定に比べて、機械学習の適用によって識別率がサングラス着用時には最大 5.2 %、サングラス+マスク着用時には最大 13.7 % 向上したことが示された。このことは、機械学習の適用が識別率の向上に効果があったことを示している。しかし、マスク着用時には識別率が最小で 2.3 % 低下するなど、有効性が見られない場合もあった。

表 2 機械学習を適用した実験結果。

	サングラス	マスク	サングラス+マスク
$PREV_{ML}$	93.7 %	97.1 %	97.1 %
$+HOG_{ML}$	93.7 %	92.6 %	93.7 %
$+撮影時間_{ML}$	96.0 %	94.9 %	<b>97.7 %</b>
$+曜日_{ML}$	93.7 %	93.1 %	96.6 %
$+HOG \& 撮影時間_{ML}$	93.7 %	92.6 %	93.7 %
$+HOG \& 曜日_{ML}$	92.6 %	92.6 %	88.6 %
$+撮影時間 \& 曜日_{ML}$	<b>96.6 %</b>	92.6 %	<b>97.7 %</b>
$+All_{ML}$	92.6 %	94.9 %	92.6 %
$Best_{Scoring}$	91.4 %	<b>99.4 %</b>	84.0 %

### 4.4.2 考察

実験の結果を詳細に解析したところ、コンテキスト情報を含むすべての特徴の中で、目に関する特徴量ももっとも重要な要素であり、精度の向上に起因していることが分かった。<sup>(注1)</sup>スコアリングに基づく手法では、顔特徴

(注1) : 具体的には、本データにおいては、両目の特徴量のみで識別した場合、その精度は 88 % であった。

の重みを衣服特徴と比べ、もともと大きく設定しており、両目の特徴量が極端に重要視され、補足的に他の特徴量が扱われた結果、高い精度が得られたのだと考えられる。

一方、機械学習では汎化を行うため、スコアリングとは異なり重みの値が比較的万遍なく設定される。そのため、顔特徴が取得できない場合には、機械学習による重みづけが手動での重みづけよりも効果があったと考えられる。そこで、目の情報が取得できる場合はスコアリング判定を、そうでない場合は機械学習を適用するなど、特徴量の取得状況によって手法の使い分けを行うことで、人物識別全体の精度向上につながると考えられる。

## 5. おわりに

本研究では、HOG で取得した衣服特徴と撮影時間・曜日から取得した時間特徴を新たなコンテキスト情報として人物識別に利用することで、顔の一部に隠れが生じても高い精度で人物識別が行えるようにした。

本手法では、山口らの手法をベースに、HOG に基づき得られた衣服特徴と撮影時間・曜日から取得した時間特徴を新たにコンテキスト情報として加えた。顔特徴と衣服特徴、時間情報の特徴量を抽出した後に類似度のスケールを統一し、機械学習を行った。機械学習により自動で重み付けが行われた類似度を利用することで人物識別の精度向上を図った。

実験では顔の一部、また、顔全体に隠れが生じた場合を想定し、本手法を用いることで手法の有効性を確認した。実験結果より、新たに追加したコンテキスト情報が一定の条件下における人物識別に対して有効性が確認された。また、特に顔特徴が全く取得できない場合において、機械学習による自動での重みづけの有用性が確認された。

今後の課題としては、より頑健なシステムを目指して、衣服領域検出の改良や、統合手法の改善、他のコンテキスト情報の追加が必要である。本研究ではテストデータが学習データのどれにも当てはまらない場合は部外者であるという判定を行っていない。そこで、閾値を設けることで部外者を棄却するなどの対応も必要になってくる。また、現在は時間情報として撮影時間と曜日をコンテキスト情報として利用しているが、撮影時間から取得可能な特徴量である入退室の時間帯、滞在時間等がコンテキスト情報として有効であるかを検証していく必要がある。さらに、取得できた特徴によって手法の使い分けを行うことで人物識別全体の精度向上につながると考えられるため、最適な手法の使い分けについて検証する必要がある。

## 文 献

[1] 井尻 善久, S. Lao, 村瀬 洋. “実用化に向けた顔画像処理技術”, 電子情報通信学会技術報告書, 電子通信学会,

PRMU2009-237, pp.107-112, 2010.

[2] A. C. Gallagher and T. Chen. “Using Context to Recognize People in Consumer Images”, *IPSN Transactions on Computer Vision and Applications*, Vol.1, pp.115-126, 2009.

[3] 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人物識別”, *日本知能情報ファジィ学会誌 知能と情報*, Vol.23, No.2, pp.13-22, 2011.

[4] P. Viola and M. Jones. “Robust Real-time Object Detection”, *Second International Workshop on Statistical and Computational Theories of Vision-Modeling, Learning, Computing, and Sampling*, pp.1-25, 2001.

[5] R. Lienhart, A. Kuranov and V. Pisarevsky. “Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection”, *MRL Technical Report*, 2002.

[6] 元吉 大介, 嶋田 和孝, 榎田 修一, 江島 俊朗, 遠藤 勉. “ロボットとの対話のための発話推定に関する事例研究”, *画像の認識・理解シンポジウム (MIRU2008)*, 2008.

[7] S. Watanabe and N. Pakvasa. “Subspace method in-pattern recognition”, *Proc. 1st Int. J. Conf on Pattern Recognition*, Washington DC, pp.2-32, 1973.

[8] N. Dalal and B. Triggs. “Histograms of Oriented Gradients for Human Detection”, *In Proc. IEEE Conf. on Computer Vision and Pattern Recognition 2005*, Vol.1, pp.886-893, 2005.

[9] Y. Freund and R. E. Schapire. “A decision-theoretic generalization of on-line learning and an application to boosting”. *In Computational Learning Theory: Eurocolt '95*, pp.23-37, Springer-Verlag, 1995.

[10] J. R. Quinlan. “C4.5: Programs for machine learning”, Morgan Kaufmann (1993).

[11] S. R. Garner. “WEKA: the Waikato environment for knowledge analysis”. *In: Proceedings of the New Zealand Computer Science Research Students Conference*, pp.57-64, 1995.

[12] 山下 隆義, 藤吉 弘巨. “特定物体認識に有効な特徴量”, *情報処理学会 研究報告 CVIM 165*, pp.221-236, 2008.