

コンピュータ理工学特別研究報告書

題目 顔と人物を検出するための
特徴量と認識率に関する実験的考察

学生証番号 9 4 7 0 4 6

氏 名 西浦 直輝

提出日 平成 25 年 2 月 5 日

指導教員 蚊野 浩

京都産業大学
コンピュータ理工学部

要約

本研究の目的は、画像認識における、対象物に適した特徴量と識別関数の関係を明らかにすることである。今回は顔画像と人物画像を使用し、顔クラスと非顔クラス、あるいは人物クラスと非人物クラスを識別した。特徴量として、濃淡値・勾配値・4方向面特徴・HOG特徴の4つを比較した。識別関数として、線形判別分析で高次元データを一次元化し、その特徴量を閾値で判別する方法と、サポートベクターマシン（SVM）を用いて判別する方法を比較した。認識率の推定は交差確認法を用いた。各特徴量と識別関数の組み合わせで認識率を比較した。顔画像を500枚、非顔画像を500枚、人物画像を500枚、非人物画像を500枚用いて実験を行った。顔画像の場合、特徴量に4方向面特徴を用い、識別関数にSVMを組み合わせると、最も良い結果になった。誤認識は1枚のみで、認識率は0.999であった。人物画像の場合、HOG特徴と線形判別分析を組み合わせることで誤認識が1枚も検出されず、認識率は1.000であった。

目次

第1章 序論	・・・ 4
1. 1 本研究の背景と目的	・・・ 4
1. 2 論文の構成	・・・ 5
第2章 画像認識と特徴量について	・・・ 6
2. 1 画像認識の概要	・・・ 6
2. 2 クラスと学習	・・・ 6
2. 3 特徴抽出	・・・ 7
2. 4 特徴選択	・・・ 7
2. 5 研究に用いた特徴量	・・・ 7
2. 5. 1 濃淡値	・・・ 7
2. 5. 2 勾配値	・・・ 8
2. 5. 3 4方向面特徴	・・・ 9
2. 5. 4 HOG 特徴	・・・ 10
2. 5. 5 Haar-like 特徴	・・・ 11
2. 6 識別関数	
2. 6. 1 線形判別分析	・・・ 12
2. 6. 2 サポートベクターマシン	・・・ 15
第3章 認識対象物の検出実験	・・・ 17
3. 1 実験用のデータセット	・・・ 17
3. 2 顔画像の正規化	・・・ 17
3. 3 交差確認法による認識率の計算	・・・ 18
3. 4 線形判別分析の場合	・・・ 19
3. 5 サポートベクターマシンの場合	・・・ 20
3. 6 Haar-like 特徴を使用した場合	・・・ 20
第4章 実験結果	・・・ 22
4. 1 顔画像の認識	・・・ 22
4. 2 人物画像の認識	・・・ 23
4. 3 実験結果に対する考察	・・・ 23

第 5 章 結論	• • • 2 4
5. 1 成果	• • • 2 4
5. 2 課題	• • • 2 5

参考文献

謝辞

付録

1. 序論

1. 1 本研究の背景と目的

画像認識を、一般的なことばで説明すれば、画像に写っているものが何であることを説明・記述することである。コンピュータ技術としての画像認識は、画像として観察された被写体が属するクラスを判別する技術である。これを実現するために、通常、画像から特徴と呼ばれる数値やベクトルを抽出し、それらを数学関数や手続き的なプログラムで分類することが行われる。

画像認識は計算コストが大きく、複雑な処理である。しかし、コンピュータ技術の進歩により、さまざまな方面で広く利用されるようになった。例えば、製品の生産、医療、セキュリティの分野で利用されている。スマートフォンには顔認識が搭載されている機種がある。登録しておいた自分の顔を認識することでロックを解除することができる。他には、笑顔検出や顔認識オートフォーカスを搭載したデジタルカメラが製品化されている。顔認識は、比較的なじみのある画像認識技術である。

画像認識のなかでも、顔認識は歴史のある研究テーマである。たとえば、顔の認識に用いられるアルゴリズムは、主成分分析[3]、線形判別分析[4]、弾性バンチグラフマッチング[5]、隠れマルコフモデル[6]などがある。これら以外にも、認識技術の発達により次々と新しいアルゴリズムが研究、開発されている[7][8]。なお、顔認識ということばには個人同定を意味する場合と、顔検出を意味する場合がある、本論文では、画像中に存在する顔を検出する意味で用いる。

顔認識が顔だけを認識対象としているのに対し、人物認識は画像中に人間の全身が存在する時、その領域を検出する技術である。人間の全身像には、いろいろな形状がありえるが、立位や歩行状態を対象とする研究が多い[9][10]

顔認識や人物認識の研究は多いが、認識画像の特徴量と認識率の関係について明確に述べている文献は少ない。それらには、新しい特徴量や認識手法の計算アルゴリズムが述べられているが、その有効性を理解することが難しい。そこで本研究では、画像の特徴量と識別関数、認識率の関係性を調査、研究することにした。画像特徴として、基本的な画素値、エッジ特徴、4方向面特徴、および比較的最近に提案された HOG 特徴を用いた。認識アルゴリズムとして、線形判別分析に基づく手法と、SVM を対象とした。認識対象物に顔画像と人物画像を用い、特徴量と認識アルゴリズム、認識率の関係を実験的に研究した。

1. 2 論文の構成

本論文は5章で構成されている。第2章では、本研究で使用する特徴量や識別関数について述べた。第3章で提案手法や認識率の調査方法について述べた。第4章で実験方法・結果・考察について記述し、第5章で研究のまとめを記述する。

2. 画像認識と特徴量について

画像認識はコンピュータで画像中の物体や状態を認識し検出する技術である。画像は、画素と呼ばれる要素の膨大な集合である。画像認識では、画素の集合を直接認識対象とするのではなく、通常、前処理として、ある種の識別候補を記述する特徴量を取り出す。そこから意味を読み取ることに相当する処理を行う。この処理全体のことを画像認識と呼ぶ。

2. 1 画像認識の概要

画像認識は、観測された画像の特徴を用いて、あらかじめ定められたクラスにその画像を識別する処理である。画像認識を画像パターンの認識という場合がある。パターンとは模様をいう意味であるが、複数の画像が共通の特徴の組をもつ時、同じ画像のパターンであるといえる。一般に画像認識の処理の流れは図1のようになる。まずクラスごとに分類された学習用の入力画像から特徴抽出を行い、そこからクラスごとの辞書（クラスに属する画像の特徴を記録した物）を作成する。識別は、テスト画像（未知画像）から特徴抽出を行い、クラス辞書を用いてテスト画像をクラスに分ける。

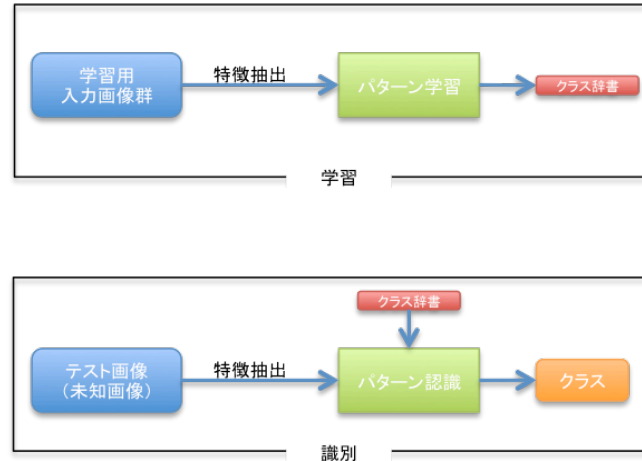


図 1 画像認識における学習と識別の流れ

2. 2 クラスと学習

クラスは、類似した特徴あるいはパターンを持つ画像の候補が属する集合のことである。学習のためのクラスをあらかじめ定めるものを教師あり学習とよぶ。そうでないものを教師なし学習とよぶ。本研究では、教師あり学習を前提としている。

2. 3 特徴抽出

画像の特徴は、画像中で他の部分と明確に区別できるような小領域である。特徴量はその明確さを数値化したものである。特徴抽出は入力画像から認識に役立つ特徴量を取り出す画像処理である。取り出された K 個の特徴量をそれぞれ x_i とし、これらを並べたものを特徴ベクトルと呼ぶ。特徴ベクトルは $x = (x_1, x_2, \dots, x_K)'$ となる。そして、特徴ベクトルで構成される空間を特徴空間という。ここで $'$ はベクトルや行列の転置を表す。本論文では特徴ベクトルを縦ベクトルで表す。

2. 4 特徴選択

画像からさまざまな特徴を抽出することができるが、それらの特徴のすべてが同じように画像認識に有効なわけではない。場合によっては、認識に有効な特徴量を特徴量全体から選択する必要がある。特徴選択には、各クラスの入力画像群をよく表現する共通の特徴を選択する方法と、クラス間をよく区別できる特徴を選択する方法の2種類がある。前者の方法に主成分分析があり、後者の方法に線形判別分析がある。

2. 5 各特徴量について

今回の実験で使用する特徴量を説明する。

2. 5. 1 濃淡値

画像の濃淡値（画素値）をそのまま特徴量として利用することができる。通常、画素値を特徴量として扱うことはないが、今回は、これを最も基本的な比較対象とした。図2は濃淡値で表現された人物画像、風景画像の例である。



図 2 濃淡値で表現された人物・風景画像

2. 5. 2 勾配値

画像の勾配は、隣接画素との画素値の変化量とその方向である。ここでは、変化量だけを扱い、その量を勾配値と呼ぶ。画像の特徴は画像信号が変化する部分に表れる。そのような箇所は勾配値が大きくなり、エッジ、輪郭と呼ばれる。

画像の勾配は次のように計算する。勾配は隣接画素との画素値の差分から計算する。座標 (x,y) における濃度の勾配を表す1次微分の値を、大きさと方向を持つベクトル量 $G(x,y)=(dx,dy)$ として表す。ここでは x 方向の微分は dx 、 y 方向の微分を dy と示す。 dx, dy はデジタル画像では $dx=f(x+1,y)-f(x,y)$, $dy=f(x,y+1)-f(x,y)$, と計算される。このとき、勾配値は $\sqrt{dx^2 + dy^2}$ である。なお、画像中での画素位置の関係を図3に示す。また、勾配値で表現された人物画像、風景画像を図4に示す。

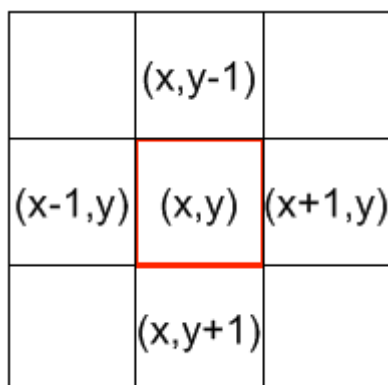


図 3 ある画素に対しての差分を求めるときの、画素の関係



図 4 勾配値で表現された人物・風景画像

2. 5. 3 4方向面特徴

4方向面特徴は各画素の勾配を4つの方向面(水平、垂直、右斜め、左斜め)に割り当てたものである。具体的には、図5に示す Prewitt オペレータで求めた各画素のエッジ勾配を、エッジ方向に基づき、4方向に分解した4枚の画像を作成する。さらに、それらの画像をガウシアンフィルタに通したのち、低解像度化した特徴である。

4方向面特徴は解像度を低下させてもエッジ方向情報を保持しており、低解像度化によりエッジの形状変動に対してロバストになる。また、低解像度化することで、特徴量の数が削減できる。本来、4方向面特徴は文字認識に有効であるとされていたが、本郷等によって、4方向面特徴を顔認識に応用し、顔認識にも有効であることが示されている[12]。図6に4方向面特徴で表現した人物画像を示す。

-1	0	1
-1	0	1
-1	0	1

水平方向

-1	-1	-1
1	0	1
1	1	1

垂直方向

図 5 Prewitt オペレータ

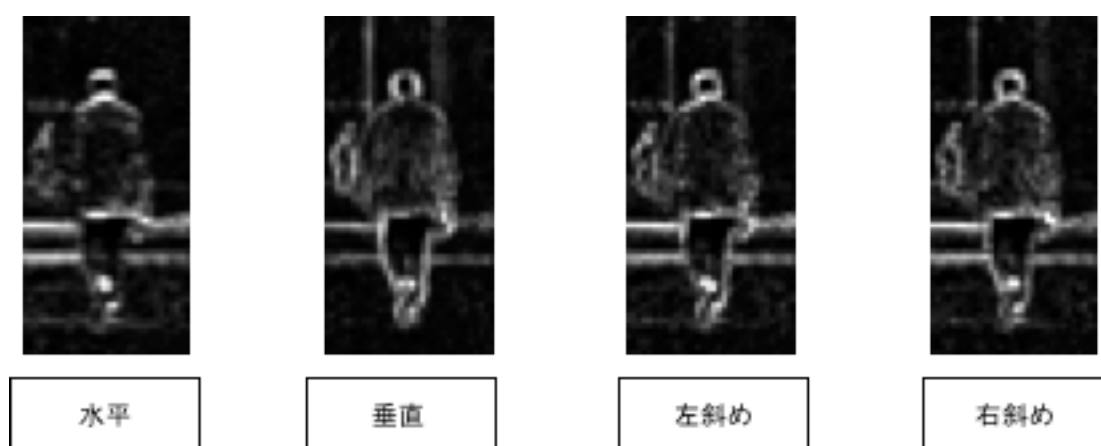


図 6 4方向面特徴に分解した人物画像

2. 5. 4 HOG 特徴

一般物体認識のための勾配ベースの特徴量として HOG (Histograms of Oriented Gradients) 特徴が提案されている。HOG は局所領域における輝度の勾配方向をヒストグラム化した特徴量である。ある大きさを持つ領域に対して特徴量を記述するため、おおまかな物体形状を表現することが可能である。人物検出や車検出などの一般物体認識に用いられる。

HOG 特徴を計算するには、ブロックとセルのサイズを決めなければならない。ここでは例として、1 セル: 5×5 画素、1 ブロック: 3×3 セルで説明を行う。はじめに局所領域をブロックとセルに分割する。次に、セル内における輝度勾配ヒストグラムを計算する。入力画像の座標 (x, y) における輝度値を $I(x, y)$ とすると、 x 方向に微分した値は $dx(x, y) = I(x+1, y) - I(x-1, y)$ 、 y 方向に微分した値 $dy(x, y) = I(x, y+1) - I(x, y-1)$ となる。これらを用いて、勾配強度 m と勾配方向 θ を求める。勾配強度 m と勾配方向 θ は以下の式により算出する。

$$m(x, y) = \sqrt{dx(x, y)^2 + dy(x, y)^2}$$
$$\theta(x, y) = \tan^{-1} \frac{dy(x, y)}{dx(x, y)}$$

算出された勾配強度 m と勾配方向 θ を用いて、勾配方向ヒストグラムを作成する。そして、ある n 番目の HOG 特徴量について以下の式 (1) を用い正規化を行う。ここで $v(n)$ は勾配方向ヒストグラム、 q はセルサイズ、 N は勾配方向の数である。分母は 1 ブロック ($q \times q$ セル) に含まれる HOG 特徴量の総和である。1 つのセルに対し複数回の正規化を行う。

$$v(n) = \frac{v(n)}{\sqrt{\left(\sum_{k=1}^{q \times q \times N} v(k)^2 \right) + \epsilon}} \quad (\epsilon = 1)$$

N : 勾配方向数

30×30 画素の画像の場合、輝度の勾配方向ヒストグラムは、 $0^\circ \sim 180^\circ$ を 20° ごとに分割すると、9 方向の勾配方向ヒストグラムになる。図 7 に HOG 特徴の計算プロセスの概要を示す。

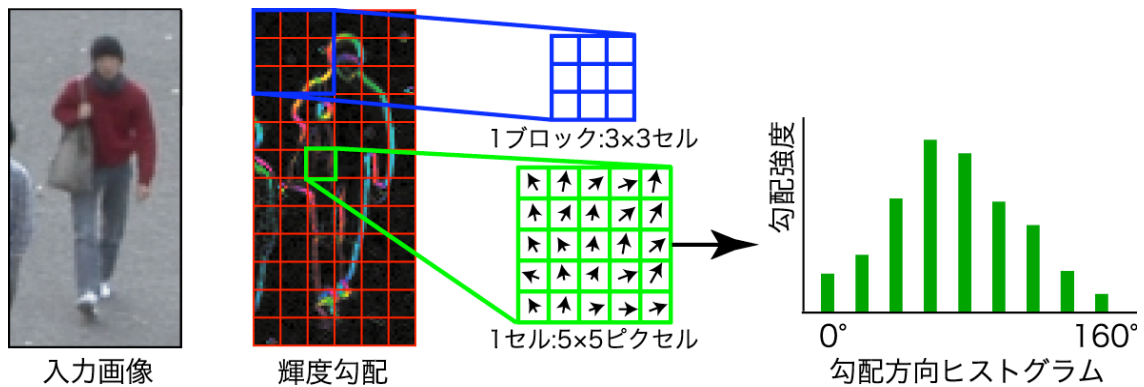


図 7 HOG 特徴の説明

2. 5. 5 Haar-like 特徴

画像における特徴量として画素値をそのまま用いると、照明条件の変動やノイズの影響を受けやすい。これに対応するために、入力画像の任意の位置に、図 8 のような矩形領域のパターンを設定し、領域間の差分を計算する。すなわち、白領域に対応する画素の和と、黒領域に対応する画素の和の差を特徴量とすることで、画素値の変動に頑健な特徴になる。これを Haar-like 特徴と呼ぶ。本研究ではあまり深く触れないが、顔画像の認識率を比較するために使用する。

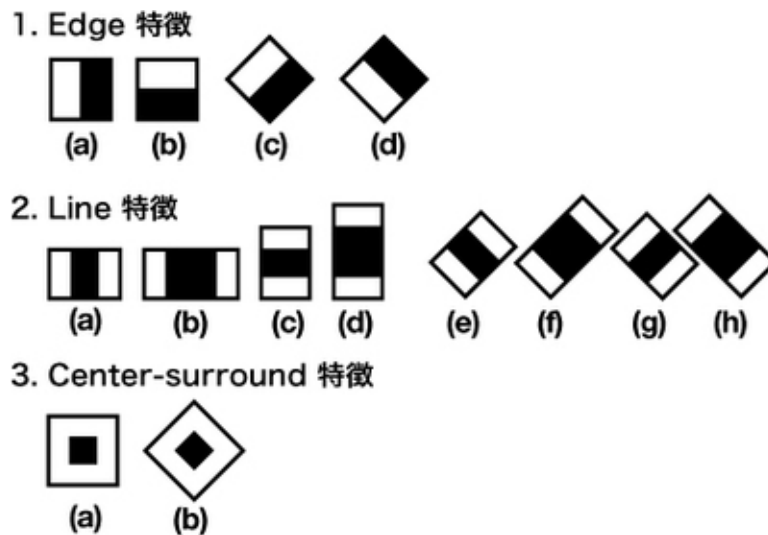


図 8 Haar-like 特徴に用いる矩形領域のパターン

2. 6 識別関数

画像認識に用いる識別関数は、画像特徴を表す特徴ベクトルを入力とする関数であり、特徴ベクトルの属するクラスを出力する。クラスを判定するための識別規則を記述するために用いる。今回は 2 クラスに識別するために線形判別分析と線形 SVM を用いる。これらが出力する一次元の数値を閾値で顔クラスと非顔クラス、人物クラスと非人物クラスに分類する。

2. 6. 1 線形判別分析

線形判別分析は、特徴量を記述した高次元データを一次元の部分空間に射影する。このとき、2 クラス間を最もよく識別する一次元空間に射影する手法である。識別の方法は、一次元に射影された数値に対し、それらが 2 クラスのどちらに属するかを、閾値により決定する。たとえば、識別する画像データが閾値に対して、正ならクラス 1、負ならクラス 2 というように分類を行う。

例として、2 つのクラスを 2 つの特徴量で識別する場合を考える。図 9 のように、クラス 1 とクラス 2 の特徴量が分布する。この 2 つの特徴量を線形判別分析で一次元の特徴量に変換すると、図 9 に示すように、最もクラス間がよく分離できる直線が選ばれる。この直線が 2 つのクラスを識別するための新たな特徴の軸、すなわち基底ということになる。二次元の特徴量をこの直線に射影して、そのヒストグラムをクラスごとにグラフ化したものを図 10 に示す。2 クラスを分類する決定境界の直線は、図 10 に示すように、2 クラスのヒストグラムが交わる交点になる。この値を閾値として、2 クラスに分類する。

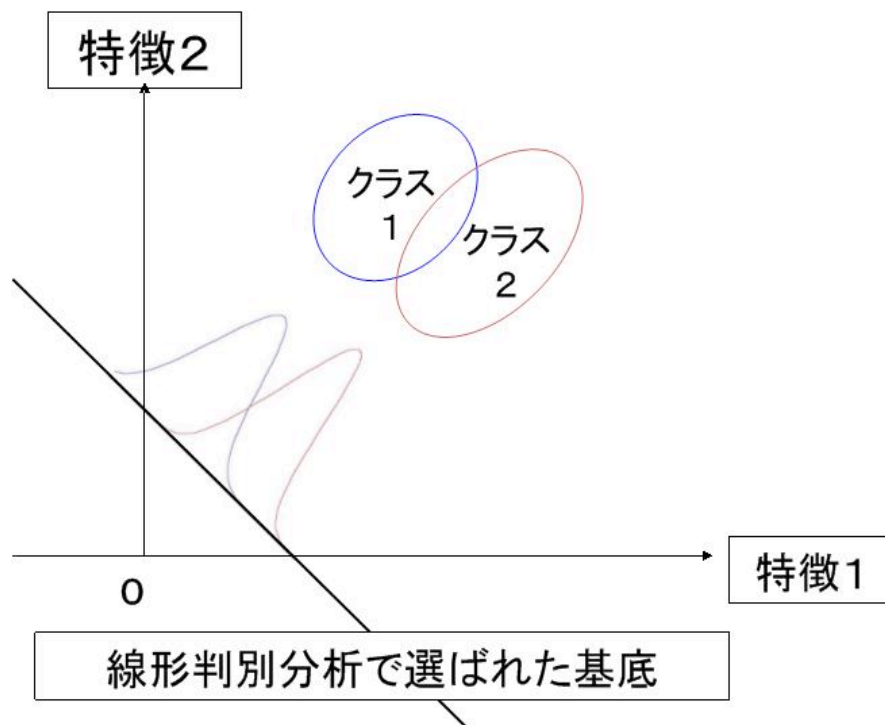


図 9 2 クラスを識別する基底を線形判別分析によって計算し、
クラスごとに特徴量を射影して一次元化した特徴のヒストグラム分布

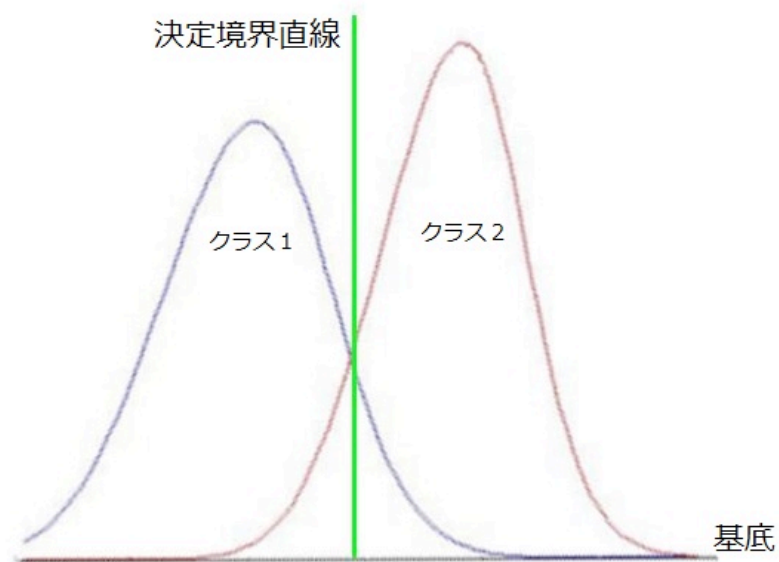


図 10 2 クラスを分類するための閾値

線形判別分析のための基底を計算する手順を説明する。クラス c のパターン数を $N^{(c)}$, 平均値を $M^{(c)}$ とし、全サンプル数を N , その平均ベクトルを M としたとき、全サンプルの分散共分散行列 S は次式のようにになる。

$$S = \frac{1}{N} \sum_{n=1}^N (x_n - M)(x_n - M)^T$$

クラス c の分散共分散行列 $S^{(c)}$ は以下のようにになる。

$$S^{(c)} = \frac{1}{N^{(c)}} \sum_{n \in 1}^N (x - M^{(c)})(x - M^{(c)})^T$$

クラス内分散共分散行列 S_w とクラス間分散共分散行列 S_b は以下のようにになる。

$$S_w = \frac{1}{N} \sum_c N^{(c)} S^{(c)}$$

$$S_b = \frac{1}{N} \sum_c N^{(c)} (M^{(c)} - M)(M^{(c)} - M)^T$$

ただし、 $\sum_c$ は全クラスの和を表し、 $S = S_w + S_b$ の関係がある。

ここで、クラス間の分離度を大きくするためには、 $S_b S_w^{-1}$ が大きくなるように基底を d 個選ばばよい。これは $S_b S_w^{-1}$ の固有ベクトルを求める問題と同じになる。

$$S_b S_w^{-1} u_j = \lambda_j u_j$$

固有値 λ_j が大きい順に対応する固有ベクトル u_j から d 個選ぶと、 $(u_1, u_2, \dots, u_d)$ が求める。これが線形判別分析のための基底となる。この基底を用いた特徴空間では、クラス間が識別されやすい特徴量に変換されている。

2. 3. 2 サポートベクターマシン

SVM は、2クラス分類問題を解く学習機械として、Vladimir N Vapnik により提唱されたものである。SVM が2クラス識別機として優れたモデルである理由は、訓練サンプル集合を最適に識別するために、「マージン最大化」という超平面決定基準が設けられている点と、カーネル学習法により非線形の識別問題に対応

している、という2点である。学習過程はラグランジュ乗数法を用いることにより、最適化問題の一種である凸二次計画問題で定式化される。ただし、学習サンプル数が増えると急速に計算量が增大するため、分割統治法の考え方をを用いた手法も提案されている。本実験では Matlab の Bioinformatics toolbox に含まれているサポートベクターマシンを使い実験を行った。

画像のパターン認識の目的は、既知の画像パターンだけでなく、未知の画像パターンに対しても正しく識別することである。このことから考えると識別面は、それぞれのクラスに属する既知の画像パターンのちょうど真ん中を通るものが良い。ちょうど真ん中を通るとは、図 11 のように識別平面からもっとも近い既知の画像パターンとの距離（マージン）を最大になるように取ることを意味する。SVM ではこのマージンを最大化する識別平面を計算するように設計されており、未知の画像パターンに対しても高い認識性能を示す。

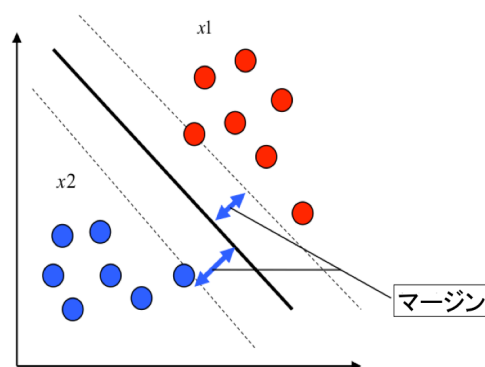


図 11 マージン最大化

非線形で複雑な識別課題を扱う場合、特徴ベクトルを非線形変換して、変換された特徴空間で線形識別を行う方法がある。これをカーネル法という。特徴空間に射影したベクトルの内積を関数 K で表すと、以下の式をつくることができる。

$$K(x_1, x_2) = \phi(x_1) \cdot \phi(x_2)$$

この関数をカーネル関数と呼ぶ。カーネル関数を求めることができると、特徴空間で ϕ の計算が必要なくなる。そして、高次元に射影しながら、実際には射影された空間での特徴の計算を避けて、カーネルの計算のみで、最適な識別関数

を構成する手法のことをカーネルトリックと呼ぶ。図 12 にカーネルトリックの考え方を模式的に示す。

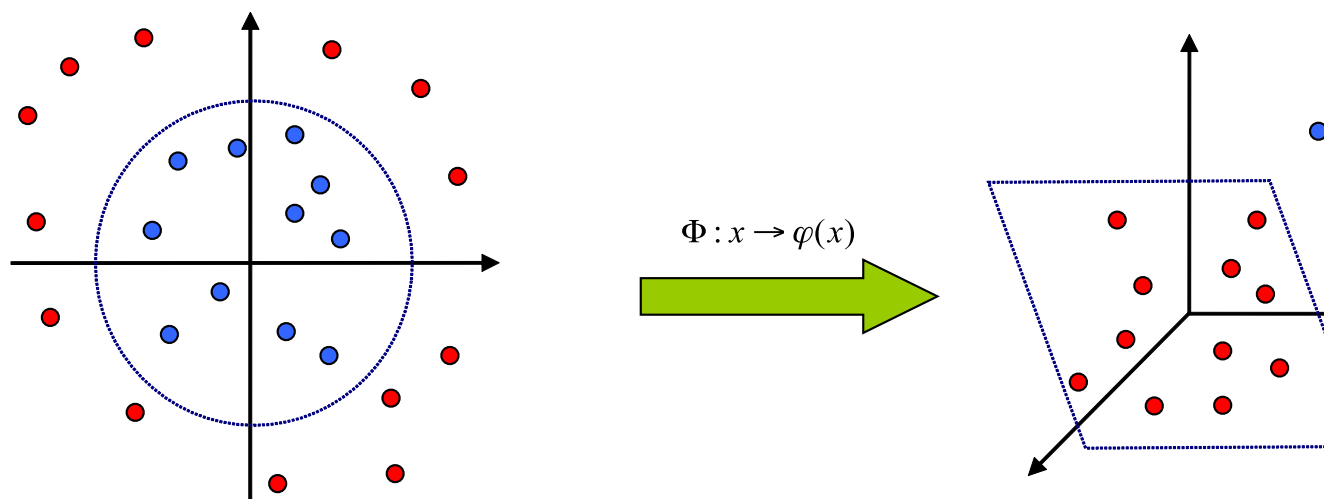


図 12 カーネルトリック

3. 認識対象物の検出実験

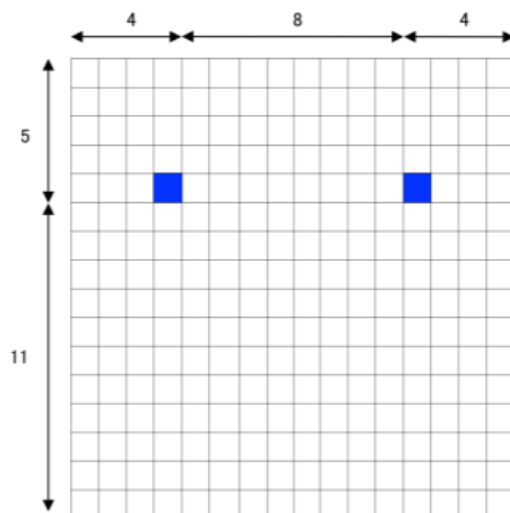
3. 1 実験用のデータセット

本実験で利用した画像データを説明する。顔画像は、 256×384 画素で顔が画像のほぼ中央にある 500 枚の濃淡画像を準備し、3. 2 に述べる方法で 16×16 画素の画像に正規化した。人物画像は 32×64 画素で、人物が画像の中央にある 500 枚の濃淡画像を準備した。人物画像は正規化を行わず使用する。顔と人物を含まない画像は、あらかじめ準備した顔と人物を含まない大きな画像に対して、ランダムに座標を決め、顔画像、人物画像の画素数に合わせて切り取った濃淡画像を 500 枚ずつ準備した。この濃淡画像から、勾配値、4 方向面特徴、HOG 特徴の抽出を行う。

線形判別分析と SVM を用いた画像認識プログラムは、Matlab2010a で作成し、実験を行った。これらとは別に、OpenCV に実装されている顔検出機能の認識率を確認する実験を行った。

3. 2 顔画像の正規化

実験に使用する顔画像を正規化するために、顔画像の目の位置を取得し、あらかじめ設定している比率関係に基づいて 16×16 画素に変換する。図 12 にその比率関係を示す。この処理を行うことで、画素数は 256 画素になる。したがって、画素値を特徴とする場合の次元は 256 次元になる。勾配値、4 方向面特徴、HOG 特徴を顔画像の認識に用いる場合には、これに合わせるために、すべて 256 次元とした。HOG 特徴の場合、1 セルを 4×4 画素、1 ブロックを 4×4 セル、勾配方向の数を 16 とし、256 次元の特徴を抽出した。図 13 に顔画像の正規化前と正規化後の顔画像を示す。



顔画像を 16×16 画素の画像に変換する際に、目の位置を参考にして正規化を行う。
左上の画素を $(1, 1)$ とし、目の位置を $(4, 5)$ $(12, 13)$ と指定して行った。

図 12 顔画像の比率関係

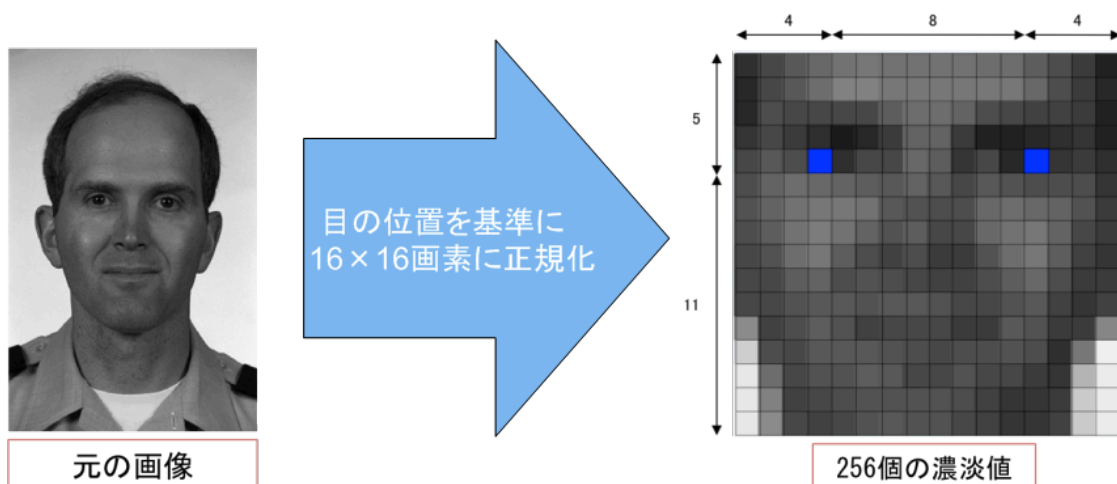


図 13 顔画像の正規化

人物画像は正規化を行わなかった。対象画像の画素数が $24 \times 64 = 2048$ 画素があるので、識別する際の特徴ベクトルの次元を 2048 次元とした。HOG 特徴の場合、1セルを 8×8 画素、1ブロックを 8×8 セル、勾配方向の数を 32 として、2048 次元の特徴を抽出した。

3. 3 交差確認法による認識率の計算

訓練とテストに使えるデータは限られている。良い学習モデルを作るために、準備した学習データをできるだけ多く使いたい。しかし、そうするとテスト用のデータが少なくなってしまう、認識率の推定誤差が大きくなってしまうという問題が生まれる。この問題を避けるために交差確認法（cross validation）が利用される。

この方法は標本群から 1 つの事例だけを抜き出してテスト事例とし、残りを学習事例とする。全事例が一回ずつテスト事例になるように検定を繰り返す。すべてのデータが学習用、テスト用として使用されるので性能評価の信頼性の精度が向上する。

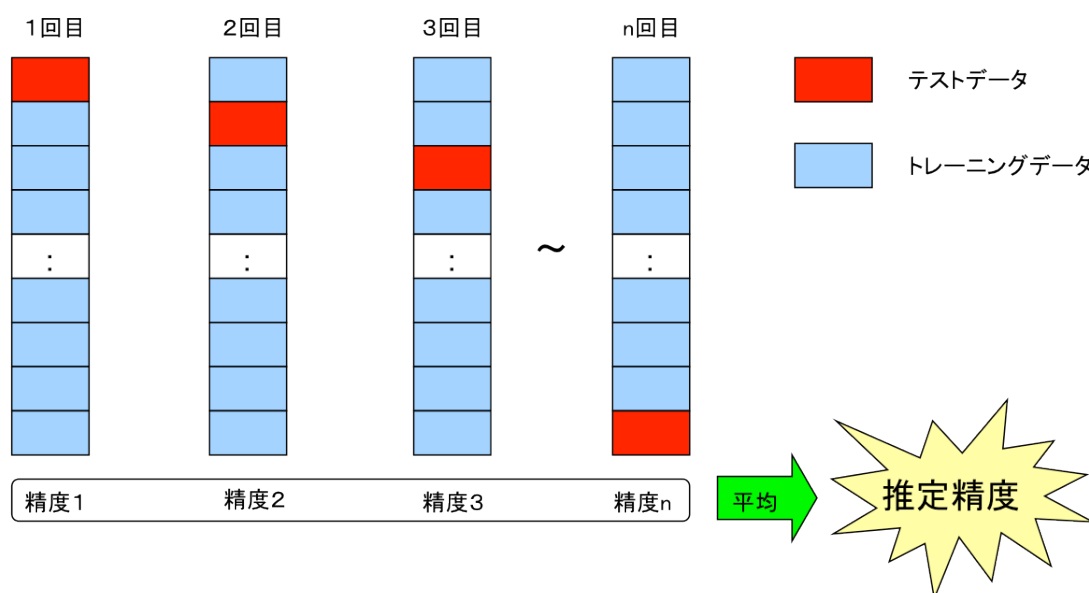


図 15 交差確認法の手順

3. 4 線形判別分析の場合

実験に用いる画像を読み込み、各画像の特徴量を計算する。次に交差判定に用いる画像をのぞいた 999 枚の学習用画像を用いて、判定を行うための基底を求める計算処理を行う。ここでは、クラスのパターン数、平均、分散、クラス内分散共分散、クラス間分散共分散を計算する。この計算により基底を求めることができ、多次元の特徴量を一次元の特徴量に変換する。基底に 2 クラスの特徴量を射影して、そのヒストグラムの交点を閾値とする。この閾値をもとにテスト用画像の識別を行う。この一連の流れをテスト用画像を変えながら、1000 回繰り返す。認識率は以下の式で求める。

$$\text{認識率}(Accuracy) = \frac{(\text{全画像数} - \text{誤認識数})}{\text{全画像数}}$$

3. 5 SVM の場合

画像を読み込み、各クラスの特徴量の総和をとる。1000 枚の画像を顔（人物）と非顔（非人物）のクラスに分けるため、500 枚ずつデータを振り分けてラベルをつける。999 枚を学習用画像、1 枚をテスト用画像とする。999 枚の学習用画像で SVM を学習させる。学習した SVM に 1 枚のテスト用画像を与えて判定を行う。テスト画像を変えながら、同じことを 1000 回繰り返す。認識率は以下の式で求める。

$$\text{認識率}(Accuracy) = \frac{(\text{全画像数} - \text{誤認識数})}{\text{全画像数}}$$

3. 6 Haar-like 特徴を使用した場合

Haar-like 特徴を用いた顔画像の検出機能が OpenCV に搭載されている。OpenCV の顔検出機能は Haar-like 特徴を用い、Adaboost と呼ばれる方法で学習したカスケード接続型の識別関数によって実現されている。今回、その顔検出率を調査した。顔を検出するにあたって 16×16 画素の画像では、画像サイズが小さすぎるため、一枚も顔として認識することが出来なかった。そのため、顔画像を正規化する前の元画像を使用した。顔が正しく検出できた場合、顔の位置に赤い円が表示される（図 16）。円が 2 つ以上表示される、もしくは何も表示されなかった場合は正しく認識していないということなので、誤認識として数える（図 17）。非顔画像の場合は、何も検出されないことが正しい（図 18）ので、円が表示されると誤認識として数える（図 19）。認識率は以下の式で求めることができる。

$$\text{認識率}(Accuracy) = \frac{(\text{全画像数} - \text{誤認識数})}{\text{全画像数}}$$



図 16 正しく認識ができた例

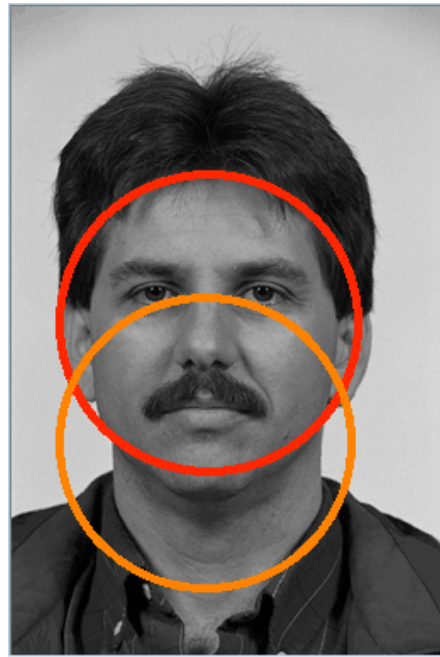


図 17 正しく認識ができなかった例



図 18 正しく認識ができた例



図 19 正しく認識ができなかった例

4. 実験結果

4. 1 顔画像の認識

線形判別分析を用いた顔画像の認識率を表 1 に示す。表 1 からわかるように、4 方向面特徴を用いた場合がもっとも良い結果となった。HOG 特徴と 4 方向面特徴は、似た性質の特徴を抽出しているが、HOG 特徴がより高度なものであるから、認識率でも上回ると想像したが、そのようにはならなかった。今回の実験では、HOG 特徴の次元数を 4 方向面特徴の次元数に揃えたので、その優位性が明確にできなかったものと考えられる。

表 1 線形判別分析による顔画像の認識率

線形判別分析				
特徴量	濃淡値	勾配値	4 方向面特徴	HOG 特徴量
認識率	0.790	0.905	0.991	0.979

SVM を用いた顔画像の認識率を表 2 に示す。どの特徴量についても、線形判別分析よりも良い結果になっている。特徴量に 4 方向面特徴を用いた場合がもっとも良く、誤認識数は 1 個であった。HOG 特徴量も線形判別分析を用いた場合よりも良い認識率になった。しかし、4 方向面特徴を上回らなかった。この結果から、識別関数としての性能は、SVM が線形判別分析よりもはるかに優れていることが確認できた。

表 2 SVM による顔画像の認識率

SVM				
特徴量	濃淡値	勾配値	4 方向面特徴	HOG 特徴量
認識率	0.944	0.968	0.999	0.988

Haar-like 特徴を用いた顔画像検出プログラムの認識率を表 3 に示す。この実験は、OpenCV に含まれている顔検出機能の性能を確認し、上記 2 つの結果と比較するために行った。結果は、誤検出数 60 個（顔画像：9 個、非顔画像：51 個）で認識率は 0.940 だった。顔を検出する関数であったためか、非顔画像での誤検出が圧倒的に多くなるという結果になってしまった。非顔画像での誤検出を除くと、認識率は 0.982 になる。これは、線形判別分析と勾配値を組み合わせた場合より認識率が高く、SVM と HOG 特徴量を組み合わせの場合と、ほぼ同等の

認識率を示している。

表 3 OpenCV による顔画像の認識率

OpenCV の顔検出機能		
認識対象物	顔・非顔	顔のみ
認識率	0.940	0.982

4. 2 人物画像の認識

線形判別分析を用いた人物画像の認識率を表 4 に示す。人物認識では服の色や体の向き、姿勢などさまざまな不確定要素が含まれているためか、濃淡値や勾配値では、顔画像の認識よりも悪い認識率になった。しかし、4 方向面特徴と HOG 特徴量は 1.000 という認識率を示した。輝度勾配から方向ヒストグラムを算出して特徴量とするため、人物の輪郭を明確に抽出できたのではないかと考えられる。

表 4 線形判別分析による人物画像の認識率

線形判別分析				
特徴量	濃淡値	勾配値	4 方向面特徴	HOG 特徴量
認識率	0.725	0.861	1.000	1.000

SVM を用いた人物画像の認識率を表 5 に示す。最も良い認識率を示したのは HOG 特徴であった。他の特徴量も認識率が 0.91 を超えており全体的に良い結果である。しかし、4 方向面特徴と HOG 特徴を SVM で認識する性能が、線形変別分析より悪い結果になったのは意外であった。

表 5 サポートベクターマシン（人物画像）

SVM				
特徴量	濃淡値	勾配値	4 方向面特徴	HOG 特徴量
認識率	0.910	0.910	0.988	0.993

4. 3 実験結果に対する考察

今回の実験から以下のように判断することができる。

- 勾配値は、濃淡値よりも優れているが、4 方向面特徴と HOG 特徴よりは劣っている。
- 顔と人物を認識するための特徴量としては 4 方向面特徴と HOG 特徴が優れている。この 2 つの性能はほぼ同じである。
- この 2 者の比較において、4 方向面特徴は顔の検出に優れており、HOG 特徴は人物の検出に優れている傾向が見られた。
- SVM の認識性能は線形判別分析よりも明らかに優れている。濃淡値のように劣った特徴を用いても、勾配と線形判別分析を組み合わせた場合よりも、高い認識性能を示した。
- OpenCV の顔検出機能は、勾配値と SVM の組み合わせ程度の性能であった。
- 顔認識と人物認識の難易度を比較すると、顔画像を対象とした平均認識率は 0.946 であるのに対して、人物画像を対象とした平均認識率は 0.923 であった。今回の課題では人物認識のほうが難しい。

HOG 特徴は、比較的最近考案された特徴で、人物検出において評価が高い特徴量である。しかし、今回の実験では、4 方向面特徴と比べて、明確な性能差がなかった。その理由を次のように考察した。

- 顔検出においては、特徴量の次元数が 256 次元と少なく、HOG の性能を十分に発揮することができなかった。
- 人物検出においては、特徴量の次元数 2048 に対して、実験に用いた画像数が人物 500 枚、非人物 500 枚と少なかった。そのため、HOG と 4 方向面特徴の性能差を明確に算出することができなかった。

SVM の認識性能が、線形判別分析よりも優れていることは 2.3.2 で述べた SVM の基本的な性質から明らかであるが、今回の実験によってそれを確認した。SVM が線形判別よりも優れているのは次の点である。

- 学習パターンに対する識別平面のマージンを最大化するため、非学習パターンに対して、高い認識性能が期待できる。
- 非線形識別が必要な特徴量に対して、カーネル関数による線形化を行うため、高い認識性能を達成できる。

5. 結論

5. 1 成果

本研究は顔認識と人物認識を対象として、4種類の特徴量と2種類の識別関数を組み合わせ、その認識率を調査した。

顔画像を認識する場合、特徴量に4方向面特徴を用い、識別関数にSVMを組み合わせると、最も良い結果となった。誤認識は1000枚中1枚のみで、認識率は0.999であった。OpenCVに実装されているHaar-like特徴を用いた顔画像の認識率と比較しても、良い認識率を出した。

人物画像を認識する場合、特徴量に4方向面特徴もしくはHOG特徴を用い、識別関数に線形判別分析を組み合わせると、最も良い結果となった。画像の誤認識が1000枚に対して1枚も検出されず、認識率は1.000であった。

今回の課題で顔認識と人物認識の難易度を比較すると、人物認識のほうが困難な問題であると言える。

5. 2 課題

認識率を計算するために交差確認法を用いた。このプログラムをMatlabで実装した。Matlabはfor文やif文などの実行速度が遅い。開発したプログラムにはfor文やif文が多く使われていたため、(256次元の特徴量)×1000枚の場合、結果出力まで約10分、(2048次元の特徴量)×1000枚の場合、結果出力まで約3時間半を要した。交差確認法の実装法を再検討する必要がある。

今回実験した特徴量と識別関数以外にも、画像認識に有効なものが知られている。特徴としては、SIFT特徴やGabor特徴、Jouin Haar-like, Joint HOGなどがある。また、識別関数としてブースティング、AdaBoost、Real AdaBoostなどがある。これらについても調査・検討することが重要である。

参考文献

- [1] 石井健一郎,「わかりやすいパターン認識」,オーム社,1998 年.
- [2] 「デジタル画像処理」,CG-ARTS 協会,2006 年.
- [3] 佐野鋭,「パターン認識における主成分分析-顔画像認識を例として-」,統計数理第 49 巻,第 1 号, pp. 23-42, 2001 年.
- [4] 高野智章 山内寛紀,「12-1 Gabor 特徴量と線形判別分析を用いた顔認識 (第 12 部門ヒューマンインフォメーション 3)」,映像情報メディア学会冬季大会講演予稿集, 12-1-1, 2007 年.
- [5] SONG G, AI H, XU 「G, Hierarchical Direct Appearance Model for Elastic Labeled Graph Localization」, Proc SPIE, 5286 巻 1 号, pp. 139-144, 2003 年.
- [6] 坂口竜己, 大谷淳, 岸野文郎,「隠れマルコフモデルによる顔動画像からの表情認識」,テレビジョン学会誌, Vol. 49, No. 8, pp. 1060-1067, 1995 年.
- [7] 岩井儀雄, 勞 世竝, 山口修, 平山高嗣,「画像処理による顔検出と顔認識」,情報処理学会研究報告. CVIM, pp. 343-368, 2005 年.
- [8] 本郷, 安本, 丹羽, 山本,「適応的判別空間生成による顔認識手法の提案」,電子情報通信学会技術研究報告, 101, pp. 55-62, 2001 年.
- [9] 生形徹, 有江誠, アレサンドロ・モロ, 寺林賢司, 梅田和昇,「距離画像セグメンテーションに基づくリアルタイム人物検出」, 第 12 回計測自動制御学会システムインテグレーション部門講演会講演論文集 (SI2011), 3A2-6, 2011 年.
- [10] 安藤慎吾, 忤 小軍, 鈴木章,「映像モニタリングのための人物姿勢推定技術」, NTT 技術ジャーナル, 19(8), pp. 21-24, 2007 年.

謝辞

本研究を進めるにあたり、指導教員の蚊野浩教授から、さまざまなご指導を賜りました。ここに感謝の意を表します。

付録

本研究を行うために開発した Matlab プログラム

1. faceNormalize.m

顔画像を正規化するプログラム

2. pctEdge.m

画像から勾配値を抽出するプログラム

3. pct4Direction.m

画像から 4 方向面特徴を抽出するプログラム

4. HoG.m

画像から HOG 特徴量を抽出するプログラム

5. face_human_LDA.m

特徴量として濃淡値・勾配値、識別関数として線形判別分析を使用し、顔画像、または人物画像の認識率を求めるプログラム

6. face_human_SVM.m

特徴量として濃淡値・勾配値、識別関数として SVM を使用し、顔画像、または人物画像の認識率を求めるプログラム

7. direction_face_LDA.m

特徴量として 4 方向面特徴、識別関数として線形判別分析を使用し、顔画像の認識率を求めるプログラム

8. direction_human_LDA.m

特徴量として 4 方向面特徴、識別関数として線形判別分析を使用し、人物画像の認識率を求めるプログラム

9. direction_face_SVM.m

特徴量として 4 方向面特徴、識別関数として SVM を使用し、顔画像の認識率を求めるプログラム

10. direction_human_SVM.m

特徴量として 4 方向面特徴、識別関数として SVM を使用し、人物画像の認識率を求めるプログラム

11. face_humanLDA_HOG.m

特徴量として HOG 特徴量、識別関数として線形判別分析を使用し、顔画像、または人物画像の認識率を求めるプログラム

12. face_humanSVM_HOG.m

特徴量として HOG 特徴量、識別関数として SVM を使用し、顔画像、または人物画像の認識率を求めるプログラム