

独立成分分析を用いた顔画像からの属性推定

村本 晴紀[†] 篠沢 佳久[‡]

[†] [‡] 慶應義塾大学 理工学部 管理工学科 〒223-8522 横浜市港北区日吉 3-14-1

E-mail: [†] h_muramoto@ae.keio.ac.jp, [‡] shino@ae.keio.ac.jp

あらまし 本研究においては、顔画像からの個人認証において有効な識別方法を活用し、その属性（性別、年齢層）を推定する手法の構築を目的とする。顔画像認識においては、主成分分析によって抽出された固有顔を用いた認識手法が確立している。しかし主成分分析によって抽出された固有顔は、顔画像全体の特徴量を示すものであり、顔画像を構成する局所的な特徴量を用いた方が、性別・年齢層推定においては、より推定精度の向上を図ることができるものと考えた。そこで本研究においては、独立成分分析（ICA: Independent Component Analysis）を用いて、顔画像の局所的な構成要素を表現する特徴量（基底ベクトル）を抽出し、これを用いて性別および年齢層を推定する手法を提案する。特に独立成分分析を行なう場合、その対象画像（観測信号）として個人ごとの顔画像ではなく平均顔を用いて行なうことを提案する。そして収集した顔画像データベースを通じて評価実験を行なった結果、性別推定においては、固有顔を用いた場合と比較して、独立成分分析を用いた場合、推定精度の向上を図ることができた。一方、年齢層推定において独立成分分析を用いた場合、固有顔を用いた場合とほぼ同じ精度の結果を得ることができ、提案手法の有効性を示すことができた。

キーワード 顔画像認識, 属性推定, 独立成分分析, 固有顔, 平均顔

1. はじめに

近年、顔画像を対象とした個人認証技術についての研究がさかんに行なわれている。一方で、不特定多数の顔画像から性別や年齢などの属性を推定する研究も行なわれている。本研究においては、顔画像を対象とした個人認証技術を利用して、不特定の顔画像を対象として性別・年齢層推定の手法を考案することを目的とする。

性別推定においては、男性と女性という二つのクラスを推定する、また年齢層推定においては、例えば20代、30代といったようにあらかじめ年齢層ごとにいくつかのクラスに区別しておき、該当する年齢層のクラスを推定するということが一般的である。しかしながら、人間の顔は同一の性別を見ても多種多様であり、男性と女性という二つのクラスに分類することは困難である。特に年齢層については、個人差が大きく、その傾向がさらに顕著である。

個人認証において、有効な識別方法の一つとして固有顔（Eigenface）が挙げられる[1][2]。固有顔とは、顔画像に対して主成分分析を行ない、抽出した固有ベクトルのことである。固有顔を用いての顔画像の個人認証は、抽出された固有ベクトルによって張られる部分空間との最小距離基準に基づいて行なわれるのが一般的である[3]。

特に個人認証においては、この固有顔を用いた場合、高い精度を得られることが報告されている[4]。さらに個人認証に対してだけでなく、この固有顔を用いて性別、年齢層推定を行なった結果、高い推定精度が得られることも報告されている[5]。しかしながら主成分分析によって抽出される固有顔は、顔全体の特徴量を表現するものであり、性別、年齢層推定においては、顔画像を構成す

る局所的な特徴量を用いた方が、より推定精度の向上を図ることができると考えた。

そこで本研究においては、独立成分分析（Independent Component Analysis）を用いて、顔画像の局所的な構成要素を表現する基底ベクトルを抽出し、これを用いて属性推定を行なうことを試みる。独立成分分析（以下ICAと略す）とは、それぞれ独立した信号が重なり合った複数個の観測信号を観測し、これらの観測信号から元の独立した信号を分離する問題として扱われ、音声分離などさまざまな分野に応用されている。特に画像の特徴抽出にICAを利用した場合、ICAによって抽出された基底ベクトルは、人間の一次視覚野におけるものと類似していることが報告されており[6]、これを顔画像に適用した場合、顔画像を構成する局所的な特徴量が抽出できることも報告されている[7]。

本研究においては、顔画像から輝度特徴（グレースケール）、離散コサイン変換（Discrete Cosine Transform）、ガボール変換によって特徴を抽出し、これらの抽出された特徴を観測信号としてICAを行ない、基底ベクトルを求める。そして未知の顔画像に対しては、この基底ベクトルの線形結合によって、顔画像を復元し、その際に求めた重みベクトルを特徴ベクトルとして扱い、性別および年齢層を推定する手法を提案する。そして顔画像データベースを収集し、評価実験を通して、固有顔を用いた方法[5]との比較検証を行ない、提案手法の有効性の評価を行なう。

2. 顔画像データベース

本研究においては、表1に示すように顔画像データベ

ースを収集した。顔画像は正面から撮影された画像としている。収集した顔画像は男性 400 名、女性 400 名（合計 800 名）である。年齢については、表 1 に示すように 4 つの年齢層に分けた。また各年齢層において、その境目における判定は困難であるものと考え、境界付近の年齢は上下を 5 歳ずつ重ねることで、どちらに分類されてもよいことにした。

表 1 顔画像データベースの内訳

年齢	男性	女性
15～29 歳	100 枚	100 枚
25～44 歳	100 枚	100 枚
40～59 歳	100 枚	100 枚
55～80 歳	100 枚	100 枚

顔画像データベースの半数は、ICA によって基底ベクトルを抽出するための学習用サンプルとして用い、残り半数は評価実験のための評価用サンプルとして用いる。

3. 独立成分分析による基底ベクトルの抽出

提案手法の流れを図 1 に示す。図 1 の左図には性別、年齢層ごとの顔画像から ICA を用いて基底ベクトルを抽出するまでの流れを、右図には抽出した基底ベクトルを用いて性別、年齢層の推定を行なう流れを示す。

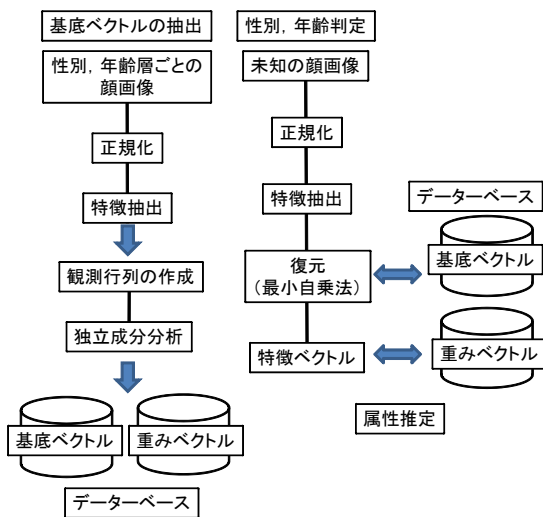


図 1 提案手法の流れ

図 1 の左図に示すように性別、年齢層ごとの顔画像については、それぞれの顔画像に対して正規化処理、特徴抽出（グレースケール特徴、二次元離散コサイン変換、ガボール変換）を行なう。次に ICA を行なう場合、個人の顔画像から抽出した特徴を観測信号として用いるのではなく、ランダムに複数枚の顔画像を選択し、それらの顔画像から抽出した特徴の平均値（グレースケール特徴

の場合は平均顔に相当する) を観測信号として扱い、ICA を行なう。ICA によって求められた基底ベクトルおよび重みベクトルはそれぞれデータベースに保存する。

3.1 正規化処理

性別、年齢層ごとに収集した顔画像について、まず顔領域のみの切り出し処理を行なう（図 2）。切り出す領域は、鼻を中心として、髪や首の領域は対象としないようにした。

切り出した顔領域については、さらに正規化処理によって、一定の大きさ（画素数 32×32）とした後、グレースケール変換を行なう。

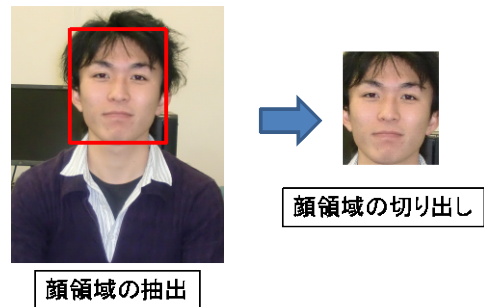


図 2 顔領域の切り出し処理

3.2 特徴抽出

正規化処理後の顔画像から特徴抽出を行なう。本研究においては、グレースケール特徴、二次元離散コサイン変換（Discrete Cosine Transform, 以下 DCT 変換と略す）による DCT 係数、ガボール変換による特徴抽出を行なった。これらの特徴は、顔画像を用いた個人識別や表情認識などで利用されている。

二次元 DCT 変換

二次元 DCT 変換について述べる。大きさ $M \times M$ の二次元画像信号 $f(i, j)$ ($i=1, 2, \dots, M, j=1, 2, \dots, M$) に対する二次元の DCT 係数を $F(k, l)$ ($k=1, 2, \dots, M, l=1, 2, \dots, M$) とすると、DCT 変換は以下の (1) 式で表わされる。

$$F(k, l) = \sum_{j=0}^{M-1} \sum_{i=0}^{M-1} f(i, j) \Phi_k(i) \Phi_l(j) \quad (1)$$

$$\Phi_k(i) = \begin{cases} \frac{1}{\sqrt{M}} & (k=0) \\ \sqrt{\frac{2}{M}} \cos \frac{(2i+1)k\pi}{2M} & (k=1, 2, \dots, M-1) \end{cases} \quad (2)$$

本研究においては、グレースケール変換後の画素値を $f(i, j)$ として (1) 式による DCT 変換を行なっている。また計算量の削減のため、顔画像（画素数 32×32）を 8×8 のブロックごとに分割し、各ブロックに対して独立

に DCT 変換を行なうようにした。以上の過程で抽出した DCT 係数の低周波成分が、表情認識に有効であるとの報告もされている[8]。

ガボール変換

ガボール変換について述べる。ガボール変換を行なうためのフィルタ G_θ は以下の (3) 式で表わされる。

$$G_\theta(x, y) = \cos\left\{\frac{2\pi}{T}(x \cos \theta + y \sin \theta)\right\} \times S \times e^h \quad (3)$$

$$h = -\frac{1}{2} \left\{ \left(\frac{x \cos \theta + y \sin \theta}{\partial x / 2} \right)^2 + \left(\frac{-x \sin \theta + y \cos \theta}{\partial x / 2} \right)^2 \right\} \quad (4)$$

(3) (4) 式中の θ は、ガボールフィルタの方位を示す。本研究においては、ガボールフィルタの方位 θ は 0 度と 90 度を用いる。入力画像 (大きさ $M \times M$) の座標 (x, y) のグレースケール変換後の画素値を $f(x, y)$ とした場合、ガボール変換は、画素値とガボールフィルタとの畳み込み演算によって行なわれる。ガボール変換は以下の (5) 式で表わされる。

$$R_\theta(x, y) = f(x, y) \otimes G_\theta(x, y) \\ = \sum_{i=-r}^r \sum_{j=-r}^r f(x+i, y+j) * G_\theta(x+i, y+j) \quad (5)$$

(5) 式中の \otimes は畳み込み演算を示す。 r はガボールフィルタの半径を示す。ガボール変換によって明暗 (もしくは暗明) のコントラストが変化する部分 (主として輪郭線) が抽出される。以上の過程によって抽出されたガボール変換による特徴は、顔画像の向きの推定に有効であるとの報告もされている[9]。

以上、本研究においては、グレースケール特徴、DCT 係数、ガボール変換による特徴を ICA の観測信号として扱う。特徴の次元数 T は、いずれも 1024 次元 (32×32) である。抽出した特徴は、性別、年齢層ごとにそれぞれデータベースに保存する。

3.3 観測信号の生成

本研究においては、ICA の観測信号として各個人の顔画像から抽出した特徴を直接用いるのではなく、複数枚の顔画像をランダムに組み合わせ、抽出した特徴の平均値を観測信号として用いることを試みる。顔画像は人によって千差万別であり、特定の性別、年齢層に属する個々の顔画像がそのカテゴリーを代表しているとは必ずしも言えないからである。

そこでランダムに K 枚の顔画像を選択し、それぞれの顔画像から特徴を抽出した後、 K 個の特徴の平均値 (グレースケール特徴の場合は平均顔) を ICA の観測信号とする。図 3 には、表 1 に示した顔画像データベースからランダムに 20 枚ずつの顔画像を選び ($K=20$ の場合)、グレースケール特徴において男性、女性の平均顔を作成し

た一例を示す。他の特徴 (ガボール変換, DCT 変換) についても同様に、観測信号としては K 個の特徴の平均値を用いることにする。



図 3 平均顔 ($K=20$ の場合)

3.4 基底ベクトルの抽出

次に ICA を用いて基底ベクトルを抽出する方法について述べる。観測信号として、学習用サンプルの顔画像から抽出した N 個の特徴を \mathbf{x}_i ($i=1, 2, \dots, N$) とする (3.3 節で述べたように、ランダムに選択した K 枚の顔画像から抽出した特徴の平均値である)。そして互いに独立な N 個の基底ベクトル (原信号) を \mathbf{y}_j ($j=1, 2, \dots, N$) とした場合、観測信号 \mathbf{x}_i は基底ベクトル \mathbf{y}_j の線形結合として表わされる ((6) 式および (7) 式)。観測信号と基底ベクトルの次元数はともに特徴数と同様に 1024 次元である。

$$\mathbf{x}_i = \sum_{j=1}^N a_{ij} \mathbf{y}_j \quad (6)$$

$$\begin{pmatrix} \mathbf{x}_1^t \\ \mathbf{x}_2^t \\ \vdots \\ \mathbf{x}_N^t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{y}_1^t \\ \mathbf{y}_2^t \\ \vdots \\ \mathbf{y}_N^t \end{pmatrix} \quad (7)$$

(6) 式および (7) 式中の $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{iN})^t$ ($i=1, 2, \dots, N$) を重みベクトルと呼ぶ。また観測行列を \mathbf{X} 、基底行列を \mathbf{Y} 、重み行列を \mathbf{A} と行列で表記すると、(7) 式は (8) 式のように表わされる。

$$\mathbf{X} = \mathbf{A} \mathbf{Y} \quad (8)$$

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)^t$$

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)^t$$

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N)^t$$

ICA においては、基底行列 \mathbf{Y} が統計的に独立になるように $\mathbf{Y} = \mathbf{W} \mathbf{X}$ で表される分離行列 \mathbf{W} を求める。本研究においては基底行列 \mathbf{Y} と分離行列 \mathbf{W} を求める手法として FastICA[10]を用いた。

ICA の結果、分離行列 \mathbf{W} から、重みベクトルは $\mathbf{A} = \mathbf{W}^{-1}$ より求められる。そして重みベクトル \mathbf{a}_i と基底ベクトル \mathbf{y}_j より、(6) 式に示すように観測信号 \mathbf{x}_i の復元が可能となる。すなわち重みベクトル \mathbf{a}_i は観測信号 \mathbf{x}_i の特徴として扱うことができる。

性別推定を行なうための基底ベクトルを抽出する場合は、男性用、女性用ともに $N/2$ 個ずつの観測信号を用意する。 \mathbf{x}_1 から $\mathbf{x}_{N/2}$ までの観測信号については男性の顔画像から抽出した特徴とし、 $\mathbf{x}_{N/2+1}$ から \mathbf{x}_N までは女性の顔画像から抽出した特徴として観測行列 \mathbf{X} を作成する。そしてこの観測行列 \mathbf{X} に対して、ICA を行ない基底行列 \mathbf{Y} と分離行列 \mathbf{W} を求める (図 4 上図)。性別推定のために抽出した基底ベクトルの一例を図 5 に示す。

従って、求められた重み行列 \mathbf{A} (分離行列 \mathbf{W} の逆行列) において、 \mathbf{a}_1 から $\mathbf{a}_{N/2}$ までの重みベクトルは男性用の特徴として、 $\mathbf{a}_{N/2+1}$ から \mathbf{a}_N までは女性用の特徴として用いることが可能となる (図 4 下図)。

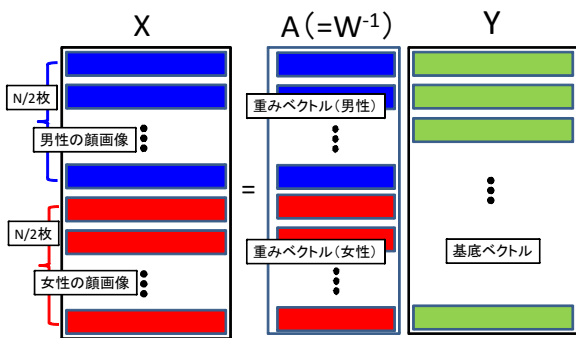
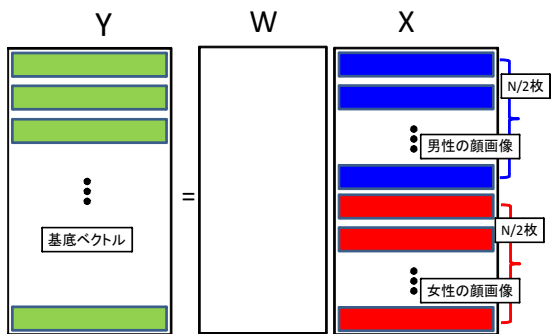


図 4 顔画像への独立成分分析の適用 (性別推定)

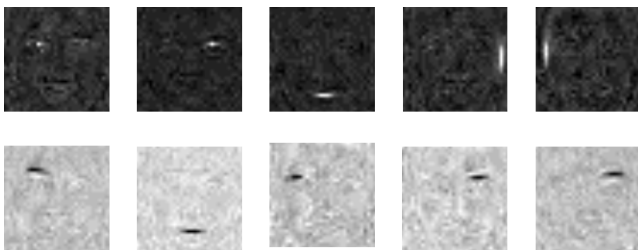


図 5 基底ベクトルの一例

そして年齢層推定の場合は、表 1 に示すように 4 つの年齢層ごとに $N/4$ 個ずつの観測信号を用意する。 \mathbf{x}_1 から $\mathbf{x}_{N/4}$ までの観測信号については、15 歳から 29 歳の顔画像から抽出した特徴として、 $\mathbf{x}_{N/4+1}$ から $\mathbf{x}_{2N/4}$ までは 25 歳から 44 歳の顔画像から抽出した特徴として、 $\mathbf{x}_{2N/4+1}$ か

ら $\mathbf{x}_{3N/4}$ までは 40 歳から 59 歳の顔画像から抽出した特徴として、 $\mathbf{x}_{3N/4+1}$ から \mathbf{x}_N までは 55 歳から 80 歳の顔画像から抽出した特徴として観測行列 \mathbf{X} を作成する。

そして同様に観測行列 \mathbf{X} について ICA を行ない、基底行列 \mathbf{Y} と分離行列 \mathbf{W} を求める。求められた重み行列 \mathbf{A} において、 \mathbf{a}_1 から $\mathbf{a}_{N/4}$ までの重みベクトルは 15 歳から 29 歳の特徴として、 $\mathbf{a}_{N/4+1}$ から $\mathbf{a}_{2N/4}$ までは 25 歳から 44 歳の特徴として、 $\mathbf{a}_{2N/4+1}$ から $\mathbf{a}_{3N/4}$ までは 40 歳から 59 歳の特徴として、 $\mathbf{a}_{3N/4+1}$ から \mathbf{a}_N までは 55 歳から 80 歳の特徴として用いることが可能となる。

以上、求められた重みベクトル、基底ベクトルはそれぞれ別々のデータベースに保存する。

4. 性別・年齢層の推定方法

性別、年齢層の推定については、図 1 の右図に示すように、未知の顔画像に対して、正規化処理、特徴抽出を行なう。次に基底ベクトルとの線形結合によって、その顔画像を復元する。その際に求めた重み係数を顔画像の特徴として扱い、データベース中の学習用パターンの重みベクトルと最近傍法によって類似度を求め、属性の推定を行なう。

4.1 基底ベクトルによる復元処理

未知の顔画像 \mathbf{U} から抽出した特徴を \mathbf{p} とする。また学習用パターンから求めた基底ベクトルを \mathbf{y}_j ($j=1,2,\dots,N$) とする (3 章で述べたように、性別推定を行なう場合は、性別推定のために抽出した基底ベクトルを、年齢推定を行なう場合は、年齢推定用の基底ベクトルを用いる)。ここで、未知の顔画像から抽出した特徴 \mathbf{p} を基底ベクトル \mathbf{y}_j によって復元する。すなわち特徴 \mathbf{p} を (9) 式に示すように基底ベクトル \mathbf{y}_j の線形結合によって表わすことが可能な重み係数 w_j ($j=1,2,\dots,N$) を求める。

$$\mathbf{p} = \sum_{j=1}^N w_j \mathbf{y}_j \quad (9)$$

ここで重み係数 $\mathbf{w}=(w_1, w_2, \dots, w_N)^t$ は (10) 式に示すように最小二乗法によって求める。ここで \mathbf{w} を特徴ベクトルとして扱う。

$$\mathbf{w} = (\mathbf{Y}^t \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{p} \quad (10)$$

すなわち特徴ベクトル \mathbf{w} と基底ベクトル \mathbf{y}_j によって、未知の顔画像から抽出した特徴 \mathbf{p} を復元することが可能となり、特徴ベクトル \mathbf{w} は顔画像 \mathbf{U} の固有の特徴として用いることができる (図 6)。

図 6 基底ベクトルによる復元処理

4.2 最近傍法による属性推定

属性推定を行なう場合、最小二乗法によって求めた特徴ベクトル \mathbf{w} を顔画像 \mathbf{U} の特徴として用い、これとデータベースに保存されている重みベクトルと最も類似しているものを照合する。すなわち性別推定の場合、学習用パターンから求めておいた全ての重みベクトル \mathbf{a}_i ($i=1,2,\dots,N/2$ の場合は男性、 $i=N/2+1,N/2+2,\dots,N$ の場合は女性) と特徴ベクトル \mathbf{w} との類似度を求め、最も類似度の高い重みベクトルに付与されている属性を推定結果とする(最近傍法)。類似度については(11)式に示すCOS尺度を用いる。

$$\cos \theta = \frac{\mathbf{w}'\mathbf{a}_i}{\sqrt{\|\mathbf{w}\| \|\mathbf{a}_i\|}} \quad (11)$$

年齢層推定についても同様の過程にて行なう。

5. 評価実験

以上、これまでに述べてきた提案手法について評価実験を行なう。性別、年齢層の推定は、同時に行なうのではなく、性別、年齢層ごとに別々に行なう。

5.1 性別の推定

表1に示す顔画像データベースのデータから、男女それぞれ半数(200枚ずつ)を学習用サンプル、残り半数(200枚ずつ)を評価用サンプルとして評価実験を行なった。ただし学習用と評価用を交換した交叉検証(2-Cross Validation)を行なった。

3章で述べたように、ICAを行なう場合、ランダムに選択した K 枚の顔画像から抽出した特徴の平均値を観測信号とする。その際、観測信号を作成するために用いる画像数を $K=1,5,10,15,20$ と変化させ性別推定を行なった。 $K=1$ の場合は、抽出した特徴をそのまま用いることを意味する。また観測行列を作成する際の観測信号数(基底ベクトルの個数) N については、10個から100個まで10個ずつ変化させた(男性、女性ともに $N/2$ 個ずつの観測信号を用いて観測行列 \mathbf{X} を作成する)。特徴として、グレースケール特徴を用いた場合について、学習用サンプルの結果を表2に、評価用サンプルの結果を表3に示す(表中の数値は識別率であり、単位は%である)。

表2および表3より、最も高い精度は、学習用パターンの場合、99.4% ($K=10, N=100$) であり、評価用パターンの場合、84.5% ($K=5, N=70$) であることが分かる。観測信号を作成するために用いる画像数 K については、個人の顔画像から抽出した特徴をそのまま観測信号とする場合 ($K=1$) よりも、複数個の特徴を平均化した場合 ($k>1$) の方が高い精度が得られることが分かる。

また観測信号数(基底ベクトル数) N については、学習用パターンの場合、 N の値が大きくなれば、明らかに

精度は高くなっていくことが分かる。評価用パターンについては、学習用パターンほど顕著ではないが、 N の値が大きいくほど、精度が高くなっていく傾向にある(ただし、 $N=100$ の場合が必ずしも最も精度が高いというわけではない)。

表2 性別推定の結果(学習用サンプル)

N	K=1	K=5	K=10	K=15	K=20
10	75.9	80.0	83.3	81.1	82.6
20	77.4	85.6	86.8	88.3	86.4
30	76.6	90.0	90.8	90.1	89.5
40	80.4	91.8	93.0	92.3	91.3
50	82.6	93.0	96.5	94.9	94.1
60	85.5	96.8	96.1	96.9	95.6
70	84.6	96.6	97.4	97.5	95.4
80	85.6	98.0	98.4	98.1	96.9
90	86.4	98.0	98.8	98.0	98.4
100	89.3	99.0	99.4	99.0	97.9

表3 性別推定の結果(評価用サンプル)

N	K=1	K=5	K=10	K=15	K=20
10	74.0	77.6	78.3	78.0	78.8
20	74.8	76.8	77.3	79.1	79.3
30	75.0	78.6	80.1	80.4	78.4
40	74.6	78.0	79.6	79.5	79.8
50	75.5	80.6	80.6	80.5	79.6
60	75.9	79.6	81.8	79.0	80.1
70	75.5	84.5	81.3	82.4	79.3
80	75.6	80.9	80.3	80.3	81.6
90	76.1	79.3	79.8	80.1	81.6
100	75.1	82.8	81.3	79.9	80.5

次に提案手法の有効性を評価するため、固有顔[5]および個人識別法による性別推定との比較を行なう。個人識別法とは、データベース中に含まれる全ての学習用データと推定したい顔画像を照合し、最も似ていると判定された人物の性別、年齢層を識別結果とする方法(最近傍法)である。特徴にはグレースケール特徴、DCT係数、ガボール変換(方位は0度、90度)を用いた。個人識別法の識別関数にはユークリッド距離を用いる。すなわち未知の顔画像と比較して、最もユークリッド距離の小さい顔画像の属性を認識結果とする。固有顔による識別方法においては、識別関数として複合類似度[3]を用いている。

また提案手法においても、特徴としてDCT係数、ガボール変換を用いた場合について、グレースケール特徴の場合と同様に観測信号を作成するために用いる画像数 K 、観測信号数(基底ベクトル数) N を変化させ実験を行なった。そして評価用サンプルに対して、それぞれの特

徴において、最も正解率が高い結果を表 4 に示す。表 4 中の ICA の結果において、括弧内は最も正解率が高かった場合の画像数 K および基底ベクトル数 N を示している。また固有顔と個人識別法による結果も合わせて、表 4 中に示す。

表 4 性別推定の結果のまとめ

	ICA	固有顔	個人識別法
グレースケール	84.5% ($K=5, N=70$)	79.6%	82.3%
ガボール (0 度)	81.3% ($K=10, N=40$)	82.0%	83.8%
ガボール (90 度)	79.6% ($K=15, N=100$)	78.6%	78.6%
DCT	66.3% ($K=10, N=50$)	79.0%	82.0%

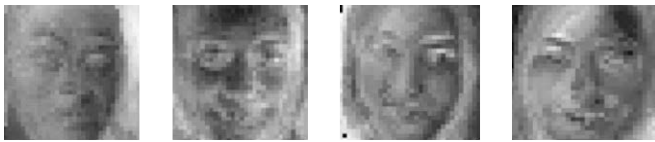


図 7 固有顔の一例[5]

表 4 より、提案手法においては、特徴としてグレースケール特徴を用いた場合、固有顔および個人識別法よりも高い精度が得られた。特に固有顔による手法と比較して、大きく精度の向上が見られる。図 7 に同じ学習用パターンから抽出した固有顔の一例を示す。

図 7 より固有顔の場合は、顔画像全体の特徴量が抽出されていることが分かる。一方で図 5 に示した ICA による基底ベクトルの場合は、顔画像を構成する局所的な特徴量が抽出されている。性別推定においては、ICA の基底ベクトルのような局所的な特徴量を用いた方が推定精度の向上を図ることができる。

次にグレースケール以外の特徴について考察する。ガボール変換 (90 度) においては、他の手法より精度の向上を図ることができた。一方で、ガボール変換 (0 度) および DCT 変換においては、他の手法よりも精度が下回る結果となった。

以上のことから、顔画像からの性別推定において ICA による基底ベクトルを用いることは有効であり、特にグレースケール特徴のような原画像に近い特徴を観測信号として用い、基底ベクトルを抽出し、この基底ベクトルによって推定した場合、高い精度が得られることが分かった。すなわち顔画像を ICA の観測信号の対象として、目や口といった顔を構成する部品を特徴として抽出し、その特徴を用いて認識を行なった方が高い精度が得られることを示している。そのためガボール変換や DCT 変換のように、変換後の特徴を観測信号の対象として、ICA

を行なうのは有効ではない。

5.2 年齢層の推定

年齢層の推定についても同様に、表 1 に示す顔画像データベースのデータから、それぞれ半数 (50 枚ずつ) を学習用サンプル、残り半数 (50 枚ずつ) を評価用サンプルとして実験を行なった (同じく学習用サンプルと評価用サンプルを交換した交叉検証 (2-Cross Validation) を行なった)。

性別推定と同様に、観測信号を生成する際に用いる画像数を $K=1, 5, 10, 15, 20$ と変化させ年齢層の推定を行なった。また観測信号数 (基底ベクトルの個数) N については、 $N=40, 80, 120, 160$ と変化させた (表 1 の各年齢層ごとで $N/4$ 個ずつの観測信号を用いて観測行列 X を生成する)。特徴として、グレースケール特徴を用いた場合、学習用サンプルの結果を表 5 に、評価用サンプルの結果を表 6 に示す (表中の単位は%)。

表 5 性別推定の結果 (学習用サンプル)

N	K=1	K=5	K=10	K=15	K=20
40	48.3	68.4	70.3	69.8	70.9
80	54.1	81.1	84.5	85.5	83.4
120	60.4	87.6	88.1	87.9	87.5
160	67.8	89.4	90.1	90.0	89.1

表 6 性別推定の結果 (評価用サンプル)

N	K=1	K=5	K=10	K=15	K=20
40	34.9	33.5	32.6	35.8	37.0
80	35.8	35.9	36.3	33.4	35.3
120	38.8	35.6	37.0	34.0	37.1
160	34.3	34.0	35.1	36.9	34.9

表 5 および表 6 より、最も高い精度は、学習用パターンの場合、90.1% ($K=10, N=160$) であり、評価用パターンの場合、38.8% ($K=1, N=120$) であることが分かる。観測信号を作成するために用いる画像数 K については、学習用パターンの場合、 K の値が大きいほど精度が高くなるが、評価用パターンの場合、 K の値に大きく影響を受けていないことが分かる。

観測信号数 (基底ベクトル数) N については、性別推定と同様な傾向にあることが分かる。すなわち学習用パターンの場合、 N の値が大きくなれば、明らかに精度は高くなっていくことが分かる。評価用パターンについては、学習用パターンほど顕著ではないが、 N の値が大きいくほど、精度が高くなっていく傾向にある。

また性別推定と同様に、特徴として DCT 係数、ガボール変換を用いた場合についても行なった。画像数 K および基底ベクトル数 N をグレースケール特徴と同様の条件

として行なった。評価用サンプルに対して、それぞれの特徴において、最も正解率が高い結果を表7に示す。また固有顔および個人識別法による年齢層推定の結果も合わせて表7に示す。

表7 性別推定の結果のまとめ

	ICA	固有顔	個人識別法
グレースケール	38.8% (K=1,N=120)	38.0%	34.5%
ガボール (0度)	38.4% (K=10,N=160)	40.8%	34.8%
ガボール (90度)	34.4% (K=10,N=10)	35.1%	31.6%
DCT	30.1% (K=1,N=20)	38.0%	34.5%

表7より、DCT変換を除く他の特徴においては、個人識別法の結果より高い精度が得られた。一方で、固有顔の結果と比較すると、グレースケール特徴の場合のみ高い精度が得られた。グレースケール特徴以外の特征において、高い精度が得られなかった理由としては、性別推定の場合と同様にガボール変換やDCT変換のように、変換後の特徴を観測信号の対象として、ICAを行なうのは有効ではないためである。

また表7より固有顔においては、ガボール変換（方位は0度）の場合、最も高い40.8%の精度が得られていることが分かる。性別推定と異なり、固有顔より高い精度が得られなかった理由としては、年齢層推定の場合、年齢層を代表する平均顔を作ることが難しいことが挙げられる。これは、提案手法においては、グレースケール特徴、DCT変換の場合、観測信号を作成するために用いる画像数をK=1として、個人の顔画像から抽出した特徴をそのまま観測信号として用いた場合の方が、高い精度が得られていることから明らかである。

以上のことから、本研究において提案するICAを用いた性別、年齢層の推定方法の有効性を示すことができた。特に性別推定においては、従来手法よりも高い精度が得られることを明らかにすることができた。

6. まとめ

本研究においては、独立成分分析（ICA）を用いた顔画像からの性別、年齢層推定の手法を提案し、収集した顔画像データベースを用いて評価実験を行なった。

その結果、性別推定においては、固有顔および個人識別法（最近傍法）を用いた手法と比較して、ICAを用いて性別推定を行なうことは有効であること、またICAの観測信号として、複数枚の顔画像を選択し、それぞれの顔画像から特徴を抽出した後、それら特徴の平均値を用いた場合、高い精度が得られることを示すことができた。

また観測信号に用いる特徴としては、グレースケール特徴のような原画像に近い特徴を用いた方が精度が高く、ガボール変換やDCT変換後の特徴に対しては高い精度を得ることは困難なことが明らかとなった。

一方、年齢層推定においてもICAを用いた場合、固有顔を用いた場合とほぼ同程度の精度が得られた。観測信号としては、性別推定と同じく、グレースケール特徴のような原画像に近い特徴を用いた方が精度が高く、その他の特徴については高い精度は得られなかった。また観測信号としては、複数枚の顔画像を利用する効果は性別推定ほど顕著には生じなかった。

今後は、ICAを用いて抽出した基底ベクトルについて、どの基底ベクトルが顔画像の属性推定において有効であるかを分析していくと同時に、観測信号に用いる特徴や識別関数の改良を行ない、より属性推定の精度の向上を図っていく予定である。

文献

- [1] M.Turk and A.Pentland : Eigenfaces for recognition, Journal of Cognitive Neuroscience, Vol.3, No.1, pp.71-86, 1991.
- [2] M. Turk and A. Pentland : Face recognition using eigenfaces, IEEE Proc. Conf. on Computer Vision and Pattern Recognition, pp.453-458, 1993.
- [3] 飯島泰蔵, パターン認識理論, 森北出版,1989.
- [4] 村瀬洋 : CVCV-WG 特別報告 : コンピュータビジョンにおける技術評論と将来展望(VI) : 固有空間法による画像認識, 情報処理学会研究報告, CV97-9, pp.59-66, 1995
- [5] 諏訪瑛, 篠沢佳久 : 固有顔を拡張した顔画像からの性別・年齢層推定, 画像の認識・理解シンポジウム (MIRU2010), IS1-4, pp.53-58, 2010
- [6] B.Olshausen and D.Field : Wavelet-like Receptive Fields Emerge from a Network That Learns Sparse Codes for Natural Images, Nature, Vol.381, pp.607-609, 1996.
- [7] 梅山, 赤穂, 菅生 : 教師付き独立成分分析の提案とその顔画像分析への応用について, 信学技法, NC99-2, pp.9-16, 1999.
- [8] 肖, チャンドラシリ, 田所, 尾田 : 2-D DCTとニューラルネットワークを用いた顔画像の表情認識, 電子情報通信学会論文誌 A, J81-A, No.7, pp.1077-1086, 1998.
- [9] 山根, 松本, サイモン, 伊藤 : 顔の Gabor ウェーブレット特徴量の推定, 信学技報, PRMU2000-217, pp.119-124, 2001.
- [10] A.Hyvärinen and E.Oja : A Fast Fixed-point Algorithm for Independent Component Analysis, Neural Computation, Vol.9, pp.1483-1492, 1997.