

研究テーマ 性別および年齢の推定に関する研究

研究者	安本 護	財団法人ソフトピアジャパン	雇用研究員
	本郷仁志	財団法人ソフトピアジャパン	雇用研究員
	輿水大和	中京大学	兼業研究員
	橋本恭貴	メディアドライブ株式会社	共同研究員
	鎌田一男	メディアドライブ株式会社	共同研究員

フェーズ I

1 研究の概要

円滑なコミュニケーションを確立するためには相手を知ることが必要であり、中でも性別や年齢を知ることが重要である。人間は、初対面の人物に対しては顔や髪型、服装等から性別や年齢を推測し、その結果に応じて対応しようとする。計算機と人のインタフェースにおいても、このような戦略は有効となると思われる。

本研究では、性別・年齢別の平均顔をプロトタイプに用いた性別・年齢推定手法を提案する。平均顔データは、15才から64才までを5才間隔で区切った男女別のクラスに対して作成する。入力顔と平均顔との特徴点間の距離を求め、距離最小を与える平均顔が代表する性別・年齢を推定結果とする。HOIP 顔画像データベースに収録された300名分の正面顔データを用いた実験により得られた推定率は、性別が約83%、年齢が約27%となった。

2 研究の目標

顔画像を用いた性別や年齢の認知、推定に関する研究は広く行われているが、性別判定と広範囲の年齢推定を同時に試みた研究は見当たらない。そこで15才から64才迄の人物を性別・年齢別に均等に収集したHOIP 顔画像データベースを用いて、顔画像を用いた性別・年齢推定の可能性を検討する。また、コンピュータによる推定結果との比較のためにも、人間による顔写真を用いた性別・年齢推定の心理実験を行う。本研究では、性別・年齢別の平均顔を基準として作成し、入力顔データと比較することにより、性別と年齢を同時に推定する手法を提案する。

3 実施内容

3.1 主観に基づく性別・年齢推定

コンピュータによる性別・年齢の推定結果との比較のために、人間による顔画像を用いた心理実験を行った。心理実験に用いる顔画像は、HOIP 顔画像データベースを用いて作成した。HOIP 顔画像データベースに関しては本報告書「顔画像データベースの構築」に詳細に記述されている。

3.1.1 顔から得られる性別・年齢特徴

表1および表2は、従来より得られている顔に表れる性差または加齢の特徴に関する知見を整理した物である。ここに示した特徴は、顔の骨格や顔筋に基づく形状情報と、しわやシミ、ソバカスのような皮膚面上の色・テクスチャ情報に二分できる。

3.1.2 心理実験

顔から得られる情報を利用して、人間がどの程度性別や年齢を推定できるかを知ることが興味深い。また、この結果を本研究で提案する性別・年齢推定手法が与える結果と比較することには意義がある。そこで、顔写真から性別・年齢を推定する二種類の心理実験A及びBを実行した。心理実験AとBでは、図1に示すように実験刺激として用いるカードが異なる。どちらのカードもHOIP

表1 顔の各部位にあらわれる性差

部位	女性	男性
額	暗い（前髪のため）	
眉	細い	長い、太い、濃い
眼	大きい、ぱっちり	細い、眉と近い
鼻	幅が狭い	幅が広い
頬	面積が大きい	ひげがある
口	上唇が薄い、唇の色が明るい	大きい、唇厚い
全体	色が明るい、輪郭が滑らか	長い、輪郭がゴツゴツ

表2 顔の各部位にあらわれる加齢特徴

部位	形状	色・テクスチャ
額	しわ、こめかみのくぼみ、額や眉間の表情じわ、生際の後退	額・眉間の表情じわあたりがくすむ
眉	細い、部分的に抜けてまばら	色が薄く見える
眼	上まぶたのくぼみ・かぶさり、目尻の下がり・しわ	眼の周りがくすみやすい、眼の下のくま、まつ毛が薄くなる
頬	くぼみ、たるみ、しわ	しみ、そばかす、くすみ
口	唇の輪郭がはっきりしない、口角が下がる	唇の色が濃くなる、くすみ
顎	たるみ、二重顎	



(a) 実験Aに用いたカード



(b) 実験Bに用いたカード

図1 心理実験に用いた刺激画像の例

顔画像データベースに収録している正面顔の白黒画像であるが、実験Aの刺激画像には、頭髪と衣服が推定に及ぼす影響を除去するためのマスクを行った。一方、実験Bの刺激画像には手を加えなかった。

実験の被験者は、全員日本人で、男性13名、女性8名からなる。被験者には、300枚のカードについて性別と15才から64才を5才間隔で区切った年齢区間を推定して、20個のクラスに分類するように教示した。実験は、予想される困難さの順にA、Bの順で実施した。実験A及び実験Bの結果を表3に示す。 n は両実験において評価対象となった延べ人数を示す。HOIP 顔画像データベー

スに収録された顔画像と本実験の被験者とを完全に独立に収集することができなかつたため、被験者がカードの人物を既知な場合があった。このような場合には、そのカードを実験結果に含めないように指示した。その結果、実験 A と実験 B で n の値は異なることとなった。 h_g は性別推定の推定率を示す。 r_g は実際の性別と推定結果との間のピアソンの積率相関である。 h_a は、推定年齢区間と実際の年齢区間が同一となった時を正解とする推定率である。 r_a は実年齢区間と推定年齢区間のピ

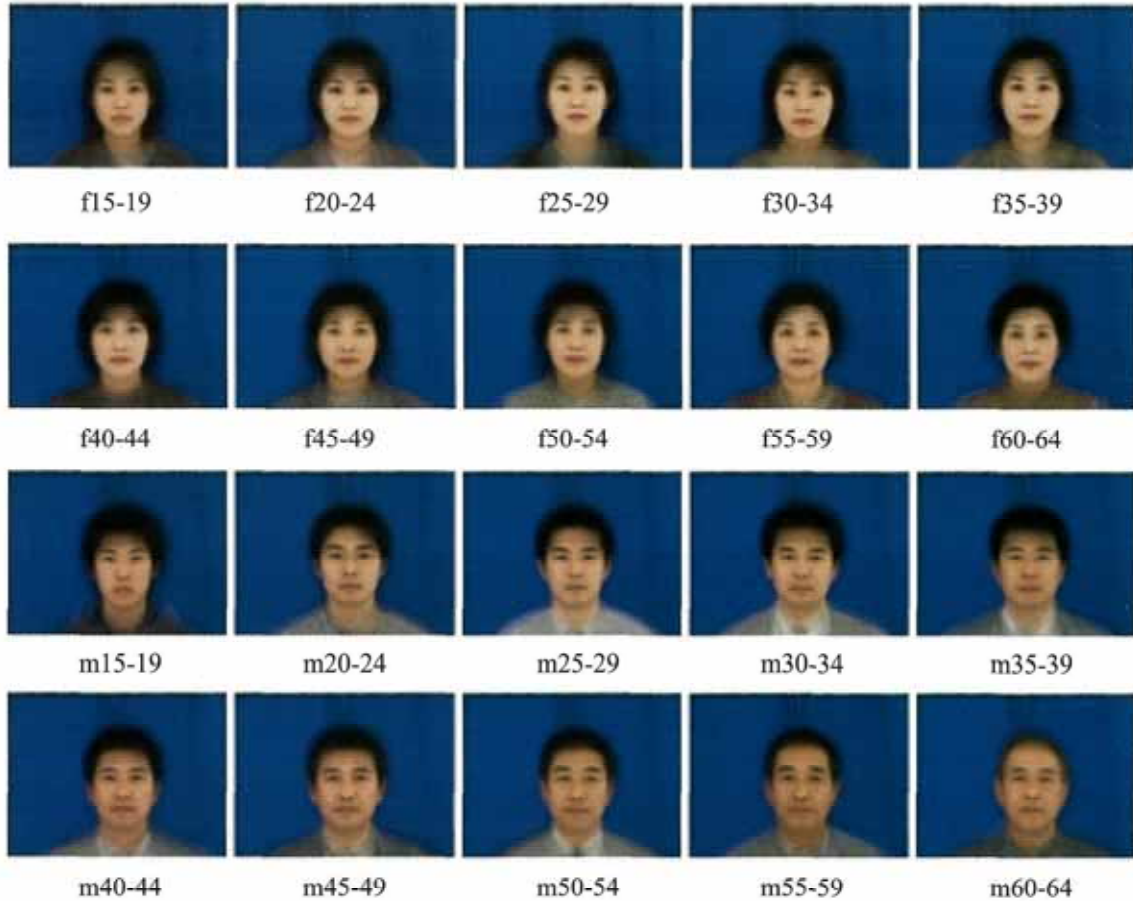


図 2 性別・年齢別平均顔

表 3 心理実験の結果

	n	h_g (%)	r_g	h_a (%)	r_a	m_e	σ_e
A	5978	98.2	0.96	30.8	0.88	-1.71	6.93
B	5887	99.7	0.99	35.5	0.91	-1.04	6.13

アソン積率相関である。 m_e と σ_e は年齢推定誤差 e_a の平均と標準偏差を示す。年齢の推定率が低いのは、年齢区間が 5 才とかなり狭いためである。前後の年齢区間も正解に含めると年齢推定率は実験 A で 72.5%であり、実験 B で 78.4%であった。実験 A の推定結果は実験 B よりやや劣るものの、性別推定、年齢推定ともに非常に高い相関が見られた。心理実験 A の結果は、顔から得られる情報だけで性別・年齢が推定可能なことを示唆するものである。

3.2 性別・年齢推定手法

3.2.1 性別・年齢別平均顔

特定の集団に属する人の平均顔を作成すると個々人の特徴は失われるが、集団に共通に存在する属性は保存される。また、保存された属性を強調することにより顔の属性に関連する要素を明らかにできる。平均顔のもつこのような特徴は、性別、年齢、魅力の研究に有効なツールとなる。

HOIP 顔画像データベースから性別及び 5 才年齢区間が同一である 15 人ごとについて生成した平均顔を図 2 に示す。ここで、図 2 のキャプションは、1 文字目が性別 (f: 女性、m: 男性)、以降が 5 才間隔の年齢区間を表している。平均顔の作成には、FaceFit と平均顔ツールを使用した。図 2 は、平均顔の特徴を非常によく示している。即ち、一人の人間が徐々に年老いていくようにも見えるし、家族の写真のようにも見える。

3.2.2 特徴点

図 2 の平均顔には、性別とその年齢区間を代表する特徴が現れている。そこで、入力顔に最も近い平均顔を探索することにより、性別及び年齢を推定する手法を提案する。推定結果は、探索された平均顔が代表する性別及び年齢とする。

性別・年齢を現す要素には形状特徴とテクスチャ・色特徴があるが、本研究では形状特徴だけでどこまで推定可能かを追求する。即ち、図 3 に示す特徴点に基づく入力顔と平均顔の距離を用いる。この特徴点は、FaceFit の顔モデルで用いられる頂点から選択したもので、顔輪郭: 38 点、両眉: 14 点、両眼: 24 点、鼻: 11 点、口: 18 点の合計 105 点からなる。

3.2.3 大域的正規化

入力顔と平均顔では、位置、サイズ、傾き等に違いがあるため、そのままでは両者の距離を測れない。そこで、入力顔を平均顔毎に正規化する処理を用いる。このような正規化にはアフィン変換がよく用いられるが、ここでは、より柔軟な非線形変換を用いる。この非線形変換は、原画像の矩形枠頂点 P_1, P_2, P_3, P_4 を任意の座標へ移動することにより指定するものである。移動後の各頂点をそれぞれ P'_1, P'_2, P'_3, P'_4 とすると、原画像上の任意の点 P は次式により点 P' にマッピングされる。

$$P' = \frac{a_x a_y P'_1 + a_x y P'_2 + x y P'_3 + x a_y P'_4}{wh} \quad (1)$$

ここで、 x と y は P の座標であり、 w と h はそれぞれ原画像の幅と高さとし、 a_x と a_y をそれぞれ $a_x = w - x$ 、 $a_y = h - y$ とする。

正規化の基準には、式 (2) に示す入力顔の特徴点 P_k と平均顔の特徴点 P_{ak} との 2 乗距離和 d を用い、 d を最小とする 4 頂点を決定する。

$$d = \sum_k \|P_{ik} - P_{ak}\|^2 \cdot \quad (2)$$

式 (2) では、 k を制御することにより顔の任意部位について 2 乗距離和を求められるが、大域的な正規化では全特徴点が対象となる。図 4 に非線形変換による大域的正規化の例を示す。

3.2.4 評価関数

性別や年齢の特徴は、表 1 及び表 2 に示したように顔の各部位に局所的に現れる。即ち、大域的な正規化後の d は性別及び年齢の推定に有効な顔形状の局所的差分を強く反映すると考えられる。顔の部位毎に推定に及ぼす影響が異なることを考慮して、式 (3) の評価値 E を定義する。 E を最大と

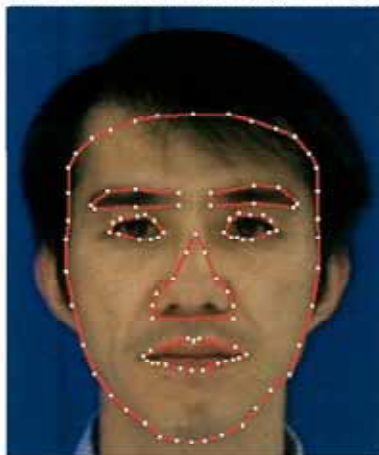


図 3 顔特徴点

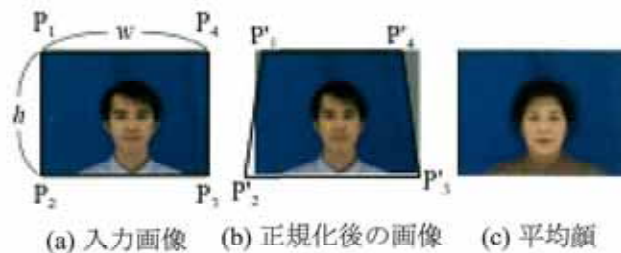


図 4 非線形変換を用いた大域的正規化

するものを推定結果と決定する。

$$E = \frac{1}{1 + (w_o d_o + w_b d_b + w_e d_e + w_n d_n + w_m d_m)^2} \quad (3)$$

ここで、 d_o 、 d_b 、 d_e 、 d_n 、 d_m は式 (2) を用いて求めた部位ごとの 2 乗距離和であり、それぞれ、顔輪郭、両眉、両眼、鼻、口に対応する。即ち、式 (3) は、各部位の 2 乗距離にそれぞれの重み係数を乗じて統合したものである。重み係数 w_o 、 w_b 、 w_e 、 w_n 、 w_m を最適化して最良の推定性能を得るために以下の式を定義する。

$$e_{gj} = \hat{g}_j - g_j \quad (4)$$

$$e_{aj} = \hat{a}_j - a_j \quad (5)$$

$$p_j = \begin{cases} 0 & (\text{if } e_{gj} = 0) \\ 1 & (\text{else}) \end{cases} \quad (6)$$

$$p_j = \begin{cases} 0 & (\text{if } e_{aj} = 0) \\ 1 & (\text{else}) \end{cases} \quad (7)$$

$$p_j = \begin{cases} 0 & (\text{if } e_{gj} = 0 \wedge e_{aj} = 0) \\ 1 & (\text{else}) \end{cases} \quad (8)$$

$$p_j = e_{aj}^2 \quad (9)$$

$$p_j = e_{gj}^2 + \left(\frac{e_{aj}}{5} \right)^2 \quad (10)$$

ここで、 g_j 及び a_j は被験者 j の実際の性別と年齢を示す。性別により、 g_j の値は $g_j=0$ (男性) または $g_j=1$ (女性) を取る。 a_j は年齢区間を区間の中央値で代表させたものである。例えば、被験者 j の年齢区間が 20-24 才の時、 $a_j=22.5$ となる。 \hat{g}_j および \hat{a}_j は、それぞれ g_j と a_j の推定値を示す。

即ち、式 (4) 及び式 (5) は、それぞれ性別と年齢の推定誤差である。式 (6) から式 (10) で定義する p_j は、被験者 j の推定結果に基づくペナルティである。重み係数は、複数の被験者をトレーニングデータとして推定した結果のペナルティ p_j の総和が最小となるように決定する。式 (6) は性別推定結果の一致、式 (7) は年齢推定結果の一致、式 (8) は性別及び年齢推定結果の一致に関するものであり、一致するとき 0、一致しないとき 1 を返す。また、式 (9) は年齢推定誤差の 2 乗、式 (10) は年齢推定誤差の 2 乗と性別推定誤差の 2 乗の和を返す。ここで、年齢推定誤差は隣り合う年齢区間との差が 1 となるようにした。なお、性別推定誤差の 2 乗は式 (6) と等価であるため省略した。

4 結果

4.1 実験と考察

性別・年齢推定実験では、テストデータとプロトタイプデータの両方に HOIP 顔画像データベースから取り出した正面顔データ 300 枚を使用した。プロトタイプデータは、性別及び年齢区間が同一であるクラスの被験者 15 名について顔特徴点を平均して得る。但し、テストデータをそれ自身が属するクラスのプロトタイプと比較する際は、テストデータがプロトタイプ作成に含まれないように該当する被験者を除いた 14 名分の特徴点データからプロトタイプを作り直した。

重み係数最適化のトレーニングデータにも HOIP 顔画像データベースを用いた。ペナルティ p_j

表 4 重み係数

Opt.	w_o	w_b	w_e	w_n	w_m
None	1.0	1.0	1.0	1.0	1.0
Eq. (6)	0.188	1.537	0.945	1.333	0.341
Eq. (7)	0.904	1.915	1.431	0.951	0.866
Eq. (8)	0.112	1.382	1.001	2.180	0.997
Eq. (9)	0.035	1.126	0.994	1.026	0.294
Eq. (10)	0.045	0.989	0.949	1.016	0.311

表 5 提案手法による性別・年齢推定結果

Opt.	hg (%)	rg	ha (%)	ra	me	σ_e
None	77.7	0.57	25.3	0.52	-0.55	14.27
Eq. (6)	83.7	0.69	26.0	0.64	0.20	12.57
Eq. (7)	79.0	0.59	27.0	0.55	0.18	13.96
Eq. (8)	83.0	0.67	27.0	0.55	-0.18	13.85
Eq. (9)	82.0	0.65	26.7	0.64	0.42	12.39
Eq. (10)	83.3	0.68	26.3	0.64	0.47	12.41

は重み係数の関数ではないため、最適化処理に関数の微分を用いる方法は適用できない。そこで、(パウエルの) 方向集合の手法を使用した。式 (6) から式 (10) に定義したペナルティ p_j に基づいて最適化処理を行った結果得られた重み係数を表 4 に示す。また、このときの推定結果を表 5 に示す。比較のために示した最適化を行わない時の重み係数は、すべて 1.0 とした。

重み係数の最適化により、表 5 の各項目に改善が見られることがわかる。性別一致を目的とする式 (7) を用いた場合、性別推定率は表 5 中最高の 83.7% となった。年齢の一致を目的とする式 (8) を用いたとき、年齢推定率はやはり表 5 で最高の 27.0% となった。しかし、これは性別及び年齢両方一致を目的とする式 (8) の年齢推定率と変わらず、しかも性別推定率はあまりよい結果とならなかった。年齢推定誤差の 2 乗最小を目的とする式 (9) を用いた場合は、年齢の推定率は最高とならなかったが、年齢推定誤差の標準偏差は最小となった。性別及び年齢の両方の推定結果をバランスよく得るには、式 (8) あるいは式 (10) がよい結果を示した。このように推定結果は、ペナルティの選び方において意図した通りの傾向を与えるものであった。

顔の輪郭に関する重み w_o は、式 (7) の場合を除いて、他の重みに比べてはるかに小さくなった。これは、少なくとも本手法では、顔輪郭の形状情報が性別・年齢推定にそれほど関与していないことを示唆している。しかしながら、表 1 及び表 2 に示したように、人間の性別や年齢の認知に関する従来研究からは、顔輪郭にも性別及び年齢に関する特徴が現れることが示唆されている。本研究における顔輪郭は顔部位の中で最も大きく、また最も多くの特徴点を含んでいるため、大域的正規化によってこのような不一致が発生した可能性がある。

これらの推定結果は、比較のため行った心理実験で得られた人間の性別・年齢認知能力には及ばなかったが、顔の形状情報を用いて性別及び年齢を推定することが可能である見込みを示すものである。人間は顔の形状に基づく特徴だけでなく、他の様々な特徴を統合して性別・年齢推定を行っている。このような人間の情報処理手法を参考に、皮膚上のしわや色といった特徴を本研究の成果と統合すれば、推定性能を向上できると思われる。

フェーズ II

1 研究の概要

フェーズ I における研究では、顔画像から性別と年齢を推定する可能性が示された。しかし、その手法は、目尻、目頭、眉、口の両端、顔の輪郭などの顔の特徴となる数多くの点を正確に入力する必要があり、実利用が困難な手法と思われる。そこで、顔画像から抽出した四方向面特徴と線形判別分析を用いて、実利用を目指した推定手法を検討する。本研究では、線形判別分析で学習した判別器をベースとし、性別・年齢のクラス分類の方法の違いによる推定率を比較する。HOIP 顔画像データベースに含まれる 300 人の正面顔画像に対して実験を行ったところ、性別推定では、男女混在分類において 92% の推定率を得た。年齢推定においては、男女と性別で分け、年齢を 10 歳区分で学習することで、学習人物に対しては約 70%、未知人物に対しては約 40% の推定率が得られた。

2 研究の目標

フェーズ I の研究において提案された推定手法は、推定に必要な特徴量を正確に与える必要があり、応用システムへの実装は困難と思われる。そこで、実装が容易なアピアランススペースの特徴量を用いた推定手法の提案を行う。また、統計的な判別処理を用いて、性別と年齢の推定に有効なクラス分け手法の検討を行い、精度の向上を目指す。

3 実施内容

3.1 顔画像における性別と年齢の特徴

顔画像の性別と年齢の推定において、顔画像におけるいかなる特徴量を用いて、いかなるアルゴリズムで推定するかが課題となる。

年齢変化に強く関与する特徴量として、幼年期から成年期にいたるまでに顔が発達することから形状情報が挙げられる。特に、顔の長さとの比率、例えば目と目の距離が大きく変化する。よって、物理的情報から識別が可能である。しかし、本研究で対象とする多くの人物は、顔が発育した思春期以降となり、そのような大きな変化は期待できない。

HOIP 顔画像データベースから作成した性別・年齢別平均顔 (図 2) を用いて 15-19 歳の平均顔を 60-64 歳の平均顔と比較すると、顎が細くなっていることが確認できる。現在の若者と一昔前の若者と顔の形状を比較すると顎が細くなっていると言われており、同様の結果が得られた。フェーズ I での研究により、年齢別に生成した平均顔間に差があることから、皮膚のたるみなどの加齢による変化に加え、世代による差もあることが予想される。また、額を顔の一部と捉えると、髪が生え際の後退から、顔領域内の顔部品の位置関係に影響する。つまり形状情報が大きく変わるようになる。よって、成年においても形状情報が利用できることが示唆される。

年齢に関与するその他の特徴量としては、加齢によるしわ情報が挙げられる。多くの関連研究では、年齢推定にしわ情報から年齢推定を行っている。老化に伴いしわが増加する傾向から、しわが年齢を推察する判断基準のひとつになるといえる。しかし、年齢に依存する“しわ”を検出することは困難である。しわを検出するために、しわの強調や検出パラメータの調整が必要となる。また、しわを撮影するだけの高解像度の顔画像が必要となる。

一方、性差としては、目や眉、唇などの顔部品や骨格などに差が現れることが知られている。例えば、女性は目が大きく、眉が細く、上唇が薄い。男性は、眉が濃く、口が大きく、唇が厚い。性別を推定するにあたり、目や口の特徴差は有効であることが示唆されている。また、男性は顔が長く、女性は丸みを帯びているなど、顔の輪郭や造形も性別を推定するのに有効な特徴と思われる。

以上により、顔全体の形状を表す特徴と、目や口の詳細な形状の特徴、しわの出現度合を表す特徴を反映するそれぞれの特徴量を抽出し、総合的に評価することが重要と考える。

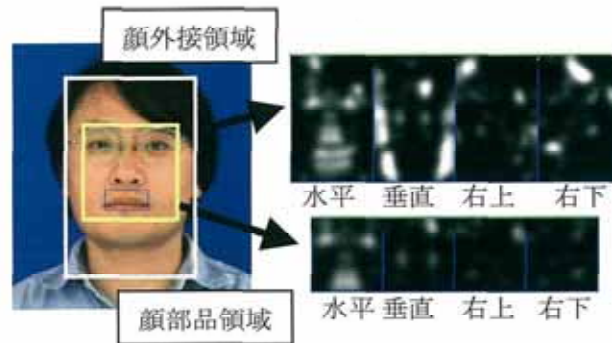


図5 特徴量抽出領域と四方向面特徴

3.2 特徴量抽出

本研究では、「顔画像認識に関する研究」や「顔向き推定に関する研究」などで用いている四方向面特徴（以降、FDF）と線形判別分析（以降、LDA）による手法の性別・年齢推定への適用を試みる。FDFは、エッジ勾配の方向に応じて、勾配値を四方向に分離する。よって、しわの出現頻度や勾配の度合が特徴量として反映されると考える。また、FDFの各方向面はモザイク特徴と同様に平面で構成されるため、顔の輪郭や造形などの大まかな形状情報を保持している。なお、FDFの解像度は 8×8 となる。

先にも述べたように、性別・年齢を推定するためには目や口の性差による特徴の差と形状特徴の両方を適切に抽出する必要がある。そこで、顔全体の特徴を捉えるため顔の輪郭を含む顔外接領域と顔部品に注目した目と口を含む顔の顔部品領域とからそれぞれ特徴量FDFを抽出する。

顔領域の検出には、肌色基準値法により顔領域を抽出した。なお、服の色によっては首または胸の一部が含まれる人物がいるため、顔領域の横幅1に対して縦方向が最大1.6の比率で切り出した。顔部品抽出には、検出した顔領域からテンプレートマッチングと弛緩整合法を組合せた手法により抽出を行った。これらの手法に関しては、「顔部品検出と注目推定に関する研究」で詳細に記述する。本手法を用いて顔画像データ300名に対して抽出実験を行ったところ、学習データセットでは1名、テストデータセットでは3名において誤検出された。誤検出された顔部品の位置はマニュアルで修正した。

図5に特徴量を抽出した両目、口と、特徴量を抽出する顔外接領域と顔部品領域を示す。顔部品領域は、目や口の検出誤差や個人による顔の傾きなどを考慮し、また眉領域を含むようにするため領域を上下方向に固定量の拡張を行った。実験では、学習データが各人1枚のため、学習パターンを増やすために各特徴量抽出領域を上下左右に伸縮させ、21パターンの特徴量を抽出した。

顔外接領域からは、顔全体の大まかな特徴を捉え、顔部品領域は、顔の中における顔部品の位置情報を無効にする代わりに、詳細に顔部品の情報を抽出する。顔外接領域と顔部品領域から四方向面特徴を抽出することで、顔の形状情報としわ情報の両方を有効に抽出することとなる。

3.3 判別クラスの構成

性別・年代層推定にSVMを用いた方法が提案されている。SVMは判別性能が優れた学習モデルのひとつであり、2クラスの判別器を構成するための学習法である。そのため性別を推定するにはよいが、年齢を推定するには判別器を階層的に構成するなど工夫が必要となり、詳細な年齢推定は困難と思われる。

LDAとFDFによる手法によって「顔向き推定に関する研究」では、顔方向毎に50名の顔画像を学習し未知人物に対しても学習人物と同等の推定率を得た。つまり、顔方向推定において汎化性を得ることができたと推察される。LDAは、学習データにおけるクラス間分散を最大に、クラス内分散を最小にするように判別空間を求めるマルチクラスの学習モデルである。

年齢・性別推定においても、同様に性別・年齢別にクラス分けを行い、クラス間で性別や年齢を判別する特徴量を学習できれば、性別・年齢の推定が可能になると思われる。したがって、年齢・性別を推定する際に、顔画像をどのようなクラス分けをすることで安定した推定が行えるのか探る必要がある。つまり、目的に応じて識別に有効な特徴量と適切なクラス分類が重要となる。

表6 性別・年齢推定のための分類方法

クラス分類方法 (性別・年齢区分)		クラス数
性別推定	男性/女性	2
年齢推定	男女別: 5歳間隔	各10
	男女別: 10歳間隔	各5
性別・年齢同時推定	男女混在: 5歳間隔	20
	男女混在: 10歳間隔	10

表7 性別推定結果

推定率 (%)		学習人物	未知人物
顔外接領域	男性	98.0	87.3
	女性	92.7	83.3
	平均	90.3	85.3
顔部品領域	男性	97.3	90.0
	女性	98.7	90.7
	平均	94.2	90.3

クラスの分類方法としては、表6に示したような分類が考えられる。性別推定は、男女の2クラスに分類する。年齢推定においては、先ず性別推定を行い、その結果に基づいて年齢推定を行うことができる。男性と女性で年齢推定の基準が異なると思われるので、性別に学習することで、より適切に推定できる可能性がある。さらには、男女混在したまま性別・年齢でクラス分けすることで、性別・年齢を同時に推定する方法も可能である。

実験では、各クラス分けによる識別精度を比較実験する。なお、年齢区間として、5歳間隔はデータ収集と同様に分類し、10歳間隔では、15-24、25-34、35-44、45-54、55-64歳と分割した。

4 結果

4.1 実験方法

実験では、HOIP顔画像データベースの正面顔画像を用いた。顔画像データベースでは、各人につき2回の撮影を行っているため、1回目に撮影した画像を学習データとして、2回目の撮影データをテストデータとして用いた。

まず、性別推定実験では、年齢順に並べて奇数番目と偶数番目の2グループ各150名に分類した。一方を学習データ(学習人物)とし、他方をテストデータ(未知人物)として、判別実験を行った。次に、学習人物と未知人物を入れ替えて同様に実験を行った。

年齢推定実験では、データセットの各クラスを均等に5セットに分割し、4セット(80%)を学習人物、残りの1セット(20%)を未学習人物とし、Leave-one-out法に基づき評価した。

4.2 結果と考察

性別判定実験の結果を表7に示す。顔部品領域の特徴量を用いた方が、顔外接領域より高い推定率を得た。これは、目や口などの顔部品に性差が現れやすいことから、顔部品に注目した顔部品領域から抽出した特徴量を学習した方がより適切に性別を分類できたことが示唆される。性別で識別率を比較すると、顔外接領域の場合、男性の方が高く、顔部品領域では逆に女性の方が高くなった。女性は髪型により顔の輪郭が変動しやすく、このような結果になったと推察される。男女混在方法による性別推定結果については後で述べる。

次に、顔外接領域と顔部品領域の年齢推定結果をそれぞれ表 8 に示す。年齢推定に関しては、全体的に年齢区分が細かいほど推定率は低下した。学習人物においては、約 70% 以上の推定率を得たが、未知人物に対しては大きく低下した。未知人物に関して、第 2 位までに正解が含まれる累積正解率を括弧内に示した。その場合、顔部品領域を用いて、10 歳区分で識別した場合に男女ともに 60% を越えた。

特徴量領域に関しては、顔部品領域の方が若干、識別率が高かった。年齢推定においても男女による識別率が逆転する傾向が見られた。このことは性別判定と同様に女性の髪型による顔の輪郭の変化が考えられる。

男女混在によるクラス分類方法と比較すると、想定したように男性・女性と性別毎に年齢分類クラスを構成する方がよい結果が得られた。識別するクラス数が少ないことも原因の一つと考えられる。

以上より、年齢推定においては、未知人物に対する汎化性が十分得られなかったことが示された。このことは、本実験で用いている特徴量に年齢を識別するための際だった優位性がないことが一因として考えられる。心理実験による 15 歳の許容範囲で約 70% の推定率からも推察されるように、顔から年齢を表す定量的な特徴量が少なく、個人差が大きいことが示唆される。

男女混在方法による性別推定結果を表 9 に示す。性別の判定方法は、識別されたクラスの性別を結果とした。ここでは 10 歳区分で分類した場合の未知人物に対する推定結果を示す。

顔部品領域において、92% の識別率を得た。これは、先の性別判定実験で得た 90.3% を上回っている。このことは、男性、女性を 1 つのクラスでまとめて学習するより、顔部品においても年齢差があるため、年齢別に分けてそれぞれ学習する方が、適切な分類ができることが示唆される。つまり、男性を 1 つのクラスで表現するのではなく、10 代、20 代、30 代と複数のクラスで表現することになり、加齢による特徴変化を分離できたことが原因の一つと思われる。

本研究で提案した手法では、人間の性別・年齢推定能力には及ばなかったが、提案手法による性別及び年齢の推定が可能である見込みを示すものである。しかし、改善点も多くさらなる改善が必要と思われる。

例えば、FDF は変動にロバストな特徴量であるため、LDA により顔の大まかな造形を学習したと思われる。よって、細かなしわなどの特徴は抑制されたと考えられる。また、一般的に学習人物

表 8 年齢推定結果

分類	顔外接領域		顔部品領域	
	学習人物	未知人物 (第2位)	学習人物	未知人物 (第2位)
男性 : 5歳区分	82.3	19.3 (36.7)	77.0	20.7 (40.0)
女性 : 5歳区分	73.2	20.0 (34.7)	74.5	22.7 (36.0)
男性 : 10歳区分	84.0	39.3 (62.0)	74.7	36.7 (61.3)
女性 : 10歳区分	74.2	35.3 (54.7)	78.5	40.0 (62.7)
混在 : 5歳区分	75.3	20.3 (33.7)	68.9	20.3 (34.7)
混在 : 10歳区分	73.8	35.7 (56.0)	68.8	33.3 (59.0)

表 9 混在方法による未知人物に対する性別推定

未知人物の推定率 (%)		顔外接	顔部品
混在10歳区分	男性	92.7	92.0
	女性	86.7	92.0
	平均	89.6	92.0

と未知人物とで推定率に差が表れたが、特徴量、学習データのバリエーション不足、識別器の構成などが原因と考えられる。

今回は、15歳未満の子供および65歳以上の高齢者はデータベースに収録していないこともあり対象外としたが、同じ枠組みで学習させることで識別可能と考える。ただし、思春期以前の子供ならび70歳以上の高齢者の顔には性差がはっきりしないため、識別が困難になることが予想される。

人間は様々な特徴を総合的に判断していることが考えられる。これらの問題は、性別・年齢を総合的に判断する枠組みが必要になると思われる。

フェーズ III

今後の取り組み

フェーズI、フェーズIIの研究において、顔画像から性別と年齢を推定する技術開発の可能性が示された。また、フェーズIIで提案された手法は、システムへの実装を意識した手法であり、実際に本プロジェクトで試作した「自動マーケティング情報システム」のデモシステムに実装し、動作確認を行っている。詳細については「注目情報とマーケティングに関する研究」に記述されている。

マンマシンインタフェースへの応用例としては、顔画像から推定された属性に応じて、操作画面や操作体系を切り替える手法が考えられる。具体的には、推定結果が子供であれば、漢字に読み仮名を付加するような表示や、高齢者と推定した場合には文字を大きく表示するようなインタフェースが有用と思われる。このようなユーザ適応型のインタフェースは、ヒューマンフレンドリな機器の実現に重要な役割を担っている。

今後の課題としては、性別・年齢推定の両方において、精度と頑健性の向上が挙げられる。特に、実環境においては、顔方向、照明、表情等の多様な顔画像の変化が考えられるため、これらの変化に対する汎化性の評価と向上が重要である。