

## 前編

## 顔画像処理技術の動向

労 世 竝 オムロン (株) 技術本部

山口 修 (株) 東芝 研究開発センター

## 顔画像処理の概観

近年、デジタルカメラやビデオカメラなどのデジタル画像機器の普及とコンピュータの処理能力の向上に伴って、顔画像のデジタル処理を手軽に行うことができるようになってきた。顔画像処理技術は大きく2種類に分けることができる。まずは、顔を理解するための技術として、顔検出、顔トラッキング、個人識別、性別年齢推定、表情推定、顔の疲れや眼気の推定の技術が広く研究されている。また、顔を表現する技術として、美肌補正、美白補正、赤目補正、表情合成、自動似顔絵作成、デジタル化粧などの研究開発も急速に広がりを見せている。

本稿では、2回にわたってこれらの最新動向を紹介する。1回目は顔を理解するための要素技術の動向、2回目はアプリケーションについて紹介する。

## □顔画像処理の重要性

顔画像処理技術が重要視されるのは主に2つの理由がある：まず、顔は日常生活の中で非常に重要で、顔からさまざまな情報を取り出すことができる。また、顔の見た目は姿勢や表情などによって変化するため、パターン認識の研究対象として学術的に興味深い。

我々は日常生活の中で相手の顔を見てさまざまな判断を行う。顔から得ている情報は、大きく2種類に分類できる。個人の身元や性別年齢を識別するための情報と人の内面的な精神と身体状態に関する情報である。いずれも人と人とのコミュニケーションの中できわめて重要な情報である。たとえば、数十年ぶりに旧友と再会する場合、思い出すのに頼れるのはやはり顔である。顔を見てその人の性別や年齢を判別する能力がなくなれば、生活に大きな支障をきたすに違いない。喜怒哀楽が一番よく分かるのも顔である。言葉やジェスチャーも、人と人とのコミュニケーションの中で非常に重要な役割を果たしているが、いずれも民族性や地域性によって共通性が乏しく、通じないどころか、逆に感じることもすらある。それに対し顔の表情は、幸い万国共通であり、世界中の人が

同じように笑顔を感じ取ることができる。「目は心の窓」ということわざがあるが、顔も心の窓と言っても過言ではない。

人の顔は両目、鼻、口の順に上から下へ万人共通で配置されている。一方、個人それぞれには異なる顔がある。この2つの特徴を有する顔はパターン認識の研究対象としては興味深い。一般的に、コンピュータはバーコードや印刷文字など決まった形の物体を認識するのが得意であるが、形状が変化に富んだ物体の認識は難しい。顔は一般的な非剛体の3次元物体と同様、照明環境や向きによって見え方が大きく変化するだけでなく、表情変化や加齢によっても顔の形状とテクスチャが変化する。このように、類似するが見え方が変化するものを見つけ、見分けることはコンピュータにとって苦手なことである。したがって、顔画像処理の研究は他のパターン認識、機械学習の研究にも大きな影響を与える重要な研究テーマと言える。

日本では顔画像処理の研究開発が非常に盛んであり、顔画像処理の重要性が広く認知され、特に応用分野においては世界をリードしていると言える。

## □顔画像処理の歴史

顔画像処理技術がどのように発展してきたのかについて、簡単に紹介する。

顔画像をコンピュータで処理する研究は40年の歴史がある。1960年代後半から70年代前半まで京都大学坂井研究室ですでに顔に関する研究成果が発表されていた。それ以来、いくつかの顔研究ブームが起きている。まず、90年代初めにTurkらによって発表された固有顔が火付け役になって、顔認識の研究が本格的に始まり、パターン認識、コンピュータビジョン関連の研究者の注目を集めた。90年代後半では、USCのMalsburgのグループがFace Recognition by Elastic Bunch Graph Matching (EBGM)を発表し、ベンチャー企業を立ち上げ、顔認識の実用化が近いと感じさせる顔認識技術を作り上げた。ほぼ同時期に、主にMITのSungらや

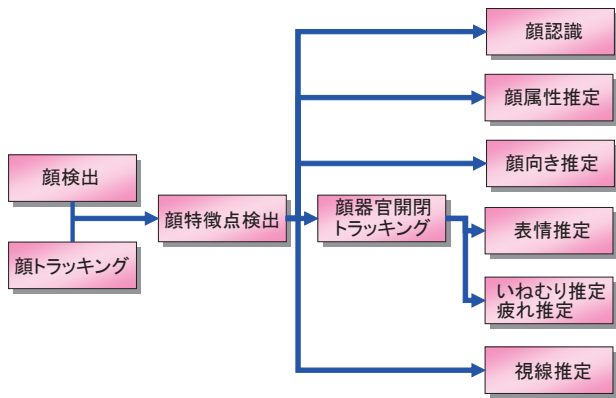


図-1 顔画像処理の要素技術例



図-2 顔の多様性

CMU の Rowley らによって、画像の中で顔の位置を特定する顔検出の技術が開発された。21 世紀始め、Viola と Jones が考案した高速な顔検出技術が顔画像処理のリアルタイム処理を実現させ、顔画像処理の実用化に向けた大きな一歩を踏み出した。

日本国内では、特に顔画像処理の実用化に関する研究開発が活発に行われている。古くからコンピュータで作る似顔絵の研究が盛んで、中京大の興水研究室や電通大の金子研究室などが長年続けてきた。近年、オムロンやデジタルカメラ各社は顔検出や笑顔認識の技術を開発し、世界に先駆けてデジタルカメラで実用化に成功している。

## 実用化が進む顔画像処理技術

従来、画像の中から高速に顔の位置を見つける顔検出技術が顔画像処理技術のネックになっていた。顔検出の高速なアルゴリズムが確立したことによって、実用化が急速に広がり、多岐にわたる顔画像処理技術の実現に至った。

### □顔画像処理技術の構成

顔を理解するための顔画像処理技術として、最も重要で基礎的な技術は顔検出技術である。画像の中から漏れなく、誤検出なく、すばやく顔の位置を正確に特定する必要があるからである。顔検出した後に動画処理が必要な場合は顔や頭部のトラッキングも必要となることが多い。

顔の位置が特定できた後に、顔の性別、年齢を含めた属性推定や個人識別を行うためには、顔の各器官の特徴点の検出が重要となる。特徴点の検出精度が後の認識結果に大きく影響を与える。また、顔向き推定も重要な技術の 1 つである。特に顔向きのロバスト性を要求する場合はまず顔向きの正確な推定が必要である。表情や居眠りの推定には、静止画で行う場合と、より高い精度が要

求されるために動画で行う場合がある。動画で行う場合は顔の器官の動きを検出、トラッキングする必要がある。図-1 に顔画像処理の技術構成例を示す。

### □顔検出

デジタル写真や映像が急速に普及している中で、顔検出技術の役割がますます重要になってきた。たとえば、コンピュータで顔認識を行う場合は画像の中からどこに顔があるかをまず検出する必要がある。また、顔認識を使った監視システムの中では動画から顔をほぼリアルタイムに検出することが要求されている。

### 顔検出の難しさ

コンピュータで顔をロバストに検出する難しさは顔の多様性に由来する。人には個人差があり、性別、年齢、人種などの違いによって、顔の見え方はかなり違う。また、同じ人の顔でも、顔の向き、照明環境と表情によっても見え方が違う。実アプリケーションの場合、顔の大きさ、回転角度も変化していることが多いので、さらに複雑になる(図-2)。

### 従来の顔検出技術とその課題

人はまったく努力せずに、自然に顔を見つけることができる。しかしながら、コンピュータで顔を検出するには画像の中で順番に領域を切り出して、その領域が顔かどうかを判定することによって顔の位置を特定する(図-3)。そのため、膨大な数の領域の判別が必要となり、高速な判別器が重要になる。

90 年代後半 CMU の Rowley らが開発したニューラルネットワークによる顔検出技術は学習技術を使った顔検出のフレームワークを確立し、当時は最も性能の良い顔検出技術として知られていたが、計算速度に関しては実時間処理には及ばなかった。

2000 年頃、Viola と Jones が高速に顔検出を行う手法を提案し、実時間での顔検出を可能にした。この手法は近年最も脚光を浴びたパターン認識手法の 1 つであ

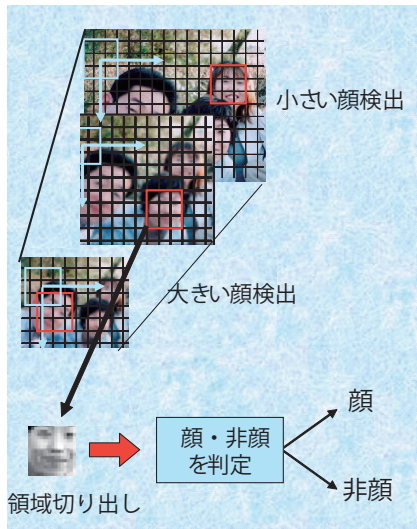


図-3 領域を切り出して顔かどうかを逐一判別することで顔を検出する方法

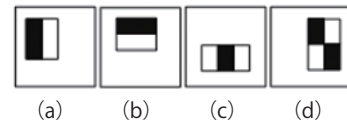


図-4 Viola と Jones が提案した 4 種類の Haar タイプの特徴量。隣接する白い矩形領域の平均輝度と黒い矩形領域の平均輝度の差で表す。

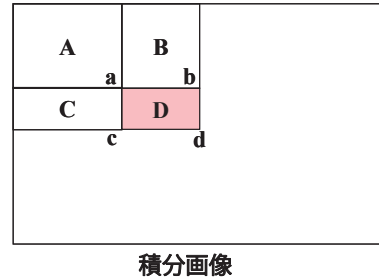


図-5 元画像中のある領域の輝度合計を矩形の大きさによらず 3 回の加減算だけで求めることができる。  
( $D=d-c-b+a$ )

る。従来の顔識別器は入力された画像がどんなものであってもすべて同じ処理を行って顔かどうかを判別する。彼らは画像の中の多くの部分は顔ではなく、ほとんどの領域は明らかに顔に似ていないことに注目し、これらの領域においてはより単純な計算で高速に顔ではないと判断できることを巧妙に利用して高速化を実現した。彼らの方法には 3 つ重要な貢献がある：

1. 積分画像を用いて、単純で解像度に関係なく高速演算できる Haar タイプの特徴量を導入した。
2. AdaBoost アルゴリズムを導入して上述した高速に計算できる特徴量を使った弱判別器を選定し、線形結合で強判別器を構築する方法を提案した。弱判別器は単純な閾値演算でバイナリの出力を決める。
3. 計算量の少ない判別器と計算量の多い判別器を直列に順番につなぐ構造の顔検出器を提案した。計算量の少ない識別器で非顔と判断された場合は途中で計算を打ち切ることで、画像の中で大部分の領域での計算を減らし、高速な顔検出を実現した。

ここで、Viola らが提案した特徴量について紹介する。学習手法などに関しては文献 1) を参照していただきたい。

Viola と Jones が提案した Haar タイプの特徴量は矩形の中の領域の平均輝度差である (図-4 における白い領域と黒い領域との平均輝度差)。Viola と Jones は 4 種類の Haar タイプの矩形特徴量を提案している。これらの特徴量の良いところは領域の大きさによらず、任意領域の平均輝度値は積分画像を使って 3 回の足し算、引き算と 1 回の割り算で高速に計算できることである (図-5)。ここで、積分画像の画素  $S(x, y)$  と元画像の画

素  $I(i, j)$  の関係は次のように定義される：

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

Viola らの手法は非常に高速で、画期的であったが、正面顔以外、横顔も対応する必要がある場合、性能的、速度的に不十分であった。実アプリケーションに応用すると、さらなる高速化や検出性能の向上とハードウェア化しやすくするためのメモリ使用量の削減などが要求される。

#### 実用化に向けた改良

実アプリケーションにおいて、以下のようなニーズがある：

- 写真印刷で逆光や白とびの顔の輝度を自動補正する場合、カメラの持ち方などによって、顔が上向きとは限らないため、回転した顔も検出する必要がある。
- 顔の向きは正面だけでなく、横顔の対応も必要。
- 組み込み機器に搭載するためにはプログラムの ROM と RAM の使用量を削減する必要がある。

これらのニーズに答えるために、さまざまな研究成果が発表されている。多く報告されているのは特徴量に対する改良である。たとえば Lienhart らが ICIP2002 で発表した Haar タイプの特徴量を斜めの領域に対応する試みや、Li らが ECCV2002 で発表した横顔検出のために設計された隣接の拘束条件を外したより自由に配置できる矩形特徴量がそれに当たる (図-6)。

Huang らのグループでは、Haar タイプの特徴量よりもより高速に計算でき、しかも識別能力が高い“SGF 特徴量” (Sparse Granular Feature) を提案している。

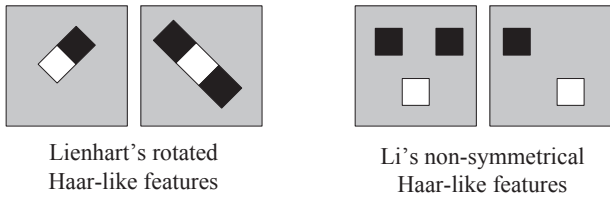


図-6 Haarタイプ特徴量への改良の試み

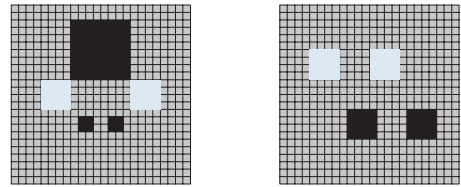


図-8 学習した結合粒子特徴量の例 黒は $\alpha = -1$ , 白は $\alpha = 1$ を示す

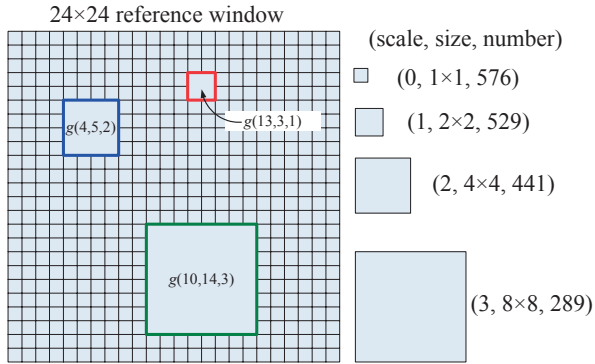


図-7 SGF 特徴量 (Sparse Granular Feature) の概念

$$F(\pi) = \sum_i \alpha_i g_i(\pi; x, y, s), \alpha_i \in \{-1, +1\}$$

ここで、 $\pi$  は入力画像の濃淡データ、 $g_i(\pi; x, y, s)$  は Granule “粒子” で  $x, y$  は位置、 $s$  はサイズを表すパラメータ (図-7, 図-8) である。 $g$  はその領域の平均輝度とし、元画像をスケール  $1/2^s$  に縮小した画像をあらかじめ生成しておけば、各粒子  $g$  の値は直接各  $1/2^s$  に縮小した画像の画素値から得られるため、計算量が少ない。

このような SGF 特徴量は速度的にも性能的にも優れていることが実証できている。また、計算が単純なため、ハードウェアでの実現がしやすいメリットがある。ほかにも学習手法や、識別器の構造などについての改良があるが、詳細は文献2)を参照いただきたい。

### □顔特徴点検出

次に、個人の特徴や属性情報を正確にとらえるためには、顔器官の特徴点を正確に抽出する必要がある。これは、後段の処理の性能に大きく影響するため、安定性と高い精度が要求される。たとえば、公開顔データベースに対して顔部品位置を手動入力したものが Ground Truth データとしていくつか公開されており、個人の識別に関して特徴点を自動検出したものと手動入力したものとを比較を行った場合、認識精度に数%の違いがあることが知られている。顔特徴点検出については、一般的な顔画像処理では、両目の位置を基準として特徴量を求める方法が多いことから、目領域検出の研究が中心となっている。しかしながら、顔の向きの変動の考慮や、よ

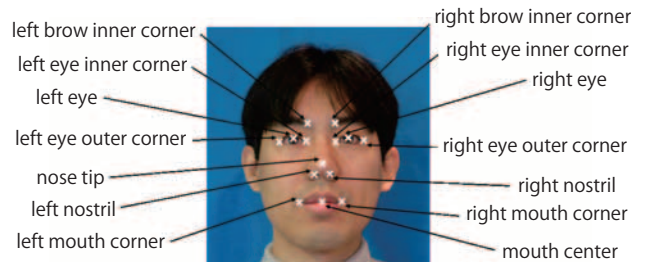


図-9 多数の顔特徴点の検出例

り個人性を反映させるためには、複数の特徴点を利用することが必要となる。

### 多種多数の特徴点検出

多数の特徴点を求める一例として、図-9に示した目、鼻、口などの各種の特徴点位置の検出を行う方法がある。これは、Fukui らが提案した円形分離度フィルタや Harris のコーナー検出を利用した特徴点検出法と部分空間法を利用したパターン照合による検証の組合せによって多数の特徴点を検出する。

さらに多種の特徴点を求めるために、顔全体の情報を使う有名な方法として、Cootes らの Active Shape Model (ASM) や Active Appearance Model (AAM) の方法がある (たとえば文献3))。特徴点の配置を表現するために座標系列に対して主成分分析を行い、固有値の大きい基底ベクトルを用いて、線形結合により形状を表現する (ASM)。同様に、見え方についても線形結合モデルを統合し (AAM)、特徴点の検出を実現している。この方法をベースとするいくつかの方法があるが、多数の特徴点が同時に見つかるもののモデルを顔にフィッティングさせる処理には、多数回の繰り返し計算が必要のため、リアルタイム処理は難しい。また、多数の学習データが必要であるため、学習データに含まれていない人物によってはフィッティングがうまくいかない。たとえば、図-10右のように幼児の顔や表情、顔の方向によっては収束しない場合が存在する。

そこで、正解位置からずれた際の特徴量に着目し、この特徴量から回帰計算によって正解位置への移動量を推



図-10 Active Shape Model での検出例

定し、ASMを拡張した方法をMIRU2008にて木下らが提案している。このようにずれた点の情報を用いた検出方法は、目領域の検出においてYuasaがMVA2002で導入しており、最近ではECCV2008でもLiangらが同様の考え方を利用して多数の特徴点検出を実現している。また、さらなる拡張により、任意の顔の部分領域からの相対的な位置関係を利用するアプローチとして、大域的な形状の制約をもたずに正解位置との相対的な方向情報を持つ拡張局所テンプレートを用いる方法がある。近似最近傍探索によって類似した領域を検出し、各局所テンプレートからの正解方向を集積することにより、各々の特徴点を求める方法をFGR2008にてKozakayaらが提案している。

## □顔認識

顔認識処理では、求めた顔特徴点位置を用いて、個人を特定するための個人特徴量を各人物ごとに生成する。

一般に顔認識はFeature-basedアプローチとAppearance-basedアプローチとに分けられる。Feature-basedの場合は、検出された特徴点間の配置の関係や特徴点付近のローカルな特徴記述子を作成する。一方、Appearance-basedアプローチでは、顔全体の輝度情報を主として取り扱う。検出された顔特徴点に基づいて、サイズ、向きなどを3次元的に正規化した顔パターンの切り出しを行う。これには、照明変動の影響を緩和するために、画像の変換や部分空間への射影計算などによる特徴抽出が行われる(図-11)。なお、個々の方式や文献についてはサーベイ<sup>7)</sup>も参考にされたい。

### 特徴抽出と識別

照明変化や画像変形に対しての頑健性を向上させるために、ローカルなフィルタリングにより特徴量を求めるものとしては、複数の解像度と方位を変化させたGabor-Wavelets Filterを畳み込んだ結果を特徴量とするものやLocal Binary Pattern (LBP) と呼ばれる周辺画素との相対的な大小の輝度差パターンを符号化するもの、またそれらを組み合わせたLocal Gabor Binary Patternなども顔認識に採用されている。

多くの認識法では、特徴量の次元削減や分離性能を向上させるために、PCA(主成分分析)、LDA(線

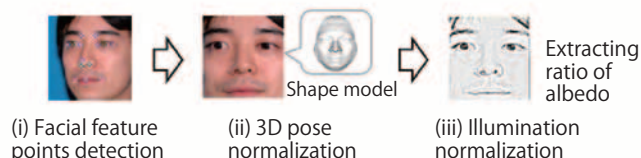


図-11 アピアランススペース手法における顔画像の正規化例

形判別分析)、ICA(独立成分分析)などの特徴変換を行い、それぞれの基底が張る空間内での距離を個人の特徴量の相違度として利用する方法がある。これらはEigenfaces/Fisherfacesなどとして有名であるが、局所的なパターン間の隣接関係の構造が取り入れられていない。そこで、元の空間で近いデータは変換先の空間でも近い距離に埋め込む方法であるLaplacian Eigenmapを利用して、コンパクトな記述を目指したLaplacianfacesがHeによって2005年のPAMIで提案されている。これは、上述の2つの方法と比較しても低次元でかつ良い識別性能を示している。

また、従来のPCAでは画像を1次元のベクトルとして扱うのに対し、画像の2次元行列として、原画像から直接画像共分散行列を作成し、その固有ベクトルを画像特徴量の抽出に利用する2DPCAが提案されている。PCAに比べて、画像特徴量の抽出にかかる演算が高速といった利点があるとされている。顔の場合はさらに変動要因が複雑に絡むため、光源方向や向きなどが変化した複数の画像、動画の系列が存在する。このようなモードの拡張を考えると、それらを直接操作するためにはテンソルとして扱い、N-mode SVDを利用して解析を行うTensorfacesもさらに高性能化を進めるものとして提案されている。

特徴量を抽出した後、個人の識別のためには、統計的パターン認識手法が適用される。空間内での最近傍点を求めるNearest Neighbor法、ベイズ推定を用いた判別、Support Vector Machineなどを用いる場合もある。ただし、登録・削除が頻繁に行われる可能性のあるアプリケーションでは、学習に時間がかかることや、メモリが大量に必要となることがあり、実用的ではない場合がある。一方、各個人ごとに部分空間を構成する部分空間ベースの識別手法もある。この手法は各個人の情報をコンパクトに表現でき、辞書の更新も効率的に行うことができるという利点があり、実用化もなされている<sup>4)</sup>。

近年の顔認識の発表論文では、複数の弱識別器の組合せによって高い認識率を得ようとするアンサンブル学習を顔認識に適用したものが多く見られる。特徴量をさまざまに変化させた識別器を組み合わせるものや、同じ特徴量でも学習サンプルや学習パラメータを変化させたも

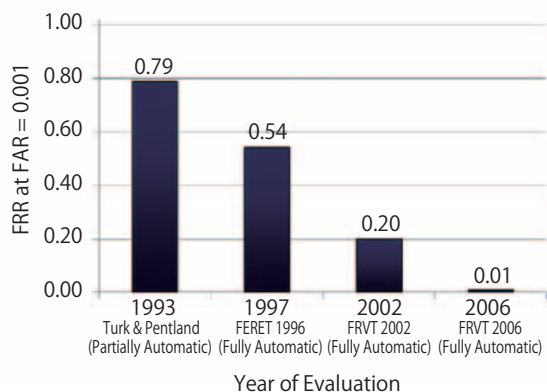


図-12 顔認識の精度の推移 (FERET, FRVT における当時の最先端アルゴリズムのエラー率. 文献4)より引用)

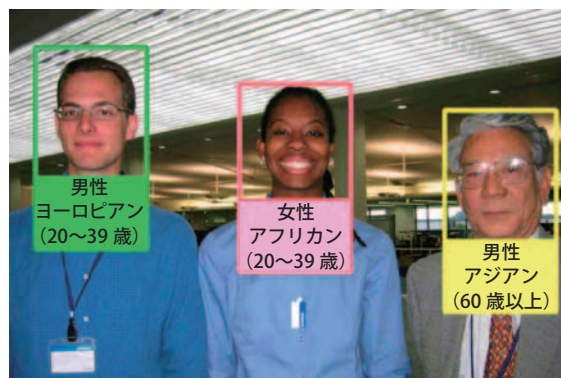


図-13 顔属性推定の例

の、顔部品からの相対位置の情報を組み合わせるものなどがあるが、いずれも従来の識別法に比べて高い認識率が得られている。しかし、部分的な隠れや強いノイズへの耐性など、さらなる課題へ対応するためには、従来の特徴量のアンサンブルによる組合せだけでは対応しきれず、新たな顔情報の表現法の開発が必要である。たとえば、新たなアプローチとして、L1 ノルム最小化を利用したスパースネスに着目した特徴表現などが採用されている<sup>6)</sup>。AdaBoostによる学習により特徴選択が行われたように、学習理論の応用でさらに新しい特徴抽出法などが出てくる可能性がある。

### 顔認識性能のベンチマーク

認識の性能ベンチマークについては、米国標準技術研究所 (NIST) が主導して行っているベンダーテスト、グランドチャレンジがある。これらは、1993年から1997年まで実施された FERET (The Facial Recognition Technology) プログラムに端を発する。DARPA (Defense Advanced Research Products Agency) のプロジェクトとして開始され、セキュリティ確保などのための自動顔認識技術を開発することが目的となっていた。その後 NIST に引き継がれ、数回のコンテストが行われている。2006年のベンダーテスト (FRVT2006) では、10カ国22の組織が参加した<sup>5)</sup>。高解像度の画像を用いた認識では、他人受入誤り率0.1%時の本人排除誤り率1%を達成しており、指紋認識と同程度の認識精度となってきた(図-12)。このテストには、国内からも東芝が相互部分空間法をベースとしたアプローチを用いて参加している。

2004年から2006年に行われた Face Recognition Grand Challenge (FRGC) では、テキサス大学の取り組みとしてコンピュータのアルゴリズムと人間の顔識別能力との比較も行われた。人間に対して同一人物もしくは他人の組合せの2枚の写真を見せ、その2枚の人物が同じであるかどうかの試行を繰り返して回答させる。同

様に複数のアルゴリズムに回答させ、それぞれのエラー率について解析した。その結果、3つのアルゴリズムが人間の識別能力を上回っており、撮影環境や状況を限定すると人間の能力を超えるとの報告がある。

しかし、これらのテストではいわゆる証明写真を撮影した画像に対する性能であり、監視カメラなどの画像から特定の写真の人物を見つけ出すといったタスクに関してはまだまだ性能が低い。2008年からは虹彩認識と組み合わせたマルチバイオメトリクスグランドチャレンジ (MBGC) が開催され、大学、企業、共に参加しており、さらに困難なタスクへの取り組みが見られる。

### 顔属性・表情推定技術

顔画像処理技術の中で、顔検出、顔認識のほかに、顔から性別、年齢などの属性を推定する技術、顔の向きや視線を測定する技術、顔の表情(喜怒哀楽、疲れを含む)を推定する技術も幅広い応用が期待でき、近年研究開発が活発になってきた。

#### 顔属性推定

顔属性推定は主に性別、年齢、人種推定に関する研究開発が報告されている。図-13に一例を示す。

顔属性推定の中で、古くから研究されているのは性別推定で、2クラスに分類するため、ニューラルネットワークの応用に適している。たとえば、Golombらは1990年に SexNet という2層のニューラルネットワークを提案し、30×30の画像サンプルで男女を推定する論文を発表している。その後、MoghaddamとYangが FG2000 で SVM を使った性別推定を行っている。彼らの研究では人でも判別しにくい8×6の極小サイズの画像でも、SVMがある程度の性能を出せることを発見した。ViolaとJonesがAdaBoostを使った顔検出の論文を発表以降はAdaBoostを使った性別推定の研究も増えた。

一般的に、顔属性推定システムは次のように構築で

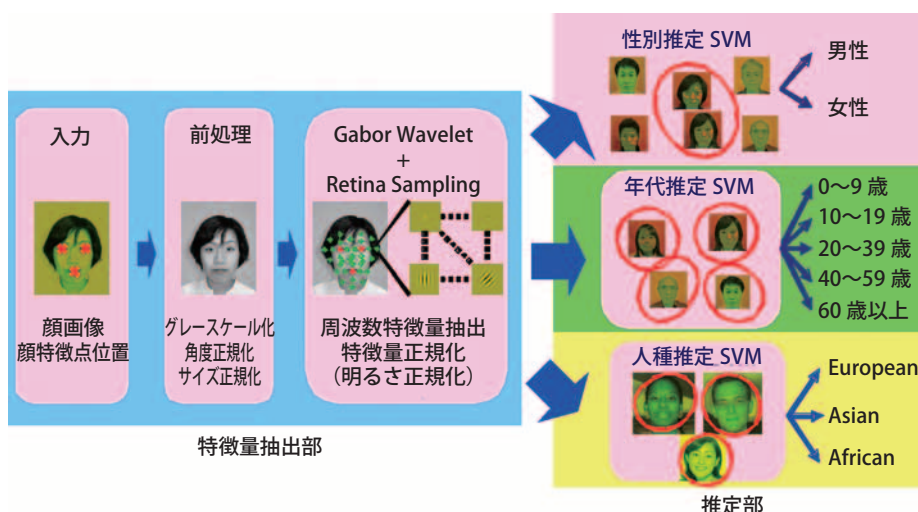


図-14 性別・年齢・人種推定のシステム例

きる：

1. サンプル画像の収集
2. 大きさを正規化した画像を作る
3. 特定場所で抽出する特徴量を定める
4. 特徴量抽出
5. サンプル画像で識別器を構築

図-14に性別、年齢、人種推定アルゴリズムの一事例を示す。このシステム構成では主に2つの部分に分けられる：(1) 特徴量抽出部、(2) 推定部。ここで、特徴量はGabor、識別器は性能が良く、使いやすいSVMを使っているが、実アプリケーションで、速度の要求が高い場合は、Haarタイプの計算量の少ない特徴量とBoostingで学習した識別器を使ってもよい。

### 表情推定

性別・年齢推定より人の内面的な状態を計るのは表情推定技術である。人間の表情は多種多様で、しかも微妙な変化が多いため、コンピュータにとっては最も苦手な対象かもしれない。表情推定の研究は長年多くの発表があるが、実用的になっているものは少ない。

ここで、表情推定の1つの特例として、笑顔推定技術を紹介する。さまざまな表情の中で、笑顔は最も重要な表情と言えよう。対象を笑顔に限定した理由は主に2つある。：(1) 笑顔の推定は他の表情推定よりも実用的な価値が高い。(2) 学習に十分なサンプルデータが取得しやすい。一例としてオムロンで開発した技術では、(1) 笑顔の度合いを数値で表せること、(2) 1枚の静止画で推定できること、(3) 個人別に学習する必要がなく、あらゆる人の笑顔度を推定できること、を目指して以下のように構築した：

1. 笑顔と非笑顔の学習サンプルデータを集める。
2. 目や口周辺領域の Haar タイプの特徴量を抽出。
3. 笑顔データと非笑顔のデータを使って識別器を学

習。識別器はFigueiredoが2003年PAMIで発表した手法を参考に約1万枚の画像を用いて学習を行った。

このシステムは1顔あたりの笑顔推定時間は約5msと速く、顔検出特徴点検出の計算時間を含めても約1フレームあたり40msほどしかかからない (Pentium4 3.2GHz PC)。

## 実応用における課題と対策

顔画像処理技術の実用化においては、言うまでもなく、性能を良くすることが重要であるが、そのほかにアプリケーションによって、ハードウェアの制約（主にメモリ使用量）と実行スピードを考慮しなければならないことが多い。

### □実応用における制約

パソコンをプラットフォームとするアプリケーションの場合、メモリの使用量はあまり問題にならない。計算速度も速いため、特に工夫しなくても実時間処理も可能な場合が多くなってきた。しかしながら、組み込み機器をプラットフォームとした場合、処理速度やメモリ使用量を気にしながら開発を行う必要がある。

たとえば、携帯電話の場合、ARM系のCPUを採用しているものが多い。多くのアプリケーションと共用するため、メモリの使用量を最小限に減らす必要がある。目安としてROMとRAMの使用量は数十から数百KBに押さえる必要がある。たとえば、オムロンで開発した顔認識モジュールのROMは500KB、RAMは300KBにて実装されている。

## □アルゴリズムの小型、高速化

組込み機器でも使えるようにするにはいくつかの点で工夫することが必要となる。たとえば、特徴量の次元数を減らすことで速度とメモリ使用量を減らすことができる。特徴量の計算を整数演算に置き換えることで、速度を上げることができる。

よく使う手法としては以下のようなものがある：

1. 必要な顔サイズの最小化
2. 特徴量のビット数の最小化
3. Course-to-Fine の探索で高速化
4. 画像分割，逐次処理による RAM 使用の削減
5. 整数演算化，固定小数化による速度向上
6. マシン依存の並列計算(SIMD)

## □ハードウェアによる実現

顔画像処理のハードウェアによる実装例として，顔検出の LSI がある。デジタルカメラで顔検出を行う場合，速度が非常に重要になる。そのため，リアルタイムでの検出を可能にするハードウェアでの実装が必要になる。いかに少ないゲート数で高速な検出器を実装できるかはハードウェア設計者の腕の見せ所と言える。設計において，重要なポイントは2つある：

1. 大量演算処理の並列化

データを分割して並列処理するか，識別の演算を分解して並列処理するかを最適に行う。

2. メモリアクセスの効率化

並列処理と同時に考慮する必要があるのはメモリアクセスの効率化。回路規模を減らすためには内部で使うメモリを最小限にすることが有効で，共有メモリを設けるのが良いが，そのアクセスのタイミングを上手に設計しないと他の処理のメモリアクセスとバッティングしてしまうことが起こり，計算速度の低下を招くことがある。

## 顔画像処理の今後の発展について

今後の顔認識処理についての，いくつかのトピックをまとめる。

- 顔認識の性能向上

実用的な環境での評価が必須であり，MBGC においても，正面顔から横顔，静止状態から自由歩行，といったものだけでなく，屋内から屋外へとといったさらに照明変動が大きな状況への対応が必要なタスクが評価対象となっている。また，大規模データによる評価に加えて，表情変化，加齢変化などの変動モデルの検討も重要である。

- 入力環境の変化

最近では，ハイビジョン (HD) 撮影も家庭用ビデオで行われ，高解像度画像が容易に入力される環境が整いつつある。今後は監視カメラなどもハイビジョン化する。また，移動する対象を確実に捉えるために，フレームレート，シャッタースピードについても考慮する必要があり，高フレームレートを実現する市販カメラも登場していることから，さらに安定した画像入力が可能となることが期待される。一方で，単位時間あたりのデータが増加するため処理のさらなる高速化の要求も増すこととなる。

- 表情，属性認識の高度化

笑顔検出など特定の表情に特化したものではなく，より多様な表情の認識も期待されている。コンピュータが人とのコミュニケーションの中で，より状況に即したきめ細かなシステムの対応やインタラクティブな処理を行う上では，高度な表情認識と属性認識が，人物の意図を理解するための重要な技術となる。

次回，顔画像処理を用いた各種のアプリケーションについて解説する。

### 参考文献

- 1) Viola, P. and Jones, M. : Rapid Object Detection using a Boosted Cascade of Simple Features, In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, USA (2001).
- 2) Huang, C., Ai, H., Li, Y. and Lao, S. : High Performance Rotation Invariant Multi-View Face Detection, IEEE PAMI, Vol.29, No.4, pp.671-686 (2007).
- 3) Edwards, G., Taylor, C. J. and Cootes, T. F. : Interpreting Face Images using Active Appearance Models, In IEEE Conf. on Automatic Face and Gesture Recognition 1998, pp.300-305, Japan, (1998).
- 4) 福井和広，山口 修：部分空間法の理論拡張と物体認識への応用，情報処理学会論文誌コンピュータビジョンとイメージメディア，Vol.46, No.SIG15 (CVIM 12), pp.21-34 (2005).
- 5) Phillips, P. J., Scruggs, W. T., O'Toole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L. and Sharpe, M. : FRVT 2006 and ICE 2006 Large-Scale Results, FRVT 2006 Report, NISTIR 7408 (2007).
- 6) Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. and Ma, Y. : Robust Face Recognition via Sparse Representation, IEEE PAMI Vol.31, No.2, pp.210-227, (2009).
- 7) 岩井儀雄，勞 世雄，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, Vol.2005, No.38, pp.343-368 (2005).

(平成 21 年 2 月 3 日受付)

勞 世雄 (正会員) lao\_shihong@omron.co.jp

1984 年中国浙江大学電気工学卒業，1992 年京都大学電気工学研究科博士後期課程退学，オムロン (株) 入社。現在同技術本部コアテクノロジセンターで顔画像処理の技術を統括する専門職。

山口 修 (正会員) osamu1.yamaguchi@toshiba.co.jp

1994 年岡山大学大学院工学研究科情報工学専攻修士課程修了，同年 (株) 東芝入社。現在，(株) 東芝研究開発センターマルチメディアラボラトリー主任研究員，コンピュータビジョン，顔画像認識の研究開発に従事。