

## オムロンにおける顔センシング技術の研究と応用

川出雅人

オムロン株式会社 センシング研究所

顔は非常に多くの情報を発信している。それを視覚で理解することにより、より良い機械と人の関係を築く事を目指して、視覚による顔センシング技術を研究開発している。顔検出・顔器官検出・顔照合・性別年代推定などの技術により、セキュリティ・エンタテインメント/コミュニケーション・インタフェースなど様々な応用が可能である。

### Vision-based Face Sensing Technologies and Applications in Omron Corporation

Masato Kawade

Sensing Technology Laboratory, OMRON Corporation

Human faces include rich useful information of the individual. The goal of Omron's vision-based face sensing technologies is to build a more comfortable man-machine relation by understanding such information of human faces using computer vision technologies. With the development of technologies such as face detection, face component extraction, face recognition and gender/age estimation, kinds of applications in various systems like security, entertainment, and communication are possible.

#### 1 はじめに

私達が直接人とコミュニケーションする時、視覚から得られる情報はコミュニケーションに大きな役割を果たしている。どこに何人いるか？、誰？、若い？、女性？、怒っている？、眠そう？、疲れている？、ハンサム？、何か話している？、など相手の顔を見て様々な情報を同時に推定している。そしてそれらの情報に応じて、例えば高齢者に対しては大きな声でやさしく話しかけ、外国人に対しては言語を変えて話すといった対応を意識的もしくは無意識的に行っている。「目は口ほどに物を言い」と言われるが、「顔は口ほどに物を言い」と言っても過言ではないと考える。これと同様に、機械が視覚によって人の顔をセンシング・理解することにより、従来の人が機械に合わせるインタフェースを革新し、機械が人に合わせるインタフェースが可能であると考えられる。機械が人を見守り、人の属性・適性・能力・状態・好みに応じて最適なサービス・インタフェース・価値・情報を提供していくことが可能である。このような機械が人に合わせるインタフェースを目指して、オムロンでは視覚による顔のセンシング・理解技術の研究開発を進めている。

## 2 顔センシング技術の要素技術と流れ

顔センシング技術の流れを図1に示す。静止画または動画から顔の位置と角度を検出する。次に、顔器官上の特徴点(目・口などの中心点や端点など)を検出し、顔照合を行ったり、性別・年代などの属性推定を行う。

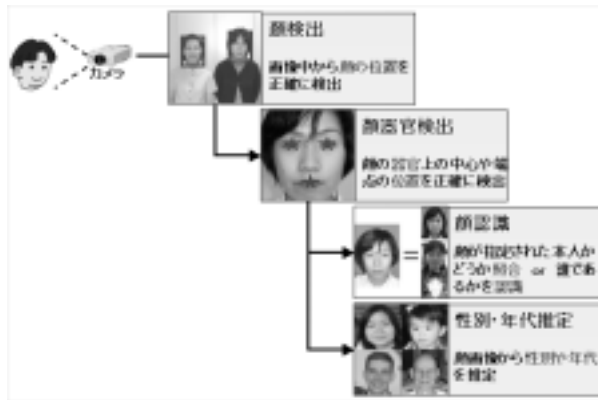


図1. 顔センシング技術の要素技術と流れ

## 3 静止画360°複数顔検出技術

顔画像の360度全方向の回転角に対応した複数の顔検出器を並列処理させることで検出性能と処理速度を両立させた。更に誤検出削減ネットワークを構築することにより、誤検出を大幅に削減した。また、検出速度向上の補助手段として、色情報を使った非顔領域マスク作成アルゴリズムを前処理に使った。これによって360度回転対応顔検出の主な課題に対処した。全体の構成を図2に示す。

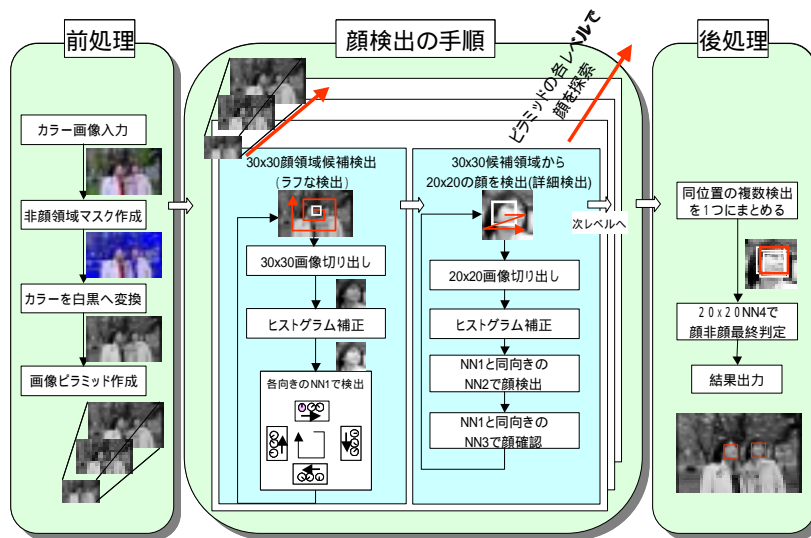


図2. 静止画360°複数顔検出技術の流れ

本システムは主に以下の4部分で構成されている：

(1) 前処理部：入力したカラー画像から非肌領域マスクの作成を行い、顔探索範囲を限定することにより、速度を向上した。また、カラー画像を白黒画像に変換し、1/1.2の比例で縮小した画像ピラミッドを作成することにより、20x20ピクセル～画像サイズ最大の顔のサイズまで探索可能にした。

(2) 360度対応顔検出器部：顔検出のメイン部分はそれぞれ30度の回転角に対応した顔検出器を12個用意することで構成される。各角度に対応した検出器はそれぞれの回転角を有する顔画像を使って学習させたニューラルネットワークで実現した。

(3) 誤検出除去部：複数のニューラルネットを使って、総合的に顔かどうかのチェックを行った。

(4) 後処理部：検出結果の重なりを確認し、同じ位置に複数回顔を検出した場合は、検出結果を一つにまとめる。

画像サイズが320x240動画において、検出する顔サイズを20x20～画像サイズに設定し、360度対応した場合でも10fpsの処理ができる。静止画においても同検出条件で画像サイズ320x240時の平均計算時間は0.1秒（Pentium4 3GHz PC）。検出性能は5843枚のポートレート写真の検出率において、約95%である。顔検出結果を図3に示す。



図3. 顔検出結果

## 5 顔特徴点検出技術

### 5.1 ガボールウェーブレット特徴

ウェーブレット変換を2次元の局所フィルタとして顔画像に応用した。ガボールウェーブレットのカーネルは式1のように表される。ここで  $\sigma$  はガウス関数の標準偏差である。また  $k$  を変化させることでガボールフィルタの周波数と方向成分を制御することができる。ガボールフィルタの例を図4に示す。また、顔画像からガボールフィルタによって特徴空間に展開した特徴画像を図5に示す。本フィルタは所望の方向、周波数に対するエッジの抽出能力と、方向、周波数の可制御性が利点である。

$$\psi_{k,\theta}(x,y) = \frac{k^2}{\sigma^4} \exp\left[-\frac{k^2(x^2+y^2)}{2\sigma^2}\right] \cdot \left\{ \exp[ik(x\cos\theta + y\sin\theta)] - \exp\left[-\frac{\sigma^2}{2}\right] \right\} \quad (1)$$

式1. ガボールウェーブレットのカーネル

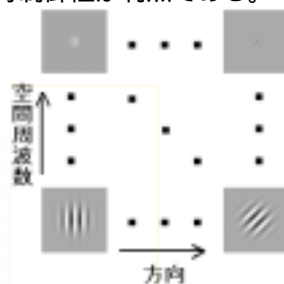


図4. ガボールフィルタの例

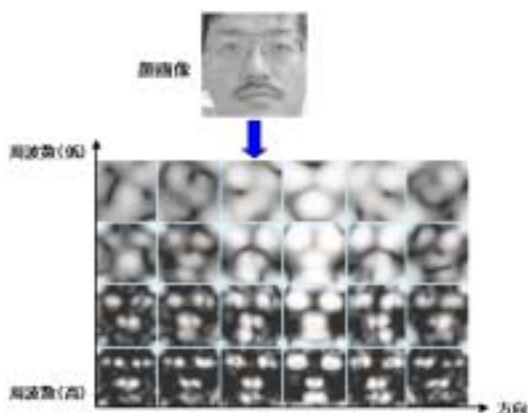


図5. ガボールフィルタの顔画像への適用例

### 5.2 グラフマッチング

次にガボールウェーブレット特徴をどのように顔器官特徴点抽出に適用するかについてグラフマッチングと呼ばれる手法について述べる。本手法では顔画像上にいくつかの特徴点を設定している。それらの特徴点は目や鼻や口などの顔器官の配置に基づいてお互いの位置関係を拘束し合っており、これをグラフと呼ぶことにする。図6に一例を示す。グラフ中の特徴点における特徴量は式1の  $k$  を変化させて周波数と方向成分を変化させたガボールフィルタカーネルと特徴点の周りの顔画像を畳み込むことで計算できる。つまり特徴量はその特徴点における顔器官の形状的な大きさや方向性を表現している。

このグラフの形状は顔特徴点の正確な位置抽出のために最適な形状が模索されている。特徴点抽出において課題となる顔画像のバリエーション変化に対応するための打ち手として、性別・年齢層・表情・眼鏡の有無など複数の標準的な顔画像から計算された複数の特徴量をグラフに登録したバンチグラフと呼ばれる構造を採用し、検出精度を向上させている。(図7)

以上のようにして作成したバンチグラフを顔特徴点の辞書として入力画像から抽出されたガボールウェーブレット特徴画像に対してサーチを行い、式2で表わされる類似度を用いて顔特徴点を求める。



図6. グラフの例

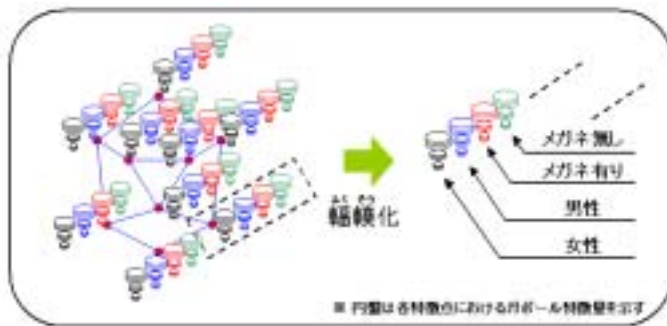


図7. バンチグラフの構造

$$S_{ij} = \frac{\sum_k a_{1k} a_{2k}}{\sqrt{\sum_k a_{1k}^2 \sum_k a_{2k}^2}} \quad (2)$$

式2. 類似度

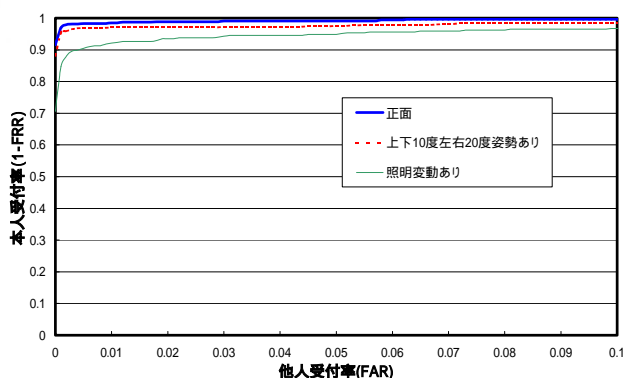


図8. 認証結果(ROCカーブ)

## 6 顔照合技術

抽出された特徴点位置において計算されるガボールウェーブレット特徴量を用いて辞書データとの比較照合を行う。登録したい顔画像からあらかじめ抽出された特徴量と入力された顔の特徴量に関して、それらの類似度が最も高い登録画像において、その値がしきい値を越えた場合に同一人物であると判断する。

120人の顔画像データベースによる認証結果を図8に示す。このグラフはROCカーブ(Receiver Operating Characteristic Curve)と呼ばれ、認証アルゴリズムの精度を表すもので横軸にFAR(他人排除エラー)、縦軸にFRR(本人受け入れエラー)をプロ

ットしたものである。正面顔の場合は FAR1% に対し FRR1% であり、上下 10 度左右 20 度の姿勢があっても FRR3% 以下と高い精度を維持しており実用精度に達していると言える。照明変動がある場合は若干の低下が見られるので、入力光学系パラメータや明るさの最適化を行うことが課題となる。

## 7 性別・年代推定技術

図 9 に性別と年代を推定する処理の概要を示す。入力された顔画像から推定に必要な特徴量をガボールウェーブレット変換+ Retina サンプルングで抽出し、その特徴量を非線形拡張したサポートベクタマシンによる識別器に与え、性別と年代を推定している。性別推定は、1240 人の DB で 90.9% の推定性能であった。年代推定は、幼年 (10 歳未満) が 98.4%、青少年 (10~19 歳) が 88.2%、青壮年 (20~59 歳) が 89.6%、高年 (60 歳以上) が 89.4% であった。処理速度は Pentium -1GHz の CPU において、1 枚あたり約 0.5 秒であった。

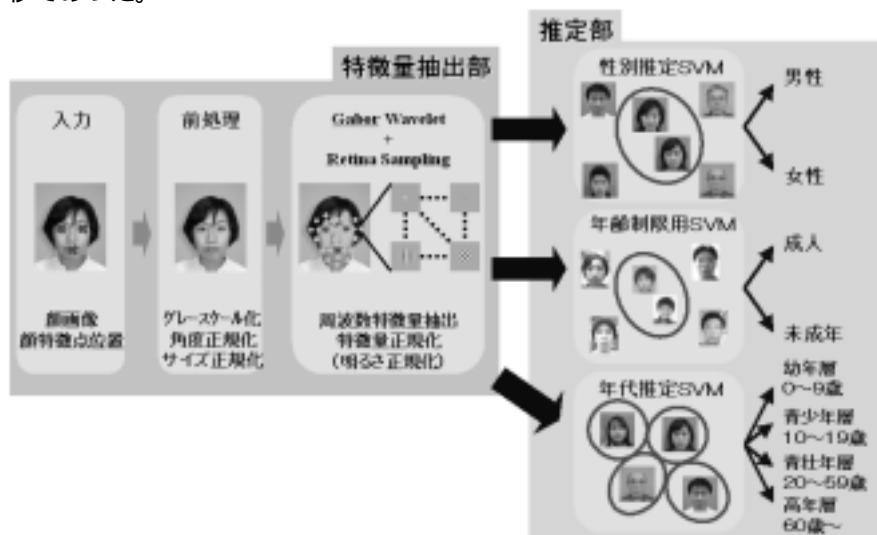


図 9. 性別・年代推定技術の処理概要

## 8 顔センシング技術の応用例

顔センシング技術の応用として、以下の 4 つの分野への適用が可能であると考える。

- 「顔で遊ぶ」エンタテインメント
- 「顔で話す」コミュニケーション
- 「非接触で手軽に個人認証」セキュリティ
- 「気の利いたインタフェース」ユニバーサルデザイン・ヒューマンインタフェース・自動マーケティング

### 8.1 エンタテインメント分野での応用例

似顔絵自動生成: 入力された顔画像より、顔の特徴を抽出し、平均からずれている特徴を誇張することにより、似顔絵の自動描画を実現した。

写真シール自販機: 顔を検出し、そこから肌色領域を推定することにより、

肌の荒れ・ソバカス・しみ・毛穴などを目立たなくするスムージング処理を行う自動美肌補正や、髪領域を自動抽出してその色を変更するヘアカラーシミュレーションや、顔の部位を認識したタブレットなどによるお化粧品ペンにおいて例えば口紅ペンが肌についても色がつかないユーザサポート機能などを実現した。

カメラ付携帯向け顔画像エンタテインメントサービス「おかもネット」：カメラ付携帯で撮影した顔画像を送ることにより、人相占い、有名人の中で誰に似ているか、その人の顔のタイプに合ったお化粧のアドバイス、顔上の自動的な落書き、顔の器官の特徴量で競い合うランキングなどを実現した。

#### 8.2 コミュニケーション分野での応用

バーチャル顔画像通信：顔の3軸の向き、目や口の開閉状態を符号化伝送し、3DのCGで再現することにより、少ない伝送量で実時間に近い顔画像伝送が可能になる。

#### 8.3 セキュリティ分野での応用

顔認証入退室管理システム：登録者であれば、顔を見せるだけで電子錠を開錠することが出来るシステム。

徘徊者保護支援システム：老人ホーム等の出入り口にカメラを設置し、出て行かれる方を顔で認識し、徘徊者の方が出て行かれるときは介護者の方に通報を行い、安全に付き添うことをサポートするシステム。

写真照合システム：消費者金融など運転免許証やパスポート等の顔写真と本人の照合が必要な時、顔照合を使って自動で本人照合を行う。

自動顔特徴分類・検索システム：登録された顔の特徴量（目の間隔、目の釣り上り具合、口の厚さ等）を分布から分類し、次に顔の特徴量の度合いを指定することによりその特徴を持つ顔を検索することが出来るシステム。

監視顔認証システム：出入り口などにおいて、監視カメラの動画画像で顔認証を行うことにより、犯罪者等特定人物の検出や、入場を許可されていない人物の検出のサポートを行う。

PCログインセキュリティシステム：PCにログイン時、login名・passwordのかわりに顔を見せるだけで簡単にログインできる。また離席時など他人がのぞいたら画面をロックさせることも出来る。

指紋認証とのマルチバイオメトリクスによるインターネットセキュリティ：インターネット上の個人データにアクセスする時、顔と指紋の両方をインターネット上のバイオメトリクス認証局で認証することにより、より高いセキュリティでアクセスできる。

#### 8.4 ユニバーサルデザイン、ヒューマンインタフェース、自動マーケティング分野での応用

年代別最適インタフェース：お年寄りには大きな文字や大きな音で、幼児にはひらがなで表示するような、その年代の能力などに最適化したインタフェース。

年齢判別インタフェース：ある特定の年齢以上でないと稼動しないようなインタフェース（お酒や外食の自販機、アダルトや暴力シーンのTVなど）。

性別・年代別情報提供インタフェース：性別や年代別に合った情報やサービスなどを自動的に表示するようなインタフェース。

自動顧客分析：顧客の年代・性別を自動推定することにより、どのような顧客がどのようなものを購入しているかなどのマーケティング情報を得る。



## 9 課題と今後の展開

顔センシング技術の研究開発を進め、上記のようなアプリケーションへの応用を進めてきた。このような応用を進めていくには、技術的に以下のような課題がある。

- ・耐環境性向上
  - 大きな顔の向きの変化への対応
  - 大きな照明の変化への対応
  - 大きな表情変化や口・目の開閉等への対応
  - 化粧、髭、眉・頬にかかる髪等の変化への対応
  - メガネに対するロバスト性向上等
- ・さらなる精度向上
- ・経年変化への対応
- ・小型化、組込み化、チップ化
- ・プライバシー保護対策
- ・他バイオメトリクスとの融合
- ・なりすまし防止のさらなる精度向上
- ・処理速度の向上、リアルタイム処理対応

以上の課題に取り組むとともに、図 10 のような要素技術の展開も考えている。



図 10. 要素技術の展開

## 参考文献

- [1] 勞ほか：360度回転対応高速顔検出システム，第9回画像センシングシンポジウム，pp.511-514，2003
- [2] 小鶴ほか：監視カメラ向け実時間顔検出・認識システムの開発，OMRON TECHNICS，pp.32-36，VOL.43，No.1，2003
- [3] 瀧川ほか：顔画像による自動性別・年代推定，OMRON TECHNICS，pp.37-41，VOL.43，No.1，2003
- [4] 川出：3.4 顔，“これでわかったバイオメトリクス” オム社雑誌局，pp.59-71，2001