

高速・高精度を両立する顔検出技術

“共起性”に着目して検出性能を向上

デジタルカメラなどで撮影された画像、あるいは監視カメラの映像に映っている人物の顔を自動検出する技術を開発しました。

検出精度の向上と高速化(省処理化)を両立する方式の開発を進めた結果、一般のパソコン(PC)に比べて計算能力が低い機器にも組み込む可能性が開けてきました。これにより、不審者を発見する監視システム、自動的に顔にフォーカスし露出を調整するデジタルカメラ、人の顔を見て会話してくれるロボットなど、多様な用途への展開が期待されます。

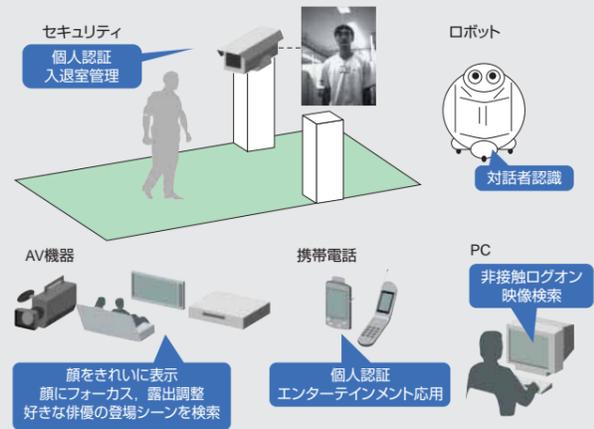


図1. 顔検出技術の用途 — セキュリティ機器だけでなく、AV機器や携帯電話など多様な用途への展開が期待されます。

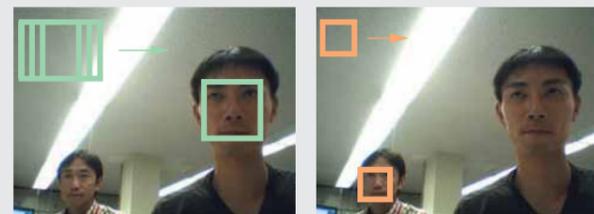


図2. ウィンドウ走査に基づく顔検出処理 — 大きいウィンドウを使って大きく写っている顔を検出し(左)、ウィンドウを縮小して小さい顔を検出します(右)。

顔画像解析技術の重要性

近年、テロや犯罪に対する社会不安の増大を背景に、セキュリティ確保のための新技術に対する関心が高まっています。中でも、人物の顔を対象とした画像解析技術が注目されています。

顔画像を扱う利点は、指紋などと異なり非接触で観測できるため、利用者に何らかの操作を要求しないで済むという利便性の高さ、装置を触る必要がないため清潔感があること、また、不正が起きたときには、画像データの履歴を後から追跡調査しやすいことなどがあります。顔検出は画像中の顔の位置と大きさを知るための技術で、“誰の顔か”を判定する顔認識を行う前に必要となる基盤技術です。

計算機によって自動的に顔を検出する技術には、解決すべき課題がいくつ

かあります。まず、顔には人種、性別、年齢の違いによる個人差があるうえ、撮影環境の照明条件やカメラの特性によって見た目が大きく変化します。このような様々な変動を吸収しつつ、安定して顔を検出できるようにする必要があります。また、デジタルカメラや携帯電話などPCに比べて計算能力が低い機器への組み込みに対するニーズが高いため、精度が良いだけでなく、計算量を小さくする必要があります。そこで東芝は、顔検出技術において高速化と高精度化の両立が鍵と考え、開発を進めています。これにより、セキュリティ分野だけでなく多様な用途への展開を目指しています(図1)。

顔検出の基本アルゴリズム

画像中の顔を検出するには、図2のように、画像をくまなくウィンドウ走査

します。各走査位置において、ウィンドウ内に顔が含まれるか否かを“識別器”により判定します。大きさの異なる顔を検出するには、ウィンドウを拡大又は縮小して再度走査します。

識別器は、注目しているウィンドウ内に顔があるかないかを判定します。判定のための識別ルールは、図3のように事前に収集しておいた多数の人物の顔画像と顔以外のサンプル画像から学習します。個人差や照明変動などの変動は吸収し、両者の違いを際立たせるような効率的で正確な識別ルールをいかに見いだすかがポイントとなります。

“共起性”に着目した識別器の学習

小さい計算量で正確に顔画像と顔以外の画像を見分ける識別器を作るために、顔の局所領域から特徴を計測

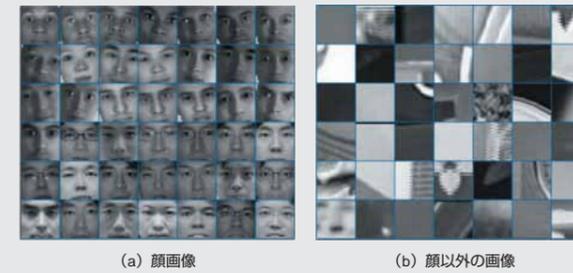


図3. 識別器の学習に用いるサンプル画像 — 顔画像と顔以外の画像の2クラスに分けておきます。

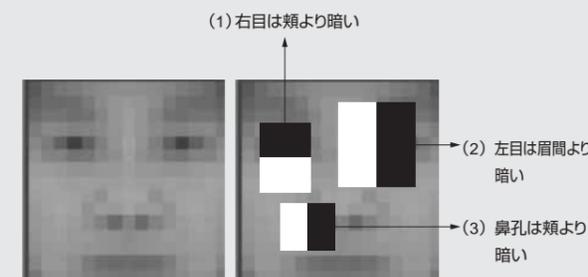


図4. 新方式の原理 — 顔の構造を反映した特徴(1)~(3)を同時に観測できるか(共起性)を評価します。

します。図4に特徴の具体例を示します。隣接する白い長方形と黒い長方形の組が三つあります。それぞれの組はHaar Wavelet 特徴と呼ばれ、白い部分と黒い部分ではどちらの領域が明るいか(暗いか)を調べます。明るさの大小関係しか見ないので、照明条件の変動や肌の色の違いに影響を受けにくい性質が得られます。

この Haar Wavelet 特徴を組み合わせることで、大幅な性能向上を達成しました。図4のように3組の特徴を配置すると、(1)右目は頬(ほお)より暗い、(2)左目は眉間(みけん)より暗い、(3)鼻孔は頬より暗い、という顔の構造的な特徴が備わっているかどうかを評価できます。このように、複数の性質が同時に満たされることを共起性といいます。顔以外の任意の画像を観察しても、図4のような共起性を満たす

画像はほとんど見つからないため、効率よく顔か否かを判定できるのです。

図4のような特徴の組合せや各長方形の位置や大きさは、Boosting という学習アルゴリズムを用いて決定します。具体的には、いろいろに位置や大きさを変えた Haar Wavelet 特徴を生成し、その特徴の様々な組合せの中で学習サンプルを識別する確率ももっとも高い組を選びます。ただし、選んだ組合せだけでは正しく識別できないサンプルも残るので、今度は識別できなかったサンプルをうまく見分けられるような別の特徴の組を選びます。これを何度も繰り返すことで、識別誤りの少ない強力な識別器ができます。

従来方式では共起性を利用せず、1個ずつ特徴を選択していました。図5は、従来方式と新方式の性能比較結果です。合計で同じ1,000個の特徴

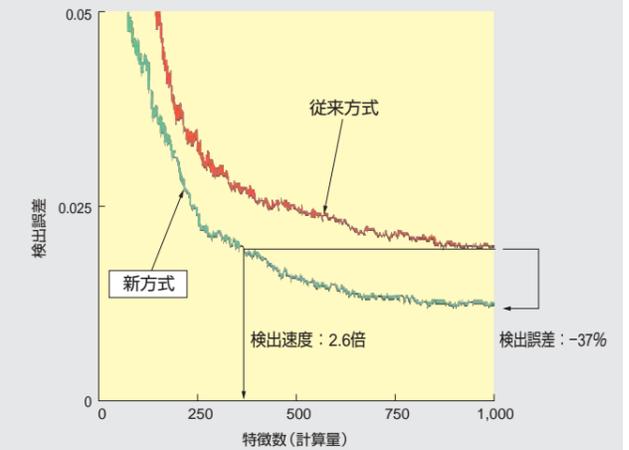


図5. 従来方式との性能比較 — 同じ計算量かけた場合、検出誤差は37%削減され、2.6倍の速度で従来と同等の精度が得られています。



図6. 顔検出結果の例 — 屋外で撮影された画像やサングラスをかけた画像でも正しく検出できています。

を使った場合(同じ計算量とした場合)、新方式は検出誤差が37%削減されました。また、従来方式で検出誤差が最小となる精度を、新方式では半分以下の特徴の数で達成できることがわかりました。この場合の検出速度は2.6倍になりました。図6は、顔検出結果の例です。屋外で撮影された画像など、難しい照明条件下でも正しく検出できるようになりました。

顔検出技術の将来展望

今後は、近づいてくる歩行者の顔を認証してドアを自動開閉させるシステムなど、様々な用途に今回開発した顔検出技術を適用していきます。

三田 雄志

研究開発センター
マルチメディアラボラトリー研究主務