

顔検出器のスイープ範囲限定による高速化とその評価

Speed up by Limiting Sweep Area for Face Detector

青木 崇浩†
Takahiro Aoki中島 克人†
Katsuto Nakajima

1 はじめに

人物の顔領域を高速かつ正確に検出する技術は、人物監視システムや表情認識システムなどの顔認識を応用する各種システムの基礎であり、活発に研究されてきた。

近年、ハードウェアの小型化・多機能化により顔検出器を応用したオートフォーカス機能を搭載したデジタルカメラやビデオカメラも普及してきた。しかし、専用ハードウェアなどに負担をかけない、より高速で高精度な顔検出器の追求は続いている。顔領域の検出技術は多数提案されているが、概ね、経験則に基づく手法[1]と統計学習に基づく手法の2つに分けられる。前者は色やエッジ等の情報を用いる手法で、後者は輝度分布を用いる手法でありViolaらが提案した手法[2]が有名である。

我々は、検出精度と高速性で定評のある[2]の手法をベースに、その検出精度を同程度以上に維持し、かつ、より高速化を目指すため、色特徴量を用いた検出手法の組み合わせ方法と前フレームでの検出結果を用いた検出位置推定方法の組み合わせ方法を評価している。今回の評価によって、色特徴量と前フレームの検出位置情報を用いたスイープ範囲制限手法は、一定の検出率を維持しながら、探索ウィンドウの増倍率1.1の時に於いて、既存手法よりもフレーム当たり約30msec速くなり、高速化効果が大きいことが示された。

2 探索範囲限定手法

2.1 概要

Violaらが提案したHaar-Like特徴とカスケード型識別器を用いた検出手法[2]は、シーン中の顔領域全てを検出するために検出対象画像全体を走査する必要がある。しかし、検出対象画像内で顔領域となる割合は通常限られているため、前フレームで顔検出が成功した領域を「顔が存在する可能性が高い領域」として、しばらくはその領域のみを探索することにより処理時間削減と誤検出率の低下を実現する。また、顔の色特徴量を用いて顔検出器の走査範囲を限定する事も考えられる。今回はそれらの手法の高速化などへの効果を評価した。

2.2 色特徴量を用いた手法

Violaらの検出手法ではイメージの濃淡を特徴として用いているため、顔ではないものを誤検出する可能性が高い。そこで本手法では、誤検出の低減と処理時間の短縮を狙い、色特徴量を用いた前処理によって顔検出器の走査範囲を限定する手法を採用した。色特徴量を用いた走査範囲限定手法の処理の流れを図1に示す。

顔検出対象画像内の各画素を肌色らしいかそうでないかの2種類に分類する。その分類結果の二値画像の積分画像を生成し、顔検出対象画像内を走査するサブウィンドウ内に存在する肌色らしい画素の割合を積分画像から求める。Violaらのカスケード型識別器は1つのサブウィンドウにおける非顔検査は非常に高速とされるが、肌色らしい画素数の検査は、サブウィンドウの4隅に対応する積分画像上の4点を読み出し、加減算を行うことで求められるため、より高速である。

本稿では、肌色らしいという検査判断は、HSV表色系で色相Hの値の0~40の範囲[3]とした。色相Hの肌色らしい範囲を図2に示す。また、髪やノイズ等によって顔の一部が隠れるため、肌色領域から顔の輪郭位置を正確には予想できないため、事前実験の結果に基づき、その集合の外接矩形の1.2倍の大きさを持つ矩形を探索範囲とする。

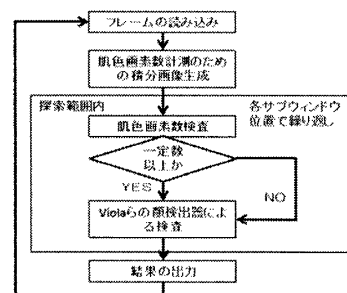


図1. 色特徴量を用いた処理の流れ

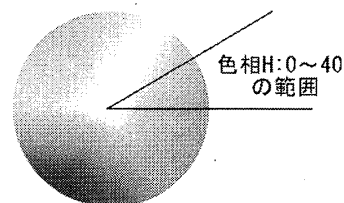


図2. 色相Hの肌色範囲

2.3 前フレームの検出位置情報を用いた手法

顔が1フレーム間に移動する距離は、カメラ内の人の移動速度等に依存する。例えば人の移動速度を時速10kmとした場合、これは1/30秒間に9cmの移動となる。30fpsで撮影中の静止カメラの映像の中で、人がカメラの撮影方向に対し直角方向に移動する場合に移動量が最も大きくなり、それは人の顔幅(約20cm)の0.45倍となる。そこで、左右上下の移動を考慮し、前フレームで検出された顔領域の1.9倍の高さと幅を持つ範囲を1フレーム間で移動可能な範囲とし、次フレームにおける顔検出器の探索範囲として限定する。複数の顔が検出された場合は、それぞれに同様の探索範囲を設定する。ただし、新たにカメラに映る人が増加した場合を想定し、一定期間に一度

†東京電機大学大学院未来科学研究科
Graduate School of Science and Technology for Future Life,
Tokyo Denki University

(今回は10フレーム毎), 検出対象画像全体を探索することで, 検出漏らしを防ぐ。処理の流れを図3に示す。

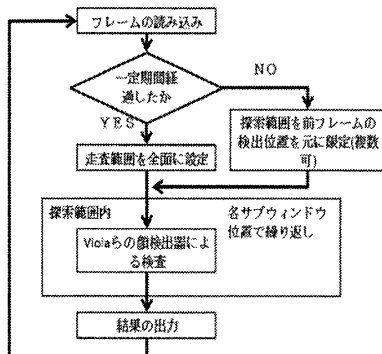


図3. 前フレームの検出位置情報を用いた処理の流れ

2.4 組み合わせ手法

色特徴量と前フレームの検出位置情報を用いる手法のどちらが有効かはシーンに影響される。色特徴量を用いた手法では, 背景に肌色らしいと判断される画素が多く分布する場合は, 探索する領域を余り限定できず, むしろ, 色特徴による検査がオーバーヘッドとなって速度低下を招く可能性が高まる。また, 前フレームの検出位置情報を用いる手法では, 実際に顔の動く方向に関わらず, 全ての方向に探索範囲を広げてしまうため, 無駄な走査が含まれる。これらの問題を改善し, より安定した処理時間と検出精度を実現するために, 両手法の組み合わせ手法を試した。組み合わせ手法の処理の流れを図4に示す。

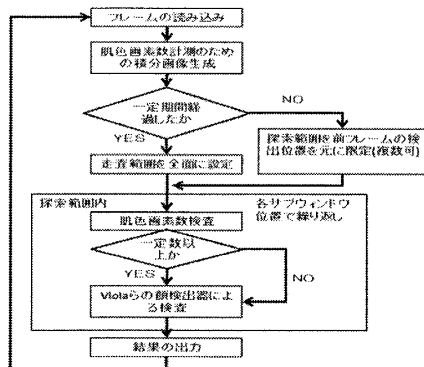


図4. 組み合わせ手法の処理の流れ

3 実験と評価

3.1 実験概要

OpenCVのサンプルに付属されている正面顔識別器 haarcascade_frontalface_default.xmlを用い, QVGAサイズの動画像(解像度320×240pixel, 30fps)を対象に正面顔検出の実験を行った。探索ウィンドウのスケール変化率は1.1倍として, 検出率と誤検出率と平均処理時間の評価を行った。また, 比較のためにViolaらの検出法のみを用いた顔検出の評価も同様に行った。計測には動作周波数3.0GHz, メモリ4.0GBのPC(OSはWindows 7)を用いた。

3.2 評価

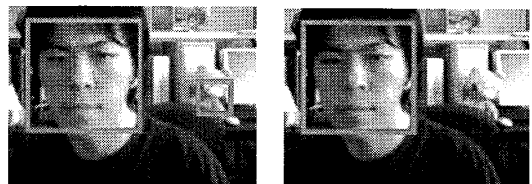
前節の環境の下, Violaらの検出法のみを用いた場合と各探索範囲限定手法の結果を表1に示す。

表1. 比較結果

手法	検出成功率	誤検出率	平均処理時間
Violaらの手法	94.61%	10.60%	57.3msec
色特徴量を用いた手法	93.27%	1.78%	25.1msec
前フレームの検出結果を用いた手法	91.81%	2.41%	25.7msec
組み合わせ手法	90.67%	1.78%	22.1msec

なお, 正面顔が画面内に存在している間に顔を検出した回数をフレーム数で割ったものを検出成功率とし, 顔領域以外を検出した場合や顔領域の一部のみを検出した回数をフレーム数で割ったものを誤検出率とした。

実験の結果, 探索範囲を限定するなどの手法でも実時間での処理を達成しており, 誤検出率も低下も確認できた。これは, 顔領域以外を走査する頻度が低下したためである(図5参照)。色特徴量を用いた限定手法では, 実験に用いた動画像によって処理時間が大きく変わり, 前フレームの検出結果を用いた限定手法では, 動画像の違いによる処理時間の差は小さかった。また, 実験に用いた全ての動画像において, 前フレームの結果を用いた限定手法よりも, 組み合わせ手法の方が高速で, 誤検出も少なかった。ただし, 前フレームの結果を用いた限定手法や組み合わせ手法においては, カメラに映る人が増加する場合において, 映り始めてから一定フレーム数の間はそれを検出できないため, 検出成功率は3~4ポイント低くなっている。色特徴量による手法での検出成功率の低下は, 肌の露出が少ない顔に対して起こりえるが, 最小限に留まっている。



(a) Violaの検出器のみ (b) 組み合わせ手法
図5. 誤検出を防いだ例(右のぬいぐるみ部分)

4 おわりに

本稿では, 色特徴や前フレームでの検出結果に基づいて探索範囲を限定し, 一定の検出率を維持しながらも高速化する手法を比較評価し, 速度向上の効果を定量的に示すと共に誤検出率の低減効果も示した。

今後の課題として, 検出成功率の低下を最小限にすること, また, 顔の部分隠れに対する適切な探索範囲指定などがある。

参考文献

- [1] 永田, 川口, "複雑な背景をもつカラー画像からの顔検出", 信学技報, PRMU, 99(448), pp. 77-82, 1999.
- [2] P.Viola, M.Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", Proc. of CVPR, pp. 511-518, 2001.
- [3] J.Sherrah, S.Gong, "Skin Colour Analysis", University of Edinburgh, 2001.