

モーション特徴を用いた顔認証システムのための 顔位置補正法に関する一検討

A Face Position Compensation Method for a Face Authentication System by Motion Characteristics.

佐々木 大輔[†]
Daisuke Sasaki

梶川 嘉延[‡]
Yoshinobu Kajikawa

1 はじめに

近年では情報漏洩事件や不法侵入による窃盗事件の被害件数が増加しており、巷では身体的特徴あるいは行動的特徴を利用した生体認証技術の研究が盛んに行われている。その中でも顔認証は完全非接触であるため、認証対象者の心理的抵抗が少なく、容易に行えるメリットがある。しかし従来の顔認証技術の多くでは、静止画像より抽出した特徴のみを用いて個人認証を行っているため、認証精度の限界が懸念される。また認証対象が生体から直接提示されているか否かを確認するため、温度センサ [1] や複数のカメラ [2] を利用している認証システムもあり、利便性やコストの面でマイナス要因となる。そこで本研究では従来の認証技術に人間が言葉を発する時の表情筋の動きを特徴に加えることで、生体検知及び認証精度の向上が期待できると考え、顔のモーション特徴を用いた顔認証システムの構築を目指す。しかしモーション特徴を得るためには、フレームごとに正確な顔画像を検出する必要があり、OpenCV の顔検出機能 [3] では顔の傾きに脆弱であるため不十分といえる。従って本稿では、肌色領域の分布及び目の位置情報を用いて顔の傾き角度及び位置を補正する方法を提案し、この方法を用いて顔検出を行い、その有効性を検証する。

2 OpenCV の顔検出機能

OpenCV の顔検出機能による処理結果の例を図 1 に示す。OpenCV の顔検出機能は、図 1(a) 及び (b) に示すように正面正立顔に対して精度は高いが、傾いた正面顔に対して精度は著しく低下し、また同図 (c) に示すように背景を誤って検出する問題を抱えている。従ってこれらの問題点を改善しなければ、個人認証精度に影響を及ぼすことになる。

3 提案手法

ここでは顔の傾きに頑健な顔検出方法を提案する。提案手法の概要を図 2 に示す。提案手法では、肌色判定及び顔判定の二方法を用いて入力フレーム画像内に顔が存在するか否かを判定する。そして顔を検出したならば、次に目の検出を行い、両目の重心座標をそれぞれ求める。そして最後に、この重心座標を用いて正規化顔画像を抽出する。また顔検出が成功した場合、現フレームの傾き角度を用いて次のフレーム画像を補正する。各判定及び目検出の方法について以下の小節で詳しく述べる。

3.1 肌色判定

ここでは入力したフレーム画像内に肌色領域が存在するか否かを判定する方法について述べる。肌色判定では修正 HSV 変換及び YCbCr 変換を用いる。これらの変換式を以下にそれぞれ示す。

$$\begin{aligned} H &= \arccos \left\{ \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} \right\} \\ S &= \sqrt{\left(\frac{(R-G)+(R-B)}{2}\right)^2 + 3\left(\frac{G-B}{2}\right)^2} \\ V &= \frac{R+G+B}{3} \end{aligned} \quad (1)$$

[†] 関西大学大学院工学研究科

[‡] 関西大学システム理工学部



図 1 OpenCV の顔検出機能

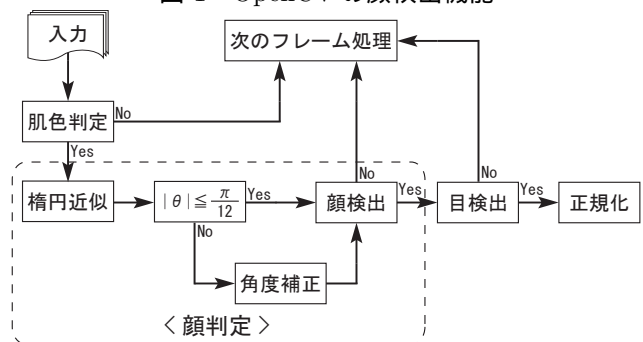


図 2 提案手法の流れ

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ Cb &= -0.169R - 0.331G + 0.500B \\ Cr &= 0.500R - 0.419G - 0.081B \end{aligned} \quad (2)$$

入力画像内に顔が存在する場合、修正彩度 S は顔の肌色領域に対して高くなるが、照明の影響により影となった肌色領域に対しては低くなるため、首領域を鮮明に抽出できない場合が生じる。一方、色差 Cr は赤成分を多く含む領域に対して高くなるため、肌色領域を抽出するのに適しているといえる。次にこれらの特徴を利用した肌色判定の手順を以下に示す。

- i 入力画像に対して修正 HSV 変換及び YCbCr 変換を行い、修正彩度 S と色差 Cr に対して大津の方法を用いて二値化を行う。なお、低彩度領域が抽出されることを避けるため、経験に基づき S の閾値の下限を 0.04 とした。
- ii S 及び Cr の 1-画素 (濃淡度 255) の総数 N_S, N_{Cr} をそれぞれ求め、両者の比 N_S/N_{Cr} を計算する。 N_S/N_{Cr} は、入力画像内に顔が存在する場合 1 に、一方、入力画像内に顔が存在しない場合 0 に近づくことが期待される。
- iii その比が任意の閾値以上ならば顔判定へ、閾値以下ならば次のフレーム処理へ進む。なお、本稿では経験に基づき、ここでの閾値を 0.4 とした。

3.2 顔判定

ここでは 3.1 で得られた色差 Cr の二値画像を用いて、入力画像内に顔が存在するか否かを判定する方法について述べる。まず顔判定の手順を以下に示す。

- i Cr の二値画像に対して穴埋め処理及びラベリングを行い、面積最大となるラベルを顔領域とする。
- ii i で求めた顔領域の輪郭に対して楕円近似を行い、求めた楕円の短軸と水平軸との成す角 θ を求める。
- iii 条件式 $|\theta| < \frac{\pi}{12}$ を満たすならば iii へ、満たさないならば入力画像に対して角度補正を行った後 iii へ進む。

iv OpenCVの顔検出機能を用いて、顔を検出できたならば目検出へ、顔を検出できなかったならば次のフレーム処理へ進む。

なお、iiiにおける条件式は3.3で述べるSSRフィルタサイズに基づいて算出した。次に二つの判定例を図3に示す。これより楕円近似による角度補正が有効であることが分かる。しかしこの方法は肌色領域に依存するため、髪型や装飾品等の影響により推定精度が低下することが懸念される。

3.3 目検出

ここではSSR(Six Segmented Rectangular)フィルタ [4]を用いた目検出方法について述べる。SSRフィルタを図4に示し、フィルタサイズは経験に基づき以下のように定義する。

$$\begin{aligned} w_1 + w_2 + w_3 &= 0.8L \\ w_1 : w_2 : w_3 &= 2 : 1 : 2 \\ h_1 + h_2 &= 0.45L \\ h_1 : h_2 &= 1 : 2 \end{aligned} \quad (3)$$

但し、 L はOpenCVにより出力される正方形の顔領域の一边の長さである。OpenCVにより検出した顔領域内に対して6セグメントに分割された矩形窓を走査し、各分割セグメント S_i 内の平均濃度 \bar{S}_i を算出する。この過程においてIntegral Imageを用いることで計算コストを大幅に削減できる。次にこの矩形窓の S_1 及び S_3 内に両目が入るように条件式を設定する。黒は濃度0であるので、次の条件式が成立することが期待できる。

$$\begin{aligned} \bar{S}_1 < \bar{S}_2 \quad \text{and} \quad \bar{S}_1 < \bar{S}_4 \\ \bar{S}_3 < \bar{S}_2 \quad \text{and} \quad \bar{S}_3 < \bar{S}_6 \end{aligned} \quad (4)$$

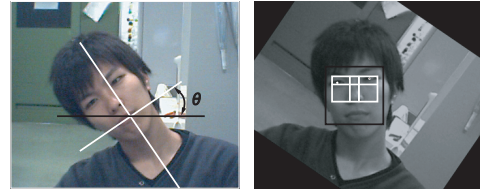
しかしこの条件式だけでは複数の候補点が生じるため、候補点を絞るため次式に示す評価関数を導入する。

$$G = 2 \times (\bar{S}_1 + \bar{S}_3 - \bar{S}_2) - \bar{S}_4 - \bar{S}_6 \quad (5)$$

S_1 及び S_3 内に両目が入る場合、 \bar{S}_1, \bar{S}_3 は0に近づき、隣接する S_2, S_4, S_6 は肌色領域となるため、 $\bar{S}_2, \bar{S}_4, \bar{S}_6$ は顔領域において最大となることが期待される。従って複数の候補点が生じた場合、関数 G の値が最小となる点を最適点とする。次に最適点における S_1 及び S_3 の矩形領域内の画素に対して、それぞれ二値化及びラベリングを行う。オブジェクト数が一つの場合には正確に目を検出できていることが期待される。一方、オブジェクト数が二つ以上の場合には各ラベルの重心座標及び縦連結画素数に基づいてノイズ除去を行い、目を検出する。

4 実験結果

ここではサンプル動画像に対して、3で述べた提案手法を用いて顔検出を行い、判定及び検出処理の精度について検討を行った。実験条件を表1に、提案手法の判定率及び検出率を表2に示す。なお、表2に示すサンプルA,B,Cには異なる被験者の動画像であり、Dは背景のみの動画像である。また肌色判定率及び顔判定率はフレーム画像が正しく判定されているか否かを示す。従って背景フレームが肌色領域あるいは顔を含むと判定された場合、誤判定されたことになり、判定率は低下する。表2より、全サンプルの各処理において9割以上の精度が得られ、特にA,B,Cのサンプルでは肌色判定が最も精度は高く、逆に目検出精度は低くなっていることが分かる。これには二つの要因が考えられる。一つ目は、顔判定を通過した入力画像にのみ目検出を行うため、目検出精度がOpenCVの顔検出精度に依存することである。二つ目は、前フレームで顔検出に失敗した場合、現フレーム処理では前フレーム情報を利用できないため、楕円近似による角度補正の精度に依存することである。従って楕円近似により求めた傾き角度に含まれる誤差量が、SSRフィルタにおける顔の傾き角度の検出範囲を超えた場合に条件式(4)を満たす目の候補点が存在せず、目検出精度が低下したと考えられる。次にDのサンプルでは肌色判定の精度が低下していることが分かる。これは照明の影響により物体が肌色に変色するためである。しかし、このような場合も顔判定で誤りを検出できるため問題ないと考えられる。



(a) $|\theta| \geq \frac{\pi}{12}$



(b) $|\theta| < \frac{\pi}{12}$

図3 顔判定の例

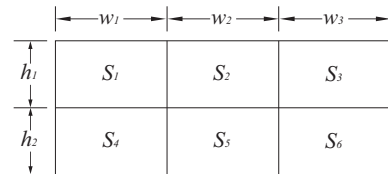


図4 SSRフィルタ

表1 実験条件

Num. of Sample	4
Frame Rate	30[fps]
Frame Size	352 × 288[pixel]
Image Type	True Color
Video Compression	None

表2 実験結果

サンプル	フレーム数	肌色判定率	顔判定率	目検出率
A(被験者1)	77	100 %	100 %	100 %
B(被験者2)	202	100 %	96.1 %	96.1 %
C(被験者3)	311	100 %	98.7 %	94.5 %
D(背景のみ)	142	93.7 %	100 %	

5 まとめ

OpenCVの顔検出機能は傾いている正面顔に対して検出精度が著しく低下する。本稿では肌色領域の楕円近似と目の位置情報を用いた顔位置補正法を提案し、サンプル動画像に対して提案手法を用いて顔検出を行い、判定及び検出処理の精度について検討を行った。その結果、肌色判定率は非常に高く、目検出率は低くなった。これには二つの要因が考えられる。一つ目は目検出率がOpenCVの顔検出機能の精度に依存することである。二つ目は前フレームで顔検出に失敗した場合、楕円近似による角度補正の精度に依存することである。しかし全サンプルの各処理において9割以上の精度が得られているため、提案手法は有効であるといえる。今後は正規化顔画像内より変動部位におけるモーション特徴を抽出し、その有効性について検証を行っていく予定である。

参考文献

- [1] 安達和隆, 亀山博史, 河田耕三, “顔認証装置および顔認証方法”, 特開 2007-11456
- [2] 早坂昭裕, 柴原琢磨, 伊藤康一, “顔の部分構造を用いた3次元顔認証の検討”, 第22回信号処理シンポジウム, No. P4-4, pp. 741-746, Nov 2007.
- [3] <http://opencv.jp/>
- [4] 川戸慎二郎, 千田康隆, 鉄谷信二, “SSRフィルタを用いた眉間の検出”, PRMU2002-207, Vol. 102, No. 651, pp. 41-46, Feb 2003.