

表情分析のための適応型顔部品形状検出手法の検討

成富 洋平[†] 坂本 博康[‡] 廉田 浩[‡]

[†]九州大学大学院 芸術工学府 [‡]同 芸術工学研究院 〒815-8540 福岡市南区塩原 4-9-1

E-mail: [†]naritomi@gsd.design.kyushu-u.ac.jp, [‡]{ sakamoto, kadota }@design.kyushu-u.ac.jp

あらまし 各種の刺激に対する人の表情変化分析には、顔位置検出後に目や口などの顔部品形状を計測する必要がある、そのためのよりロバストな輪郭形状検出手法を提案する。従来からの輪郭検出法である、「空間フィルタによるエッジ検出処理」や「色相分布による領域分割処理」に対して、ここでは各顔部品の内、目と唇に特化した条件で両処理を選択的に融合したエネルギー関数を構成し、輪郭線を近似する外接多角形のエネルギーを最小化して形状の最適化を行う。また色相による領域分割では多少の照明環境変化にも対応できるように閾値設定を適応型にする。これらの結果、高性能な顔部品形状検出が可能になった。

キーワード 顔部品検出、色相による領域分割、エッジ検出、輪郭エネルギー関数

A Study on the Adaptive Shape-Extraction Method of Facial Parts for the Facial-Expression Analysis

Youhei NARITOMI[†], Hiroyasu SAKAMOTO[‡] and Hiroshi KADOTA[‡]

{ [†] Graduate School, [‡] Faculty } of Design, Kyushu University, Minami-ku, Fukuoka, 815-8540 Japan

E-mail: [†]naritomi@gsd.design.kyushu-u.ac.jp, [‡]{ sakamoto, kadota }@design.kyushu-u.ac.jp

Abstract For the analysis of human facial-expression change to various stimuli, it is necessary to measure the shapes of facial parts, such as eyes or mouth. For this purpose, a new robust shape-extraction method is proposed. We are focusing on the eyes and the lip, and have composed dedicated energy functions to estimate their shapes, using the conventional two methods: "edge-detection filtering" and "segmentation by color," selectively. By minimizing the circumscribed polygon energy, we can realize an optimum shape-fitting procedure. In the sub-procedure of segmentation by color, an adaptive threshold-generation technique is used to correspond with different illumination conditions. The experimental results show its better performance.

Keywords Facial-Parts Detection, Segmentation by Color, Edge Detection, Contour Energy Function

1. まえがき

画像中の人の顔認識技術は、セキュリティ、ネットワークを介してのマルチコミュニケーション支援、知的なマシンインターフェース、或いはロボットによる人の識別など、広義のマシンビジョン応用分野で非常に重要である。この内、顔の位置検出に関しては近年多くの研究がなされ、或る程度環境条件が整っていれば、良好な認識率で処理が実行できることが確認されていて、幾つかの応用では実用化されている[1],[2]。更に、顔検

出を含む知的な画像処理プログラムは、オープンソースのライブラリとして公開もされている[3]。

これらの検出アルゴリズムは画像に対して各種特徴量を計算し、多くの学習サンプルを使って事前に各特徴量に対して誤検出の確率が最小になる判別閾値を決め、その閾値の組み合わせで検査サンプルに対して検出処理を実行するものである。

一方、顔位置検出に続き表情分析などを行うための顔部品識別では、比較的にまかな個別の顔部品の形状・輪郭を抽出する必要があり、上記のような確率・統計的なアプローチよりは、アビアラ

ンスペース（特徴点の位置を検出するなど）の処理が重要となる。従来この処理に関しては Wavelet 変換など空間フィルタでエッジ検出し直接輪郭線を抽出する手法や色相による領域分割手法などが行われているが[4][5][6][7]、環境変化に対する安定性が必ずしも十分ではなかった。

今回我々は、顔のカラー画像中から顔部品の形状を抽出する手法として、空間フィルタによるエッジ検出と色相での領域分割とを各部品に応じて選択的に併用し、輪郭検出精度を向上させる方法を提案する。また、この色相での領域分割処理では、照明条件変化による誤認識を少なくするため、画像に合わせて色相分類の閾値を変化させる適合型閾値方式を採用する。

提案方式の概略処理手順は以下ようになる。

(1)最初に、顔位置検出処理を、既存技術(OpenCV の関数など)を使って行い、その情報により顔部品の探索範囲を絞り込むと同時に、その顔画像の色分布の特徴もその場で分析する。(2)次に、空間フィルタによるエッジ検出と色相での領域分割処理を行う。この色相領域分割処理では先に求めた各顔画像特有の色分布に合わせた閾値を設定する。(3)次に、事前に求めている顔部品平均形状、上記エッジ検出情報、色相による領域分割の領域境界情報の三種類を部品ごとに異なる重み付けで組み合わせ、各部品輪郭形状を近似する多角形の適正度を評価するエネルギー関数を構成する。(4)最後に、このエネルギー値を最小化するような近似多角形形状を求め、それで形状の計測を行う。

本稿では、上に述べた各概略手順を説明し、本方式の実験・評価結果を記述する。

2. 顔位置の検出と平均顔部品形状取得

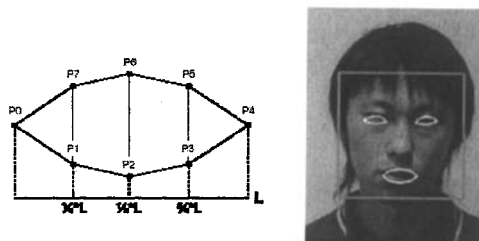
今回の検討は、1 人の被験者の表情変化を自動的に計測するシステム構築のために行われていて、使われる顔画像は、VGA サイズ(640x480)のカラー画像で、被験者の頭髮の一部を除く顔全体が画面内に入り、かつ左右の目の各中央間の間隔が 50 画素程度以上あり、顔の向きとしては、画像平面内回転角 ϕ 、左右の回転角 θ は何れも $\pm 15^\circ$ 程度以内、また視点の仰角はほぼ水平を想定している。また、幾つかの顔付属物（眼鏡や口ひげなど）は条件付で識別可能なものもあるが、今回は十分に検討していない。また、屋外で直射日光照射状態の画像も対象としていない。

2.1. 画像内の顔検出と平均顔部品形状

画像中の顔位置と概略顔サイズ検出のために、ここでは公開画像処理ライブラリである OpenCV 中の顔検出関数 `cvHaarDetectObjects()` を使う[3]。この関数では、基本的に AddaBoost アルゴリズム[2]により顔検出が行われているので、比較的高速で安定した顔検出が可能である。但し、現在公開されている識別器データを使った場合、顔の面内回転(ϕ)に対しては適用範囲がやや狭く($\pm 20^\circ$ 程度以内)、大きい回転角の顔検出のためには追加前処理が必要である。

まず事前準備として、100 枚の正面水平向きの顔画像集合[8]を用意し、これに対して図 1(a)に示すような輪郭を近似する 8 角形の頂点座標を手動で求めておく。次に各顔画像に顔検出処理を適用すると、顔中心座標位置(C_x, C_y)および顔の幅(F_x)と高さ(F_y)が求まる。これらの値を使って、各頂点座標値を式(1)に示すように規格化し、その分布と平均値を求める。図 1(b)に平均部品形状を示す。但し、規格化された平均座標値をこの顔画像の実座標値に変換して表示している。

$$x^* = (x - C_x) / F_x \quad y^* = (y - C_y) / F_y \quad (1)$$



(a) 顔部品輪郭近似 8 角形 (b) 顔部品平均形状

図 1 顔部品輪郭近似 8 角形とその平均形状

2.2. 顔部品の探索範囲の決定

各顔部品近似 8 角形頂点の規格化座標値の分布から、部品ごとに分布を全て含むように探索範囲(長方形領域)を決定する。また、各顔画像の標準肌色分布を取得するため、目と唇領域を除く鼻周辺の領域も設定する。目、唇、鼻に対応する長方形領域の対角頂点座標（左上 $x, y: L, B$; 右下 $x, y: R, T$ ）は各々式(2)-(4)で表される。また、これらを実画像で示すと、図 2 のようになる。

(y 座標は下が正)

$$\begin{aligned} L &= C_x - 0.85r; & B &= C_y - 0.5r \\ R &= C_x + 0.85r; & T &= C_y + 0.07r \end{aligned} \quad (2)$$

$$\left. \begin{aligned} L &= Cx - 0.6r; & B &= Cy + 0.35r \\ R &= Cx + 0.6r; & T &= Cy + 0.9r \end{aligned} \right\} \quad (3)$$

$$\left. \begin{aligned} L &= Cx - 0.7r; & B &= Cy - 0.05r \\ R &= Cx + 0.7r; & T &= Cy + 0.3r \end{aligned} \right\} \quad (4)$$

但し, $r = (Fx + Fy)/4$ (顔半径)



図2 目・唇・鼻周辺の探索範囲を示す長方形

3. 顔画像の色相での領域分割とエッジ検出

カラー顔画像を目視した場合、普通の肌部分と唇や目(顔部品)の部分とは、色相や輝度が異なっていることがわかるが、この事実を使って顔部品領域の分割・抽出を行う。また、並行して簡単な空間微分フィルタによるエッジ検出処理も行う。

3.1. 顔を構成する画素の色相分布

式(5)により顔画像中の各画素色情報を RGB から YCbCr 空間に変換し、色相を分離した状態で Cr-Cb 平面(以降色相平面と呼ぶ)上で分布を調べると、図3(a),(b)に示すように通常肌部分と顔部品部分とは色相平面の異なる領域に分布していることが分かる。

$$\left. \begin{aligned} Y &= 0.3R + 0.59G + 0.11B \\ Cb &= B - Y \\ Cr &= R - Y \end{aligned} \right\} \quad (5)$$

従って、もし通常の肌部分の分布と顔部品部分の分布との間に分離のための適切な境界線を引くことができれば、その線により顔部品部分の画素が判明し領域分割・顔部品輪郭検出が可能になる。しかし、図3の(a)と(b)から分かるように、画像の照明条件などが異なると、最適な境界線の位置が画像毎に異なり、事前学習などで求めた固定的な境界線で識別を行うと、誤検出が非常に多くなり実用に適さないものになる。そこで我々は検査顔画像毎に通常の肌色画素の色相・輝度分布を求め、それを基準に境界線を求める手順を考え、それを組み込んだ領域分割手法を提案する。次にその手法について説明する。

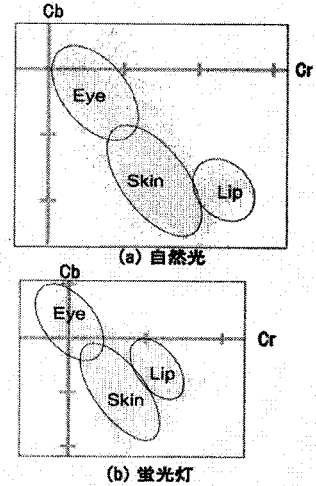


図3 光源による顔画素の色相平面内分布変化

3.2. 色平面内識別線による適応型領域分割

提案手法では、(1)各顔画像中の通常肌部分(目・唇部分以外の鼻周辺)画素の色相・輝度空間内分布を画像毎に独立して求め、(2)これにより顔の通常肌と顔部品部の各画素の色分布とその識別境界線を決定し、(3)顔部品を含む領域の画素の色相空間内分布に対してこの識別境界線を使って顔部品の画素を選別し、画像に適応した領域分割を行う。各処理手順を説明する。

(1)鼻周辺を通常肌色のみを含む部分と仮定してこの領域の全画素に対して色相平面での分布パラメータを計算する。計算量を圧縮するため個々の画素ではなく、色相平面を 3×3 の升目に分け、各升目に入る画素数のヒストグラムを作成し、同時にその重心点の座標値($Cr0, Cb0$)を計算する。次に二次元ヒストグラムの分布に対して主軸と副軸および各軸に対する標準偏差を求める。主軸方向標準偏差 σu 、副軸方向標準偏差 σv 、 $\alpha (\alpha_{1,2,3,4}) = 1.65 \sim 4.0$ とした時、図4に示すように、分布の重心点を中心にし、長径を $2\alpha \cdot \sigma u$ 、短径を $2\alpha \cdot \sigma v$ とする楕円を描き、一旦この楕円に外接する長方形を境界線の0次近似とする。また、Cr-Cb座標系から、u-v座標系への変換は、u軸とCr軸との角度を θ とすると式(6)となり、境界線内部は式(7)で表される。

$$\left. \begin{aligned} u &= (Cr - Cr0) \cdot \cos \theta + (Cb - Cb0) \cdot \sin \theta \\ v &= -(Cr - Cr0) \cdot \sin \theta + (Cb - Cb0) \cdot \cos \theta \end{aligned} \right\} \quad (6)$$

$$-\alpha_1 \cdot \sigma u \leq u \leq \alpha_2 \cdot \sigma u \quad -\alpha_3 \cdot \sigma v \leq v \leq \alpha_4 \cdot \sigma v \quad (7)$$

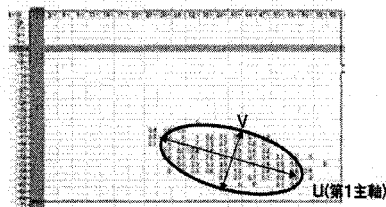


図4 通常肌画素の色平面分布とそのパラメータ

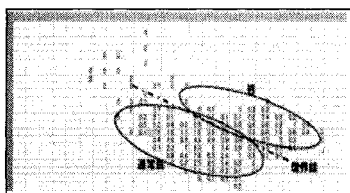


図5 唇を含む領域画素の色平面ヒストグラム例

(2)鼻付近と同様に、目を含む長方形と唇を含む長方形で画素の色分布のヒストグラムを作成する。例えば、唇を含む長方形では図5に示すように、通常肌の色分布：分布(A)と、それとは異なる位置に中心がある顔部品部分の色分布：分布(B)の両方を含む。これら色平面内の適切な識別境界線で分離する。

まず唇領域抽出に関しては、唇を含む長方形内の全画素に対して、色相平面上のヒストグラムを作成し、式(7)で示される重心点(Cb0,Cr0)点付近の画素部分を除く。即ち混合領域のヒストグラムから、通常肌分布、分布(A)を引いた残りの分布(これを分布(B*)とよぶ)を求める。但し、この場合の式(7)の α は各々 $\alpha_{1,2,3,4} = 2, 3, 3, 1.65$ で、更に開口している場合も考慮し、輝度値の低い画素も通常肌色画素とは別にする。この分布(B*)が唇画素の分布：分布(B)とほぼ等しいと考え、その重心点(Cr1,Cb1)と標準偏差などのパラメータをヒストグラムから求める。そして、分布(A)と分布(B*)各重心間を結ぶ色相平面の線分上で両重心からのマハラノビス距離が等しくなる点を通り両方の分布を均等に分ける線分を求め、これを識別境界線とする。この線によって画素の所属を識別し画像平面上で領域を分割する。

目領域の抽出に関しても、類似の手法を適用する。但し、この場合の α は各々 $\alpha_{1,2,3,4} = 2, 3, 3, 3$ である。更にこの場合幾分複雑になる。まず、目領域中白目部分と通常肌の部分は色相平面で識別し、黒目と通常肌や白目との識別はY-Cr平面で行

う。また、式(2)から得られる目を含む長方形は目を完全に含むように左右の幅を多少広めに、高さも多少大きめに取っているため、顔の範囲外の左右の背景や髪の毛および眉毛も含む場合が多く、これらも通常肌以外の色相画素として検出される。これらの誤検出領域は顔中心からの位置関係を使って除去する。

(3)顔画像中、通常肌以外(鼻や耳を除く顔部品領域)の色をもつとして検出された画素をそのままプロットしても、本来の部品領域に未検出画素が点在する場合が多いので、これの穴埋めを行い、完全な連結領域にするため、一種の平滑フィルタ処理を行う。平滑化は画面を 3×3 画素の正方形の升目に分けてその単位で行う。もし、升目の9画素中顔部品対応画素が5個以上ある場合、その升目全体が対応部品領域に属し、4個以下の場合通常肌領域とする。この処理により本来の部品領域外の誤検出画素もほぼ除去され、唇に関しては、面積最大の連結領域1つを残して他を全て除去する。目の領域についても左右を合わせて2個以上の完全連結領域を形成する。それらをまず、顔の中心座標から左右を分離し、かつ顔半径 r から算出される目領域面積の最低予想値($r^2/128$)を超えるものを左右ともに選び、もしそれが左右各々で複数あった場合、その中心(重心)点と顔の中心点との距離が最も近いものを選択することで左右各々の目領域を抽出できる。

3.3. 空間微分フィルタによるエッジ検出

顔部品の中で目については色相領域分割だけでは形状検出が困難な場合があるので、エッジ情報も併用する。目の輪郭線は縦方向の要素が少ないので、縦方向のノイズを抑えることも考慮して、図6に示すような、水平と $\pm 45^\circ$ の線分を検出する3種類のフィルタパターンを使いエッジ検出を行った。検出閾値はエッジ強度絶対値分布の(平均値+標準偏差値)を使った。

色相による領域分割とエッジ検出の両方の結果を重ね合わせて表示した顔画像の例を図7に示す。

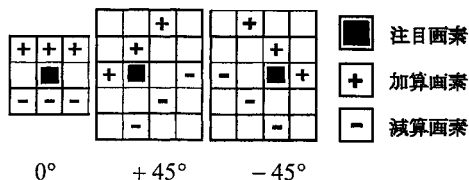


図6 エッジ検出用フィルタパターン3種類



図7 領域分割+エッジ検出の結果画像例

4. 輪郭近似多角形取得とエネルギー関数

4.1. 顔部品輪郭形状検出の特徴

前節までの処理手順で抽出された、目・唇などの顔部品対応領域および輪郭を含むエッジ情報を重ね合わせると照明条件（光線方向・照明色温度など）が良好な場合は、比較的容易に各部品の輪郭線を取得することができる。しかし、証明条件が多少悪くなった場合は、輪郭線の一部しか検出できていない場合も多く、しかも部品によって異なった検出のされ方をする。その定性的な特徴をまとめると、

(1)目では、エッジ検出処理の方がより多くの輪郭線を検出できる。中でも目の上辺および左右の両端は確実に検出できる。目の下辺については黒目の下付近は確実にエッジが検出されるが、それ以外は正確でない場合もある。

(2)目の下辺の形状は、まぶたの開閉などの表情変化があっても、大きくは変化しない。

(3)唇では逆に、色相による領域分割処理で求めた領域は照明条件が多少悪くても精度よく唇領域と一致している。

(4)領域分割処理で求めた領域の境界部分でかつエッジ検出もされている部分は、各部品の輪郭の一部である可能性が非常に高い

以上のことから、領域分割処理で求めた領域境界付近でかつエッジ検出された画素を α 、領域内の画素を β 、その境界付近で α 以外の画素を β^* 画素、エッジ検出された画素を γ 画素、 α 以外の γ 画素で α 画素と連結し同じ側の部品境界を構成している画素を γ^* とすると、目の輪郭検出には、 α 画素を主に、 γ^* 画素を補助的に使う $\{\alpha, \gamma^*\}$ の組み合わせ、唇の輪郭検出には、 α 画素を主に、 β^* 画素を補助的に使う $\{\alpha, \beta^*\}$ の組合せが有効ということになる。

4.2. 輪郭近似8角形抽出とエネルギー関数

図1(a)に示したような8角形で各顔部品の輪郭を近似する。この8角形の頂点を主要なものとして補助的なものに分類する。

主要頂点：左右両端：P0(x0,y0), P4(x4,y4)

中央上下：P2(y2), P6(y6)

補助頂点：上記以外：P1(y1), P3(y3), P5(y5), P7(y7)

但し、()内は真の未知変数で合計10個

顔部品輪郭検出の処理は2段階からなる。

(1) まず、各部品の概略位置を求める。この段階では、各部品の平均形状の近似8角形に外接する長方形をHaar識別器として、目に対しては、 γ 画素との相関が最も高い位置、唇に対しては β 画素との相関が最も高い位置を決定し、各部品形状検出処理の初期位置とする。ここで得られた左右の目の傾斜角を平面内回転角 ϕ として、この角度は固定化する。また、 $\{\alpha, \gamma^*\}$ 或いは $\{\alpha, \beta^*\}$ のすべての画素で近似8角形の頂点座標の最適化を行うと計算量が大きくなるので、概略位置にあるHaar識別器をx方向に1.2倍程度拡張したものを更にx方向に8等分して、各小領域で $\{\alpha, \gamma^*\}$ 或いは $\{\alpha, \beta^*\}$ の代表点を決め $\{Q_i\}$ とし、これらを使って次の段階のエネルギー関数を構成する。但し、Q0, Q1は各々左右両端代表点である。

(2)各部品に対する10個の未知変数を式(8)のようなエネルギー関数を最小化するものとして決定する。但し、LQ0, LQ1は各々P0-Q0間とP4-Q1間の距離、LQi ($i>1$)は点Qiと8角形の最短辺までの距離、Yiは平均形状とP0, P4から推定される標準高さ、 ω_i は重み付けで通常は $\omega_i=1$ 、目に関しては $\omega_{1,2,3}=8$ である。これは目の下辺輪郭形状がほぼ固定化されていることによる。

$$E = 8(LQ_0^2 + LQ_1^2) + 4 \sum_{i>1} LQ_i^2 + \sum_{i \neq 0,4} \omega_i (y_i - Y_i)^2 \quad (8)$$

このエネルギー関数は、未知変数間の独立性が比較的高いので、変数ごとの最小化を順に行うことで、比較的少ない繰り返しで、収束解を求めることができる。

5. 抽出結果と性能の概略評価

5.1. 顔部品形状輪郭検出の結果

図8に2種類の顔部品輪郭検出例を示す。(a)の例では比較的良好な照明環境下での画像のため、色相による領域分割だけでも目・唇ともに高い精度で輪郭検出が可能であるが、(b)の例では、色相で

の領域分割だけでなくエッジ検出と合わせた本方式で輪郭検出が可能になっている。

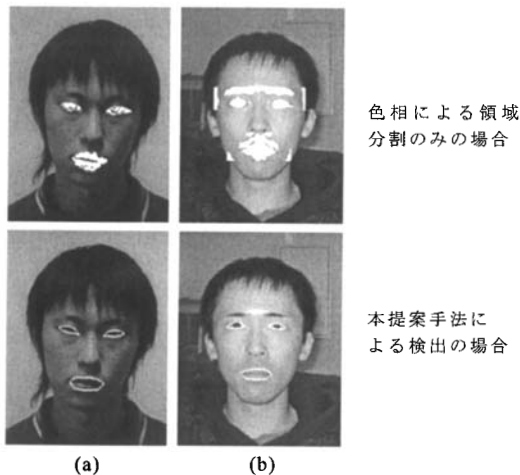


図8 顔部品検出結果例

5.2. 顔部品輪郭検出性能

独自に用意した照明条件が少しずつ異なる顔画像約30枚に対して、顔部品輪郭検出処理を行った。提案方式で求めた部品の水平長・垂直高(中央部)を、目視によって求めた値(一応これを正解値とする)と比べた差分の分布を調べ、提案方式の領域抽出性能を評価した。この結果、各顔部品形状の水平長に対する誤差は約±7%、垂直高に対する誤差は約±12%の範囲に入っていることが分かった。この値は、現在検討している表情解析応用にとっては問題ない範囲であると考えられる。

また、処理時間としては、標準的なデスクトップPC(Pentium4:3GHz, 主記憶1GB)を使い、チューニングを施していないプログラムで実行した場合、「画像読み込みから、顔検出、色相領域分割まで」が約300ms、後半の「エッジ検出から代表点決定とエネルギー最小化・結果画像の表示まで」が約150msを要した。動画像でのリアルタイム処理を目指す場合、同一人物の顔をトラッキングするのであれば、第2フレーム以降は前フレームで求めた各種情報も利用すれば、前半の処理時間は1/4~1/2程度に短縮される。従って更に高速なPCや専用の処理プロセッサ等を使うと、毎秒5~15フレームの動画像に対しても半リアルタイム処理が可能である。

6. まとめ

知的なヒューマンマシンインターフェース(HMI)応用で重要な機械による顔表情識別や分析を行うための、新しい顔部品形状検出手法を提案した。

前処理として必要な顔検出は、公開されている技術(OpenCV)を用いる。この結果の位置情報から、顔部品の探索範囲を絞り込むとともに、顔各部分の画素の色・輝度分布を求めて、この分布の情報を使って色相による画素の識別と部品形状の抽出を行う。即ち、顔の通常の肌部分、目の部分、唇の部分各々の色相平面(Cr-Cb平面)内のヒストグラムからその画像に適した判別閾値を推定し、領域を分割して各顔部品の形状を抽出する。目の形状検出では、これだけでは不十分な場合があるので、エッジ検出情報も融合したエネルギー関数を構成し、これを最小化することで、各部品の輪郭線を近似する多角形形状を最適化する。

この手法を使って照明条件の異なる画像サンプルで実験を行い、顔部品形状検出誤差は概ね7-12%以内、処理時間は単独画像で約450ms、動画に対して連続処理を行えば平均で300msも可能という、比較的良好な結果が得られた。

文 献

- [1] M-H.Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.24, No.1, pp34-58, Jan. 2002.
- [2] P.Viola and M.Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. of IEEE Conf. CVPR, pp.297-336, 1999.
- [3] <http://sourceforge.net/projects/opencvlibrary/> "Open Computer Vision Library" Top Page,または、<http://opencv.jp/> "OpenCV 日本版" Top Page
- [4] 廉田浩, 若谷彰良, "セグメント自己適合型テンプレートによる顔検出手法の検討(2)" 信学技報, Vol.106 No.230, pp.9-16, PRMU2006-78, 2006年9月
- [5] 黒田勉, 渡辺富夫, "HSV表現法に基づく顔画像の唇抽出法," 日本機械学会論文集 C, Vol.61, No592, pp.150-155, 1995年2月
- [6] 横山太郎, 呉海元, 谷内田正彦, "色彩画像からの顔の発見と顔部品の同定," 情報学研報, CVIM100-11, pp.85-92, 1996
- [7] 村田彰, 羽飼直記, 本郷仁志, 加藤邦人, 山本和彦, "形状変化に強い目・口領域抽出方法の提案," 信学技報, PRMU97-159, Nov.1997
- [8] (財)ソフトピアジャパン顔画像データベース・制作メディアドライブ(株) <http://www.hoip.jp/>