

# 曲線近似による表情部品の実時間パラメータ抽出 Realtime parameter extraction of facial expressions by curvilinear approximation

叶 冠峰†  
GuanFeng Ye

伊藤 昭‡  
Akira Ito

寺田和憲‡  
Kazunori Terada

## 1. はじめに

対面のコミュニケーションでは、言語情報のみならず、顔の表情など様々な非言語情報が相互に伝達され、非常に重要な役割を果たしている。顔や顔表情の自動認識は、コンピュータビジョン分野において非常に重要な課題となっている。近年、多くの研究者は機械と人間とのコミュニケーションをめぐる表情認識、個人認証などの様々な研究を展開させている。これらの実現のためには、前提としての顔器官の輪郭線を正確かつ安定に抽出する事が不可欠なと思われる。

本稿ではまず、顔動画画像から画像処理手法によりリアルタイムで顔と顔器官特徴点の検出をする手法を述べる。次に、曲線近似を用いて、眉、目、口の顔器官の輪郭線を自動的に安定して抽出する手法を述べる。最後に、本手法の評価を実際の人物の表情動画を用いて行った結果を述べる。

## 2. 輪郭線抽出の流れ

顔器官輪郭線抽出手法の手順は、図1に示すように4つの部分から成る。ここでは、顔器官の輪郭線抽出の流れを、入力画像から輪郭線の生成までの各過程及びそれぞれの抽出手法について詳細に述べて行く。

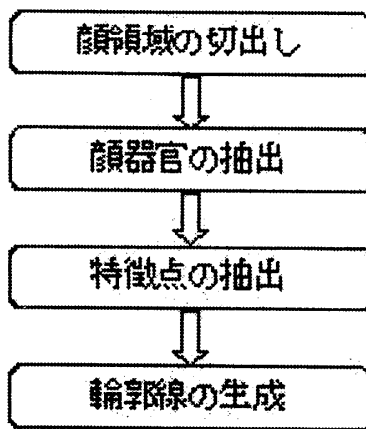


図1: 輪郭線抽出の流れ

### 2.1 顔領域の切出し

原画像の中から顔の領域を適切に切り出すことは、その後の顔器官抽出にかかる処理時間の短縮や、背景による影響の軽減につながる。本稿では、肌色値の分布を利用して、以下に示す簡単な方法を用いて顔領域を切り出す。

†岐阜大学大学院工学研究科  
‡岐阜大学工学部

- (1) RGB表色系の入力画像を YCbCr 表色系に変換する。
- (2) Cb と Cr の値を適度な閾値にして 2 値化することによって肌色領域を求める。
- (3) 二値画像の中から粒状の小さな領域はノイズとして除去する。残った領域の中で最も大きい領域を顔領域と判定する。

以下に RGB 表色系から YCbCr 表色系への変換を示す。

$$Y = 0.29900 \times R + 0.58700 \times G + 0.11400 \times B$$

$$Cb = -0.16874 \times R - 0.33126 \times G + 0.50000 \times B + 128$$

$$Cr = 0.50000 \times R - 0.41869 \times G - 0.08131 \times B + 128$$

以下は本稿で採用した肌色の範囲である。

$$108 < Cb < 122$$

$$132 < Cr < 155$$



原画像 二値画像

図2: 顔領域の抽出

### 2.2 顔器官の抽出

顔領域をもとに以下の顔器官抽出を行う。抽出する対象は、表情を表すのに最も重要な役割を果たす眉、目、口の3部分である。万人に共通である顔器官の位置や大きさの割合などの情報を用いて顔を分割することで、各器官の検索範囲を縮小することができる。また、眉、目、口それぞれ顔器官の抽出はその領域内で閉じた処理となり、単独で抽出可能になる。これにより処理時間を短縮したり、検出率を向上したりすることができるという長所がある。顔の分割は以下の通りを行う。

- (1) 顔の横幅を基準として、目の領域を測る。
- (2) 目領域の中で、次の章で述べる目の抽出アルゴリズムを用いて両目の位置を求める。
- (3) 左右の目の位置関係を基準として、他の器官の位置を推測する。図3の数字は両目の間の距離を1とした時の割合である。

#### 2.2.1 目の抽出

下記の二つ条件をもとに目の検出を行う。

- (1) 目領域は白色と黒色とで構成される。
- (2) 眼球は円である。

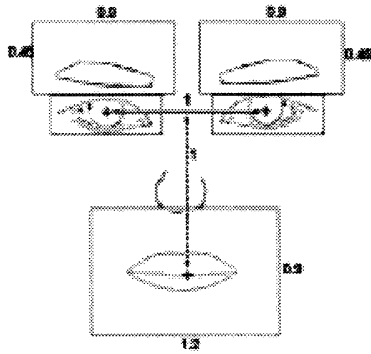


図 3: 顔の分割

RGB 表色系を YCbCr 表色系に変換すると、黒色 (0,0,0) と白色 (255,255,255) とともに (X, 128,128) になる。従って、Cb、Cr の値が 128 の近くにあることを利用して白色と黒色を一括検出することができる。本稿では、 $118 < Cb, Cr < 138$  として、目の候補 1 組を決める。

黒目の検出には、ハフ変換が利用できる。ハフ変換の精度を上げるには細かいエッジ情報が必要である。そのため、Canny エッジ検出法を適用してエッジ画像を作る。Canny エッジ検出法が、他のエッジ検出法と異なる部分は、2つの異なる閾値 (強いエッジと弱いエッジの検出) を使い、弱いエッジも強いエッジに関連している場合にエッジとして出力される点である。この後、検出されたエッジ画像に円検出のためのハフ変換を適用して眼球を検出し、目の候補ペアを決める。

このようにして検出された目の候補ペアの中で、中心座標が最も近いものを目として抽出する。

### 2.2.2 眉の抽出

眉領域にある以下の2つの条件に合うものを眉として判定する。

- (1) 色は黒い。
- (2) 横幅は縦幅の2倍以上である。

ただし、前髪と眉が繋がっていると、眉か前髪かの区別がつかなくなってしまい、正しく検出はできない。

まず、濃淡画像から黒い物体を探し出す。そのためには、YCbCr 系の Y 値を用いる。黒色の探索範囲は以下の通りである。

$$0 < Y < 0.5 \times (\text{領域内の } Y \text{ 平均})$$



図 4: 眉の抽出

### 2.2.3 口の抽出

ここでは色抽出法による口の検出を行う。RGB 値が次の条件を満たしているものを唇色 (赤色) と定義する。

$$(R > 10.0) \wedge (G/R < 0.5) \wedge (B/R < 0.5)$$

この判定条件は R 値に対する G 値及び B 値の比を用いることで照明などの環境の変化に伴う影響を受けにくくなっている。

図 5 に、口矩形領域から唇の色分布を持つ領域を抽出した結果を示す。図から検出がうまくできていることが確認できる。

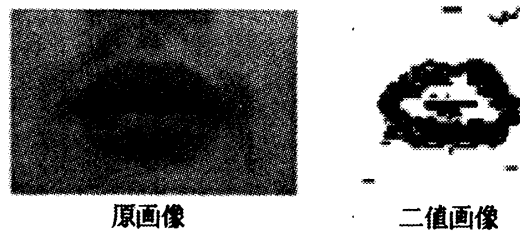


図 5: 口の抽出

### 2.3 特徴点の抽出

像処理による表情認識の手法の一つとして、顔器官特徴点の動きを追跡することによって表情を推測する方法がある。そのためには、表情をうまく表現している特徴点を正確に検出できるかどうかが鍵になる。

本稿では図 6 に示すような 22 個の特徴点を抽出することにした。右眉・左眉について各 5 点、右目・左目・口について各 4 点の計 22 点である。

右目・左目・口の特徴点是对应する二値画像の中にある最も左端、右端、上端、下端の 4 つピクセルを特徴点とする。

また、眉の特徴点是对应する二値画像を四等分にして、両端含む、得られた五つピクセルを眉特徴点とする。

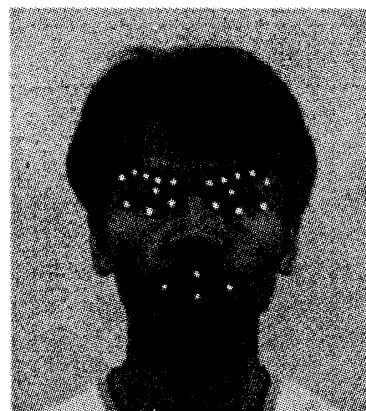


図 6: 特徴点

眉	左眉・右眉各5点
目	左目・右目各4点
口	4点
合計	22点

表 1: 特徴点

曲線	関数
L1	$Y = Y_L + high_1 \times (1 - \frac{(0.75 \times (width_1) - (X - X_L))^2}{(0.75 \times (width_1))^2})$
L2	$Y = Y_R + high_2 \times (1 - \frac{(0.75 \times (width_2) - (X - X_R))^2}{(0.75 \times (width_2))^2})$
L3	$Y = Y_L - high_3 \times (1 - \frac{(X_D - X)^2}{width_3^2})$
L4	$Y = Y_R - high_4 \times (1 - \frac{(X_D - X)^2}{width_4^2})$

表 3: 口の曲線関数

2.4 輪郭線の生成

エッジ画像から物体の輪郭線を取り出すことができるが、画像の撮影環境に大きく影響されるため、エッジが弱くなることが多い。これら弱いエッジから構成される不明瞭な輪郭の検出は不安定であり、完全な輪郭を検出することが難しい。一方、われわれ人間の顔器官の輪郭線は不規則ではなく、幾つかの曲線から成ると思われる。そこで、顔の特徴点より曲線関数を用いて近似した輪郭線を構築することができる。

2.5 曲線関数の定義

目・口の輪郭線は各4本曲線の組み合わせで構成する。各曲線は次式で定義される。

眉の輪郭線は5個の特徴点を繋ぐ直線で構成する。各直線関数は次式で定義される。

$$Y = Y_{(n-1)} + (Y_n - Y_{(n-1)}) \times \frac{X - X_{(n-1)}}{X_n - X_{(n-1)}}$$

3. 実験

3.1 入力画像

喜び・怒り・恐怖・驚き・嫌悪・悲しみなどの六表情を表す被験者の顔をビデオカメラを用いて撮影し、動画画像を得る。顔器官の特徴点をはっきり見えるように、頭頂から肩までが2/3以上の画面を占めるように大きさを調整する。入力画像は1秒間に30フレームのカラー動画画像であり、解像度は320×240ピクセルである。

3.2 実験方法

被験者に「無表情→喜び→無表情→怒り→無表情→恐怖→無表情→驚き→無表情→嫌悪→無表情→悲しみ→無表情」の順で六表情を一度ずつ表出させ、表情の表出過程をビデオカメラで撮影する。得られた動画画像をシステムの入力画像として、顔器官の輪郭線検出を行う。

3.3 実験結果

本手法により顔器官輪郭線を抽出した結果を図9に示す。出力動画画像から見ると、本手法により、無表情から各々の表情への変化過程に従い、六表情のいずれも顔器官の輪郭線を正しく検出できたことが分かった。

又、六表情のうち、無表情と喜び表情の画像について、動画画像から100枚ずつの画像をランダムに取り出し、それに対して抽出された輪郭線が正しいかどうかを他の人に主観的に評価してもらった、その結果を表4に示す。無表情画像に対して78%の検出率、喜び表情に対して75%の検出率が得た。目と口の抽出に失敗した画像のなかに、口角、眼高のところに影が映るため、失敗しているものが多かった。これは光線の影響を受けやすいことが原因と考えられる。また、眉の抽出に失敗した画像のなかに、前髪が眉にかかるほどに下がるため、失敗しているものが多かった。頭髪などの影響を受けやすいことが原因と考えられる。

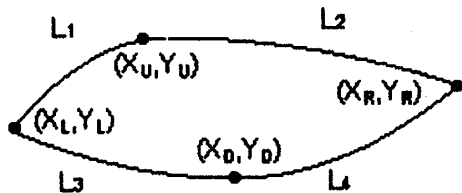


図 7: 目の曲線関数

曲線	関数
L1	$Y = Y_L + high_1 \times (1 - \frac{(X_U - X)^2}{width_1^2})$
L2	$Y = Y_R + high_2 \times (1 - \frac{(X_U - X)^2}{width_2^2})$
L3	$Y = Y_L + high_3 \times (1 - \frac{(X_D - X)^2}{width_3^2})$
L4	$Y = Y_R + high_4 \times (1 - \frac{(X_D - X)^2}{width_4^2})$

表 2: 目の曲線関数

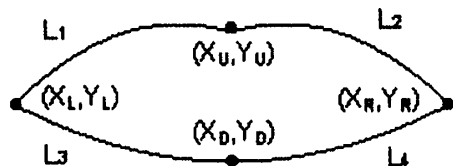


図 8: 口の曲線関数



無表情



喜び

驚き



怒り

嫌悪



恐怖

悲しみ

図 9: 実験結果

	画像数	検出数	検出率
無表情	100	78	78%
喜び	100	75	75%

表 4: 実験結果

#### 4. おわりに

顔の表情は人間同士のコミュニケーションにおいて非常に重要な情報であり、相手の表情から内面状態を読み取ることにより、円滑なコミュニケーションを実現している。人間と機械とのより良いコミュニケーション方法

を確立するためには、この人間の動的な表情変化を機械が認識し、人間の内面状態を推定する必要があると考えられる。そこで、本稿では、その基礎研究として、目や口など顔器官輪郭線の抽出を実現する手法を述べた。

今後、撮影環境に左右されない抽出手法を開発し、さらに、特徴量から表情認識を行うシステムの開発が今後の重要な課題である。

#### 参考文献

- [1] J.Canny: "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.8, No.6, Nov.1986.
- [2] 川口 剛, モハメッド リゾン, 日高 大輔: "ハフ変換と分離度フィルタによる人物からの両目の検出," 信学論 (D-II), vol.J84-D-II,no.10,pp.2190-2200,Oct.2001
- [3] 小田 正志, 船曳 信生, 東野 輝夫: "Deformable Template マッチング法による眼の輪郭線抽出に関する研究", 信学技報 PRMU2003-140 150, pp.1-6
- [4] 関岡 哲也, 横川 勇仁, 船曳 信生, 東野 輝夫, 山田 用弘, 森 悦秀: "関数合成による唇輪郭抽出法の提案," 信学論 (D-II), vol.J84-D-II, no.3, pp.459-470, Mar 2001
- [5] 福井 和広, 山口 修: "形状抽出とパターン照合の組み合わせによる顔特徴点抽出," 信学論 (D-II), vol.J80-D-II no.8,pp.2170-2177,Aug 1997
- [6] 下田 宏, 国弘 威, 吉川 榮和: "動的顔画像からのリアルタイム表情認識システムの試作," ヒューマンインタフェース学会論文誌 Vol.1, no.2, 1999
- [7] 小林 宏, 木村 貴宣, 高橋 久憲: "Windows 上の顔器官認識標準 API の開発," 信学技報 HCS2001-29, pp.23-30, Oct,2001