

## 目と口の動きの追跡による笑顔の分類

片岡 祐介\* 岩口 浩章\* 佐治 斉\*

\*静岡大学大学院 情報学研究科

432-8011 静岡県 浜松市 城北 3-5-1

本論文では、より細かな表情認識を目的として、笑顔に着目し、快・不快・社会的の3つの笑顔に分類する方法を提案する。まず、目・口・眉の構成要素の追跡を行い、次に、表情筋の6グループの動作と追跡結果を適合させる。最後に、どの表情筋の動作により各笑顔が表出されるかを観察することにより、笑顔の分類基準を検討する。

### Classification of smiling face by tracking motions of the eyes and mouth

Yusuke Kataoka\* , Hiromasa Iwaguchi\* , Hitoshi Saji\*

\*Graduate School of Information , Shizuoka University  
Johoku 3-5-1 , Hamamatsu , Shizuoka 432-8011 , Japan

In this paper, we propose the method of classifying a smile into pleasantness, unpleasantness, and sociability for recognition of facial expressions in detail. First, we track the facial components, such as the eyes, mouth and eyebrows. Second, we match the tracking result with the motions of the six groups of facial muscles. Finally, we investigate the criterion for classifying the smile by observing which muscle motion creates the smile.

#### 1 まえがき

人が互いにコミュニケーションを図る際、相手の顔や身ぶりを見て、それが持つ意味を無意識のうちに理解している。顔が発している情報には、顔構成要素の位置や大きさによる個人的特徴や、心の内の感情を間接的に表す表情などがある。このような情報をコンピュータが取得することによって、人とコン

ピュータとの自然なコミュニケーションが実現できることから、顔画像を用いて人物識別や表情認識を行う研究が盛んに行われている。

各種の表情を分類する際に用いられる手法は古くから考察されているが、静止画像を用いた大まかな分類(喜怒哀楽、驚き、恐れなど)手法が多い。人はコミュニケーションの際に、より粒度の細かな分類を行っていると考えられるが、その分類は表情の静的な要因

のみからは困難と思われる。また、種々の感情表出の中でも、コミュニケーションの場においてとりわけ重要な役割を果たすのが「笑顔」である。特に日本人は様々な笑顔を使い分け、また見分けることができると言われている。

顔画像を用いた表情認識の手法としては、顔特徴点の動きから AU(Action Unit) を求め認識する方法 [1]、顔全体のポテンシャル場の変化量を用いる方法 [2]、オプティカルフローから顔の皮膚の動きを検出し認識する方法 [3] ~ [8]、空間周波数を用いる手法 [9],[10] などがある。いずれも顔表皮の動きから表情の特徴を抽出するものである。

オプティカルフローによる認識手法では、顔面筋の位置によって顔を部分領域に分割しておき、それぞれの領域でオプティカルフローを求め、それを顔面筋の収縮としてとらえ表情分析を行っている。しかしながら、このような顔全体の大局的な動きをとらえる手法では、目の瞬きや、発話による口の変形などのような個々の顔構成要素の細かな動きである局所的な動きをとらえることはできない。

さらに、笑顔に関する研究としては、喜びにおける笑いの表情の識別に顔全体の時間的な変化が影響することを示したものの [11]、目と口の動きの時間的なずれをとらえるもの [12] などがあるが、いずれも顔構成要素の細かい動きを考慮していない。

本研究では、表情の中でも特に「笑顔」に注目して認識を行う。まず、顔動画像を研究対象とし、独自の手法で目・口・眉の各構成要素の動きを追跡する。動きをとらえた後、表情筋の変化を見分け、入力顔動画像がどの笑顔であるかを分類する。今回は各種の笑顔の特徴を見いだすことを目標とした。2 節では各構成要素の追跡手法、3 節では笑顔の分類、4 節では実験と考察について述べる。

## 2 構成要素の追跡

ここでは表情筋の影響を受けるとされる目・口・眉の追跡方法について述べる。追跡は目・口・眉に設置された特徴点の動きを各動画像のフレーム毎でそれぞれ行うこととする。ただし、動画像の第一フレームについては、目・口・眉の特徴点の初期位置を手動で与えるものとする。本研究において、各フレーム間の対応は次の手順で行われる。

(1) 第一フレームに特徴点を手動で与え、その近傍をテンプレートとして保存する。

(2) 次のフレームに前フレームで生成したテンプレートを用いて、マッチング処理を行い、求められた特徴点の位置から再び、テンプレートを生成し、保存しておく。

(3) 以上の処理を最終フレームまで繰り返す。

### 2.1 テンプレートマッチング

テンプレートマッチングは、対象物のパターン画像(テンプレート画像)を探索領域内で動かして探索し、最も良く照合する領域を算出する処理である。本研究では、時刻  $t$  における画像中から特徴点を複数抽出し、これが時刻  $t + \Delta_t$  における画像のどこに照合するかを求める。時刻  $t$  の特徴点に対応する画像  $I_t$  上の画素値をテンプレートとし、時刻  $t + \Delta_t$  の画像  $I_{t+\Delta_t}$  上でマッチングを行う。ここでは、式 (1) に示す相違度  $S$  が最小となる位置を求める。

$$s = |I_t(x, y) - I_{t+\Delta_t}(x', y')| \quad (1)$$

$$S = s(R) + s(G) + s(B) \quad (2)$$

( $R, G, B$  は、RGB 表色系における  $R$  成分、 $G$  成分、 $B$  成分)

以下の小節で、テンプレートマッチングに用いる顔面各部位のモデルについて説明する。(なお、左目の左端点と右目の右端点を通る直線を  $X$  軸、両端点の重心位置を原点と

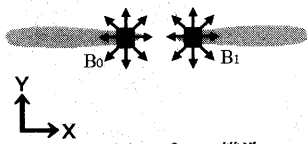


図 1: 眉モデルの構造

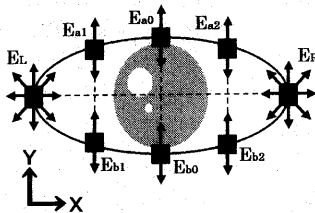


図 2: 目モデルの構造

し、この原点を通り、 $X$  軸に垂直な直線を  $Y$  軸とする.)

## 2.2 眉の追跡

眉モデルの構造を図 1 に示す。左眉の内側先端の特徴点を  $B_0$ 、右眉の内側先端の特徴点を  $B_1$  とする。これらには大きな変動はないと仮定し、前述の式 (1)、(2) を用いて、画像  $I_t$  上のマッチングポイントの近傍を、画像  $I_{t+\Delta t}$  上で探索するテンプレートマッチングにより対応付けを行う。

## 2.3 目の追跡

目モデルの構造を図 2 に示す。左目と右目は同様の構造である。特徴点は目の両端  $E_L, E_R$ 、上瞼の  $E_{a0}, E_{a1}, E_{a2}$ 、下瞼の  $E_{b0}, E_{b1}, E_{b2}$  である。瞼上の各点の  $X$  軸方向の初期位置は、 $E_L, E_R$  間の距離を四等分したところに、左から順に配置されているものとする。瞼上の点の対応付けは次の手順で行われる。

(1) 眉と同じように、画像  $I_t$  上の目の両端点のマッチングポイント  $E'_{Lt}, E'_{Rt}$  を用い、その近傍を画像  $I_{t+\Delta t}$  上において探索するテンプレートマッチングにより対応付けを行い、両端点  $E_{Lt+\Delta t}, E_{Rt+\Delta t}$  を求める。

(2) 求めた特徴点  $E_{Lt+\Delta t}, E_{Rt+\Delta t}$  の座標を  $(x_L, y_L), (x_R, y_R)$  とすると、瞼上の各点の  $x$  座標は次のように表される。

$$x_{a0} = \frac{3x_L + x_R}{4} \quad (3)$$

$$x_{a1} = \frac{x_L + x_R}{2} \quad (4)$$

$$x_{a2} = \frac{x_L + 3x_R}{4} \quad (5)$$

( $x_{b0}, x_{b1}, x_{b2}$  も同様)

(3) 次に瞼上の各点の追跡を行う。例として上瞼の追跡を考える。本研究では目の境界線を楕円形で近似し、以下の楕円の方程式に上瞼の 3 つの特徴点が存在すると仮定する。

$$\frac{(x - x_c)^2}{a^2} + \frac{(y - y_c)^2}{b^2} = 1 \quad (6)$$

(ここで、 $(x_c, y_c)$  は  $E_L$  と  $E_R$  の重心点の座標)

変形すると次の式になる。

$$y = \pm b \sqrt{1 - \frac{(x - x_c)^2}{a^2}} + y_c \quad (7)$$

ここで、先に求めた両端点間の距離から  $a - |\frac{x_R - x_L}{2}|$  となる。また、 $x$  は瞼上の各点  $x_{a0}, x_{a1}, x_{a2}$  と対応している。ここで、変数  $b$  を少しずつ変化させ、画像  $I_t$  で得られたテンプレートの位置を変えていき、3 つそれぞれのテンプレートと画像  $I_{t+\Delta t}$  の間で相違度  $s_{a0}, s_{a1}, s_{a2}$  をそれぞれ求め、それらの総和  $S = s_{a0} + s_{a1} + s_{a2}$  が最小となる  $y$  座標をマッチングポイントとする。(瞼下の追跡も同じように行う。)

## 2.4 口の追跡

口モデルの構造を図 3 に示す。目と同じように特徴点は口の両端  $M_L, M_R$ 、上唇の境界線の  $M_{a0}, M_{a1}, M_{a2}, M_{a3}, M_{a4}$ 、下唇の境界線の  $M_{b0}, M_{b1}, M_{b2}, M_{b3}, M_{b4}$  である。各点の追跡は次のように行われる。

(1) 眉と同じように、画像  $I_t$  上の目の両端

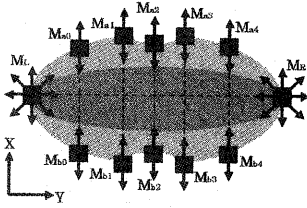


図 3: ロモデルの構造

点のマッチングポイント  $M_{Lt}, M_{Rt}$  を用い、その近傍を画像  $I_{t+\Delta t}$  上において探索するテンプレートマッチングにより対応付けを行い、両端点  $M_{Lt+\Delta t}, M_{Rt+\Delta t}$  を求める。

(2) 求めた特徴点  $M_{Lt+\Delta t}, M_{Rt+\Delta t}$  の座標を  $(x_L, y_L), (x_R, y_R)$  とすると、臉上の各点の  $x$  座標は次のように表される。

$$x_{a0} = \frac{5x_L + x_R}{6} \quad (8)$$

$$x_{a1} = \frac{2x_L + x_R}{3} \quad (9)$$

$$x_{a2} = \frac{x_L + x_R}{2} \quad (10)$$

$$x_{a3} = \frac{x_L + 2x_R}{3} \quad (11)$$

$$x_{a4} = \frac{x_L + 5x_R}{6} \quad (12)$$

( $x_{b0}, x_{b1}, x_{b2}, x_{b3}, x_{b4}$  も同様)

(3) 次に口境界線上の各点の追跡を行う。口は目のような楕円形を想定していない。しかし、テンプレートの探索移動量として楕円形を考えると、マッチングが可能となる。例として上唇における境界線上の追跡を考えると、画像  $I_{t+\Delta t}$  上での探索後の特徴点の  $y$  座標は次のように表される。

$$y = y' \pm b \sqrt{1 - \frac{(x - x_c)^2}{a^2}} \quad (13)$$

( $y'$  は画像  $I_t$  上でのマッチングポイント  
 $x_c$  は  $M_L$  と  $M_R$  の重心点の  $X$  座標)

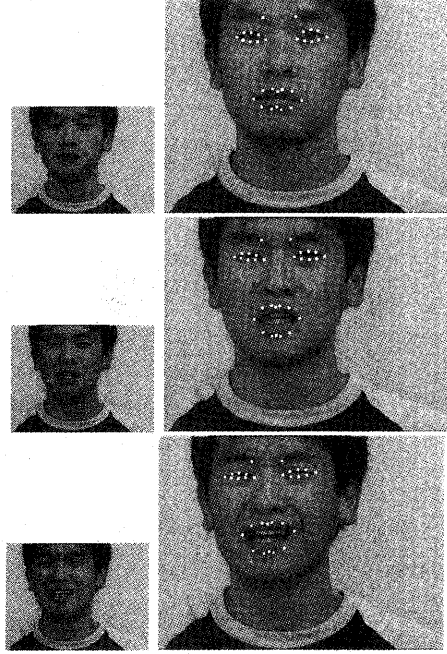


図 4: 追跡結果 (元画像と追跡画像)

ここで、先に求めた両端点間の距離から  $a = \frac{|x_R - x_L|}{2}$  となる。また、 $x$  は境界線上の各点  $x_{a0}, x_{a1}, x_{a2}, x_{a3}, x_{a4}$  と対応している。ここで、変数  $b$  を少しずつ変化させ、画像  $I_t$  で得られたテンプレートの位置を変えていき、それぞれのテンプレートと画像  $I_{t+\Delta t}$  の間で相違度  $s_{a0}, s_{a1}, s_{a2}, s_{a3}, s_{a4}$  をそれぞれ求め、それらの総和  $S = s_{a0} + s_{a1} + s_{a2} + s_{a3} + s_{a4}$  が最小となる  $y$  座標をマッチングポイントとする。(下唇における境界線上の追跡も同じように行う)

## 2.5 追跡結果

眉・目・口の各特徴点の追跡例を図 4 に挙げる。

## 3 笑顔の分類

この節では具体的な笑顔の分類方法について説明する。

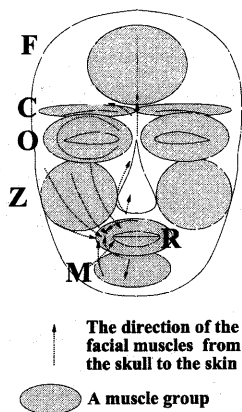


図 5: 表情筋群の構造

### 3.1 笑顔の分類定義

笑顔の分類については、自発的な喜び(快)の笑顔と、愛想笑いなどの社交的な笑顔とに分類されることが多い。本研究では、幅広く笑顔を取り扱うために、泣き笑い、苦笑いに代表されるような負の感情を伴った笑顔を3番目の類として導入し、次の3つの基本的な笑顔に分類する方法を活用する [12].

- 快の笑顔 (喜びの笑顔など)
- 不快の笑顔 (泣き笑い, 苦笑いなど)
- 社交的な笑顔 (愛想笑い, 作り笑いなど)

### 3.2 顔面筋の6区分

本研究では、表情のパターンを形成する主要な表情筋群の動きをより簡明にするために、顔面筋を6つに区分し、次のように分類した [13]. 顔を上部の額と眉, 中部のまぶたと頬, 下部の口と顎の6区分に分けて、それぞれ F, C, O, Z, R, M の6つの略記号で呼ぶこととする。顔の略図を図5に示す。F(額)には額にしわをつくる前頭筋があり、C(眉)には眉を動かす皺眉筋、O(眼)には眼を開閉す

る眼輪筋、Z(頬)には頬に位置して協同して上唇を上下している頬骨筋、上唇挙筋、笑筋のグループ、R(口)には口を開閉する口輪筋、M(顎)には顎に位置して協同して下唇を上下しているオトガイ筋、下唇下制筋、口角下制筋のグループが存在している。なお、鼻根筋は皺眉筋と協同して作用しているのでC(眉)のグループにまとめ、鼻中隔下制筋は上唇の上下に協同して作用しているためZ(頬)のグループにまとめた。このような区分により、表情筋を、表情パターンを形成するのに必要な最低限の筋群にまとめることができる。

### 3.3 構成要素の追跡と6区分との対応付け

ここでは前節で述べた顔の6区分と前章で述べた顔構成要素の追跡の対応付けについて説明する。以下に示す変位とは、第一フレームの特徴(手動で初期位置を与えたところ)と比べてどれだけ変化したかを表すものとする。F, C, O, Z, R, M それぞれの表情筋群に対する特徴量は以下の通りである。

- F(額)
  - 左眉の内側先端の  $B_0$ 、右眉の内側先端の  $B_1$  のY軸正方向の距離の変位
- C(眉)
  - 眉上の2点  $B_0, B_1$  間の距離
- O(眼)
  - 目の境界線の周囲長
- Z(頬)
  - 口の両端点  $M_L, M_R$  のY軸正方向の距離の変位
- R(口)
  - 口の境界線の周囲長
- M(顎)
  - 口境界線上の2点  $M_{b0}, M_{b4}$  のY軸負方向の距離の変位

これらの特徴量を用い、実験を行った。



(a) 快 (b) 不快 (c) 社会的

図 6: 撮影画像例

## 4 実験

### 4.1 環境

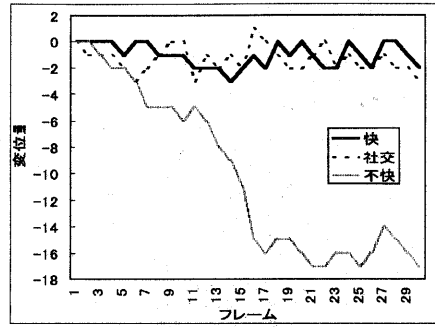
本研究では汎用のビデオカメラを用い、顔の撮影を行った。フレーム間隔は30分の1秒、画像サイズは320×240画素である。被験者には無表情な状態から各々の笑顔に変化するという動作をさせた。これを5名の人物に数回ずつ行わせた。この画像を1フレーム間隔で用いる。この方法から撮影された画像を図6に示す。

### 4.2 実験結果

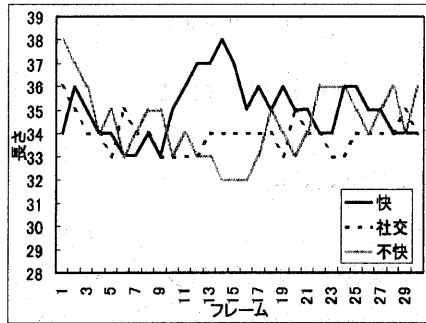
撮影された画像に対し、追跡・特徴量計算の処理を行った結果について説明する。図7,8はそれぞれの特徴量の時間的な変化を表したグラフの一例である。各グラフの横軸はフレーム、縦軸は前節で定義した特徴量である。(特徴量の単位は画素である)

各々のグラフから次のことが読み取れる。

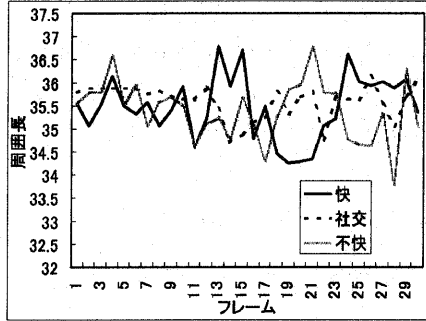
- F(額)
  - 不快の笑顔の場合、大きく下がる。
- C(眉)
  - 快の笑顔の場合、眉間が大きく広がる。
- O(脛)
  - 社会的な笑顔の場合、変動が小さい。
- Z(頬)
  - 笑顔の表出時にほぼ同時に動き始め、3種類の笑顔を区別できる特徴は見られなかった。



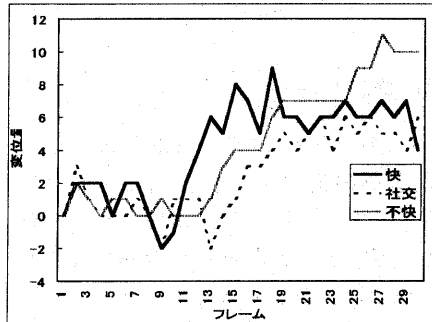
(a) F



(b) C



(c) O



(d) Z

図 7: 各特徴量の時間的な変化-1

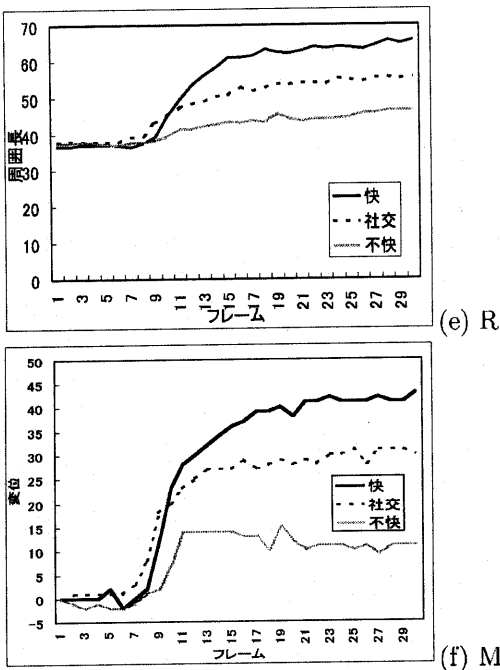


図 8: 各特徴量の時間的変化-2

- R(口)  
不快の笑顔の場合、変動が小さい。
- M(顎)  
Rと同じく、不快の笑顔の場合、変動が小さい。

他の被験者に対しても、同様であった。

#### 4.3 結論

実験結果から、本研究では笑顔の分類基準を次のように定めることとした。

- 快の笑顔  
眉間が大きく広がる (C)。
- 不快の笑顔  
眉が大きく下がる (F)。  
口周囲長の変動が小さい (R)。顎方向への変動が小さい (M)。

- 社会的な笑顔  
目周囲長の変動が小さい (O)。

## 5 まとめ

撮影された笑顔画像に対し、目・口・眉の追跡を行い、表情筋の6区分のモデルを適用することで笑顔の分類基準を考察することができた。

最後に、今後の課題としては、分類基準の数値化、笑顔の判別実験とその評価、他の表情認識モデルとの比較、などが挙げられる。

## 参考文献

- [1] 崖昌石, 原島博, 武部幹, "顔の3次元モデルを用いた顔面表情の分析", 信学論 (D-2), vol.J74-D-2, no.6, pp.766-777, June 1991.
- [2] 松野勝弘, 李七雨, 辻三郎, "ポテンシャルネットとKL展開を用いた顔表情の認識", 信学論 (D-2), vol.J77-D-2, no.8, pp.1591-1600, Aug. 1994.
- [3] 石川貴博, 矢崎和彦, 世良元, 森島繁生, "顔面筋肉モデルに基づく表情トラッキングと再合成", 信学技報, PRMU97-135, Aug. 1997.
- [4] 大塚尚宏, 大谷淳, 中津良平, "連続出力確率密度分布を用いたHMMによる動画からの複数人物の表情認識", 信学論 (D-2), vol.J80-D-2, no.8, pp.2129-2137, Aug. 1997.
- [5] K.mase, "Recognition of facial expression from optical flow", Trans.IEICE, vol.E74, no.10, pp.3474-3483, 1991.
- [6] I.A.Essa and A.P.Pentland, "Facial Expression recognition using a dynamic model and motion energy", IEEE Proc.of 5th Int.Conf.on Computer Vision, pp.360-367, 1995.
- [7] I.A.Essa and A.P.Pentland, "Coding, analysis, interpretation, and recognition of facial expressions", IEEE Trans.Pattern.Anal.Machine Intell. vol.19, no.7, pp.757-763, 1997.
- [8] M.J.Black and Y.Yacob, "Recognizing facial expressions in image sequences using lo-

- cal parameterized models of image motion," Int.J.Computer Vision, vol.25, no.1, pp.23-48, 1997.
- [9] 坂口 竜巳, 大谷 淳, 岸野 文郎, "Wavelet 変換による顔動画像の特徴抽出と表情変化への応用", 信学技報, PRU94-147, 1995.
- [10] 坂口 竜巳, 森島 繁生, "画像の2次元離散コサイン変換を利用した実時間顔表情認識", 信学論 (D-2), vol.J80-D-2, no.6, pp.1547-1554, June 1997.
- [11] C.M Leonard, K.K.S.Voeller, and J.M.Kuldau, "When's a smile a smile? Or how to detect a message by digitizing the signal", Psychological Science, vol.2, no.3, May 1991.
- [12] 西尾 修一, 小山 謙二, "目と口の動きの時間的差異に基づく笑いの分類基準", 電気情報通信学会論文誌, vol.J80-A No.8
- [13] 古川 正重, "美しい表情の秘密", アテネ社, 1995.