

表情認識のための顔器官および表面からの特徴量の抽出

志賀 泰博[†] 海老根 秀之[†] 中村 納[‡]
gashi@swan.cc.kogakuin.ac.jp

[†] 工学院大学大学院 工学研究科 情報学専攻 [‡] 工学院大学 工学部 情報工学科

〒 163-8677 東京都新宿区西新宿 1-24-2

あらまし

人間の表情解析は、顔画像の知的符号化や人間と機械との友好的なマン・マシンインタフェースを構築する上で重要である。本報告では、表情を認識する際に必要となる特徴量の抽出について述べる。顔器官のうち目領域は、表情変化を捉えるためには特に重要である。本報告では、目の開閉度と瞬きの新しい検出方法について述べる。また、口領域では、垂直および水平の各方向への唇の開閉度を検出する。さらに、平常時には比較的エッジ成分が少なくその動きが捉えにくい頬領域からの有効な特徴量の抽出手法についても報告する。

On Extraction of the Feature Amounts from Facial Parts for the Recognition of Expressions

Yasuhiro SHIGA[†] Hideyuki EBINE[†] Osamu NAKAMURA[‡]

[†] Graduate School of Informatics Kogakuin University

[‡] Dept. of Computer Science and Communication Eng. Kogakuin University

1-24-2 Nishishinjuku, Shinjuku-ku, Tokyo 163-8677, JAPAN

Abstract

For precise recognition of facial expressions, effective features concerned with each expression must be extracted from the specific areas on a face. In this paper, we propose a new scheme for extracting a blink and an opening and closing ratio of the mouth. Furthermore, an new algorithm for extracting feature amounts from the area, in which the edges do not exist, such as the cheek is presented. The efficacy of the proposed method is examined experimentally using several images.

1 はじめに

人間同士のコミュニケーションにおいて、意思、感情などは、言葉、表情、身振りおよび手振りなどにより伝達されている。言葉を除いては、そのほとんどは視覚的情報である。視覚的情報の中でも顔は、性格、性別、年齢、さらには意思や感情の心理的な情報まで多岐にわたり発信している。このような観点から、表情の解析は、人間と機械との円滑な意思疎通を可能とする友好なマン・マシンインターフェースを構築する上で極めて重要であると考えられる。

表情の解析には、顔部位の動きを的確に捉える有効な特徴量の抽出が必要である。顔表情を認識するための特徴抽出に関する報告は多くなされているが、顔器官(目、口および鼻等)のエッジ成分から抽出した特徴点を用い、これらの特徴点の二次元的な位置関係や動き情報を特徴量としているものがほとんどである [1], [2], [3], [4].

また、頬領域等の部位で、比較的動きの捉えにくい肌部からの特徴抽出に関してはオプティカルフローを用いたものも報告されているが、大まかな肌部の動きの方向性は考慮されているものの、頭部の動きや雑音に弱いことなどが指摘されている [5], [6]. 本報告では、目および口の開閉度および頬領域からの特徴量の新しい抽出法を提案し、その有効性について検討したのでこれについて述べる。

2 前処理

時系列で表情解析を行う場合、平常顔と表情顔との差分から顔部位の動きは捉えることができるが、頭部の動きが大きな問題となる。さらに、特徴量の抽出においても顔の大きさや顔部位の相対的な位置関係の違いが大きな問題となる。

本手法では、アフィン変換によって入力画像を平均的な顔画像へ変換している。

処理手順は大きく次の4段階からなる。

- (1) 顔表情の認識に特に必要の無い背景、髪および首などの領域を修正 HSV 表色系を用いて除去する [10].
- (2) 目、口および鼻領域を抽出する [11]. これには、顔領域を抽出処理をする際に生じるホール(色相成分の肌の色と異なる部分)と YIQ 基底 [7] の Q 成分を用いる。
- (3) 上記(2)で抽出された目と口領域を基にアフィン変換で用いる目と口の端点および鼻の頂点を参

照点として抽出する。

- (4) 上記(3)で得られた参照点を元にアフィン変換を行い顔画像を正規化する [12], [13].

この処理によって、多様な顔画像に対しても固定ウインドウを用いて、容易に顔部位の指定(図2)が可能となる。図1に入力画像と正規化後の顔画像の例を示す。

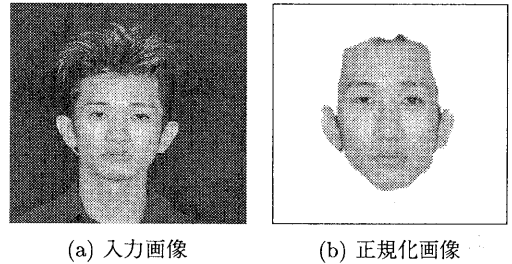


図1: 顔領域の抽出と正規化

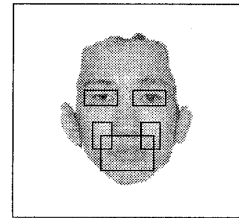


図2: 固定ウインドウ

3 特徴量の抽出

3.1 目領域の特徴量

目領域中の上まぶたや下まぶたから、その輪郭線や特徴点を安定に抽出することは照明条件等の影響で、困難な場合がある [8], [9]. 本論文では、顔領域を抽出する際に使用した修正彩度成分(図3(c))と強度成分(図3(b))を利用して目領域の動きを検出することとした。これらの画像は、式(1)および(2)によって得られる。

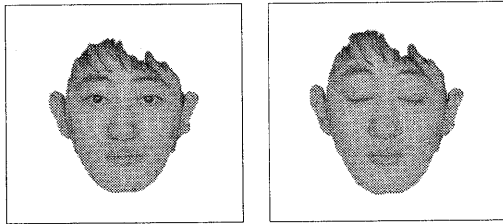
$$Qc = \sqrt{\left(\frac{2R - G - B}{2}\right)^2 + \left(\frac{\sqrt{3}(G - B)}{2}\right)^2} \quad (1)$$

$$I = \frac{R + G + B}{3} \quad (2)$$

ここで、 Q_c は色の鮮やかさを表す修正彩度、 I は色の明るさを表す強度である。

修正彩度成分と強度成分はその濃度値のダイナミックレンジは異なるものの肌、口および眉といった領域では、それらの濃度分布が類似している。しかし、図4に見られるように眼球付近だけは、この2つの画像間でそれらの濃度分布が異なる。

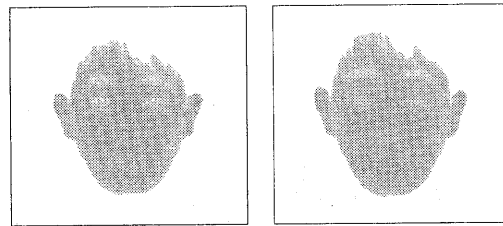
本手法では、この2つの画像の分布特性の違いを利用する。2つの成分の濃度値のダイナミックレンジの分散値を等しくすることによりそれぞれの分布を強調し、得られた分布(図5の破線と実線)の差分を求め瞬きを検出することとする。



(a) 白黒濃淡画像



(b) 強度成分画像



(c) 修正彩度成分画像

図 3: 各色成分

具体的手順は次の通りである。

- (1) 固定ウインドウ内に肌色領域がどの程度含まれているかを調べるために、各成分の濃度値の平均から、それぞれの行(走査線)ごとに分散 $\sigma_c^2(y)$ お



(a) 白黒濃淡画像



(b) 強度成分画像 (c) 修正彩度成分画像

図 4: 目領域画像

よび $\sigma_I^2(y)$ を算出する。

ここで、固定ウインドウ全体に対してその分散を求めた場合、例えばウインドウ内に眉領域が侵入した場合、肌領域だけの場合と比べ2つの画像(図4の(b)および(c))間の濃度値の差が大きくなるため、行ごとに算出することとした。

- (2) 式(4)を用いて修正彩度成分と強度成分の分散比を算出する。

修正彩度成分と強度成分の分散比を同じ値にすることにより、図5(a)に見られるような、眼球付近で色分布が異なる2つの画像に対しては、よりその違いを強調することができる。また逆に、図5(b)に見られるような、上まぶたのような色分布の相関が強い肌色領域では、より類似した分布を得ることができる。

- (3) 式(3)を用いて、修正彩度成分と強度成分の濃度値の差分を算出し、目の開閉度 M_e を検出する。

図6に見られるように、目が閉じられた場合、肌色領域が増加するため、 M_e 値は低くなり、逆に開かれた場合、色成分の異なる眼球の領域が増加するため高い値となる。

$$M_e = \frac{\sum_{y=Y_1}^{Y_2} \sum_{x=X_1}^{X_2} \{k_y(C(x,y) - \bar{C}) - (I(x,y) - \bar{I})\}^2}{S} \quad (3)$$

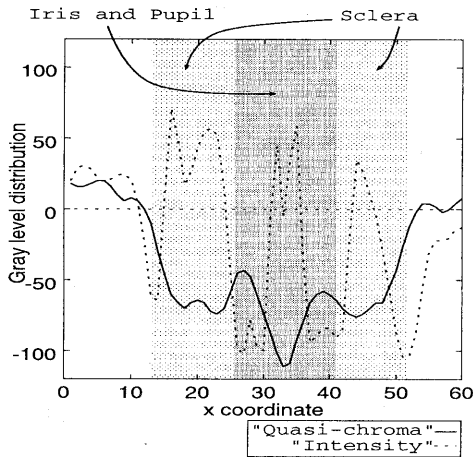
$$k_y = \sqrt{\sigma_c^2(y)/\sigma_I^2(y)} \quad (4)$$

M_e : 目の開閉度

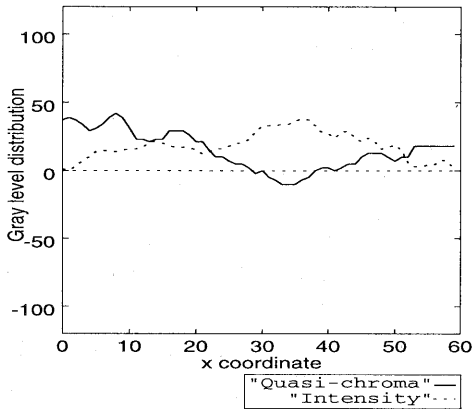
$C(x,y)$: 点 (x,y) における修正彩度成分

$I(x,y)$: 点 (x,y) における強度成分

\bar{C} : 修正彩度成分の濃度値の平均値



(a) 目の中心



(b) 上まぶた

図 5: 目領域の濃度変化

- \bar{I} : 強度成分の濃度値の平均値
- $\sigma_C^2(y)$: y 行目の修正彩度成分の濃度値の分散
- $\sigma_I^2(y)$: y 行目の強度成分の濃度値の分散
- X_1, X_2 : 固定ウインドウの x 座標の始点と終点
- Y_1, Y_2 : 固定ウインドウの y 座標の始点と終点
- S : 固定ウインドウ内の総画素数

3.2 口領域の特徴量

口領域の開閉度は、顔領域の正規化をする際に用いた YIQ 基底の Q 成分を用いて検出する。

手順は次の通りである。

- (1) 式(5)で得られた Q 成分に対してその濃度値の

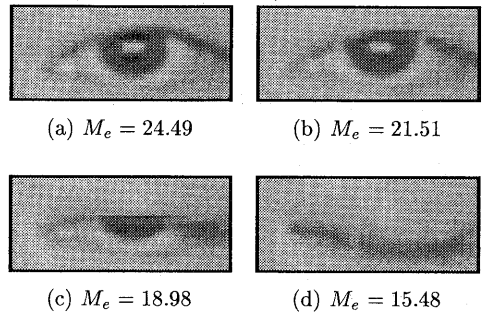


図 6: 目の開閉度

最大が 255, 最小が 0 になるよう線形変換する。

$$Q = 0.21R - 0.52G + 0.31B \quad (5)$$

- (2) 図 7(a) に示す顔画像の固定ウインドウ内の口領域 (図 7(b)) から $160 \leq Q \leq 255$ の画素で面積が最大となる領域を抽出し、得られた領域を二値化処理により図 7(c) に示す唇領域として抽出する。
- (3) 唇領域 (図 7(c)) の濃度値を水平および垂直方向に射影し、各方向の射影分布に対して式 (6) および (7) を用いて分散比を求め、唇領域の開閉度の評価値 M_{mv} および M_{mh} とする。

$$M_{mv} = \frac{\sigma_{ev}^2}{\sigma_{nv}^2} \quad (6)$$

$$M_{mh} = \frac{\sigma_{eh}^2}{\sigma_{nh}^2} \quad (7)$$

M_{mv} : 垂直方向の唇の開閉度

M_{mh} : 水平方向の唇の開閉度

σ_{nv}^2 : 平常顔での垂直方向への射影分布の分散

σ_{nh}^2 : 平常顔での水平方向への射影分布の分散

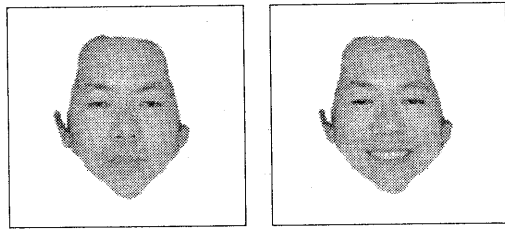
σ_{ev}^2 : 表情顔での垂直方向への射影分布の分散

σ_{eh}^2 : 表情顔での水平方向への射影分布の分散

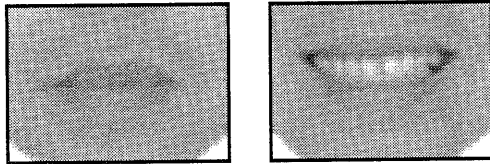
3.3 頬領域の特徴量

表情の表出に伴って頬領域がどのように変化したかを捉えるため濃度変化の方向を特徴量として抽出する。

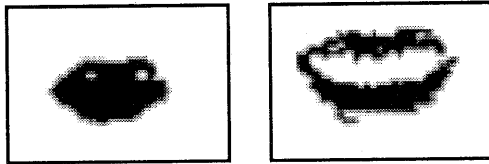
この処理は、大きく以下の 4 段階の処理から成る。



(a) 正規化画像



(b) 固定ウインドウで切り出された口領域



(c) 抽出された唇領域

図 7: 口領域

3.3.1 差分画像の作成

頬領域の動きを捉えるために平常顔画像(図8(a))から表情顔画像(図8(b))を差し引いて差分画像を作成する。この差分画像(図8(c),(d))は、2枚の顔画像から簡単に作成できるだけでなく、顔器官や肌領域の三次元的形状変化を直接的に捉えている。ここで、表情表出後に濃度値が低くなった(明るくなった)画素は正の差分値をとり、逆に高くなった(暗くなった)画素は負の差分値をとる。例えば、口領域での歯の出現は、正の差分値に捉えられ、頬領域の皺の出現は、負の差分値に捉えられるように顔表面の三次元的な動きの特徴を反映している。本手法では、皺の出現による頬領域の動きを捉えている負の差分画像を利用して方向情報を抽出する。なお、従来手法では、各顔部位に固定ウインドウを設定して、固定ウインドウ内の濃度値の二乗平均値を式(8)で求め、特徴量としていた[12],[13]。しかし、この特徴量は、スカラー量であるため、どのような変化をしたのかを評価するには限界があった。

$$\epsilon = \sqrt{\frac{\sum_{x=X_1}^{X_2} \sum_{y=Y_1}^{Y_2} (F_n(x,y) - F_e(x,y))^2}{S}} \quad (8)$$

ϵ : 平均二乗誤差

$F_n(x,y)$: 平常顔画像

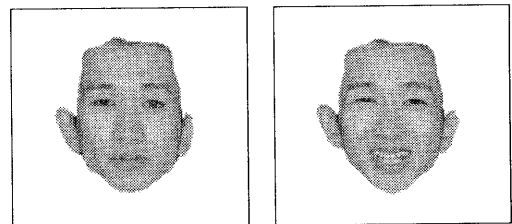
$F_e(x,y)$: 表情顔画像

X_1, X_2 : 固定ウインドウの x 座標の始点と終点

Y_1, Y_2 : 固定ウインドウの y 座標の始点と終点

(x,y) : 画素の位置

S : 固定ウインドウ内の総画素数



(a) 平常顔

(b) 表情顔



(c) 正の差分画像

(d) 負の差分画像

図 8: 差分画像の作成

3.3.2 差分値最急上昇方向の抽出

差分値の増加する方向を捉えるため、上記3.3.1で作成した負の差分画像に対して図9に示す8方向の定義に従い図10に示す各方向に窓を設定し、微分処理を実施する。方向1に対しては、 f_1 の窓(注目画素とその近傍6画素)を設定する。同様に方向2から8に窓を設定し、式(10)により微分処理を行う。求めた微分値の最大の方向を差分値最急上昇方向として求める(式(9))。

$$f_{max} = \max_n \{f'_n\} \quad (n = 1, \dots, 8) \quad (9)$$

$$f'_n = f_n - f_0 \quad (n = 1, \dots, 8) \quad (10) \quad \text{回目の確率 } P_j^{k+1}(d) \text{ は, 式 (11),(12),(13) で表される.}$$

f_n : n 方向の窓内の平均濃度値
 f'_n : n 方向の微分値
 f_{max} : 差分値最急上昇方向

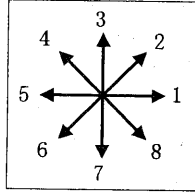


図 9: 8 方向の定義

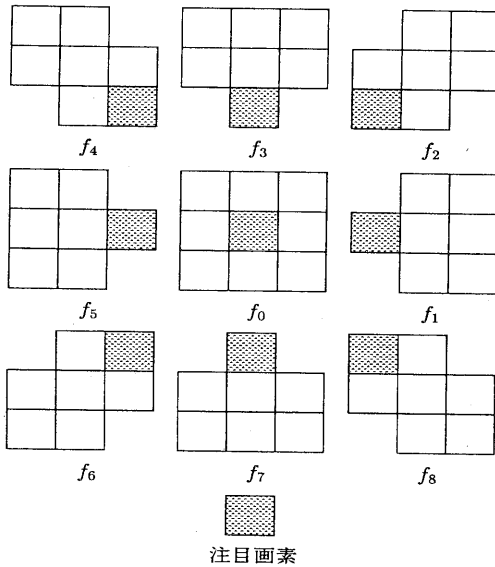


図 10: 微分フィルタ

3.3.3 方向の修正

上記 3.3.2 で得られた差分値最急上昇方向の中には、雑音等によって誤抽出される場合があるため、確率的弛緩法 [14], [15] によって、これを修正する。

本処理では、8 方向の定義に従う 8 個のラベル、および最急上昇方向が存在しないことを表すラベルの合計 9 個のラベル情報を用いて反復計算を行った。

いま、注目画素 B_j における k 回の反復計算で、その方向ラベルが d となる確率を $P_j^k(d)$ としたとき $(k+1)$

$$P_j^{k+1}(d) = \frac{P_j^k(d)\{1 + q_j(d)\}}{\sum_{d=0}^8 P_j^k(d)\{1 + q_j(d)\}} \quad (11)$$

$$q_j(d) = \sum_{j'=1}^8 W_{jj'} \left\{ \sum_{d'=1}^8 \gamma_{jj'}(d, d') P_{j'}^k(d') \right\} \quad (12)$$

$$W_{jj'} \geq 0, \quad \sum W_{jj'} = 1 \quad (13)$$

ここで $W_{jj'}$ は B_j の近傍 $B_{j'}$ からの影響を表わす重み係数である。本処理では B_j として、8 近傍の画素を用いていることから $W_{jj'} = 1/8$ とした。

確率的弛緩法を適用するにあたり、初期確率と適合係数の設定が重要となる。本処理では、差分値最急上昇方向を持たないことを表すラベルには初期確率 0.01 を与え、差分値最急上昇方向のラベルに初期確率 0.3、それ以外の方向ラベルに初期確率 $0.69/7 = 0.09857$ を与えた。また、差分値の最大方向が唯一に求まらぬ画素については、どちらの方向を取るかが不確定であると判断し、その画素が持っている 8 方向すべてのラベルに初期確率 $0.99/8 = 0.12375$ を与えた。

また、式 (12) の $\gamma_{jj'}(d, d')$ は、適合係数と呼ばれ、注目画素と近傍画素との関係を表わす。頬領域では、図 11 の模式図に示すような皺の出現が考えられる。このような動きを考慮して、表 1 に示すように注目画素と対象画素が等しい場合、または正反対の場合には、その存在確率を高めるため高い値を設定した。図 8(c),(d) に示した差分画像の左頬から求めた差分値最急上昇方向の画像 (適用前) と適用後の画像を図 12 に示す。

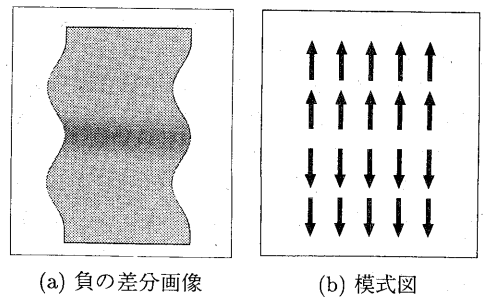


図 11: 皺のイメージ

表 1: 適合係数

d	↑	↑	↑	↑	↑	↑	
d'	↑	↘	←	↙	↓		*
$\gamma_{ij}(d, d')$	1.0	0.1	0.01	0.001	0.5	0.15	0.1

*:Don't care.

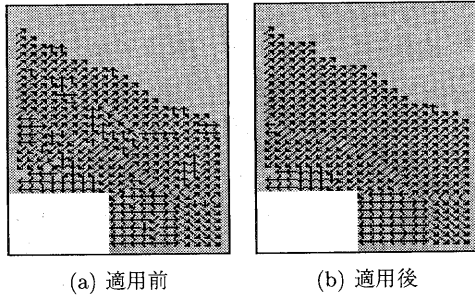


図 12: 確率的弛緩法

3.3.4 方向連続性の抽出

- (1) 弛緩法適用後の差分値最急上昇方向の分布から、皺であると考えられる図 13 に示したような画素間での方向の組合せを抽出する。

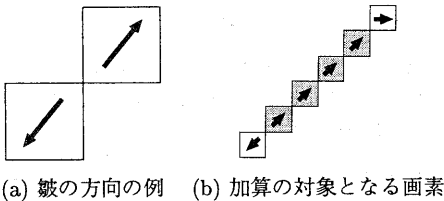


図 13: 皺の例

- (2) (1) で抽出された皺の位置から上側に続く肌部の凹から凸へ向かう三次元的形状変化を捉えるため、図 13 の網掛部に示すような皺の位置から上側に同一方向のラベルの数を加算して行く。
ここで、上側方向だけを加算の対象とした理由としては、皺の位置から下側の部分では、鬚や口領域の侵入が考えられ、また動きに個人差があり一定の方向性を捉えにくいとの判断からである。
- (3) 加算したラベルの総数 n を固定ウインドウの総

画素数 S で割ることにより正規化し、この値を方向連続度 C として算出する。

$$C = \frac{n}{S} \quad (14)$$

4 実験結果と考察

提案した手法を用いて目領域について、目が開いた状態から閉じた状態までの連続する画像 (被験者 31 名, 両目で 397 枚) を用いて実験を行った。結果を表 2 に示す。実験の評価は、目が閉じるにつれて特徴量が減少するものを成功とし、連続画像中で 1 フレームでも目が閉じつつあるにもかかわらず特徴量が増加したものは失敗と判定した。開閉度の検出に失敗した画像は、図 14 のようにインターレススキャンの影響がほとんどであったが、比較的ゆっくりとまぶたが動いている画像については、高い抽出率が得られた。

表 2: 瞬き検出の実験結果

右目	左目	平均
27/31	24/31	50/62
(87.1%)	(77.4%)	(80.6%)



図 14: インターレススキャンによる影響

口領域に関しては、幸福および驚きの各表情 20 枚の平常顔と表情顔について実験を行った。幸福の評価は、式 (6) に示した分散比が水平方向に増えていけば成功であるとした。また、驚きの評価では、垂直方向への分散比が増えていけば成功とした。この結果を表 3 に示す。2 つの表情ともに、前処理の際に使用した Q 成分が正確に抽出されなかった画像に関しては失敗が認められた。

頬領域に関しては、幸福、怒り、悲しみ、驚きの各表情 20 枚について方向連続性の抽出実験を行った。図 15 に示されるように、幸福の表情では出現した皺から上に続く肌部の凹から凸への連続性が捉えられているため評価値が高くなっている。一方、驚きの表情で

表 3: 口領域の開閉度の実験結果

幸福	驚き
17/20 (85.0%)	17/20 (85.0%)

は、皺がほとんど出現しないため連続性の評価値は低くなっている。このことから、従来用いていた二乗平均値では区別のつかなかった動きの方向性を捉えることが可能であることが分かった。

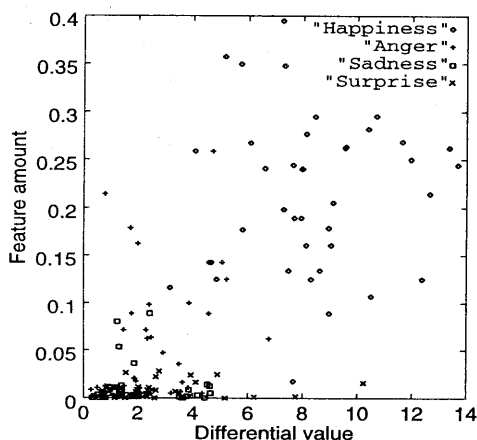


図 15: 頬領域の実験結果

5 まとめ

目領域に関しては、インターレススキニングの影響で瞬きの抽出率が若干低下することが確認された。しかし、他の肌部と異なった色情報を有することから瞬きだけでなく表情表出に伴う目の開閉度を捉えることが可能であると考えられる。頬領域に関しても、スカラー量だけでは区別の付かない方向性を捉えることができた。口領域に関しては、唇領域を抽出する際のQ成分の閾値を安定かつ自動的に決定する必要があると考えられる。

参考文献

- [1] 間瀬, “表情の自動認識”, 映像メディア学会誌, Vol.51, No.8, pp.1136-1139, 1997.
- [2] 松野, 李, 辻, “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, 1994.

- [3] Yaser Yacoub and Larry Davis, “Computing Spatio-Temporal Representations of Human Face”, IEEE CVPR, Vol.1994, pp.70-75, 1994.
- [4] 大橋, 町村, 荒川, 萩原, “動的協調のための表情認識に関する研究 (居眠りの検出)”, 日本機械学会機械力学・計測制御講演論文集, Vol.B, No.98.8, pp.548-551, 1998.
- [5] 矢崎, 石川, 森島, “オプティカルフローを用いた正面顔画像からの顔面筋パラメータの自動推定”, 信学技報, PRMU97-281, pp.121-128, 1998.
- [6] ズデネク プロハースカ, 伊藤, 岡本, “動き情報を用いた表情特徴の自動抽出手法”, 信学論, D-2, Vol.J81-D-2, No.6, pp.1150-1159, 1998.
- [7] A. R. Smith, “Color Gamut Transform Pairs”, ACM-SIGGRAPH '78 Conf. proc., pp. 12-18, 1978.
- [8] 坂本, 宮尾, 田島, “顔画像中からの目の特徴点抽出”, 信学論, D-2, Vol.J76-D-2, No.8, pp.1796-1804, 1993.
- [9] 片原, 青木, “上まぶたの動きにもとづく閉眼の検出”, 信学技報, PRMU97-35, pp.81-88, 1997.
- [10] 松橋, 藤本, 中村, 南, “顔領域抽出に有効な修正HSV表色系の提案”, テレビジョン学会誌, Vol.49, No.6, pp787-787, 1995.
- [11] 海老根, 志賀, 中村, “判別分析法による特徴点の自動抽出と顔表情の認識”, 1999 信学総大, D-12-79, pp.252, 1999.
- [12] H.EBINE, O.NAKAMURA, “The Recognition of Facial Expressions Based on Production System”, IEEE CCECE'99 Conf. Proc., pp.675-682, 1999.
- [13] 海老根, 中村, “顔部位の個人的差異を排除した顔表情の認識”, 電学論 C, Vol.119-C, No.4, pp.474-481, 1999.
- [14] 中村, 後藤, 南, “方向分布パターンによる指紋画像の分類”, 信学論, Vol.J56-D, No.10, 1982.
- [15] 南, 中村, “画像工学—画像のエレクトロニクス—”, コロナ社, 1989.