

# 口周りの特徴を用いた笑い顔表情認識\*

佐々木一洋 (学籍番号 200521357)

研究指導教員：小高和己

## 1. はじめに

近年、コンピュータ技術やネットワーク技術の進展、インターネット等の普及に伴い、人々は機械に接する機会が増え、多様なメディアへのアクセスが可能となっている。しかし、人と機械のコミュニケーションにおいては、人が機械に合わせて行動しなければならない場合が多く、人にとって負担が大きい。人と人とのコミュニケーションにおいては、顔の表情によって伝えられるメッセージが全体の 55% に達するとも言われており[1]、機械が顔の表情を読み取ることができるようになれば人にとって負担の少ないインターフェースが実現され、人と機械のコミュニケーションがより容易になると考えられる。

機械による顔表情認識の研究では、静止画像や時系列画像が用いられる。表情は時間的に固定されたものではなく、表情の表出の時間的特性が感情のカテゴリー判断を左右するとの報告[2]や、目と口の動きの時間的差異が笑いの分類基準の要因になるとの報告[3]がある。このような知見によれば時系列画像からの顔表情の認識はより合理的であると考えられる。

笑いに対する機械認識の試みはこれまでいくつか行われており、カメラから入力された時系列画像を分析して笑いを検出しようとする研究[4],[5]などがある。しかし、これらは特徴点や特徴領域を手動で与えているなど完全な自動化を目指したものはなかった。

本研究では時系列画像中の笑い顔を認識するための簡便な特徴を提案し、それを口周り領域から自動抽出する手法について提案した。

## 2. 笑い顔特有の特徴の検討

笑い顔が表出する際の顔面上の輝度値の変化量の調査により、口や口の端の少し上の部分に大きな変化が起きていたことが分かったため、この領域(口周り領域)が笑い顔の認識に有効か否かを検討するための予備実験を行った。手順は以下の通りである。

[1] 無表情→笑い顔→無表情へと変化する、30 フ

レーム/秒(fps)の時系列カラー画像を輝度画像へと変換する。

[2] 時系列画像の一番始めの第 1 フレームを基準フレームとする。図 1 のように、第 2 フレーム目以降の各フレーム(2,3,⋯,k,⋯,n)の各画素( $x_i, y_j$ )に対して、基準フレームの対応する画素( $x_i, y_j$ )との輝度値の差分を求めて、差分画素値( $\Delta k$ )を決定する( $i, j$  はそれぞれ  $x$  方向、 $y$  方向の画像の画素番号)。

[3] 各フレームにおいて口周り領域内の各座標の差分画素値を合計し、各フレームにおける特徴量とする。

実験結果の一例を図 2 に示す。横軸はフレーム数、縦軸はフレームごとに合計した特徴の値である。ここでは縦軸の単位は任意である。使用した時系列画像においては①の部分で笑いが始まり、②の部分で笑いが終わっている。この結果から、①の部分で特徴量が増大し、②の部分で減少していることが分かる。このことから、口周り領域を用いることで時系列画像中で笑い顔が表出している区間を検出することが可能であり、口周り領域の特徴が笑い顔の認識に有効であることが確認できた。

## 3. 口周り領域の自動検出

口周り領域が笑い顔の認識に有効であることを 2. において示した。しかし口周り領域は顔の動きや個人間の顔の形状差などによりその位置が変化するため、画像枠に固定した領域を用いた場合には必ずしも有効に作用しない可能性がある。そこで本研究では、比較的動き変化の少ない両目をまず検出し、次いで両目の位置から口周り領域を自動で

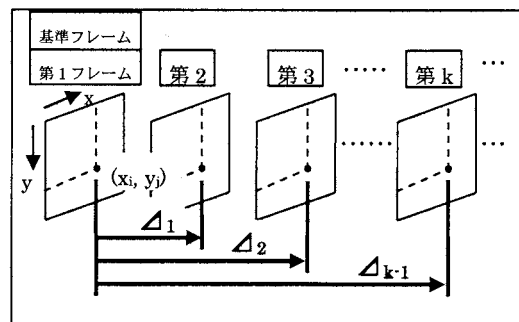


図 1 処理例

\* "Laughing face recognition using mouth and its surrounding features" by Kazuhiro SASAKI

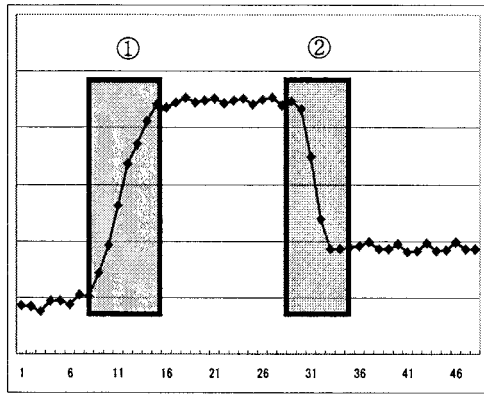


図2 予備実験結果

決定することとした。具体的には、2値化処理、収縮膨張処理、ラベリング処理などを組み合わせて実現した。

#### 4. 識別実験

4名分の学習サンプルを用いた実験から識別のための特徴量の閾値を決定し、その閾値を用いて識別実験を行った。筑波大学の学生12名の被験者を対象に撮影を行い、これらを未知サンプルとした。これらの顔画像データに対し、筆者があらかじめどのフレームが笑い顔表情でどのフレームが無表情であるかを判定している。識別は、笑い顔表情フレームと無表情フレームのみを対象とし、笑い表出の途中に関しては対象外とした。

実験結果を表1に示す。目の検出率における(10/12)は顔画像12データ中10データで目の検出に成功したことを示す。識別率における(x/y)は識別対象フレーム数y中の正識別されたフレーム数xを示す。実験の結果、目の検出に成功したデータ全体に対して97.6%の識別率が得られた。

しかし、顔画像12データ中2データで目の検出に失敗した。これは目に髪がかかっていたことが原因であった。また誤識別32フレーム中29フレームは特定の2データで発生しており、これは笑いの強さが弱かったことが原因であった。

またこの識別実験とともに、目の検出に成功した10個のデータを3名の被験者に提示し、それぞれ1~10の10段階(10が最も大きい笑い顔)で評価させた印象評価調査を行った(図3)。その結果、笑いの強さの評価値と特徴量には強い相関があることが確認できた。これは、本研究で提案した特徴量の大小を判断することで笑いの強さを判定できることを示唆しているものと考えられる。

#### 5. おわりに

本研究では時系列画像中の笑い顔表情を認識

表1 実験結果

目の検出率(P)	識別率			P×Q
	無表情区間	笑い区間	全体(Q)	
83.3% (10/12)	99.5% (550/553)	96.3% (759/788)	97.6% (1309/1341)	81.4%

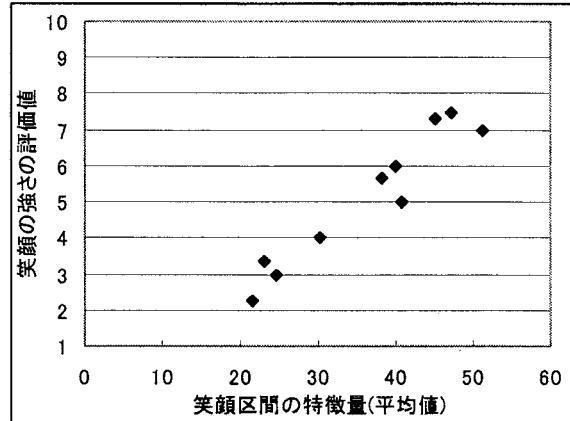


図3 笑顔の強さと特徴量の関係

するための簡便な特徴として、口周り領域における輝度値の瞬時変化量を提案した。また、両目と口との位置関係から口周り領域を自動的に決定する手法を提案した。さらに、笑い顔の強さと特徴量の大きさに強い相関が見られることを明らかにし、本研究で提案した特徴が笑い顔の強さ判定に応用できる可能性を示した。

今後は、笑い顔の強さ判定や、軽い笑いに対処するための口周り領域内での特徴の重み付け法などを検討することが課題である。

#### 文献

- [1] 黒川隆夫: ナンバーバルインタフェース、オーム社、1994.
- [2] M. Kamachi, V. Bruce, S. Mukaida, J. Gyoba, S. Yoshikawa and S. Akamatsu: Dynamic properties influence the perception of facial expressions、Perception, Vol.30, p.875-887, 2001.
- [3] 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、電子情報通信学会論文誌A, Vol.J80-A, No.8, p.1316-1318, 1997.
- [4] 吉川奈穂美, 桜井正広, 宮川道夫: 目・口領域のオプティカルフロー推定による「笑い」認識の試み、電子情報通信学会技術研究報告, PRMU2003-164, p.79-84, 2003.
- [5] 中野実代子, 満倉靖恵, 福見稔, 赤松則男, 安酸史子: ニューラルネットワークを用いた笑顔の認識、電子情報通信学会技術研究報告, HIP2003-21, p.37-42, 2003.