

高次局所自己相関特徴と多重解像度画像を利用した 顔の表情認識に関する一考察

A Study on the Recognition of Facial Expression Using Higher-order
Local Autocorrelation Features and Multiresolution Images

山田陽介*

青木直史*

山本強*

YAMADA Yosuke

AOKI Naofumi

YAMAMOTO Tsuyoshi

1 はじめに

本研究では、高次局所自己相関特徴と多重解像度画像を用いた顔表情認識の可能性について検討を行う。高次局所自己相関特徴は、平行移動に不変な性質を持つ自己相関関数を高次に拡張し、変位を 3×3 画素の局所領域に限定したものである。また、顔画像の認識における最適な解像度を求める代わりに、ウェーブレット変換を用いて複数の多重解像度画像を作成し、解像度の異なる全ての画像から高次局所自己相関特徴を抽出する。

実験では、学習データとして被験者の無表情、喜び、怒りの顔画像をそれぞれ複数枚撮影し、各表情における高次局所自己相関特徴のヒストグラムを作成する。これにより、高次局所自己相関特徴が顔の表情認識に有効であるかどうかを検討する。

2 高次局所自己相関特徴

自己相関関数はパワースペクトルと密接な関係があり、定常時系列データにおける有効な特徴のひとつである。また、平行移動に不変であるという性質があり、図形認識において無関係な変形である平行移動の影響を取り除くことが可能となる。

大津ら [1] は、画像の認識や計測に有効な基本的画像特徴量として、自己相関関数を拡張した高次局所自己相関特徴を提案し、それらの特徴を多変量データ解析手法を用いて統合して有効な特徴抽出する画像計測・認識手法を提案した。

自己相関関数の高次への拡張は、高次自己相関関数と呼ばれる [2]。参照点 $\mathbf{r}$ での対象画像の輝度値を $I(\mathbf{r})$ とすると、 N 次自己相関関数は、参照点周辺の N 個の変位 $(\mathbf{a}_1, \dots, \mathbf{a}_N)$ に対して、

$$x(\mathbf{a}_1, \dots, \mathbf{a}_N) = \int I(\mathbf{r})I(\mathbf{r} + \mathbf{a}_1) \dots I(\mathbf{r} + \mathbf{a}_N) d\mathbf{r}$$

で定義される。次数 N や変位 $(\mathbf{a}_1, \dots, \mathbf{a}_N)$ の取り方により、高次自己相関関数は無数に定義できるが、得られる特徴の数は N に対して指数関数的に膨大なものとなり、実用的ではない。そこで、次数 N を高々2までとする。また、一般に画像データでは、隣接画素間の局所的な相関の方が重要であると考えられるので、変位を参照点周辺の局所的な 3×3 画素の領域に限定する。これにより、平行移動による等価な特徴を除くと、二値画像においては25個の特徴が得られる(0次:1個, 1次:4個, 2次:20個)。同様に濃淡画像においては、1次で1個, 2次で9個の特徴が増え、35個の特徴が得られる。このようにして計算された特徴を高次局所自己相関特徴と呼ぶ。図1に35個の局所パターンを示す。ここで、灰色がかかったパターンは濃淡画像のみに適用されるパターンであり、“*”は“don't care”を表す。

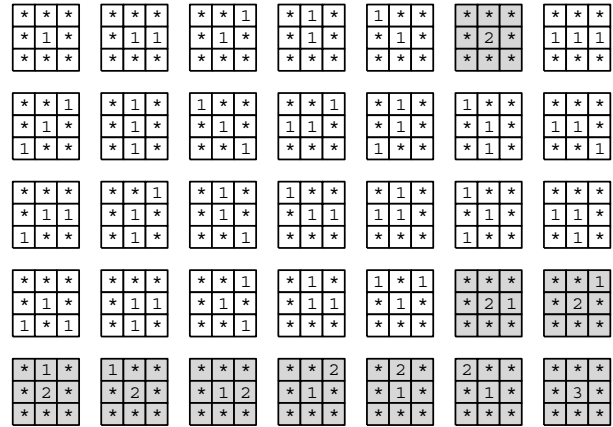


図1 高次自己相関特徴のための局所パターン

3 ウェーブレット変換と多重解像度画像

高次局所自己相関に基づく特徴は、近傍の画素の濃淡値の積を画像全体に対して足し合わせて得られる非常に局所的な特徴であるので、高解像度の画像から抽出された特徴は顔の識別のためには細か過ぎるかもしれない。つまり、認識対象に応じた最適な解像度があると考えられる。栗田ら [3] は、認識に最適な解像度を探索する代わりに、画像ピラミッドの各画像から抽出した特徴をすべて利用することを提案している。画像ピラミッドは、高解像度の画像から低解像度の画像までのいくつかの異なる解像度の画像の集合として構成されるので、画像ピラミッドの各画像から高次局所自己相関に基づく特徴を抽出すると、それらの特徴の集合には対象の詳細な情報から大まかな情報まで含まれている。しかも、対象の平行移動に関する不変性は、これらの特徴にも引き継がれる。

本研究では、ウェーブレット変換による多重解像度解析により、画像ピラミッドを作成した [4]。任意の信号 $f(t)$ は、ハールのスケーリング関数の1次結合で近似することができる。近似の精度をレベルと呼び、レベル0が最も精度の高い近似となる。スケーリング関数で近似されたレベル j の波形 $f_j(t)$ は、やはりスケーリング関数で表現される1つ下のレベル(解像度が半分となる)の近似波形 $f_{j+1}(t)$ と、ウェーブレットで表現される誤差波形 $e_{j+1}(t)$ に分解される。これは、信号を低周波成分と高周波成分の2つのサブバンドに分解することと等価である。一般に画像では、周波数成分が低域側に集中しているので、劣化の少ない低解像度の画像を生成することが可能となる。この操作を繰り返すことで、解像度の異なる多重解像度画像を生成し、画像ピラミッドを作成した。

*北海道大学大学院工学研究科

4 実験

顔の表情認識における高次局所自己相関特徴の有効性を検証するために、実験を行った。以下にその手順と結果を示す。

(1) **顔画像データの取得**: 被験者 1 名について、無表情、喜び、怒りの顔画像を学習データとしてそれぞれ撮影した。撮影枚数は各 20 枚とし、解像度は 640×480 である。撮影の際、照明条件、背景画像は同じであるとし、顔の向きを正面向きに限定した。

(2) **多重解像度画像の生成**: 撮影された各画像に対して、ウェーブレット変換を用いて 5 種類・計 300 枚の多重解像度画像を生成した。それぞれの画像の解像度は、 640×480 , 320×240 , 160×120 , 80×60 , 40×30 である。

(3) **高次局所自己相関特徴の抽出**: 生成された全ての多重解像度画像に対して、高次局所自己相関特徴を抽出した。この際、抽出した特徴量を各画像の解像度で割ることによって、特徴量の正規化を行った。

(4) **顔表情ヒストグラムの作成**: 抽出した高次自己相関特徴量に対して各表情ごとの平均値をとり、横軸を局所パターンの種類、縦軸を特徴量の平均値とする、それぞれの表情におけるヒストグラムを作成し、これを顔表情認識における基本パターンとした。図 2 に、解像度ごとの各表情におけるヒストグラムを示す。ここで、実線は無表情 (neutral), 点線は喜び (joy), 破線は怒り (angry) を表し、横軸の局所パターンは、適当に並べ番号を定めてある。また、0 次と 1 次の高次局所自己相関特徴は、2 次に比べ値が極端に小さく、ヒストグラム上では 2 次の高次局所自己相関特徴のみ示した。

(5) **考察**: 5 種類のヒストグラムを比較して定性的に言えることは、解像度を下げるにしたがってヒストグラムの振幅が大きくなり、表情ごとの差が小さくなっているように見える。個人認識においては、低解像度の画像のほうが認識率が高いという結果も出ているが、表情認識に関しては高解像度の画像のほうが認識には有効であると思われる。

5 まとめ

本研究では、高次局所自己相関特徴と多重解像度画像を用いた顔表情認識の可能性について検討を行った。今後は、顔の大きさや回転、照明条件などの変化に強い特徴量の抽出手法について検討していく予定である。

参考文献

- [1] N.Otsu and T.Kurita, "A new scheme for practical flexible and intelligent vision systems," Proc. IAPR Workshop on Computer Vision, pp.431-435, 1988.
- [2] J.A.Mclaughlin and J.Raviv, "Nth-order autocorrelations in pattern recognition," Information and Control, Vol.12, pp.121-141, 1968.
- [3] T.Kurita, N.Otsu, and T.Sato, "A face recognition method using higher order local autocorrelation and multivariate analysis," Proc. of 11th Int. Conf. on Pattern Recognition, The Hague, Aug.30-Sep.3, Vol.II, pp.213-216, 1992.
- [4] 中野, 山本, 吉田, ウェーブレットによる信号処理と画像処理, 共立出版, 東京, 1999.

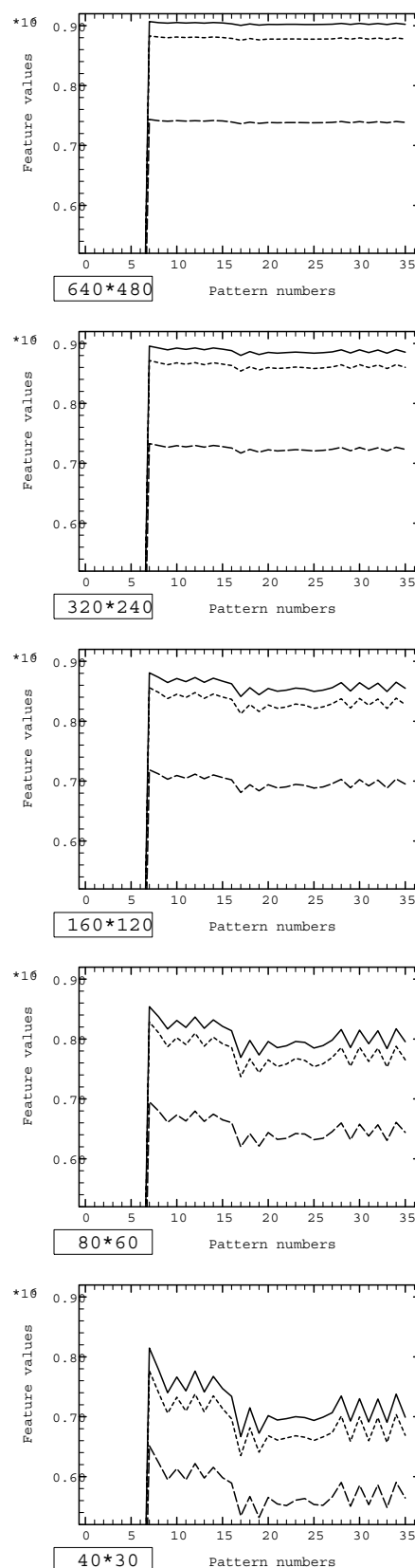


図 2 顔表情ヒストグラム