顔画像の特徴点位置情報と輝度情報を用いたクラス特徴に基づく固有空間法による表情解析

Facial Expression Analysis by Eigenspace Method based on Class features of feature-point positions and gray levels

新座良和　剣持雪子　小谷一孔
Yoshikazu SHINZA　Yukiko KENMOCHI　Kazunori KOTANI

北陸先端科学技術大学院大学 情報科学研究科
School of Information Science ,JAIST

Abstract: This paper describes a facial expression analysis by eigenspace method based on class features(EMC) using both information of feature-point positions and gray levels. This method is used since it derives eigenvectors by which we may extract facial expression information from given facial images. We try to improve the recognition rates by applying this method to the feature-point positions and gray levels.

1 はじめに

表情解析の代表的な手法として，表情筋の動きに着目した手法 [1][2]，顔部位の変化を大局的に記述する手法 [3]-[5] 等がある．これらの手法は単一の特徴量を用いており，輝度情報を用いた場合では，照明条件による影響を受けやすく，特徴点を用いた場合では，特徴点の抽出精度による影響を受けやすい等の問題がある．

複数の特徴量を組合せて用いて互いの短所を補えば，識別率の向上が期待できる．本報告では，大局的な特徴を表す輝度情報と局所的な特徴を表す顔の部位に配置した特徴点との位置情報の関係を調べ，識別率向上のためにこれらの情報の統合方法について検討を行う．

2 クラス特徴の解析手法

我々がすでに報告したクラス特徴に基づく固有空間法 (Eigenspace Method based on Class features;EMC)[6] により解析を行う．

2.1 級間分散と級内分散

F を表情クラスの集合とし，各クラス $f \in F$ に対して M_f 個のデータ (輝度情報や位置情報) が与えられているとする．$m = 1, 2, \cdots, M_f$ 個目のデータを N 次元ベクトル $\boldsymbol{x}_{fm}$ で表す．ここで，クラス f の m 個目のデータを N 次元空間中のある軸に射影したときの値を z_{fm} とし，z_{fm} の級間分散 S_B と級内分散 S_W を次式により求める．

$$S_B = \frac{1}{M}\sum_{f\in F} M_f(\bar{z}_f - \bar{z})^2 \tag{1}$$

$$S_W = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(z_{fm} - \bar{z}_f)^2 \tag{2}$$

ただし，$M = \sum_{f\in F} M_f$, $\bar{z}_f = \frac{1}{M_f}\sum_{m=1}^{M_f} z_{fm}$, $\bar{z} = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f} z_{fm}$ とする．

2.2 クラス特徴に基づく固有空間法 (EMC)

$\boldsymbol{x}_{fm}$ を K 個 $(K \leq M-1)$ の正規直交基底 $\boldsymbol{\phi}_k (k = 1, 2, \cdots, K)$ により次式で近似することを考える．

$$\widetilde{\boldsymbol{x}}_{fm} = \sum_{k=1}^{K} z_{kfm}\boldsymbol{\phi}_k + \overline{\boldsymbol{x}} \tag{3}$$

ただし，$\overline{\boldsymbol{x}} = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}\boldsymbol{x}_{fm}$ とする．従来の固有空間法では，$\boldsymbol{\phi}_k$ を $\widetilde{\boldsymbol{x}}_{fm}$ と元パターン $\boldsymbol{x}_{fm}$ との平均 2 乗誤差を最小にするように選ぶが，EMC では，理想的な近似パターンを各クラスの平均パターンと考え，$\boldsymbol{\phi}_k$ を $\widetilde{\boldsymbol{x}}_{fm}$ と各クラスの平均パターン $\overline{\boldsymbol{x}}_f$ との平均 2 乗誤差 $\varepsilon^2(\boldsymbol{\phi}_k)$ を最小にするように選ぶ．

$$\varepsilon^2(\boldsymbol{\phi}_k) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}||\overline{\boldsymbol{x}}_f - \widetilde{\boldsymbol{x}}_{fm}||^2 \tag{4}$$

ここで，$\overline{\boldsymbol{x}}_f = \frac{1}{M_f}\sum_{m=1}^{M_f}\boldsymbol{x}_{fm}$ とする．選ばれた K 個の $\boldsymbol{\phi}_k$ を使って展開した結果から得られる K 次元ベクトル $\boldsymbol{z_{fm}}$ が，解析に用いる次元圧縮された特徴ベクトルとなる．$\boldsymbol{z}_{fm}$ の k 番目の要素 z_{kfm} は次式から得られる．

$$z_{kfm} = (\boldsymbol{x}_{fm} - \overline{\boldsymbol{x}})^T\boldsymbol{\phi}_k \tag{5}$$

$\boldsymbol{\phi}_k$ は次式の固有値問題を解くことにより求まる．

$$S\boldsymbol{\phi}_k = \lambda_k\boldsymbol{\phi}_k \tag{6}$$

ただし，$S = \frac{1}{M}\sum_{f\in F} M_f(\overline{\boldsymbol{x}}_f - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}_f - \overline{\boldsymbol{x}})^T - \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\boldsymbol{x}_{fm} - \overline{\boldsymbol{x}}_f)(\boldsymbol{x}_{fm} - \overline{\boldsymbol{x}}_f)^T$ ．式 (6) の両辺に左から $\boldsymbol{\phi}_k^T$ を掛け，式 (1)(2)(5) を使うと，

$$\lambda_k = S_{Bk} - S_{Wk} \tag{7}$$

S_{Bk} と S_{Wk} は全クラス集合 F における z_{kfm} の級間分散と級内分散である．式 (7) は固有値 λ_k が級間分散と級内分散の差に等しいことを表している．ここで級間分散と級内分散の差をクラス集合 F の特徴の分離度とみなすと，λ_k が大きいほどそれに対応する $\boldsymbol{\phi}_k$ は F の特徴を良く表す固有ベクトルと考えられる．

2.3 表情識別

式 (5) と同様に，入力画像 $\boldsymbol{x}$ からを特徴ベクトル $\boldsymbol{z}$ を算出する．

$$\boldsymbol{z} = \sum_{k=1}^{K}(\boldsymbol{x} - \bar{\boldsymbol{x}})^T\boldsymbol{\phi}_k\boldsymbol{\phi}_k \tag{8}$$

ここで，式 (7) より $S_{Bk} < S_{Wk}$ を満たす $\boldsymbol{\phi}_k$ のみを使用し、その個数を K とする．そして，$\boldsymbol{z}$ と各表情クラス f の辞書ベクトル $\bar{\boldsymbol{z}}_f$

図 1: 特徴点の配置

$$\bar{z}_f = \frac{1}{M_f} \sum_{m=1}^{M_f} z_{fm} \tag{9}$$

との Euclid 距離 D を調べる．

$$D^2 = (z - \bar{z}_f)^T (z - \bar{z}_f) \tag{10}$$

D が最小となる f を入力ベクトルの識別結果とする．

3 顔画像の輝度情報と位置情報

顔画像の輝度値を輝度情報とし，特徴点の座標値を位置情報とする．表情が多数の顔面筋の選択的な収縮による顔の形状変化 (顔を構成する部位の移動や変形) により生じるため，特徴点はこれら顔の形状変化を良く表すように顔の部位 (眉，眼，口) の周りに配置し，図 1 に示すような，22 点を用いる．それぞれ x, y 座標を持つため，特徴ベクトルは 44 次元となる．22 の特徴点は手動により抽出した．

4 輝度情報と位置情報の統合

輝度情報と位置情報の特徴ベクトルに対して EMC による解析を行った結果を表情識別に用いる．式 (10) により，輝度情報による入力ベクトル z と辞書ベクトル $\bar{z}_f$ の距離 D_{gf} を求め，次式で類似度 S_{gf} に変換する [7]．

$$S_{gf} = \exp\left(-D_{gf}^2/2\right) \tag{11}$$

S_{gf} は $[0,1]$ の値をとり，1 に近い値ほど z が $\bar{z}_f$ と似ている事を表わす．また，位置情報についても同様に D_{pf},S_{pf} を求める．このスケーリングが統一された S_{gf} と S_{pf} を線形結合にし，統合類似度 S_f 求める．

$$S_f = \alpha S_{gf} + \beta S_{pf} \tag{12}$$

S_f が最大となる f を識別結果とする．各類似度の重み α, β については，以下の 2 通りの決定法を考える．

(1) 均等荷重：$\alpha = \beta = 1$ とする．
(2) 調整荷重：評価関数 $H(\alpha, \beta) = S_{c_1} - S_{c_2} - (\alpha - \beta)^2$ を最大化するように α, β を $[0,1]$ の範囲で選ぶ．ただし，c_1, c_2 はそれぞれ統合後の 1 位候補，2 位候補を表す．

5 各情報の識別特性と情報統合による表情識別実験

輝度情報と位置情報の識別特性を調べ，またこれらを統合したものについて表情識別実験を行い比較する．

5.1 実験に使用した顔画像

実験には 5 表情 (無表情，幸福，悲しみ，怒り，驚き) を各 30 枚ずつ計 150 枚の画像を用いた．各表情ごとに 1 枚を test data として選び，残りの 145 枚を training data とした．test data と training data の組を 30 組作成し実験を行った．

実験の前処理として，affine 変換を用いて，左右の瞳の中心，鼻の下の 3 点の位置を一致させ，それら 3 点を中心として 90×96[pixels] の画像領域の切り出しを行った．また，輝度情報については，照明の影響を軽減するために，輝度値の平均値が 0 となるようにした．

5.2 実験結果

輝度情報と位置情報を統合した場合の表情識別実験結果を表 1 に示す．輝度情報と位置情報を個別に用いた場合に比べこれらを統合した場合の方が識別率が向上している．したがって，輝度情報と位置情報が互いに識別に有利となるように補い合っていると考えられる．均等荷重に比べて調整荷重の識別率が低いが，test data の数を増やすことによって識別率の向上が期待できる．

表 1: 輝度情報，位置情報およびそれらの統合情報を用いた場合の各識別率

	使用情報			
	輝度	位置	輝度+位置 (均等荷重)	輝度+位置 (調整荷重)
識別率 [%]	50.0	64.7	68.7	65.3

5.3 輝度情報と位置情報の識別特性

輝度情報と位置情報の識別特性を調べるため，輝度情報と位置情報によって共に正しく識別される割合を求めた結果 36.7%となった．表 1 の結果と合わせて考察すると，輝度情報の識別率 50.0%のうち 13.3%，位置情報の識別率 64.7%のうち 28.0%は，その一方の情報によってのみ正しく識別されており，それぞれの識別特性には違いがあることが分かった．

1 つの情報では正しく識別できない場合に他の正しく識別できる情報を用いることによって足りない情報を補うことができれば識別率の向上が期待できる．

6 まとめ

顔画像の表情識別率の向上のために，輝度情報と位置情報のそれぞれの識別特性を調べ，式 (12) によって情報の統合を行った．輝度情報と位置情報を個別に用いた場合に比べ，これらを統合した場合に識別率が向上した．今後は，画像枚数を増やした場合の検討，他の情報統合方法についての検討，式 (7) による分離度の設定の良否についての検討を行う必要がある．

参考文献

[1] 間瀬 健二, 末永 康仁, “表情動画像からの感情の認識の 1 手法”, 信学技報, PRU91-24, pp.95-101, May 1991.
[2] 崔 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, Vol.J74-D-II, no.6, pp.766-777, June 1991.
[3] 松野 勝弘, 李 七雨, 辻 三朗, “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, Vol.J77-D-II, no.8, pp.1591-1600, Aug. 1994.
[4] 坂口 竜己, 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認識への応用”, 信学技報, IE94-147, PRU94-147, March 1995.
[5] 宗 欣光, 李 七雨, 辻 三朗, “部分特徴テンプレートとグローバル制約による顔器官特徴抽出”, 信学論, Vol.J77-D-II, no.8, pp.1601-1609, Aug. 1994.
[6] 黒住隆行, 新座良和, 剣持雪子, 小谷一孔, “固有空間法と重判別分析による顔画像の個人性と表情の解析”, 信学技報, CS98-122, IE98-101, pp.57-64, Dec. 1998.
[7] 前田茂則, 岡本真明, 河原達也, 美濃導彦, 池田克夫, 堂下修司, “顔画像特徴，歩行画像特徴および音声特徴の統合による個人識別”, 信学論, Vol.J79-D-II, no.4, pp.600-607, Apr. 1996.

〒 923–1292 石川県 能美郡 辰口町 旭台 1–1
北陸先端科学技術大学院大学 情報科学研究科
Phone 0761–51–1231　E-mail yshinza@jaist.ac.jp