JAIST Repository
https://dspace.jaist.ac.jp/

Title	顔表情クラスの分離に適した非線形基底を用いた固有空間法(KEMC)による顔画像の表情解析
Author(s)	小阪,洋平; 小谷,一孔
Citation	情報処理学会研究報告 : オーディオビジュアル複合情報処理, 2003(125): 17-22
Issue Date	2003-12
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/3282
Rights	社団法人 情報処理学会, 小阪洋平／小谷一孔, 情報処理学会研究報告 : オーディオビジュアル複合情報処理, 2003(125), 2003, 17-22. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2003－ＡＶＭ－43 （４）
2003／12／18

顔表情クラスの分離に適した非線形基底を用いた固有空間法(KEMC)による顔画像の表情解析

小阪 洋平† 小谷 一孔†

† 北陸先端科学技術大学院大学情報科学研究科
〒923-1292 石川県能美郡辰口町字旭台 1-1
E-mail: †y-kosaka@jaist.ac.jp, ††ikko@jaist. ac. jp

あらまし 線形写像により特徴ベクトルを選び表情を解析し，識別を行った場合，表情によっては識別困難なものがある．本稿では識別困難な表情は，互いに特徴空間上で線形分離不可能な分布をなすと仮定し，これらの表情についても良好に分離可能となる非線形写像関数を見出し，写像した特徴空間において表情を解析する．本稿ではカーネル関数によって定義される非線形写像関数で特徴空間を写像した高次元空間においてEMC(Eigenspace Method based on Class feature) を適応したKEMC(Kernel EMC) を提案し，解析結果に基づいて表情識別実験を行い，提案手法の有効性を検討する．

Facial expression analysis by Kernel Eigenspace Method based on Class features (KEMC) using non-linear basis for separation of expression-classes

Yohei KOSAKA† and Kazunori KOTANI†

† School of Information Science, Japan Advanced Institute of Science and Technology Asahidai 1-1, Tatunokuti-mati, Nomi-gun, Ishikawa, 923-1292 Japan
E-mail: †y-kosaka@jaist.ac.jp, ††ikko@jaist. ac. jp

Abstract In the facial expression recognition by analyzing feature-vectors with linear transformation, the accuracy of recognition is depending on expression-classes. The accuracy falls sharply when the feature vector of the expression-class has a distribution with difficult linear separation in the feature-space. This paper describes a new method of facial expression analysis and recognition by using non-linear transformation for separating each expression-classes. Our new method, namely KEMC, consists of the non-linear transformation defined by kernel functions for transforming higher dimensional space and EMC (Eigenspace Method based on Class features). This paper also shows experimental results of facial expression classification by KEMC.

1. はじめに

コンピュータビジョンの分野では，人間適応型の情報処理の研究が活発になっている．中でもロボットに対する応用では人に対するマンマシンインターフェース部分への活用として注目されており，介護用ロボットやエンターテイメントロボットといった人とのコミュニケーションを必要とするロボットへの応用が盛んに研究されている．特に文字や言葉のようなバーバルな情報だけでなく，顔の表情やジェスチャのようなノンバーバルな情報をコミュニケーションに利用することが試みられている．計算機により顔画像から表情情報の抽出及び表情の認識を行うことが可能となれば，ロボットが人の表情を判断したり表情を合成する事が可能となり，人とロボットのコミュニケーションに大いに役するであろう．

画像のパターン解析の基本的手法として主成分分析 (PCA) がある．PCA を顔画像の表情解析に適用した場合，得られた基底は個人性や表情，照明，背景等の成分が混合したものとなり，顔画像の表情特徴だけを精度良く抽出出来ない．これに対し黒住らは任意の顔パターンに対して表情の特徴を表す基底を導出するために，クラス特徴に基づく固有空間法 (Eigen-space Method based on Class-features : EMC) を提案し，PCA に比べ顔画像の表情を精度よく分離出来ることを示した [1]．しか

し EMC で Ekman らが提唱している基本 6 表情 [2] を解析し識別を行った場合, "Happiness", "Surprise" に比べ "Anger", "Disgust", "Sadness", "Fear" の識別率が低く, 表情クラスの分離が十分でないものがあった [3]. 本稿ではこれらの表情クラスが線形分離の適用限界を越えているものと仮定し, 線形分離が難しい表情クラスに対しても分離可能な解析手法を検討する.

線形分離困難なクラスの分離には, クラス間を複数の線形分離問題に分割して解くか, 特徴ベクトルを非線形写像して線形分離可能にする手法等が考えられる. 前者は分割数の決定に処理コストと精度とのトレードオフの関係があり, 後者は非線形関数を見出さなければならない. 本稿では表情クラスに対してより分離度の高い基底を求めるために, 任意の顔パターンの次元を拡張し, その高次元空間で表情解析において優れた表情クラス分離特性を有した EMC を用いて非線形基底を求め, その基底を基に表情を識別することを試みる. 本手法を顔表情クラスの分離に適した非線形基底を用いた固有空間法 (KEMC; Kernel EMC) と呼ぶ.

2. 固有空間法による表情解析

PCA は解析対象全てのデータに対する平均 2 乗誤差を最小とする基底を導出するため, 表情のような複数のクラスを分離する基底を導出することは難しい. EMC は各クラスの級間分散と級内分散の差を最大とする基底を導出している [1]. 本節は EMC による基底の導出法を示した後, 本稿が提案する KEMC による基底の導出法を示し, さらに表情の識別方法を示す.

解析対象である各表情クラスの集合を F とし, 各クラス $f \in F$ に対し M_f 枚の顔パターンが与えられているとする. $m = 1, 2, \ldots, M_f$ 枚目の顔パターンを画像サイズ N の各画素のパラメータを要素とする N 次元ベクトル $\mathbf{x}_{fm}$ で表す.

2.1 クラス特徴に基づく固有空間法 (EMC)[1]

$\mathbf{x}_{fm}$ を N 次元空間内のある d 番目の基底に射影したときの値を y_{dfm} とし, y_{dfm} の級間分散 S_{Bd} と級内分散 S_{Wd} を以下のように与える.

$$S_{Bd} = \frac{1}{M}\sum_{f\in F} M_f(\bar{y}_{df} - \bar{y}_d)^2 \tag{1}$$

$$S_{Wd} = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_F}(\bar{y}_{dfm} - \bar{y}_{df})^2 \tag{2}$$

ここで,

$$M = \sum_{f\in F} M_f \tag{3}$$

$$\bar{y}_{df} = \frac{1}{M_f}\sum_{m=1}^{M_f} y_{dfm} \tag{4}$$

$$\bar{y}_d = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f} y_{dfm} \tag{5}$$

$\mathbf{x}_{fm}$ を D 個 ($D \leqq M-1$) の正規直交基底 $\mathbf{w}_d = (w_{1d}, w_{2d}, \ldots, w_{Nd})^T, d = 1, 2, \ldots, D$ により $\tilde{\mathbf{x}}_{fm}$ に近似することを考える.

$$\tilde{\mathbf{x}}_{fm} = \sum_{d=1}^{D} y_{dfm}\mathbf{w}_d \tag{6}$$

主成分分析では, $\mathbf{w}_d$ を $\tilde{\mathbf{x}}_f m$ と元パターン $\mathbf{x}_{fm}$ との平均 2 乗誤差を最小にするように選ぶが, EMC 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, $\mathbf{w}_d$ を $\tilde{\mathbf{x}}_{fm}$ と各クラスの平均パターン $\bar{\mathbf{x}}_f$ との平均 2 乗誤差 $\varepsilon^2(\mathbf{w}_d)$ を最小にするように選ぶ.

$$\varepsilon^2(\mathbf{w}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}\|\bar{\mathbf{x}}_f - \tilde{\mathbf{x}}_{fm}\|^2 \tag{7}$$

この D 個の $\mathbf{w}_d$ を使って展開した結果から得られる D 次元ベクトル $\mathbf{y}_{fm}$ が解析に用いる次元圧縮された特徴ベクトルとなる. なお, $\mathbf{y}_{fm}$ の d 番目の要素は以下のように表せる.

$$y_{dfm} = \mathbf{x}_{fm}^T\mathbf{w}_d \tag{8}$$

$\mathbf{w}_d$ は次式の固有値問題を解くことにより求まる.

$$\mathbf{S}\mathbf{w}_d = \kappa_d\mathbf{w}_d \tag{9}$$

ただし, $\mathbf{S}$ と $\bar{\mathbf{x}}$ を式 (10), (11) とする.

$$\mathbf{S} = \frac{1}{M}\sum_{f\in F} M_f(\bar{\mathbf{x}}_f - \bar{\mathbf{x}})(\bar{\mathbf{x}}_f - \bar{\mathbf{x}})^T - \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\mathbf{x}_{fm} - \bar{\mathbf{x}}_f)(\mathbf{x}_{fm} - \bar{\mathbf{x}}_f)^T \tag{10}$$

$$\bar{\mathbf{x}} = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}\mathbf{x}_{fm} \tag{11}$$

ここで, 式 (9) の両辺に左から $\mathbf{w}_d$ を掛け, 式 (1), (2), (8) を使うと, 式 (12) が導ける.

$$\kappa_d = S_{Bd} - S_{Wd} \tag{12}$$

S_{Bd}, S_{Wd} は全クラス集合 F における y_{dfm} の級間分散と級内分散である. 式 (12) は固有値 κ_d が級間分散と級内分散の差に等しい事を示す. 級間分散と級内分散の差をクラス間の特徴の分離度と見なすと, κ_d が大きいほど軸 $\mathbf{w}_d$ は F の特徴を良く示す基底と考えられる.

2.2 顔表情クラスの分離に適した非線形基底を用いた固有空間法 (KEMC)

2.2.1 カーネルトリックを用いた解析手法 [4]

与えられたパターン $\mathbf{x}$ を非線形の写像 $\Phi(\mathbf{x})$ によって変換し, その空間で線形識別を行うことを考える. 例えば, 写像 Φ として, 入力パターンを 2 次の多項式に変換する写像を用いるとすると, 写像した先で線形識別を行うことは, 元の空間で 2 次の識別関数を構成する事に対応する. 一般的に, こうした非線形の写像によって変換した特徴空間の次元は非常に大きくなることが多い. これにより演算が困難となるので, 非線形解析手法 [5]-[9] では, カーネルトリックを用い, 写像先のデータを直接扱うことを避けている. カーネルトリックとは, 非線形に写像した空間での 2 つの要素 $\Phi(\mathbf{x}_1)$ と $\Phi(\mathbf{x}_2)$ の内積が

$$\Phi(\mathbf{x}_1)^T\Phi(\mathbf{x}_2) = k(\mathbf{x}_1, \mathbf{x}_2) \tag{13}$$

のように, 入力パターン $\mathbf{x}_1$ と $\mathbf{x}_2$ のみから計算出来るなら, 非線形写像によって変換された高次元空間での $\Phi(\mathbf{x}_1)$ と $\Phi(\mathbf{x}_2)$ を陽に計算する代わりに, $k(\mathbf{x}_1, \mathbf{x}_2)$ から最適な非線形写像を構成出来るとみなす考え方である [4]. このような k のことをカーネルと呼ぶ. 実用的には, k は計算が容易なものが望ましい. 例えば

多項式カーネル

$$k(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1^T\mathbf{x}_2)^p \tag{14}$$

Gauss カーネル

$$k(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(\frac{-||\mathbf{x}_1 - \mathbf{x}_2||^2}{2\sigma^2}\right) \tag{15}$$

シグモイドカーネル

$$k(\mathbf{x}_1, \mathbf{x}_2) = \tanh(a\mathbf{x}_1^T\mathbf{x}_2 - b) \tag{16}$$

などが使われている.

本稿では, 入力パターンを非線形の写像によって高次元空間に写像し, 写像先の空間において EMC により非線形基底を求めたい. 実際の計算はカーネルトリックを用いて計算量を減らす.

2.2.2 KEMC による基底の導出方法

EMC における $\mathbf{x}_{fm}$ を非線形の写像によって $\Phi(\mathbf{x}_{fm})$ に写像し, $\Phi(\mathbf{x}_{fm})$ を用い $\mathbf{x}_{fm}$ を D 個 ($D \leqq M-1$) の正規直交基底 $\mathbf{v}_d = (v_{1d}, v_{2d}, \ldots, v_{Nd})^T, d = 1, 2, \ldots, D$ により $\tilde{\mathbf{x}}_{fm}$ に近似することを考える.

$$\tilde{\mathbf{x}}_{fm} = \sum_{d=1}^{D} z_{dfm}\mathbf{v}_d \tag{17}$$

解析に対する基準関数を KEMC も EMC と同様に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, $\mathbf{v}_d$ を $\tilde{\mathbf{x}}_{fm}$ と各クラスの平均パターン $\mathbf{u}_f$ との平均 2 乗誤差 $\varepsilon^2(\mathbf{v}_d)$ を最小にするように選ぶものとする.

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}||\mathbf{u}_f - \tilde{\mathbf{x}}_{fm}||^2 \tag{18}$$

$$\mathbf{u}_f = \frac{1}{M_f}\sum_{m=1}^{M_f}\Phi(\mathbf{x}_{fm}) \tag{19}$$

この D 個の $\mathbf{v}_d$ を使って展開した結果から得られる D 次元ベクトル $\mathbf{z}_{fm}$ が解析に用いる次元圧縮された特徴ベクトルとなる. なお, $\mathbf{z}_{fm}$ の d 番目の要素は以下のように表せる.

$$z_{dfm} = (\mathbf{v}_d \cdot \Phi(\mathbf{x}_{fm})) \tag{20}$$

式 (18) に式 (17) を代入する. ただし $\bar{z}_{df}$ は式 (22) である.

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}\left|\left|\mathbf{u}_f - \sum_{d=1}^{D} z_{dfm}\mathbf{v}_d\right|\right|^2 \tag{21}$$

$$\bar{z}_{df} = (\mathbf{v}_d \cdot \mathbf{u}_f) \tag{22}$$

とする. この式 (21) に式 (19), (20) を代入すると式 (23) を得る.

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}||\bar{x}_f||^2 - \sum_{d=1}^{D}\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d \tag{23}$$

ここで, $\mathbf{S}'$ は式 (24) で与えられ, さらに $\mathbf{S}'_B$ と $\mathbf{S}'_W$ を式 (25), (26) とする.

$$\mathbf{S}' = \frac{1}{M}\sum_{f\in F}M_f\mathbf{u}_f\mathbf{u}_f^T - \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\Phi(\mathbf{x}_{fm}) - \mathbf{u}_f)(\Phi(\mathbf{x}_{fm}) - \mathbf{u}_f)^T \tag{24}$$

$$\mathbf{S}'_B = \frac{1}{M}\sum_{f\in F}M_f\mathbf{u}_f\mathbf{u}_f^T \tag{25}$$

$$\mathbf{S}'_W = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\Phi(\mathbf{x}_{fm}) - \mathbf{u}_f)(\Phi(\mathbf{x}_{fm}) - \mathbf{u}_f)^T \tag{26}$$

KEMC では $\varepsilon^2(\mathbf{v}_d)$ の値を最小にしたい. $\varepsilon^2(\mathbf{v}_d)$ は式 (23) から示すように第 1 項は $\mathbf{v}_d$ によらないので, 第 2 項を最大とすることが $\varepsilon^2(\mathbf{v}_d)$ を最小にする. このときの $\mathbf{v}_d$ は $\mathbf{S}'$ の固有値を降順に選んだときの固有ベクトル $\mathbf{v}_d$ に対応するが, $\mathbf{S}'$ には, 非線形写像関数である $\Phi(\mathbf{x})$ が存在するので直接求めることは困難である. そこで $\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d$ を求めやすいように変形する.

特徴ベクトル $\mathbf{v}_d$ は特徴空間における点集合 $\Phi(\mathbf{x}_{fm})$ の線形結合で表され,

$$\mathbf{v}_d = \sum_{f\in F}\sum_{m=1}^{M_f}a_{fm}\Phi(\mathbf{x}_{fm}) \tag{27}$$

となる. 式 (27) によって各クラスの平均である $\mathbf{u}_f$ と $\mathbf{v}_d^T$ は次のようにカーネルを用いた形に変形出来る.

$$\mathbf{v}_d^T\mathbf{u}_f = \mathbf{a}^T\begin{pmatrix} \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 1}, \mathbf{x}_m) \\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 2}, \mathbf{x}_m) \\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 3}, \mathbf{x}_m) \\ \cdots \\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{2\ 1}, \mathbf{x}_m) \\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{2\ 2}, \mathbf{x}_m) \\ \cdots \\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{f\ m}, \mathbf{x}_m) \end{pmatrix} = \mathbf{a}^T\mathbf{m}_f \tag{28}$$

式 (25), (28) から式 (29) が導ける.

$$\mathbf{v}_d^T\mathbf{S}'_B\mathbf{v}_d = \mathbf{a}_d^T\mathbf{R}_B\mathbf{a}_d \tag{29}$$

ここで,

$$\mathbf{R}_B = \frac{1}{M}\sum_{f\in F}M_f\mathbf{m}_f\mathbf{m}_f^T \tag{30}$$

同様にして, 式 (26), (28) から式 (31) が導ける.

$$\mathbf{v}_d^T\mathbf{S}'_W\mathbf{v}_d = \mathbf{a}_d^T\mathbf{R}_W\mathbf{a}_d \tag{31}$$

ここで,

$$\mathbf{R}_W = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\zeta_{fm}-\mathbf{m}_f)(\zeta_{fm}-\mathbf{m}_f)^T \quad (32)$$

$$\zeta_{fm} = (k(\mathbf{x}_{1\ 1},\mathbf{x}_m),k(\mathbf{x}_{1\ 2},\mathbf{x}_m),k(\mathbf{x}_{1\ 3},\mathbf{x}_m), \ldots,k(\mathbf{x}_{2\ 1},\mathbf{x}_m),\ldots,k(\mathbf{x}_{f\ m},\mathbf{x}_m))^T \quad (33)$$

以上の式 (29), (31) から式 (34) が導ける.

$$\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d = \mathbf{a}_d^T\mathbf{R}\mathbf{a}_d \quad (34)$$

ただし,

$$\mathbf{R} = \mathbf{R}_B - \mathbf{R}_W \quad (35)$$

このベクトル $\mathbf{a}_d$ は行列 $\mathbf{R}$ の固有値問題を解くことで求まる.

$$\mathbf{R}\mathbf{a}_d = \lambda_d\mathbf{a}_d \quad (36)$$

$\varepsilon^2(\mathbf{v}_d)$ を最小にする $\mathbf{a}_d$ は固有値 λ_d の降順に対応する $\mathbf{a}_d$ である. 求められた $\mathbf{a}_d$ と式 (20), (27) から z_{dfm} は式 (37) のように $\mathbf{a}_d$ と $k(\mathbf{x}_{fm},\mathbf{x})$ を用いることで, 入力パターンの写像を求めずに得ることが出来る.

$$z_{dfm} = (\mathbf{v}_d\cdot\Phi(\mathbf{x}_{fm})) = \sum_{f\in F}\sum_{m=1}^{M_f}a_{dfm}k(\mathbf{x}_{fm},\mathbf{x}) \quad (37)$$

2.3 固有空間法による表情識別

与えられた学習パターン $\mathbf{x}_{fm}$ より, EMC と KEMC によってそれぞれの固有空間を導出する. 式 (9), (36) から得られる, λ_d と κ_d の降順に対応する上位 d 基底 $(\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_d)$ と $(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_d)$ を用い, 入力パターン $\mathbf{x}'$ から特徴ベクトル $\mathbf{y}'$ と $\mathbf{z}'$ を写像する.

$$\mathbf{y}' = [\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_d]^T\mathbf{x}' \quad (38)$$

$$\mathbf{z}' = \left[\sum_{f\in F}\sum_{m=1}^{M_f}a_{1fm}k(\mathbf{x}_{fm},\mathbf{x}'),\sum_{f\in F}\sum_{m=1}^{M_f}a_{2fm}k(\mathbf{x}_{fm},\mathbf{x}') ,\ldots,\sum_{f\in F}\sum_{m=1}^{M_f}a_{Dfm}k(\mathbf{x}_{fm},\mathbf{x}')\right]^T \quad (39)$$

表情特徴空間における各表情のクラス f の辞書ベクトル $\bar{\mathbf{y}}_f$ と $\bar{\mathbf{z}}_f$ を以下のように定める.

$$\bar{\mathbf{y}}_f = \sum_{m=1}^{M_f}\mathbf{y}_{fm} \quad (40)$$

$$\bar{\mathbf{z}}_f = \sum_{m=1}^{M_f}\mathbf{z}_{fm} \quad (41)$$

この入力パターンと辞書ベクトルを比較し, 入力パターンの表情クラスを決定する. 比較には単純類似度 Ry_f^2 と Rz_f^2 を用い Ry_f^2 と Rz_f^2 が 1 に最近となる表情クラス f を入力パターン $\mathbf{x}'$ が所属する表情クラスとする.

$$Ry_f^2 = \frac{\mathbf{y}'\cdot\bar{\mathbf{y}}_f}{|\mathbf{y}'||\bar{\mathbf{y}}_f|} \quad (42)$$

$$Rz_f^2 = \frac{\mathbf{z}'\cdot\bar{\mathbf{z}}_f}{|\mathbf{z}'||\bar{\mathbf{z}}_f|} \quad (43)$$

3. 表情識別実験

本稿で与えた KEMC が表情解析に有効な手法であるか, またカーネル関数によって識別特性の違いが現れるかを検討する. このため KEMC と KEMC のベースとなっている EMC, さらに KEMC と同じくカーネルトリックを用いた KPCA[7] の 3 種の手法で表情を解析し識別実験を行った. また, KEMC, KPCA のカーネル関数に式 (14), (15) の多項式カーネル, Gauss カーネルを用い表情を解析し識別実験を行った.

3.1 実験内容

- KEMC, KPCA, EMC を解析手法として closed-data, opne-data に対して識別を行う
- 使用したカーネルは, 多項式, Gauss カーネルの 2 種類
- closed-data に対する識別は各カーネル関数のパラメータを実験的に選択する
- KEMC, KPCA の open-data に対する識別は, それぞれの closed-data で学習したカーネル関数のパラメータを使用する

入力パターンの選び方である, closed-data, open-data の定義は以下の通りである.

- closed-data
  入力パターンは学習パターンに含まれている
- open-data
  入力パターンは学習パターンに含まれていない

3.2 実験データ

実験に使用する顔画像は, 基本 6 表情 (Anger, Disgust, Fear, Happiness, Sadness, Surprise) を 61 人の大学院生にそれぞれ表現してもらい撮影をした. 本稿では表情を受けとる側の識別を表情識別としたいので, 撮影した顔画像を 3 人の評価者による主観評価から選別した. 評価者はランダムに表示された顔画像を一枚づつ「基本 6 表情」と「どれにも識別出来ない」の 7 つに識別し, 表現された表情と識別した表情を, 3 人の評価者のうち 2 人以上が正解した場合, 実験画像として採用した. 主観評価で選ばれた各表情の画像枚数を表 1 に示す. なお以降, 表情の略記は次のことを意味する. Anger;An. Disgust;Di. Fear;Fe. Happiness;Ha. Sadness;Sa. Surprise;Su.

表 1 主観評価により選別された各表情の枚数

An.	Di.	Fe.	Ha.	Sa.	Su.	合計 [枚]
36	37	26	49	39	55	242

主観評価により選別された画像から辞書と入力データを作成する. 辞書に用いる顔画像数を各表情の 6 割として, 残りの 4 割を入力データとした. 各表情の辞書と入力枚数を表 2 に示す.

撮影条件, 正規化条件を以下に示す. 画像データの例を図 1 に示す.

- 顔画像は MINOLTA VIVID700 で撮影した輝度画像
- 目, 鼻の位置で正規化処理を行い, 各顔の位置, サイズ, 傾きを補正
- 顔画像の解像度は $100\times100[pixels]$, パターンとして扱う際はラスタスキャン順に 10000 次元のベクトルとする

表 2 各表情の辞書と入力画像の枚数

	An.	Di.	Fe.	Ha.	Sa.	Su.	合計 [枚]
辞書の枚数	22	22	15	30	23	33	146
入力の枚数	14	15	11	19	16	22	96
合計	36	37	26	49	39	55	242

図 1 用いた画像の例 上段-左から An. Di. Fe.
下段-左から Ha. Sa. Su.

3.3 識別結果

closed-data, open-data に対する表情識別実験の結果を表 3, 4 に示す. なお表中の略記, ploy は多項式カーネル, rbf は Gauss カーネル, Av. は平均識別率を表す.

(1)KPCA と KEMC の識別率の比較

open-data に対する平均識別率は KEMC(ploy) の 77.2% が最高で, 続いて KEMC(rbf) の 65.2%, EMC の 61.9%, KPCA(poly) の 44.9%, KPCA(rbf) の 42.3% の順になっている. KPCA の open-data に対する平均識別率は, KEMC より 20～30 ポイント低いことから, 表情のような複数クラスの識別には, データ全体の分散を最大にする PCA に比べ, クラスの級内, 級間分散の差を最大とする EMC が効果的であるという結果 [1], [3], [10] が, 顔パターンを非線形写像した高次元空間においても成り立つことが分かる.

(2)EMC と KEMC の各表情の識別率の比較

EMC と KEMC の open-data に対する識別率を比較する.

- KEMC(ploy) は線形手法である EMC を内包するので, EMC に比べ識別結果が低くなることはなく, 平均識別率は 15 ポイントほど高い
- EMC は "Happiness", "Surprise" に比べ "Anger", "Disgust", "Sadness", "Fear" の識別率が低い
- KEMC(ploy) は EMC に比べ "Anger", "Disgust", "Sadness", "Fear" の識別率が高く, 特に "Anger", "Sadness" の識別率は 18～25 ポイント高い

以上から KEMC(ploy) は EMC に比べ "Anger", "Sadness" も高精度に分離出来る固有空間を導出したと考えられる.

(3)EMC と KEMC の各表情の識別画像の比較

EMC と KEMC(ploy) が識別した open-data の顔画像を比較する. 図 2 に "Anger" と "Disgust", 図 3 に "Sadness" と "Fear" における識別結果の違いを示す. なお図 2, 3 の画像は左右には, EMC と KEMC(ploy) 共に識別に成功したものを示し, 中央には, EMC が "Anger" を "Disgust" に, "Sadness" を "Fear" に識別したのに対し, KEMC(ploy) は識別を成功したものを示す.

表情動作は額の部分, 目の周囲, 口の周囲の変化から生まれ [2], 主観評価で中央の画像を "Anger" と識別した理由を評価者に問うと, 最初に目の周辺から "Anger" もしくは "Disgust" と判断し, 歯を食いしばっている所から "Anger" と識別したという回答が得られた. これらを踏まえて, 以下に図 2, 3 に対する考察を示す.

- 顔の輝度画像に対して EMC で表情を解析した場合, 第 1 基底は目と眉の変化に反応し, 第 2 基底は口の変化に反応しているが [3], 中央の画像は口の周囲の変化が少なく, 目と眉の変化から中央の画像を "Disgust" と識別した
- KEMC の基底が顔のどの部分に反応しているか不明なので断定は出来ないが, 目と眉に加えて口の周囲の変化にも反応して中央の画像を "Anger" と識別した
- 図 3 も同様に目と眉の変化に良く反応する EMC は "Sadness" と "Fear" の分離が不十分となったのに対し, KEMC(ploy) では口の変化にも反応して識別に成功した

(4)EMC と KEMC の特徴ベクトルの比較

図 4, 5 に EMC と KEMC それぞれの上位 2 基底で open-data の "Anger", "Disgust", "Sadness", "Fear" の特徴ベクトルを各々の固有空間へ写像した結果を示す. ここで x 軸は第 1 基底, y 軸は第 2 基底 の値である. 図 4, 5 から以下のことがわかる.

- EMC では各々写像された特徴ベクトルは "Anger" は "Disgust" に "Sadness" は "Fear" に近い分布となり, それぞれの表情の識別率を低下させた
- KEMC では各々写像された特徴ベクトルは "Disgust" と "Fear" が互いに近く分布しているが, "Anger" と "Sadness" の分布は良好に分離し, 高い識別率が得られた

以上から KEMC は非線形写像関数で特徴空間を写像した高次元空間で表情クラスの分離を行うことで, EMC に比べ各表情クラスの分離に適した固有空間を導出することが出来たと考えられる.

(5)KEMC(ploy) と KEMC(rbf) の識別率の比較

KEMC(ploy) と KEMC(rbf) の open-data に対する識別率を比較する. KEMC(rbf) の open-data に対する識別率は KEMC(ploy) と比較して 12 ポイント低い. KEMC(rbf) の Gauss カーネルは分散 σ^2 を持つ Gaussian Radial Basis Function(RBF) に相当する. すなわち closed-data のような辞書と同一のベクトルが入力された時は他のベクトルと大きく分離され, 100% の平均識別率を得たと考えられる. open-data では辞書と同一のベクトルが入力されることはほとんどない, Gauss カーネルの性質を生かすためには, 辞書をより多く, さらに精度の高いものが必要と考えられる. KEMC(rbf) の識別率が低い表情は "Anger", "Disgust", "Sadness", "Fear" の様に辞書枚数の少ない表情となったのはこのためであろう. これを確かめるには辞書枚数を各表情で同程度にした状態で, 解析を行う必要があり, 今後の課題である.

表 3　closed-data に対する識別率 [%]

	An.	Di.	Fe.	Ha.	Sa.	Su.	Av.
EMC	81.8	100	80	100	100	100	93.6
KEMC(poly, p=13)	100	91.0	86.6	100	100	100	96.3
KEMC(rbf, $\sigma = 3.0$)	100	100	100	100	100	100	100
KPCA(poly, p=10)	59.1	13.6	53.3	100	56.5	100	65.4
KPCA(rbf, $\sigma = 5.0$)	100	100	100	100	100	100	100

表 4　open-data に対する識別率 [%]

	An.	Di.	Fe.	Ha.	Sa.	Su.	Av.
EMC	57.1	46.6	45.5	84.2	56.3	81.8	61.9
KEMC(poly, p=13)	78.5	60	63.6	84.2	81.3	95.5	77.2
KEMC(rbf, $\sigma = 3.0$)	64.4	53.3	45.5	81.3	56.3	90.9	65.2
KPCA(poly, p=10)	35.7	6.6	36.4	63.2	50	77.3	44.9
KPCA(rbf, $\sigma = 5.0$)	0	6.6	45.5	68.4	37.5	95.5	42.3

図 2　An. , Di. における KEMC(ploy) と EMC の識別結果の違い
左から　EMC ; An. → An. : An. → Di. : Di. → Di.
KEMC(ploy) ; An. → An. : An. → An. : Di. → Di.

図 3　Sa. , Fe. における KEMC(ploy) と EMC の識別結果の違い
左から　EMC ; Sa. → Sa. : Sa. → Fe. : Fe. → Fe.
KEMC(ploy) ; Sa. → Sa. : Sa. → Sa. : Fe. → Fe.

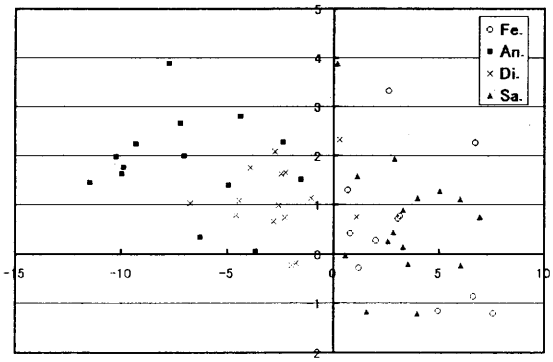

図 4　open-data に対する上位 2 基底の分布; EMC

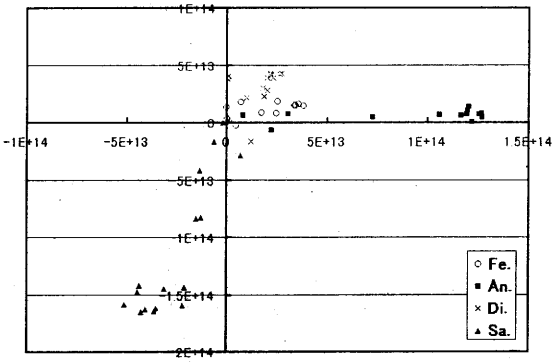

図 5　open-data に対する上位 2 基底の分布; KEMC

4. ま と め

本稿は表情に対して優れた解析特性を有する EMC において識別精度が低下していた表情クラスが線形分離の適応限界を越えていると仮定し, 線形分離が難しい表情クラスに対しても分離可能な非線形解析手法 (KEMC) を検討した. KEMC は EMC をベースに, カーネルトリックを用いて与えた.

KEMC を用いた表情識別実験の結果, EMC で識別精度の低かった "Sadness" と "Anger" で 18～25 ポイント高い識別率を得た. さらに平均識別率も約 15 ポイント高い結果を得た.

今後の課題として, Gauss カーネルを同程度の辞書数で解析を行い, 表情解析への適応性を検討することが挙げられる. さらに他の種類のカーネル関数の適用も検討したい.

文　献

[1] 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.

[2] P. Ekman, W. V. Friesesn 著, 工藤 力編, "表情分析入門-表情に隠された意味をさぐる", 誠信書房, 1987

[3] 藪井 智彦, 剣持 雪子, 小谷 一孔, "3 次元顔形状を用いた固有空間法による 3 次元固有顔の導出と表情解析", 信学技報, PRMU2002-120, HIP2002-27, pp. 7-12, Nov. 2002.

[4] 麻生英樹, 津田宏治, 村田昇, "パターン認識と学習の統計科学-新しい概念と手法", 岩波書店, pp102-106, 2003

[5] M. -H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods", Proc. of Int'l Conf. on Automatic Face and Gesture Recognition, pp. 215-220(2002)

[6] Q. Liu, R. Huang, H. Lu, and S. Ma, "Face recognition using kernel based Fisher discriminant analysis", Proc. of Int'l Conf. on Automatic Face and Gesture Recognition, pp. 197-201(2001)

[7] B. Schölkopf, A. Smola and K. R. Müller, "Nonlinera Component Analysis as a Kernel Eigenvalue Problem", Neural Computation, vol. 10, pp. 1299-1319, 1998.

[8] 前田 英作, 村田 洋, "Support Vector Machine とカーネル部分空間法", 信学技報, PRMU1998-81, pp. 59-66, 1998.

[9] C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery, Kluwer Academic Publishers, pp. 1-43, 1998.

[10] 新座 良和, 剣持 雪子, 小谷 一孔, "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", IMPS99, pp. 63-64, Dec. 1999.