原著論文

感情の次元説に基づくファジィメンバシップ関数の 顔表情画像からの推定法†

有田 繭子*・越智 晴香*・福島 重廣*・坂本 博康*

本論文は、顔表情画像から感情・表情のファジィメンバシップ関数を推定する方法を提案する。このメンバシップ関数は感情の次元説における感情空間の上で定義され、学習用の顔表情画像に対する主観評価の実験結果などを利用して、次の2つのモデルで表現される。a) 各顔画像におけるメンバシップ関数の平均と広がりパラメータを、それぞれ主観評価の平均値と標準偏差とするもの。b) 主観評価に加えて画像空間内の分布の情報を導入したもの。推定されたメンバシップ関数は、もとの関数からの形状近似における誤差などを用いて評価される。統計解析の手法としてカーネル化を含む正準相関分析を利用し、これを濃淡画像データに適用した、男女2種類の顔表情画像のデータベースを用いた実験の結果、人の主観評価と比べ最大で約2倍以下のばらつきを持つ推定結果が得られ、提案方法はこの意味で有効であることが示された。

キーワード：感情、顔表情画像、ファジィメンバシップ関数、正準相関分析

1. はじめに

コンピュータにより非侵襲に顔表情画像から人の感情を認識し計測する技術は幅広い応用が期待されており、現在積極的に研究が進められている。その応用例として身動きや会話の不自由な高齢者の介護、作業による疲労感やストレスの検出、エンターテインメントなど、多様な分野を挙げることができる。以下では簡単のために、顔表情と感情が十分に対応付け可能という仮定の下で、特に区別を要しない場合には、両者を同一視した記述を行うことにする。

感情や表情のコンピュータによる自動認識に関する研究のうち、顔表情画像を喜び、悲しみ、怒り、驚き、恐れ、嫌悪の6種または4種の基本感情のクラスに識別する手法については、多数の先行研究が行われている[1]–[4]。この基本感情クラスへの分類法は、表情に関する認知心理学の分野でカテゴリ説と呼ばれており、多くの有用な表情の自動認識手法の基礎となっている。

感情の定義に関して、心理学の分野では一般に認められた標準的な定義は存在しないといわれているが[5]、上の基本感情が広く一般に感情と呼ばれること

も多い。

このカテゴリ説に対して、感情を表す空間内に快–不快、覚醒–沈静といった数本の座標軸をとり、そこに基本感情などを配置する記述法は次元説と呼ばれ、図1のRussellの感情円環がその代表例として挙げられる[6]。

カテゴリ説と次元説の間では従来から対立的な議論が行われているが[7]、カテゴリ説には以下の欠点を挙げることができる。基本6感情をポジティブ、ネガティブな感情に分けると、前者は「喜び」と「快を含む驚き」だけであり、残り全部は後者に属するので、偏りが大きい。ここで、驚きの基本感情は快と不快の感

図1 Russellの感情円環モデル

† A Method for Estimating Fuzzy Membership Functions of Emotion Based on Its Dimensional Theory from Facial Expression Images

Mayuko ARITA, Haruka OCHI, Shigehiro FUKUSHIMA and Hiroyasu SAKAMOTO

* 九州大学
Kyushu University

情の両面を持ち、これとは反対に、悲しみを含む笑いのような、複数の基本感情の混合または中間的なものも存在することが分っている[7]。このように、基本6感情に識別できたとしても、さらにその区別や混合が必要となる場合も起こる。

他方、次元説の感情空間上のファジィメンバシップ関数を利用した感情の表現法によれば、上記の偏りの問題は起こらず、また基本感情の区別、混合の必要性に対しても比較的容易に応えることが可能になる。更に、感情や表情の現象には、その表出者や観察者のいづれにも曖昧さや個人間のばらつきが存在するので、メンバシップ関数による表現に適していると言われている。従って、この表現による感情の推定法は、その結果を上述のような分野へ応用する際、カテゴリ説に基づく結果に比べて、より高い有用性を持つことを期待できるだろう。

次元説に基づく表情や感情の推定法の従来研究はあまり見られないが、文献[8]、[9]はともに顔部品の変形などの特徴量をベースに、その感情空間へのマッピング法を示すと同時に、基本表情のカテゴリへの分類法も提案している。文献[10]では顔表情の濃淡画像をベースに、各画像に対する快-不快、覚醒-沈静の2種類の主観評価値による学習に基づいて、入力画像に対する2つの評価値を推定する方法を提案し、被験者の主観評価値のばらつきと同程度の推定精度を得ている。

一方、ファジィ理論を利用した顔表情や感情に関する従来研究として、8種の基準感情カテゴリを定義域とするメンバシップ関数を利用した顔表情画像の生成法[11]、基本6(または4)表情への分類のために、顔部品の変形の程度や基本表情の強度を定義域としてファジィ化されたメンバシップ関数を適用した方法[12]、[13]などが報告されている。しかし、次元説の感情空間上におけるメンバシップ関数を利用したものは見られない。

そこで本論文では、次元説の快-不快、覚醒-沈静の感情空間で定義されたメンバシップ関数を利用して、顔の濃淡画像ベースでこの関数を学習し推定する方法を提案する。この関数の推定結果から、その最大値の位置や関数の形状として、顔の表情と感情の状態を知ることができる。この状態には前述の基本感情の区別や混合の情報も含まれているので、本研究では基本感情への識別を最終目的とはしない。従って本手法の推定結果の評価には、従来法のような基本感情への識別率ではなく、メンバシップ関数の近似度を利用する。

ここでは2種類のメンバシップ関数のモデルを採用する。(a)被験者の顔表情画像に対する快-不快、覚

醒-沈静の2座標軸の主観評価結果だけに基づくモデル、(b)これに加えて顔表情の画像空間上における分布状況を加味したモデル。さらに前者では次の3通りのモデルを用いる。(a1)2種類の主観評価値をファジィ数としたときの平均と広がりのパラメータによる4次元ベクトル、(a2)2つの主観評価値それぞれについて、(a1)と同じパラメータを持つガウス型関数の標本値のベクトル、(a3)2種類の主観評価値について、(a1)と同一パラメータを持つ2変数ガウス型関数の標本値のベクトル。(b)では(a2)に加えて、顔画像データベースの全画像を主観評価値によって分類した後、各クラスの平均画像から測った対象画像の距離を同時に考慮する[14]。人の顔表情に対する主観的な印象に加えて、顔画像空間における客観的な距離情報に基づくファジィ化[17]を組み込むことにより、メンバシップ関数の信頼性を高め、その推定誤差の低減を期待できる。

統計の手法として本論文では、正準相関分析法(CCA)とカーネル化されたCCA(KCCA)を用い、これらの学習とテストにおいてはleave-one-out cross validation法(以降はLOO法)を用いた。実験のためには、JAFPE女性表情顔画像データベース[15]と、本学及び文献[16]等の複数の機関から収集した男性顔データベースを用いた。

以下本論文では、2章でCCAとKCCAの解析法を簡単に述べ、3章では、顔画像データベースとその主観評価実験法について述べる。4章ではメンバシップ関数の構成方法について説明する。5章で実験方法について述べ、その結果を示す。6章ではまとめを述べる。

2. 正準相関分析とカーネル正準相関分析

2.1 正準相関分析(CCA)[19]

CCAは、ゼロ平均の2変量 x_n と y_n ($n=1, \dots, N$) の線形な相関関係を解析するための手法として広く知られている。本実験では、 x_n を n 番目の顔表情画像、 y_n をそれに対応するメンバシップ関数の情報を並べたベクトルとして解析を行う。各変数の学習データは行列 $X=[x_1, x_2, \dots, x_N]$ および $Y=[y_1, y_2, \dots, y_N]$ に統合される。 X と Y のそれぞれまたは相互の共分散行列を $\Sigma_x = XX^T/N$, $\Sigma_y = YY^T/N$, $\Sigma_{xy} = XY^T/N$, $\Sigma_{yx} = YX^T/N$ とすると($\cdot^T$ は転置)、CCAの変換行列 A, B は次の一般固有方程式から求められる。

$$\Sigma_{xy} \Sigma_y^{-1} \Sigma_{yx} A = \Sigma_x A \Gamma^2 \quad (1)$$

$$\Sigma_{yx} \Sigma_x^{-1} \Sigma_{xy} B = \Sigma_y B \Gamma^2 \quad (2)$$

Γ は、(1)、(2)式の共通の正準相関値を対角要素とする対角行列である。 A, B, Γ を学習した後、テストデー

タ $\mathbf{x}_{\text{test}}$ のメンバシップ関数 $\mathbf{y}_{\text{test}}$ は、次式で推定できる。

$$\mathbf{y}_{\text{test}} = (\mathbf{A}\mathbf{\Gamma}\mathbf{B}^{-1})^T \mathbf{x}_{\text{test}} \quad (3)$$

サンプル数の不足から共分散行列 $\mathbf{\Sigma}_x, \mathbf{\Sigma}_y$ の一方または両方が特異行列となる場合も多いが、その際は主成分分析(PCA)によるCCAの分解手法を用いて、上式は

$$\mathbf{y}_{\text{test}} = (\mathbf{W}_x \mathbf{A}_x^{-1/2} \mathbf{A}\mathbf{\Gamma} (\mathbf{W}_y \mathbf{A}_y^{-1/2} \mathbf{B})^{-1})^T \mathbf{x}_{\text{test}} \quad (4)$$

と変形される。ここで、 $\mathbf{A}_x, \mathbf{A}_y$ は $\mathbf{X}, \mathbf{Y}$ にそれぞれPCAを適用したときの正固有値のみで作った対角行列であり、 $\mathbf{W}_x, \mathbf{W}_y$ は $\mathbf{A}_x, \mathbf{A}_y$ のそれぞれに対応する固有ベクトルを並べた直交行列である。 $\mathbf{A}$ と $\mathbf{W}$ の次元を減らすことにより、指定の次元まで情報圧縮を行うこともできる。

2.2 カーネル正準相関分析(KCCA) [20]

前項のCCAに対し、2変量が非線形な関係を持つ場合は、KCCAが利用されることが多い。KCCAの変換行列 $\mathbf{A}, \mathbf{B}$ は、カーネル関数 $k(\mathbf{x}_i, \mathbf{x}_j)$ を第 (i, j) 要素とするカーネル行列 $\mathbf{K}$ とメンバシップ関数の行列 $\mathbf{Y}$ のそれぞれまたは相互の共分散行列 $\mathbf{\Sigma}_k, \mathbf{\Sigma}_y, \mathbf{\Sigma}_{ky}, \mathbf{\Sigma}_{yk}$ による以下の一般固有方程式で求められる。

$$\mathbf{\Sigma}_{ky} \mathbf{\Sigma}_y^{-1} \mathbf{\Sigma}_{yk} \mathbf{A} = \mathbf{\Sigma}_k \mathbf{A} \mathbf{\Theta}^2 \quad (5)$$

$$\mathbf{\Sigma}_{yk} \mathbf{\Sigma}_k^{-1} \mathbf{\Sigma}_{ky} \mathbf{B} = \mathbf{\Sigma}_y \mathbf{B} \mathbf{\Theta}^2 \quad (6)$$

$\mathbf{\Theta}^2$ は非負のカーネル正準相関からなる対角行列である。KCCAの変換行列などを学習した後、テストデータ $\mathbf{x}_{\text{test}}$ のメンバシップ関数 $\mathbf{y}_{\text{test}}$ は次式で推定される。

$$\mathbf{y}_{\text{test}} = (\mathbf{A}\mathbf{\Theta}\mathbf{B}^{-1})^T \mathbf{k}(\mathbf{X}, \mathbf{x}_{\text{test}}) \quad (7)$$

ここで、 $\mathbf{k}(\mathbf{X}, \mathbf{x}_{\text{test}})$ は $\mathbf{X}$ の各列と $\mathbf{x}_{\text{test}}$ によるカーネル関数の N 次元ベクトルである。なお、KCCAの $\mathbf{\Sigma}_k$ は正則行列のことが多いが、次式による次元削減も可能である。

$$\mathbf{y}_{\text{test}} = (\mathbf{W}_k \mathbf{A}_k^{-1/2} \mathbf{A}\mathbf{\Theta} (\mathbf{W}_y \mathbf{A}_y^{-1/2} \mathbf{B})^{-1})^T \mathbf{k}(\mathbf{X}, \mathbf{x}_{\text{test}}) \quad (8)$$

なお本論文では、次式のガウスカーネルを用いている。

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2) \quad (9)$$

3. 顔表情画像と主観評価実験

実験には女性の顔画像(DB1)、男性の顔画像(DB2)の2つのデータベースを用いる。DB1には JAFFE データベース [15] の日本人女性10名の7表情(提供されたファイル名に従って、笑、怒、泣、驚、恐怖、嫌

悪、中立)の全213枚の画像が含まれる。DB2には、本学学生19名の4表情(笑、怒、泣、驚)と、外部機関 [16] 等から収集された41名4表情の合計144枚の画像が含まれる。全画像とも256レベルのグレースケール画像である。

これらの画像を、両眼の中心と上唇の下端中央の位置がそろうように、アフィン変換によって変形し、64×64画素のサイズにそろえた。また、全画像の平均明度とコントラストを、それぞれ平均レベル128、および明度の分散35²として正規化した。図2に各データベースの正規化画像の例を示している。

このように男女別々のデータセットの構成とした理由の第1は、撮影条件、単一人種、表情の種類と強度などのDB1(JAFFE)の高い均質性に対して、多機関から収集されたDB2では上記の均質性が非常に低いという事実が挙げられる。この事実は以降の実験結果に表れる。これに加え、筆者らによる顔表情画像の主成分分析実験の経験より、男女間には第1主成分ベクトルが逆符号となるような相違があることから、今回は別々に利用した。

上記のデータの均質性について説明する。DB1は屋内の制御された撮影条件下で、日本人女性が各表情当たり3～5段階の強度で表情表出した画像から構成されている。他方、DB2のうち各人物ごとに基本4表情がそろい撮影条件が比較的良好なのは本学学生の19名×4表情=76枚だけである。海外機関からの画像68枚は、黄色人種を中心として白人や黒人も含む41名分であるが、日本人でも見かけることの多い顔で、しかも基本4表情に分類可能な画像だけを選んだ。しかしその原画像では照明方向や強度、陰影、明度とコントラスト、画質やノイズ、分解能、背景、年齢や顔面とカメラの角度などが多様である。各人当りの基本表情数は1(24名、うち2名は強度の異なる同一表情画像が2枚)、2(11名、うち1名は同上の画像が2枚)、3(6名、うち1名は同上の画像が2枚)であり、基本4表情のそろう人物は得られていない。

図2 顔表情画像の例
(a)DB1, (b)DB2

図3 主観評価実験の画面

メンバシップ関数を構成するため、全画像について図3の画面を被験者に提示しながら主観評価実験を行った。この画面は図2の正規化以前の顔全体をDB1, DB2それぞれにランダムな順番で表示するが、被験者は左上のボタン群を用いた順／逆／位置指定の送りにより、任意の画像へ移動できる。画面右側には上から、操作説明の窓、快・不快度と覚醒・沈静度の指定用窓、被験者名の入力用窓、実験終了とデータ保存ボタン等が配置されている。

各被験者には実験開始直前に感情の次元説およびRussellの感情円環について簡単な説明を行っている。被験者は、各表情画像の快－不快、覚醒－沈静の主観評価値を、それぞれ+3から-3の整数7段階で回答した。DB1に対する被験者数は16名、DB2のそれは18名である。各画像について、主観評価値の全被験者に関する平均値とその標準偏差をメンバシップ関数のパラメータとして利用する。DB1全体の統計量を求めると、[快不快の平均値(a), 覚醒沈静の平均値(b); 快不快の標準偏差の平均値, 覚醒沈静の標準偏差の平均値; (a-快不快の評価値)の絶対値の平均値, (b-覚醒沈静の評価値)の絶対値の平均値]を並べたベクトルは $[-0.613, 0.499; 0.626, 1.049; 0.474, 0.813]$ であり、同様にDB2のものは $[-0.233, 0.620; 0.774, 1.100; 0.604, 0.864]$ であった。これらの数値にも、上述のDB1, DB2におけるデータの均質性の違いが表れていることが分かる。

4. メンバシップ関数の作成

本手法で使用するメンバシップ関数のモデルの2種類(a), (b)について、その概略を1章で述べた。メンバシップ関数としてよく利用される関数には、台

形、三角形、ガウス関数形などいろいろあるが、ここではその広がりのパラメータの計算に標準偏差を用いるので、ガウス関数の採用が妥当と思われる。以下本章では各モデルにおける具体的なメンバシップ関数の計算法について述べる。

(a1) メンバシップ関数パラメータの4次元ベクトル

各画像に対する、快－不快、覚醒－沈静の主観評価値と標準偏差の4変数を4次元のベクトルに並べて、メンバシップ関数のパラメータのベクトル y とする。

(a2) ガウス型メンバシップ関数

連続なガウス型の関数を利用し、その平均値と標準偏差として(a1)と同じ値を当てはめた曲線の標本値をとり、その最大値を1と正規化することにより、離散的なメンバシップ関数のベクトルを生成する。このとき、標本点の個数を主観評価実験における被験者の評価レベル数(7)と同じにする必要はないので、本実験では、標本点の個数として快－不快、覚醒－沈静の各軸で13, 11, 9, 7, 5の5通りの場合に対して実験を行った。13点の場合は、ガウス型関数を0.5間隔に $\{-3, -2.5, -2, \dots, +2.5, +3\}$ の点で標本化する。11, 9, 7, 5点を指定する場合の標本間隔は、それぞれ0を中心の標本点として0.6, 0.75, 1, 1.5にとられた。各軸のガウス型関数の標本値を並べてメンバシップ値のベクトル y とする。標本点数が13, 11, 9, 7, 5のベクトル y の次元は、主観評価値の2軸を合わせて、それぞれの2倍となる。全部のメンバシップ関数において、その値域は $[0, 1]$ である。

(a3) 2変数ガウス型メンバシップ関数

前項(a2)と同様の平均値と標準偏差を用いて、2変数の連続なガウス型関数を作成する。本実験の2軸の主観評価値においては最大でも0.2以下の小さな相関係数しか観測されなかったもので、ここでは簡単のために次式の無相関2変数ガウスメンバシップ関数を利用する。

$$p(x) = \exp \left[-\frac{1}{2} \left\{ \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right\} \right] \quad (10)$$

なお、顔表情の認知心理学的研究では、多数の基本表情の分類データに多次元尺度法(MDS法)を適用した結果として、快－不快と覚醒－沈静の座標軸が得られており、この2軸は直交する[7]。本実験ではMDS法を使用しないが、実験開始前の被験者への説明によって、直交に近い主観評価値が得られていると考えられる。

標本点の個数としては、(a2)と同様に各座標軸あたり13, 11, 9, 7とする。この2変数メンバシップ関数を1次元ベクトルに並べ y とすれば、その長さはそれぞれ169, 121, 81, 49となる。この場合も、全メンバシップ関数で、その値域は $[0, 1]$ である。

(b) 画像空間のデータ分布に基づくメンバシップ関数

以上では、主観評価の平均値とその標準偏差によってメンバシップ関数を作成したが、ここでは主観評価結果と顔表情の画像空間上におけるデータの分布を考慮したモデルを利用する。このようなデータの特徴空間上の分布を考慮した顔画像の個人識別の従来研究例として、顔画像データと各人の平均顔とのユークリッド距離を用いたファジィ化によりメンバシップ関数を作成して利用するものがある[17]。

本実験では、画像空間内の各顔画像と平均顔画像のユークリッド距離に加えて、主観評価値に基づく分類のクラス間距離も同時に考慮したメンバシップ関数を利用する。ここで前者の平均顔を後者の分類のクラスごとに求める必要があるため、各クラスでなるべく同数の顔画像を含むことが望ましい。いま、(a2)にならない、全被験者の主観評価の平均値における閾値間隔1で7個のクラスに分類すると、特にこの値の絶対値が大きいクラスに属するデータの最小個数が5個以下となり有効な平均顔を求められない場合が生じた。これを避けるために、5クラスに分類して処理を進めた。この分類に用いた主観評価の閾値は $[-1.8, -0.6, +0.6, +1.8]$ であり、その閾値の符号を反映させたクラスの番号を $[-2, -1, 0, 1, 2]$ とする。

$x_{ij}(i=-2, -1, 0, 1, 2, j=1, \dots, N_i)$ を第 i クラスに属する j 番目の顔画像データとする。ここで、 N_i はクラス i に属するデータ数である。クラス k に属する顔画像の平均画像を m_k とし、 x_{ij} とのユークリッド距離 $d_{ij}(k) = \|x_{ij} - m_k\|$ を求める。また、主観評価値に基づく距離として、 $s_{ij}(k) = 0.5|i - k|$ を導入する。 s の値は、そのデータが属するクラスでは0、それ以外のクラスでは主観評価値のクラスの番号の差に応じた正の値となる。以上の2種類の距離から、重み a 付きの和を次式で求める。

$$D_{ij}(k) = d_{ij}(k) + a s_{ij}(k) \quad (11)$$

主観評価の2軸について得た合計10個の上式の値からベクトル D_{ij} を作り、そのノルムを1に正規化する。正規化後の D_{ij} の要素を用いて、メンバシップ関数の値 $y_{ij}(k)$ を次式で計算する。

$$y_{ij}(k) = \begin{cases} 1, & k = i \\ -\frac{1}{\max_j D_{ij}(k)} D_{ij}(k) + 1, & k \neq i \end{cases} \quad (12)$$

このとき、データ x_{ij} が属するクラスへのメンバシップ値は1、最も帰属度の低いクラスへのメンバシップ値は0となる。快-不快、覚醒-沈静の2軸の $y_{ij}(k)$ を並べた10次元ベクトル y_{ij} をメンバシップ関数のベクトルとする。

5. CCAとKCCAによるメンバシップ関数の推定結果

前章の方法によって、4通りのメンバシップ関数のベクトル y を作成した。ここからは、CCAとKCCAによるメンバシップ関数のパラメータまたはその形状の推定実験について述べる。

両データベースに含まれる顔表情画像 x は、平均値を差し引いて0平均とし、画像のサイズ $64 \times 64 = 4096$ のベクトルとする。 x, y の学習データを並べた行列 X, Y を作成し、LOO法を用いてCCAもしくはKCCAによる解析を行う。このとき、式(1), (2)における Σ_x の次元は最大で(学習枚数-1)となる。そのため、2章で述べたPCAによるCCAの分解手法を用いて、DB1, DB2における Σ_x, Σ_k の次元を、ノイズの影響も考慮して、それぞれの最大値である211, 142に近い値の210, 140とした。また、KCCAでは式(9)のガウスカーネルを用いるが、その σ^2 の値は式(9)の分子の値に見合う値の範囲($50^2 \sim 200^2$)で実験的に良好な結果を与える12,000(DB1)と25,000(DB2)にそれぞれ設定された。後者が前者に対して約2倍の値をもつことは、分子の比較的に大きな値(画像データDB2の不均質性)を抑圧することにより、より良好な結果を得ている可能性を示唆している。

なお、(a2), (a3)および(b)のメンバシップ関数では、学習段階におけるメンバシップ関数値の範囲は $[0, 1]$ となっている。しかし、CCAもしくはKCCAによる解析においては、推定値がその範囲外となる場合もある。本実験においては、全データのメンバシップ関数の推定値の最小値は-1.3、最大値は+1.4程度のもので出現した。 $[0, 1]$ の範囲を超える場合は、メンバシップ関数の定義に従わせるために、その値を0または1に置き換えた。

以下では、メンバシップ関数ごとの推定結果について述べる。

(a1) メンバシップ関数パラメータの4次元ベクトル

図4に、DB1のメンバシップ関数の4つのパラメー

タの推定結果を示した。図4 (a), (b)の縦軸はそれぞれCCAとKCCAによる推定値を、横軸はすべて主観評価結果の値を表す。また、(a), (b) 両図ともに上段は主観評価の平均値、下段は標準偏差値に対する推定結果である。このプロットの相関係数が1に近いほど推定精度の高さを表すので、両図ごとにそれぞれ[左上, 右上, 左下, 右下]の順で相関係数の値を求めると、(a) 図で [0.922, 0.887, 0.345, 0.5737], (b) 図で [0.936,

0.905, 0.384, 0.674] が得られた。主観評価の平均値では文献[10]同様の良好な推定結果が得られたが、標準偏差では推定不可能と思われる。

表1には、両データベースの各パラメータの推定における誤差の絶対値の平均値を示した。図4でも見て取れるが、表1においてCCAとKCCAの絶対値誤差を比較すると、1項目(DB2の快不快主観評価値)を除き、両データベースともKCCAの性能の方が優れている。

また、両データベースを比較すると、DB1よりDB2のほうが大きい推定誤差となっている。この原因は、3章で述べたような両データベースにおける画像データの均質性の違いにあると思われる。つまり、DB1に含まれる女性の顔表情画像は良く制御された撮影条件下で、同一人物の同種の表情で強度の違う画像も含まれており、被写体の人数自体もDB2の60名に対して、10名と少ない。そのため、DB1においては、学習データ中にテストデータと類似度の高いデータを含むこととなり、精度が上がったと考えられる。

表1 (a1)の推定値における誤差の絶対値の平均値

評価項目	DB1		DB2	
	CCA	KCCA	CCA	KCCA
快-不快 主観評価値	0.406	0.375	0.756	0.759
覚醒-沈静 主観評価値	0.432	0.407	0.685	0.664
快-不快 標準偏差	0.157	0.146	0.177	0.164
覚醒-沈静 標準偏差	0.262	0.229	0.228	0.218

(a) CCAによる推定結果

(b) KCCAによる推定結果

図4 4次元パラメータの推定結果

図5 固有値数の増加による推定誤差の減少

また図5には、DB1に対してKCCAを適用する際の、PCAの次元数の増加による平均推定誤差の減少の様子を示した。覚醒-沈静の推定誤差の方が快-不快に比べて減少率が少し緩やかであるが、それでも固有値数150を超えても推定誤差は少しずつ減少する。線形処理であるCCAやPCAにおいては、最大固有値から40個程度を用いれば十分小さい近似誤差を達成することが知られている[10], [18]。この相違の原因は、KCCAが非線形処理手法であることと考えられる。

(a2) ガウス型メンバシップ関数

図6には、ガウス型メンバシップ関数の推定結果の例を示した。これは、図6(a)に示したDB1の2枚とDB2の1枚の顔表情画像をテストデータとした場合、それぞれ図6(b), (c), (d)に、本来の推定対象のメンバシップ関数(太実線)と、推定結果のメンバシップ関数(マーカ付)をプロットしている。各図の左列が快-不快の推定結果、右列が覚醒-沈静の推定結果を示す。それぞれのガウス型メンバシップ関数を、0.6の間隔で11点で標本化した場合を図示している。若干のピーク位置のずれが見られるものの、どの画像のメンバシップ関数の推定結果も太線を良好に近似したものが得られている。

また表2の上部(i)には、図6の3枚の画像に対するメンバシップ関数の推定結果における誤差の絶対値の和をサンプル数11で割った平均値と、データベースごとの全画像に対する平均誤差を示した。DB1を用いた推定については、絶対値誤差の平均が0.2以内に収まっている。メンバシップ関数としてとる値が[0, 1]の範囲であるので、DB1に関しては比較的良好な結果が得られていることがわかる。DB2に関しては、DB1の推定誤差より大きくなった。この原因もDB1とDB2におけるデータの均質性の違いによるものと考えられる。また、CCAとKCCAの両データベースの結果を比較すると、快-不快、覚醒-沈静の軸ともに非線形処理であるKCCAの推定精度の方が高い。2軸の推定結果を比較すると、DB2のCCAによる推定以外で、快-不快の軸の方が推定誤差は小さくなった。

更に以上の誤差の程度と比較するために、表2(ii)には、本手法における推定対象のメンバシップ関数の信頼性を被験者間の絶対値誤差として示した。この誤差は、被験者を半数のグループに分けるすべての場合(DB1では ${}_{16}C_8$ 通り)について、グループごとに平均化した2つのメンバシップ関数の絶対値誤差の平均値を求めたものである。

表2(ii)の4個の数値によって、表2(i)のうち下2行の8個のそれぞれ対応する数値を割れば、最小値は

図6 ガウス型 2×11 次元メンバシップ関数の推定結果。横軸はいずれもガウス型メンバシップ関数の標本点の座標。太実線：テストデータのメンバシップ関数、○と実線：CCAによる推定値、*と破線：KCCAによる推定値。

0.939, 最大値は1.89が得られる。従って、本手法は被験者間に見られるメンバシップ関数のばらつきの2倍以下のばらつきで推定可能であることがわかる。

以上の結果は、ガウス関数を主観評価値の0.6間隔ごとに11点で標本化したメンバシップ関数を用いた場合のものであった。図7は、他に5, 7, 9, 13点で標

本化した場合の、推定誤差の比較を示したものである。本実験では、標本数を主観評価実験の段階数である7以上にした場合の推定誤差に大きな違いは見られなかったが、逆に標本数を減らしたときに、多少誤差が減少している。しかし、KCCAについては、CCAに比べて推定誤差への影響はかなり少なかった。

表2 ガウス型メンバシップ関数の推定(11標本点数)における(i)絶対値誤差の平均値と標準偏差,(ii)被験者間のメンバシップ関数の絶対値誤差の平均値。

対象データ	快-不快		覚醒-沈静	
	CCA	KCCA	CCA	KCCA
(i) 推定実験の絶対値誤差, その平均値と標準偏差				
DB1 #2	0.0744	0.0863	0.206	0.168
DB1 #120	0.224	0.137	0.108	0.0731
DB2 #122	0.248	0.203	0.175	0.144
DB1 平均値 (標準偏差)	0.169 (0.0698)	0.125 (0.0517)	0.189 (0.0894)	0.155 (0.0705)
DB2 平均値 (標準偏差)	0.225 (0.0959)	0.201 (0.0851)	0.220 (0.0112)	0.205 (0.0970)
(ii) 被験者間の絶対値誤差				
DB1	0.108		0.165	
DB2	0.119		0.159

図7 クラス数の変化による絶対値誤差の比較(DB1). 実線: 快-不快, 破線: 覚醒-沈静, *: CCA, ○: KCCA.

(a3) 2変数ガウス型メンバシップ関数

次に、2変数ガウス型メンバシップ関数の推定例を図8に示す。図6(a)と同一の画像をテストデータとして用い、不快-快、沈静-覚醒をそれぞれ11クラスに分割している。グラフの左から右に向かって不快-快の主観評価値の-3から+3を、手前から奥に向かって沈静-覚醒の-3から+3を表す。左列が本来(推定目標)のメンバシップ関数、右列がKCCAによる推定結果である。

実際のメンバシップ関数の最大値が1であるのに対し、推定値の最大値は0.8程度と低めに出る場合が多く、全体になだらかなメンバシップ関数として推定される傾向が見られる。これは、CCAとKCCAの手法が

(a) DB1 #2

(b) DB1 #120

(c) DB2 #122

図8 2変数ガウス型メンバシップ関数の推定。図6(a)と同じ画像に対する結果。左列: テストデータの推定目標のメンバシップ関数, 右列: KCCAによる推定結果。

原理的に基底関数の重ね合わせで関数近似を行う際に、多数の関数の平均化操作を伴うことから起こると考えられる。しかし、推定された関数の形状としては類似の関数が得られている。DB1 #2とDB2 #122については不快の傾向が、DB1 #120については快の傾向が、正しく表現されている。また、不快表情の2枚について、覚醒-沈静の実際のピーク値は、DB1 #2で-3から+3のうちの+1.8、DB2 #122では-0.6となっているが、どちらの画像に関しても、推定値とピーク値の一致が見られた。覚醒-沈静の頂点の推定についても、ある程度の精度が得られていると言える。

問題点としては、(a2)のメンバシップ関数を利用した場合と比較して、(a3)では多峰性を持つメンバシップ関数として推定される場合が増えた、ということが挙げられる。その原因は推定する変数が多いからと考えられるが、このような場合は、小さい峰のメンバシップ値は0とするような処理が必要かも知れない。

表3 (i) にはメンバシップ関数とその推定値との絶対値誤差を示した。図8の3画像に対する絶対値誤差はもとより、DB1全体においても0.09以下、DB2でも約0.12以下となり、(a2)の実験結果と比較しても全体的に小さくなっている。表3 (ii) の被験者間の絶対値誤差は表2 (ii)と同じ方法で求められた。この表3 (ii)と、表3 (i)の推定結果の絶対値誤差との比率は、最小値1.06から最大値1.59の間にあり、人との比較の意味において、本推定方法は有効であると考えられる。

表3 2変数ガウスメンバシップ関数の推定 (標本点数 11²)。推定結果の絶対値誤差と被験者間の絶対値誤差。

対象データ	CCA	KCCA
(i) 実験結果の絶対値誤差, その平均値と標準偏差		
DB1 #2	0.0936	0.0651
DB1 #120	0.0624	0.0463
DB2 #122	0.158	0.106
DB1 平均値 (標準偏差)	0.0875 (0.0244)	0.0696 (0.0229)
DB2 平均値 (標準偏差)	0.121 (0.0404)	0.109 (0.0381)
(ii) 被験者間の絶対値誤差		
DB1	0.0659	
DB2	0.0759	

なお、(a3)のメンバシップ関数についても(a2)と同様に、主観評価値の標本間隔を変更した実験も行ったが、こちらも推定精度に大きな変動は見られなかった。

(b) 画像空間のデータ分布に基づくメンバシップ関数

ここからは、以上で用いた主観評価結果に加えて、顔表情の画像空間上の分布を考慮したメンバシップ関数(b)を推定した結果について述べる。(11)式中の重み係数 α の値は、1 から 3 まで0.5おきの試行の結果、両データベース、CCA, KCCA, 快-不快, 覚醒-沈静のすべての場合で総合的に最良となった2と設定した。

まず図9に、快-不快, 覚醒-沈静の両軸を5段階ずつに分けた場合のDB1の平均顔を示す。各平均顔画像には主観評価値の強度ごとの特徴が見られる。各画像のメンバシップ関数を計算する際には、これらの平均顔からのユークリッド距離を計算して利用する。次に、図10には図6 (a)と同じ3枚のテスト画像に対して推定されたメンバシップ関数を示している。更に、表4 (i) には図10の3画像に対する推定結果における絶対値誤差、および2つのデータベースに対する絶対値誤差の平均値を示す。

この実験においても、CCAよりKCCAによる推定の方が高い性能を示した。特に、DB1の快-不快の軸に関するKCCAにおいては、絶対値誤差が0.0924と0.1以下となっており、高い推定精度が得られている。(b)のメンバシップ関数を用いた場合は、DB1, DB2またはCCA, KCCAともに、快-不快の推定精度は覚醒-沈静より高い。図9を見ると、覚醒-沈静の平均顔は強い特徴を表しているものの、様々な表情が混ざったようにも見える。特に、最も覚醒度の高いクラスの平均顔は、驚き顔の要素と喜び顔の要素の2つを強く含んでいる。覚醒-沈静の軸に関しては、平均顔との距離によって特徴量を決定することが難しくなったものと考えられる。

図9 主観評価値のクラスごとの平均顔 (DB1)。いずれも左から右へ-2, -1, 0, 1, 2のクラスラベル。

(a) DB1 #2

(b) DB1 #120

(c) DB2 #122

図10 画像空間のデータ分布に基づくメンバシップ関数の推定結果。図6 (a) と同じ表情画像に対するもの。左列：不快-快の推定値，右列：沈静-覚醒の推定値。太い実線：テストデータのメンバシップ関数，○と実線：CCAによる推定値，*と破線：KCCAによる推定値。

また、本項と同様に1変数のメンバシップ関数(a2)による実験結果(表2)と比較すれば、表4のほぼ全部の絶対値誤差が少ないことが分かる。

また、表4 (ii)に示した被験者間の絶対値誤差との比率は、最小で1.06および最大で2.35となり、人におけるばらつきの約2倍以内のばらつきで推定可能なことから、この方法によっても有効な結果が得られると考えられる。

表5には、各画像に対して推定されたメンバシップ

表4 メンバシップ関数(b)の絶対値誤差。

対象データ	快-不快		覚醒-沈静	
	CCA	KCCA	CCA	KCCA
(i) 実験結果の絶対値誤差				
DB1 #2	0.147	0.0284	0.140	0.108
DB1 #120	0.0796	0.0496	0.0990	0.0549
DB2 #122	0.219	0.173	0.185	0.115
DB1 平均値 (標準偏差)	0.133 (0.0787)	0.0924 (0.0611)	0.189 (0.125)	0.141 (0.0901)
DB2 平均値 (標準偏差)	0.192 (0.123)	0.168 (0.106)	0.218 (0.125)	0.193 (0.101)
(ii) 被験者間の絶対値誤差				
DB1	0.0694		0.133	
DB2	0.0816		0.156	

表5 メンバシップ関数(b)の最大値に判別した場合の識別率(%)。

対象データ	CCA		KCCA	
	快-不快	覚醒-沈静	快-不快	覚醒-沈静
DB1	69.9	54.9	73.2	56.8
DB2	52.0	36.1	56.9	41.6

関数が最大値となるクラスに、その画像を識別した場合の識別率を示した。表5から、DB1に対するCCAとKCCAによる快-不快の識別では70%前後と比較的良好な結果を得た。しかし、その他の項目に関しては36~57%の範囲となり、良好な識別率は得られなかった。

6. おわりに

本論文では、顔表情画像から得られる快-不快、覚醒-沈静の情報を、メンバシップ関数を用いて表現し、その形状をCCAもしくはKCCAにより推定することによって、感情推定を行う方法を提案した。メンバシップ関数については、主観評価値とその標準偏差をもとにしたモデルと、それに主観評価値と特徴空間における距離情報を組み合わせたモデルを設定し、前者

に関しては、4次元のパラメータを用いるもの、1変数ガウス型関数を用いるもの、2変数ガウス型関数を用いるものの3種類について検討を行った。

4次元のメンバシップパラメータを推定する実験については、快-不快、覚醒-沈静の主観評価値の推定において、DB1における絶対値誤差が0.4程度、DB2のそれが0.7程度となり、3章に示した主観評価実験結果の統計量と比べて同程度またはより高い精度を示した。標準偏差の推定では、覚醒-沈静の軸でまずまずの精度が得られた。快-不快の軸に関しては、標準偏差自体の値が小さく、推定が困難になったものと思われる。ガウス型メンバシップ関数を用いた実験では、1変数メンバシップ関数の推定誤差が0.15程度、2変数メンバシップ関数の推定誤差が0.1程度(DB1)となり、ある程度良好な形状予測が可能であった。感情の取り扱いにおける主観評価の段階数についても検討を行った。実験結果に大きな差異は見られないものの、9段階から11段階程度の分割数の有効性がわかった。2変数メンバシップ関数を用いれば、2軸の同時評価が記述できて推定誤差も小さくなるが、多峰性を持つ関数として推定されることもあり、その取り扱いが今後の課題である。画像空間のデータ分布を考慮に入れたメンバシップ関数の場合においても、快-不快の関数の形状の推定において良好な結果を得た。

CCAとKCCAを比較すると、ほとんどの手法と両方のデータベースでKCCAの推定誤差がより低かった。しかし、データベースの均質性や画像の枚数等の影響でカーネル関数の最適なパラメータが大きく異なり、その値を実験的に求める必要性が生じる。この数値の選択法についても今後の検討が必要である。

また今後、メンバシップ関数の計算法の更なる改良と Multivariate linear regressionなどの回帰モデルを扱う他の統計手法の適用について検討していきたい。

謝辞

本研究の一部は、科学研究費補助金(基盤研究(C) 20500194)による。

参考文献

- [1] P.Ekman and W.V.Friesen, *Unmasking the face* (paperback), Malor Books, 2003.
- [2] 赤松茂: “人間とコンピュータによる顔表情の認識 [I] [II] [III] [IV]”, 電子情報通信学会誌, Vol's. 85-86, No's.9, 10, 12, 1, 2002-2003.
- [3] M.Pantic and L.J.M.Rothkrantz: “Automatic analysis of facial expressions: The state of the art”, IEEE Trans. PAMI, 22, 12, pp.1424-1445, 2000.
- [4] Z.Zeng, M.Pantic, G.I.Roisman and T.S.Huang: “A Survey of Affect Recognition Method,” IEEE Trans, PAMI, 31, 1, pp.39-58, 2009.
- [5] 大平秀樹編: 感情心理学・入門, 有斐閣, 2010.
- [6] J. A. Russell and M.Bullock: “Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults,” J. Personality Social Psychology, 48, 1290-1298, 1985.
- [7] 竹原卓真, 野村理朗 編著: 「顔」研究の最前線, 北大路書房, 2005.
- [8] 坂口竜己, 山田 寛, 森島繁生: “顔画像を基にした3次元感情モデルの構築とその評価”, 電子情報通信学会論文誌, J80A, 8, pp.1279-1284, 1997.
- [9] 石井雅樹, 他: “自己写像特性を用いた顔表情カテゴリーの抽出と感情空間マップの生成”, 電子情報通信学会論文誌, J91-D, 11, pp.2659-2672, 2008.
- [10] 谷卓哉, 他: “正準相関分析と注視特性による顔表情画像からの感情の測定法”, 日本知能情報ファジィ学会誌, 22, 1, pp.52-64, 2010.
- [11] 山下利之, 他: “ファジィ理論による表情選択モデルのヒューマンインタフェースへの応用”, 日本ファジィ学会誌, 12, 2, pp.105-112, 2000.
- [12] M.Ilbeygi and H.S.Hosseini: “A novel fuzzy facial expression recognition system based on facial feature extraction from color face images,” Engineering Applications of Artificial Intelligence, 25, pp.130-146, 2012.
- [13] N.Esau, E.Weitzel, L.Kleinjohann and B.Kleinjohann: “Real-time facial expression recognition using a fuzzy emotion model,” Proc. IEEE Int. Fuzzy Systems Conf., pp.1-6, 2007.
- [14] M. Arita, et al.: “Estimation of fuzzy membership of facial emotions using CCA,” Proc. 2012 International Workshop on Advanced Image Technology, Jan.9-10, Ho Chi Minh City, 2012.
- [15] M. J. Lyons, et al.: “Coding facial expression with gabor wavelets,” Proc. 3rd Int. Conf. Automatic Face & Gesture Recog., pp.200-205, 1998.
- [16] T. Kanade, et al, “Comprehensive database for facial expression analysis,” Proc. 4th IEEE Int. Conf. Automatic Face & Gesture Recog., pp.46-53, 2000.
- [17] Y. Liu, et al, “A new fuzzy approach for handling class labels in canonical correlation analysis,” Neurocomputing, 71, pp.1735-1740, 2008.
- [18] M.Kirby and L.Sirovich: “Application of the Karhunen-Loeve procedure for the characterization of human faces,” IEEE Trans. PAMI, 12, 1, pp.103-108 (1990).
- [19] 柳井春夫: 多変量データ解析法-理論と応用-, 朝倉書店, 1994.
- [20] 福水健次: カーネル法入門-正定値カーネルによるデータ解析-, 朝倉書店, 2010.
(2011年12月18日 受付)
(2012年 6月16日 採録)

[問い合わせ先]

〒815-8540 福岡市南区塩原 4-9-1

九州大学 大学院芸術工学研究院

坂本 博康

TEL: 092-553-4506

FAX: 092-553-4506

E-mail: sakamoto@design.kyushu-u.ac.jp

著者紹介

ありた まゆこ
有田 繭子 [非会員]

2010年九州大学芸術工学部画像設計学科卒業。2012年同大学大学院芸術工学府修了。在学中、マルチモーダルの顔画像処理の研究に従事。現在、株式会社構造計画研究所勤務。

おち はるか
越智 晴香 [非会員]

2009年九州大学芸術工学部画像設計学科卒業。2012年同大学大学院芸術工学府修了。在学中、ウルカヌス・イン・ヨーロッパプログラム参加(1年間)、および顔表情画像処理の研究に従事。現在、日本電気株式会社勤務。

ふくしま しげひろ
福島 重廣 [非会員]

1971年京都大学工学部卒業。1973年同大学院修士課程修了。運輸省電子航法研究所、京都工芸繊維大学、九州工業大学を経て、1997年九州芸術工科大学大学院教授。2003年九州大学大学院芸術工学研究院教授。現在に至る。京大工博。画像処理と医用画像工学を中心に多様な画像関連技術に関する研究に従事。IEEE、電子情報通信学会、医用画像工学会、等の会員。

さかもと ひろふみ
坂本 博康 [非会員]

1975年九州工業大学電子工学科卒業。1977年同大学大学院修士課程修了。同年同大学助手。1986年九州芸術工科大学講師。2003年同教授。現在、九州大学大学院芸術工学研究院教授。博士(工学)。計算機視覚、顔画像処理、画像計測などの研究に従事。電子情報通信学会、情報処理学会、映像情報メディア学会、各会員。

A Method for Estimating Fuzzy Membership Functions of Emotion Based on Its Dimensional Theory from Facial Expression Images

by

Mayuko ARITA, Haruka OCHI, Shigehiro FUKUSHIMA and Hiroyasu SAKAMOTO

Abstract :

This paper proposes a method for estimating fuzzy membership functions of human emotion from facial expression images. The membership functions are based on the dimensional model of emotion and are defined on the emotion space with coordinate axes of valence and arousal. We discuss two models of membership functions. a) The first model uses results of subjective evaluation experiments for learning facial expression images where mean value and standard deviation for each facial image are respectively adopted to mean and spread parameters of the corresponding membership function. b) The second model is a modification of the first model by introducing information of the distribution of the data in the facial image space. The performance of our method is evaluated by approximation error of the estimated membership functions to the original ones. As statistical techniques, canonical correlation analysis (CCA) and kernel CCA are employed and applied to gray-scale image based data. Using two facial image databases of male and female, our experimental results show valid estimation results which have at most less than twice as much dispersion of that of subjective evaluation by human observers, and in the sense of the dispersion, availability of the proposed method is indicated.

Keywords : emotion, facial expression images, fuzzy membership function, CCA(KCCA)

Contact Address : **Hiroyasu SAKAMOTO**

*Department of Communication Design Science, Faculty of Design, Kyushu University
4-9-1, Shiobaru, Minami-ku, Fukuoka, 815-8540, JAPAN*

TEL : 092-553-4506

FAX : 092-553-4506

E-mail : sakamoto@design.kyushu-u.ac.jp