

表情画像の位相特性に基づいた感情特徴空間の生成

Generation of Emotional Feature Space based on Topological Characteristics of Facial Expression Images

○石井雅樹[†], 佐藤和人[†], 間所洋和[†], 西田眞[‡]

○Masaki Ishii[†], Kazuhito Sato[†], Hirokazu Madokoro[†], Makoto Nishida[‡]

[†]秋田県立大学, [‡]秋田大学

[†]Akita Prefectural University, [‡]Akita University

キーワード: 顔画像処理 (Facial image processing), 表情認識 (Facial expression recognition), 自己組織化マップ (Self-Organizing maps), カウンタプロパゲーションネットワーク (Counter propagation networks)

連絡先: 〒015-0055 秋田県由利本荘市字海老ノ口84-4

秋田県立大学 システム科学技術学部 機械知能システム学科

石井雅樹, Tel.: (0184)27-2220, Fax.: (0184)27-2188, E-mail: ishii@akita-pu.ac.jp

1. まえがき

顔表情からの感情認識は, マンマシン間の自然な対話を実現するための一手法として有用である. 感情を表す表情パターンは人物ごとに異なるが, 表情に表出しようとする感情や表情から知覚される感情は人物によらず普遍と考える. すなわち, 表情から感情の程度を推定するためには, 物理的パラメータである個人固有の表情パターンを如何に抽出し, これらを普遍的な心理的パラメータである感情に如何に対応付けるかが重要な課題と考える.

そこで, 本論文では, 教師なし学習の自己組織化マップ (Self-Organizing Maps, 以下, SOMと表記する)¹⁾と教師あり学習のカウンタプロパゲーションネットワーク (Counter Propagation Networks, 以下, CPNと表記する)²⁾を併用することにより, 表情パターンと感情の対応付けを行い得る個人固有の感情特徴空間の生成手法を提案する.

2. 表情画像

本論文で対象とする表情画像は, 4名の被験者が人為的に表出した6基本感情 (怒り, 悲しみ, 嫌悪, 喜び, 驚き, 恐れ)^{3, 4)}および無表情とした. 各表情を動画像 (30 frame/sec, 8 bit Gray) として取得し, 静止画像 (320×240 pixel) に変換した後, 学習データとして使用した. また, 表情の表出に重要となる顔の構成要素 (眉, 目, 鼻, 口) を含む領域を処理の対象とし, 以下の手順に従い顔領域画像の抽出, 並びに, 正規化を行った. 表1に被験者ごとの学習データ数を示す. また, 図1に正規化後の顔領域画像の例を示す.

1) Haar-like 特徴⁵⁾による顔検出を行い, 80×96 pixel のサイズに正規化した顔領域画像を抽出した.

2) ノイズの除去を目的とし, メディアンフィルタを施した. また, 画像の次元数の削減及び顔領域の微小な位置ずれに対応するため, 2

Table 1 Number of training data.

Subject (ID)	Anger (An.)	Sadness (Sa.)	Disgust (Di.)	Happiness (Ha.)	Surprise (Su.)	Fear (Fe.)	Neutral (Ne.)	Total (Total)
A	454	448	533	530	473	479	605	3,522
B	284	337	323	335	325	323	599	2,526
C	493	556	592	475	455	471	599	3,641
D	574	518	457	473	459	468	605	3,554

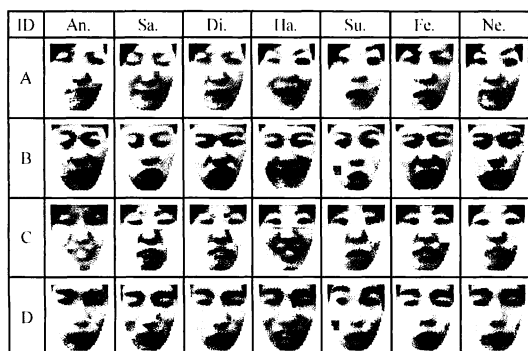


Fig. 1 Examples of facial expression image.

×2 pixel のブロック単位で粗視化処理を施した後、平滑化処理を施した。

- 3) 顔の構成要素を含む擬似輪郭を設定し、背景領域を除去した。
- 4) ヒストグラム線形変換により濃度値補正を行った。

3. 提案手法

本論文で提案する個人固有の感情特徴空間の生成法は、SOM を用いた表情カテゴリの抽出フェーズと CPN を用いた表情空間マップの生成フェーズ及び感情空間マップの生成フェーズから構成される。

3.1 表情カテゴリの抽出フェーズ

感情によって顔に表れる表情パターンは人物ごとに異なり、個人固有の特徴を有している。したがって、表情パターンの分類問題は未知のカテゴリ

リ分類問題と考えることができる。そこで、提案手法では、狭い写像空間を有する SOM (Kohonen 層が 5 ユニットの 1 次元 SOM) を階層的に用いることにより、個人固有の表情カテゴリの抽出を試みた。SOM は教師なし学習アルゴリズムであり、与えられた表情画像の位相特性に基づき、自己組織的に分類を行うことから、未知のカテゴリ分類問題に適している。また、狭い写像空間を用いることにより、SOM は表情画像の位相情報を圧縮し、学習データを大別する特徴に基づいた学習を行う。これらを階層的に繰り返すことにより、表情パターンの粒度（バラツキ）を一定とした個人固有の表情カテゴリの抽出を行い、得られたカテゴリに対して感情（6 基本感情および無表情）の意味づけを行った。表 2 に提案手法により抽出された表情カテゴリ数を示す。

3.2 表情空間マップの生成フェーズ

感情特徴空間への表情パターンの布置を考えた場合、感情の程度を表現するためには、表情パターンの類似性及び連続性を考慮したカテゴリ間を補間する情報が必要と考える。

そこで、提案手法では、広い写像空間を有する CPN (Kohonen 層が 30×30 ユニットの 2 次元 CPN) を用いることにより、前節で抽出したカテゴリ間の情報拡張を行った。前節の処理において、各表情カテゴリは 5 ユニットの 1 次元 SOM によって生成されている。したがって、学習データとして使用する 1 カテゴリの代表画像数は 5 枚 (Kohonen 層の各ユニットの結合荷重) である。ここで、CPN を採用した理由は、前節の処理により学習データの感情カテゴリが既知であるためである。なお、表

Table 2 Number of extracted categories.

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	6(7)	6(10)	8(6)	8(11)	4(13)	6(12)	0(9)	106
B	3(5)	4(1)	6(6)	6(6)	6(6)	9(6)	0(6)	70
C	8(7)	8(4)	1(5)	8(8)	5(8)	7(10)	0(2)	81
D	2(1)	4(8)	9(6)	8(10)	5(9)	8(6)	0(1)	77

(): Numbers of categories labeled as "Neutral".

Table 3 Number of extracted representative images.

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	30	30	40	40	20	30	340	530(15.0 %)
B	15	20	30	30	30	45	180	350(13.9 %)
C	40	40	5	40	25	35	220	405(11.1 %)
D	10	20	45	40	25	40	205	385(10.8 %)

(): Data compression rate from Table 1.

情パターンの類似性及び連続性に基づいた情報拡張を目的としていることから、CPNの写像空間(Kohonen層)は学習データ数よりも大きく($30 \times 30 = 900$ ユニット)、かつ、トーラス状の構造を採用した。表3に使用した学習データ数を示す。また、図2に被験者Aの表情空間マップを示す。

3.3 感情空間マップの生成フェーズ

人物ごとに表情パターンは異なっているが、表情として顔に表れる感情及び表情から知覚される感情は人物によらず普遍と考えられる。前節で生成した表情空間マップは、人物ごとに異なる物理的パラメータとしての表情パターンの固有性に着目した特徴空間と言える。一方、心理的パラメータとしての感情は人物によらず普遍であることから、特徴空間は表情空間マップとは性質が異なり、共通の指標に基づいて生成されるべきと考える。

提案手法では、感情特徴空間を生成する際の指標として心理学の分野で提案されているRussellの円環モデル^{6,7)}に着目した。Russellの円環モデルとは、表情から認知される様々な感情の意味的な対極性に注目し、各表情を「快-不快」、「覚醒度」を軸とした空間上に配置した心理空間モデルである。Russellらは、上述の心理空間上で表情刺激が円環状に配列されること、また、この円環状の配

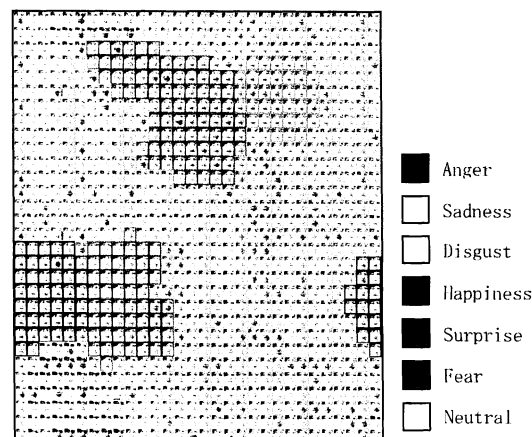


Fig. 2 Generation result of Facial Expression Map (Subject A).

列は子供から大人まで一貫して等しいことを見出した。

提案手法では、前節の処理と並行し、CPNの教師信号として、図3に示すRussellの円環モデルより得られた座標値を付加することにより、物理的パラメータ(表情)と心理的パラメータ(感情)を対応付ける感情空間マップの生成を行った。提案手法によって生成された被験者Aの感情空間マップを図4に示す。

Table 4 Number of winner units when inputting training data.

ID	A	B	C	D
Numbers of Kohonen layer units	900	900	900	900
Number of winner units	252	183	244	231
Number of extended units	648	717	656	669

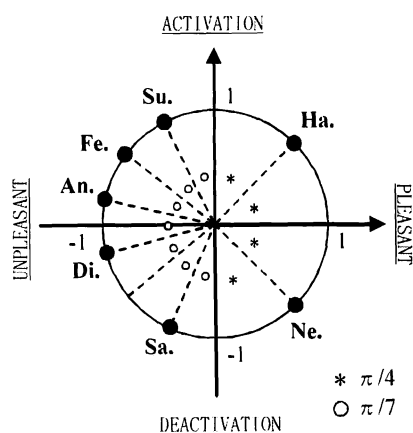


Fig. 3 The circumplex model of Russell.

4. 考察

表4に、表情空間マップに対して表3の学習データを入力した場合の発火ユニット数を示す。被験者Aに着目すると、900ユニット（30×30ユニット）のうち、発火の起こったユニットは252ユニットであった。すなわち、残りの648ユニットは、CPNの近傍学習により学習データ間を補間するユニットとして生成されたと考えられる。

図5に被験者Aの感情空間マップの詳細を示す。同図(a)は、表3の学習データを入力した際、表情空間マップ上で発火したユニット（252ユニット）を感情空間マップ上に布置した結果を示している。また、同図(b)は、表情空間マップのすべてのユニット（900ユニット）を感情空間マップ上に布置した結果を示している。図5(a)では、CPNの教師信号の座標値付近（図3）に集中して布置されているのに対し、同図(b)では、拡張生成されたユニット（648ユニット）が各座標間を補間する形で布置されている様子が認められる。

図5(c)に、同図(b)の「喜び」の領域（第1象限

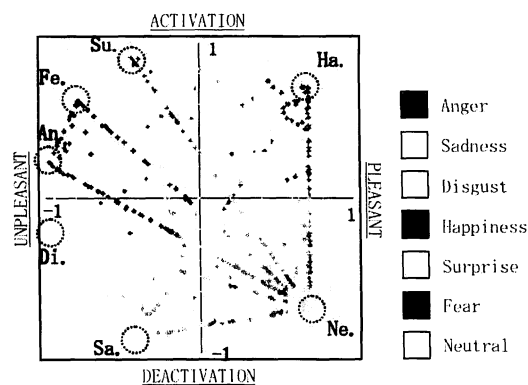
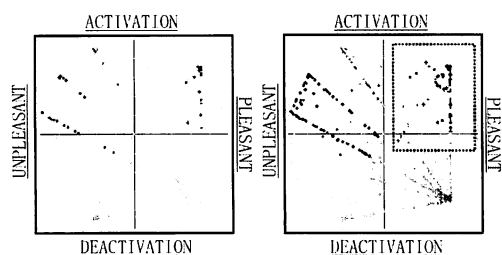


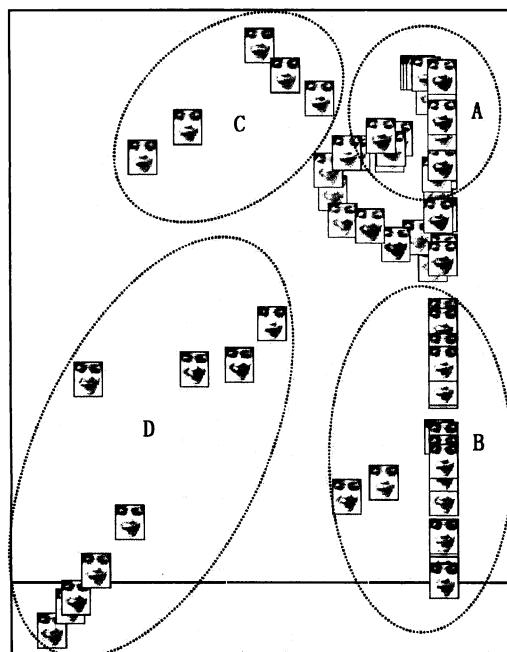
Fig. 4 Generation result of Emotion Map (Subject A).

付近)を拡大した図を示す。図5(c)の領域Aでは、CPNの教師信号として与えた座標値付近であることから、純粋な「喜び」の表情が布置されている。また、領域Bでは、領域Aと比較して「覚醒度」の値が徐々に小さくなり、「無表情」に近づいていくような表情が布置されている。同様に、領域Cでは、「快-不快」の値が徐々に小さくなり、「驚き」の感情の混在した表情が布置されており、領域Dでは、「覚醒度」及び「快-不快」の値が徐々に小さくなり、「悲しみ」の感情の混在した表情が布置されている。

以上の結果は、広い写像空間を有するCPNを用いた提案手法により、表情空間マップでは、学習データの類似性、連続性に基づいた情報の拡張が行われ、より多くの表情パターンを生成し得ること、また、Russellの円環モデルに基づく座標値をCPNの教師信号として与えることにより、感情空間マップでは、物理的パラメータとしての表情パターンに対し、心理的パラメータとしての感情の程度を表現し得ることを示唆している。



(a) 学習データの布置結果 (b) 全ユニットの布置結果



(c) 喜びの布置結果 (拡大図)

Fig. 5 Detail of the Emotion Map (Happiness of Subject A).

5. まとめ

本論文では、表情パターンと感情の程度を関連付ける個人固有の感情空間マップの生成手法について検討を加えた。

提案手法では、はじめに、SOMを用いて物理的パラメータである表情パターンを階層的に学習し、個人固有の表情カテゴリを抽出した。次に、CPNを用いて各カテゴリの表情画像を学習し、表情パターンの類似性に基づく表情空間マップを生成した。更に、CPNの教師信号としてRussellの心理空間モデルの値を付加することにより、共通の指標に基づく感情空間マップの生成を試みた。

被験者4名に対する実験結果より、感情空間マップでは表情パターンと感情の対応付けを行い得ることが確認され、提案手法は感情の推定手法として有用であることが示された。

参考文献

- 1) T. Kohonen, "Self-organizing maps," Springer Series in Information Sciences, 1995.
- 2) R. H. Nielsen, "Counterpropagation networks," Applied Optics, vol.26, no.23, pp.4979-4984, 1987.
- 3) P. Ekman, "Darwin and Facial Expression: A Century of Research in Review," Academic Press, N.Y., 1973.
- 4) P. Ekman and W.V. Friesen, "Facial Action Coding System," Consulting Psychologist Press, 1978.
- 5) R. Lienhart and J. Maydt, "An Extended Set of Haarlike Features for Rapid Object Detection," Proc. IEEE Int. Conf. Image Processing, vol.1, pp.900-903, 2002.
- 6) J.A. Russell and M. Bullock, "Multidimensional Scaling of Emotional Facial Expressions: Similarity from Preschoolers to Adults," J. Personality and Social Psychology, vol.48, pp.1290-1298, 1985.
- 7) J.A. Russell and L.F. Barrett, "Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant," J. Personality and Social Psychology, vol.76, pp.805-819, 1999.