

顔画像の位相特性に基づいた個人固有の表情特徴空間の形成に関する検討

石井 雅樹[†] 佐藤 和人[†] 西田 眞[‡] 間所 洋和[†]

秋田県産業技術総合研究センター[†] 秋田大学[‡]

1. はじめに

情報機器の機能の高度化により、知的なマンマシンインタフェースの重要性が高まっている。人間と情報機器のコミュニケーションの中で実現されていないものとして、感情のコミュニケーション、中でも、顔表情の認識が挙げられる。

従来の表情認識手法では、顔面上に設定した特徴点の移動量や顔部位の輪郭形状、表情筋の動きベクトルを特徴量とする手法が一般的である^(例えば [1])。しかしながら、これらの特徴量は個人固有であるため、認識に有効な特徴量の選定や組み合わせ、閾値の決定が被験者ごとに必要となる。

より個人に応じた表情認識を実現するためには、提示された顔画像上の位相特性の変化から、表情に関する個人固有の特徴を自動的に抽出することが必要であり、また、それらの特徴を考慮した「個人固有の表情特徴空間」を自己組織的に形成することが重要な課題と考える。

そこで本論文では、顔画像上の顔部位の幾何学的形状および相対的位置関係の変化を表情の位相特性として捉え、自己写像特性を有する自己組織化マップ(Self-Organizing Map, 以下 SOM と記す)^[2]を用いた個人固有の表情特徴空間の形成手法を提案する。

2. 対象画像

本論文で対象とした表情画像を図 1 に示す。被験者が主観的に表現した 7 表情(平常, 怒り, 嫌悪, 恐れ, 喜び, 悲しみ, 驚き)を対象とし、420 枚の画像(60 枚×7 表情)を学習データとして使用した。

代表的な表情記述モデルとして FACS(Facial Action Coding System)が知られている^[3]。FACS では、表情に関する顔部位の動き特徴が AU(Action Unit)として定性的に定義されている。AU を詳細に分類すると、顔上部(眉, 目周辺)に関連した AU と顔下部(鼻, 口周辺)に関連した AU に大別される。このことから、表情の位相特性は顔上部、顔下部といった局所領域、両者を包含した顔全体領域で固有の特徴を有していると考えられる。そこで本論文では、対象領域として顔全体領域、顔上部領域、顔下部領域に着目し、領域毎に表情特徴空間の形成を試みた。

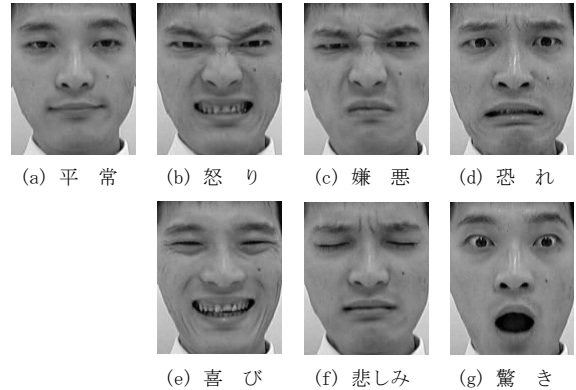


図 1 本研究で対象とした表情画像

3. SOM の階層化による表情特徴空間の形成

SOM は教師なし学習(自己学習)アルゴリズムであり、入力層とマップ層から構成される。マップ層のユニット間に結合関係は存在しないが、近傍学習により隣接するユニットの結合荷重は学習過程で互いに類似した関係へ導かれる。本論文では、マップ層が 5 ユニットの 1 次元 SOM を採用した。これは、狭い写像空間に入力群の位相情報を圧縮することにより、SOM は与えられた入力群を大別する特徴的な表情位相特性を学習すると考えたためである。

提案手法の処理手順を図 2 に示す。また、処理の詳細を以下に示す。

- (1) 420 枚の表情画像を SOM の学習データとした。
- (2) 学習データからランダムに入力画像を選択し、SOM による学習を行った。SOM への入力は画像の輝度分布である。また、学習回数は経験的に 20 万回とし、近傍領域は勝者ユニットの第一近傍で固定とした。学習率係数の初期値は 0.5 とし、学習回数が 10 万回の時点で 0.02, 最大学習回数の時点で 0 まで線形減少する構成とした。
- (3) マップ層の各ユニットの結合荷重(0~1)を 0 から 255 の値に変換し、結合荷重の可視化画像を作成した。
- (4) マップ層の隣接ユニット間における結合荷重の相関係数を算出し、最小値のユニット間を境界としてマップ層を 2 分割した。SOM は近傍学習を行うため、隣接するユニットの結合荷重は近い値へ導かれる。したがって、境界の双方に分類された入力群には、大きな位相特性の差が存在すると考えられる。
- (5) 境界の双方に分類された入力群は、独立な部分問題と考えることができる。そこで、双方の入力群を対象として再帰的に(2)から(4)の処理を

Formation of the Individual Expression Feature Space based on the Topological Characteristic on Facial Images

[†] Masaki Ishii, Kazuhito Sato and Hirokazu Madokoro (Akita Prefectural R&D Center)

[‡] Makoto Nishida (Akita Univ.)

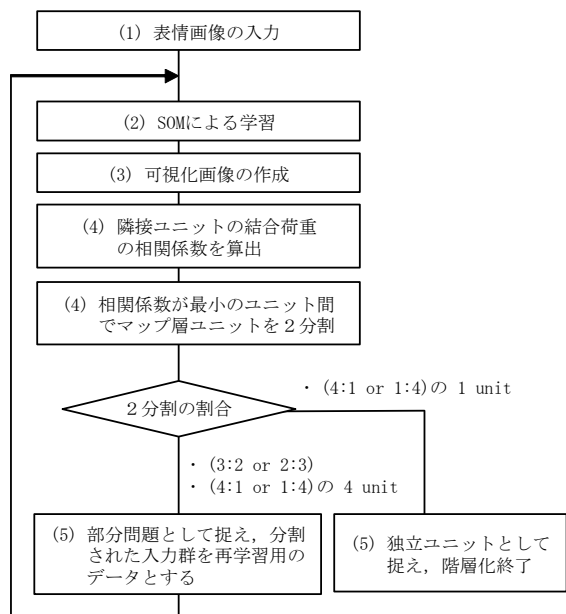


図2 提案手法の処理手順

繰り返し、SOMの階層化を行った。再帰処理を行う場合のルールを以下に示す。

- ・マップ層が(3:2), (2:3)に分割された場合、双方に分類された入力群をそれぞれ再学習用のデータとする。
- ・マップ層が(4:1), (1:4)に分割された場合、4ユニットに分類された入力群を再学習用のデータとする。1ユニットに分類された入力群は、表情位相特性が他と比較して特異性を有しているとみなし、階層化処理を終了する(独立ユニットとして扱う)。

なお、本論文では階層数を3層とした。ただし、第1層で独立ユニットが現れた場合は4層とした。

4. 提案手法の適用結果およびまとめ

顔全体領域の階層化結果を図3に示す。なお、図中の分類結果の値は、マップ層の各ユニットに対応付けられた学習データの数を意味する。

提案手法の有用性を検討するため、FACSのAUを指標とし、表情特徴空間の解析を行った。具体的には、各階層の分類境界を中心とした双方の可視化画像に着目した。また、表情特徴空間が被験者固有の特徴に基づいているかを比較検討するため、事前に図1に示した学習データを対象としてAUの解析を行った。

図3より、顔全体領域では、第1層で[驚き]、第2層で[喜び]が1ユニットに独立して分類されている。その他は、第3層以降で2分木問題として展開され、それぞれ分類されている様子が認められる。ここで、第1層の可視化画像に着目すると、同階層では主に[眉を上げる(AU1,2)], [口を大きく開く(AU27)]といった位相特性が[驚き]の分類基準として学習されていると考える。これは、

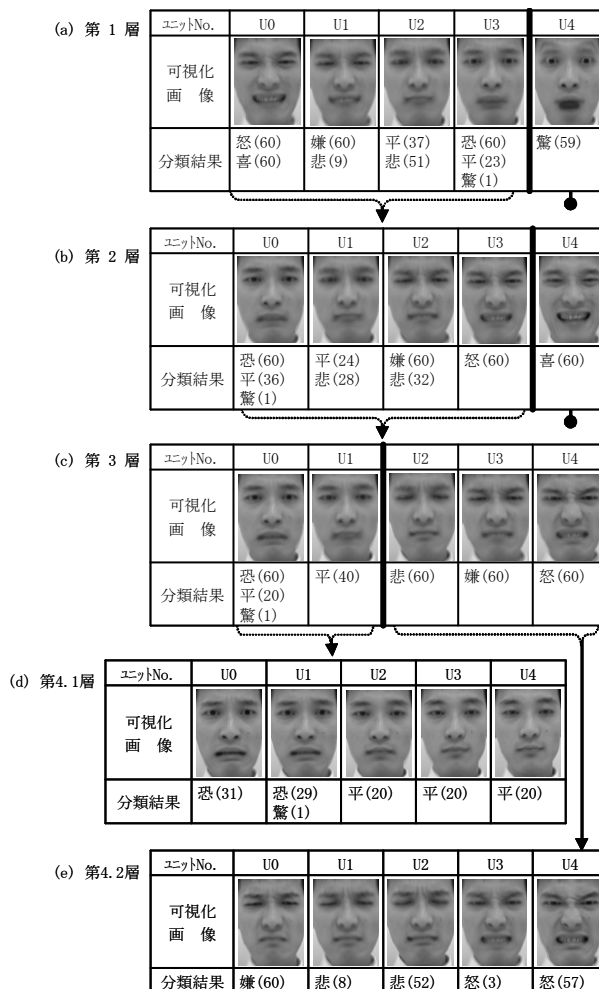


図3 顔全体領域の階層化結果

事前に行った学習データに対するAU分析結果と一致することを確認している。同様の解析を最下層まで行ったところ、被験者の表情を分類するための特徴が、2分木構造のノード部に表れていることを確認した。また、顔上部領域、顔下部領域の表情特徴空間においても同様の結果が認められることを確認した。

以上の結果は、提案手法が特別な前処理を必要とせず、顔画像上の位相特性に基づき、個人固有の表情特徴空間を自己組織的に形成可能であることを示している。また、3領域の表情特徴空間を統合することにより、より個人に応じた表情認識のための識別器を形成できることを示唆している。

参考文献

- [1] 下田, 國弘, 吉川: 動的顔画像からのリアルタイム表情認識システムの試作, ヒューマンインタフェース学会論文誌, Vol.1, No.2, pp.25-32 (1999).
- [2] Kohonen, T.: Self-organizing maps, Springer Series in Information Sciences (1995).
- [3] Ekman, P., Friesen, W.V.: Facial Action Coding System, Consulting Psychologist Press (1978)