

A3-4 人間共生ロボットのための自己組織化マップを用いた人間の顔表情認識

福井大学 工学部 知能システム工学科 進化ロボット研究室
松原 大典 (指導教員：前田 陽一郎，高橋泰岳)

1. 緒言

従来の表情認識の研究では、あらかじめ表情の辞書を作成して、認識を行うものが多く、辞書にない複数の表情が表出された複雑な表情は認識できない。本研究では、多様な表情の認識を行うため、T.Kohonen[1] が提案した人間の適応学習をモデルとした、教師なし学習の自己組織化マップ (SOM: Self-Organizing Maps) を用いた表情認識システムの構築を行う。

従来研究としては、石井ら [2] は、SOM の特性を用いて特徴に基づいた顔画像の分類を行った。また、鈴木ら [3] は、操作が容易な線形図形を用いて顔表情を記述した。

本研究の提案手法では、複雑な表情の認識を行うため、顔画像から表情表出に大きく関わる目や口の顔要素を抽出し、抽出した顔要素が表出している表情のラベル付けを行う。顔要素ごとに分けて、表情を認識することで、混在している複数の表情を分類し、総合的に最も強く表出されている表情の認識を行う。提案手法の有用性を確認するための実験を行ったので、その結果について報告する。

2. 自己組織化マップ

自己組織化マップ (SOM: Self-Organizing Maps) は、T.Kohonen[1] により提案された教師なしのニューラルネットワークアルゴリズムであり、高次元データを傾向や共通点に基づいて分類を行うことで、人間がデータを視覚的に理解できる手助けを行う。

SOM は、複数のユニットを持つ入力層と出力層の 2 層から構成される (図 1 参照)。出力層の各ユニットは、ユニット間の結合係数である参照ベクトルを持つ。入力ベクトルが与えられると、入力ベクトルと一番近い参照ベクトルを持つ出力層のユニットを勝者ユニットとし、勝者ユニットの参照ベクトルを入力ベクトルに近づけるように学習が行われる (図 2 参照)。以下に、SOM の学習アルゴリズムを示す。

1. $m_j(t)$ を時刻 t における入力層のユニット i から出力層のユニット j への参照ベクトルとし、乱数を用いて初期化する。
2. $x_i(t)$ を時刻 t における入力層のユニット i への入力とし、 $x_i(t)$ と $m_j(t)$ のユークリッド距離 $d_j(t)$ を計算する。
3. $d_j(t)$ が最小となる、出力層のユニットを探索する (このユニットを勝者ユニットと呼ぶ)。
4. 勝者ユニット付近のユニットの参照ベクトル $m_j(t)$ を、式 (1) を用いて更新する。 $\alpha(t)$ は、学習率係数である。

$$m_j(t+1) = m_j(t) + \alpha(t)(x_i(t) - m_j(t)) \quad (1)$$

5. 上述の処理を、最大学習回数まで繰り返す。

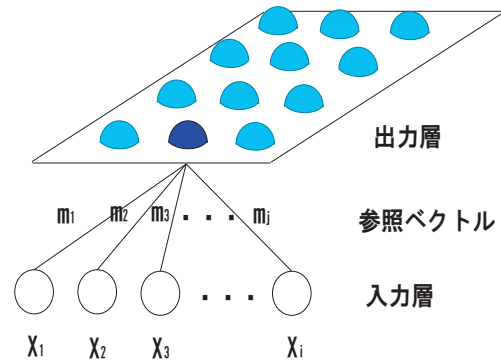


図 1 SOM の概念図

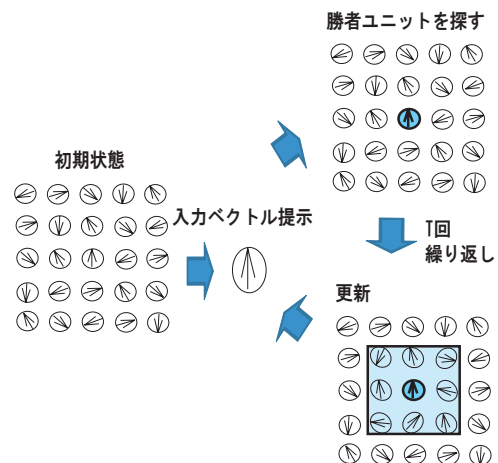


図 2 SOM の学習アルゴリズム

3. SOM を用いた顔表情認識手法

本研究で提案する SOM を用いた表情認識手法の処理概要を示す。本手法では、人種や文化、性別に関係なく認知できる普遍的な表情である喜び・驚き・恐怖・怒り・嫌悪・悲しみの基本 6 表情 (図 3 参照) と無表情の 7 つの表情の認識を行う。

まず、表情辞書である表情マップの作成を、抽出する顔要素ごとに行う。7 つの表情を認識できる表情マップを作成するため、同様に基本 6 表情の 6 つと無表情顔の日本人女性の表情顔データが収納されている JAFFE データセットを作成に用いる。JAFFE の顔表情画像から顔要素を抽出し、SOM の入力データとして用いて、2 次元の平面マップを作成する。マップを 2 次元にすることで、複数の表情の特徴を含んだユニットが作成され、認識を行う 7 つの表情以外の表情も認識できるようになると考えられる。

表情は、無表情時からの顔要素の変位によって表出されるので認識を行う表情画像の顔要素を抽出し、作成した表情マップへの入力データとする。表情マップ上で、最も一致率が高い勝者ユニットの表情を、抽出した顔要素が表出している表情としてラベル付けする。ラベル付けを行った顔要素の表情ラベルの組み合わせから、顔画像の表情を判断する。

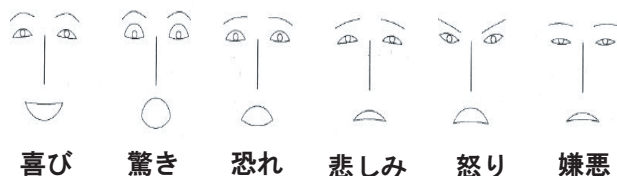


図3 基本6表情 (文献 [3] より引用)

4. 顔認識実験

顔要素のラベル付けに基づく表情認識を検証するため、日本人女性の表情データが収納されている JAFFE データセットの中から、ランダムに選択を行った表情顔画像を用いて実験を行う。顔要素の抽出は、顔表情画像からテンプレートマッチングを用いて行い、本実験では、目と口の抽出を行う。顔要素の抽出を行うのは、無表情時から形状や位置の変位によって表情を表出するためである。

認識画像から、表情マップの作成時と同様に顔要素の抽出をテンプレートマッチングを用いて行う。抽出した顔要素を作成した表情マップへ入力し、入力画像と最も一致率の高い画像のユニットである勝者ユニットの探索を SOM を用いて行い、勝者ユニットを四角で囲むようにする。勝者ユニットは、入力ベクトルと最も近似した参照ベクトルを含んでいるので、勝者ユニットの表情を入力ベクトルの表情としてラベル付けする。SOM を使用することにより、データの類似性を視覚的に確認することができる。

探索を行った勝者ユニットの表情を、入力画像の顔要素が表出している表情としてラベル付けする。表情の認識は、ラベル付けされた表情の組み合わせから行う。

ここでは、喜びの表情が表出している顔画像を用いた認識結果を図4に示す。認識画像の顔要素を表情マップに入力して勝者ユニットを探索した結果、それぞれ喜びの表情と認識されたが、認識された勝者ユニットには喜び以外の表情も含まれている。そこで、ラベル付けされた顔要素を組み立てて顔画像を作成し、顔画像が表出している表情が6基本表情か無表情、または表情が分からないかでアンケートを取った。

アンケートは、20代の学生5名を対象に行った。アンケートの結果、全ての対象者が顔画像を喜びを表出しているという回答を得た。この回答から、認識画像は喜びの表情を表出していると判断することができる。

5. 結言

本研究では、表情表出時の特徴を顔要素ごとに、SOM を使用して作成した表情マップから表情のラベル付けを行い、ラベル付けした表情の組み合わせから表情認識を行う手法を提案し、実験を行った。

表情は常に、複数の感情が混在しており認識を行うには、学習により単一の表情特徴を持つユニットだけでなく、複数の特徴を持つユニットを形成し、マッピングできる2次元マップが有用であることが実験から確認できた。本実験では、表情辞書である表情マップの作成に自己学習が可能な SOM を用いて、平面マップを構築したことで、複数の特徴が含まれるユニットをもつ表情マップを形成することができた。

顔要素への表情のラベル付けには、個人の主観が入るので今後の課題として、共通の指標に基づいた、表情マップの作成と表情マップ上の各ユニットの自律的表現認識が挙げられる。

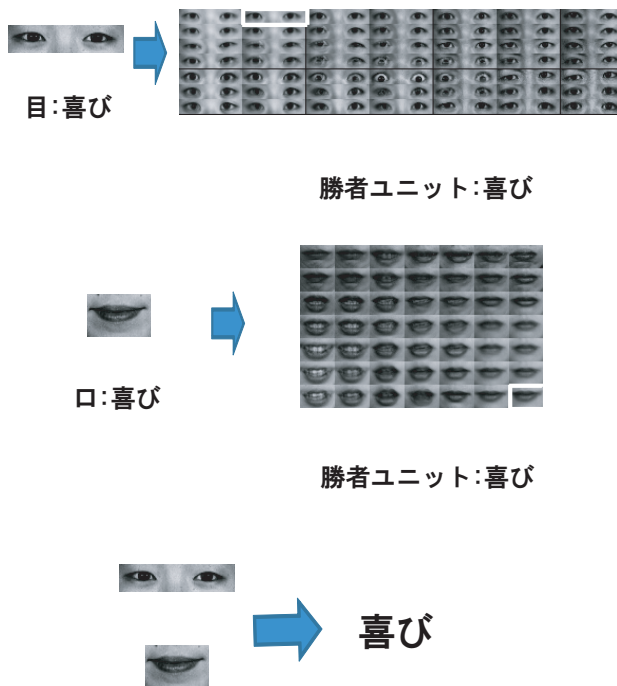


図4 勝者ユニットの探索 (2次元マップ)

参考文献

- [1] T.Kohonen, "Self-Organized Formation of Topologically Correct Feature Maps," *Biol.Cybern.*, 43, pp.59-69 (1982)
- [2] 石井雅樹, 佐藤和人, 間所洋和, 西田眞, "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, Vol.J91-D, No.11, pp.2659-2672 (2008)
- [3] 鈴木健嗣, 山田寛, 橋本周司, "ニューラルネットワークによる顔面表情の物理的パラメータと感性的パラメータの対応付け," 電子情報通信学会 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol.100, No.612, pp.7-13 (2001)