顔面表情のカテゴリー認識

RBF ネットワークによるシミュレーション

○ 山口 拓人　　浅川 伸一　　山田 寛

(日本大学大学院文学研究科)　(東京女子大学情報処理センター)　(日本大学文理学部)

key word : 顔面表情認知・RBF ネットワーク・確率密度関数

表情のカテゴリー判断は，判断に関わる心理空間における情動カテゴリーの典型的表情 (プロトタイプ) を中心に正規分布し，反応時間はプロトタイプからの距離の一次関数増加となるという実験結果が示されている (山田，2000 など) . このことから表情のカテゴリー判断は，各カテゴリーごとのプロトタイプを中心傾向とした正規性の確率密度関数に基いていることが示唆される . そこでカテゴリーごとの確率密度関数を人間の判断データから学習することが可能なニューラルネットワークを作成しシミュレーションを行うことで，人間の持つ確率密度関数を再現できると考えた . シミュレーションには中間層に確率密度関数として妥当であるガウシアン関数を基底関数に持つことができる動径基底関数 (RBF; radial basis function) ネットワークを使用し，心理学的な実験データと比較することでこの考えを検証した .

方法

学習データ : Yamada (1993) の第一実験で被験者によって生成された表情図形 216 枚 (36 × 6 感情:喜び, 驚き, 恐れ, 悲しみ, 怒り, 嫌悪) における特徴点の移動量を示す 9 つのパラメータを学習データとした . 各学習データには，先行研究で得られた情動カテゴリーをラベリングした .

ネットワークの構造 : 3 層の RBF ネットワークで入力層 9，中間層 7，出力層 6 のユニットからなる . 中間層のバイアスユニットを除いた 6 ユニットのみがガウシアン関数を持つ . 入力層と中間層間は線形結合，中間層と出力層間は重み付きの線形結合となっている . 中間層と出力層の各 6 ユニットが 6 感情の判断ユニットとなるように学習される . 心理実験結果と同一視するために RBF ネットワ ークの出力を，任意の表情であると判断される確率とみなす必要がある . このため中間層の出力に softmax 法を用いて規格化を行なった . softmax 法によれば，中間層の基底関数の値を ϕ としたときの中間層ユニットの出力 s_i は $s_i = \exp(\phi_i(x))/\sum_j \exp(\phi_j(x))$ と表現できる .

ネットワークの学習:物理的特徴量で構成される 9 次元空間内に 6 つの感情プロトタイプが表現されていると仮定し，6 つの感情ごとに 9 次元正規分布の中心と分散 (または共分散) とを算出した . 刺激 s の i 番目の中間層の出力は多次元正規分布の密度関数 $1/\sqrt{2\sigma_s^2}\exp\left\{-\frac{1}{2}|\boldsymbol{s}-\boldsymbol{s_i}|^2/\sigma_s\right\}$ を用いて表現される . ここで $\boldsymbol{s_i}$ は i 番目のユニット (すなわち i 番目の感情) の中心である . 上式は刺激 s と i 番目の感情のプロトタイプ s_i と間の距離が大きくなればなるほど，その中間層ユニットの出力 $h_i(s)$ は小さくなることを意味している . 中間層ユニットと出力層ユニットとの間の結合係数行列 $\boldsymbol{W}$ は各刺激に対応する感情の値が 1, 0 で表現された教師信号行列 $\boldsymbol{Y}$ と softmax によって規格化された中間層の出力行列 $\boldsymbol{\Theta}$ を使って $\boldsymbol{W} = \boldsymbol{\Theta}^t\left(\boldsymbol{\Theta}\boldsymbol{\Theta}^t\right)\boldsymbol{Y}$ によって解析的に求めた . ここで添字 t は行列の転置を表している .

中間層における計算の違いから 2 つのシミュレータを構築した . 1) 感情ごとの学習データにおける各パラメータの分散を用いてシミュレーションをおこなったもの (VAR) . 2) 共分散を用いてシミュレーションをおこなったもの (COV) .

シミュレーション 1 (以下 Sim.1) :学習データに使用された 216 枚から，Yamada (1993) の第二実験で刺激として抽出された 72 枚 (12 × 6 感情) の図形をテストデータとした . Sim.1 の結果を評価するために先行研究の結果と比較を行なった .

シミュレーション 2 (以下 Sim.2) : Sim.1 のテストデータで使用された図形から各感情ごとに平均した図形を作り，その図形における各パラメータの分散から，その分散の 1.0，0.5，0.1 の比率の分散を持った正規乱数を平均した図形に加えることで 180 枚 (10 × 6 感情× 3 比率) の図形を作成した . これをテストデータとしシミュレーションを行った . Sim.2 の結果を人間の判断を比較から評価するために，新たに実験を行い，その結果と比較をおこなった .

実験 : Sim.2 のテストデータとして使用された 180 枚の表情図形を紙上で提示し，被験者に表情を判断させた . 大学生の被験者 146 人，(男性 89 名，女性 54 名，M=20 . 4 歳，(*SD* 1.3)) を 3 群，50 人，48 人，48 人にわけ，各群 180 枚のなかの 60 枚について表情の判断が行なわれた .

結果

各図形に対して一番判断の値 (出力値) が高いものを，その図形に対する判断，つまり一番多くの人が選ぶであろう判断とみなしたところ，Sim.1，2 ともに高い割合でテストデータが示唆する表情として判断が行われた (VAR ,COV 共に 70 ~ 100%の正答率) . 先行研究の結果との間のカルバックライブラー の情報量 (KL 情報量) を求めたところ，1.54 (0.57) となった (VAR による結果，括弧内が COV による結果，以下同様) . Sim.2 の結果においても本研究で行った実験の結果との間の KL 情報量を求めたところ，分散比率 1.0 において 1.32 (3.02) ，比率 0.5 において 0.00 (0.00) ，比率 0.1 において 0.00 (0.10) となった . 実験の結果を正規乱数の分散の比率ごとにまとめた結果を Table1 に示す .

考察

RBF ネットワークが正しい判断を行うことができたことは，生成された確率密度関数が実際の表情認識に妥当なものであることを示すものである . また，シミュレーション 2 の結果からわかるように，プロトタイプの表情に近づくにつれ KL 情報量が低下したことから，RBF ネットワークで再現されたプロトタイプに基く表情認識が実際に行なわれていると考えられる . しかし，実験による恐れの判断の困難性，実験の正しい判断の成績が刺激の変化に関わらず上がらないといった結果が見られた . このようなシミュレーションと実際の表情認識との乖離の存在から，実験の手続きの検討と学習により獲得したシミュレーションの内部構造についての分析が必要であると思われる.

Table 1: 先行実験と本実験の結果

	正しい判断 (%)			
		実験		
表情刺激	先行実験	比率 1.0	0.5	0.1
喜び	100	100	100	100
驚き	100	90	100	100
恐れ	8	10	0	0
悲しみ	92	80	80	70
怒り	100	100	100	100
嫌悪	75	70	70	70

(YAMAGUCHI Takuto, ASAKAWA Shinichi, YAMADA Hiroshi)