

適応学習機能を有する表情特徴空間の生成と基礎評価

Generation of expression feature space and its basic evaluation with the adaptive learning functions

○山屋孝史*, 石井雅樹**, 景山陽一*, 高橋毅*, 西田眞*

○Takashi Yamaya, Masaki Ishii, Yoichi Kageyama, Tsuyoshi Takahashi
Makoto Nishida

*秋田大学(Akita University)

**秋田県立大学(Akita Prefectural University)

キーワード: ヒューマンマシンインターフェイス(Human Machine Interface),
対向伝播ネットワーク(Counter Propagation Network),
適応共鳴理論(Adaptive Resonance Theory), 追加学習(Incremental Learning),
表情認識(Facial Expression Recognition)

連絡先: 〒010-8502 秋田県秋田市手形学園町 1-1 秋田大学大学院工学資源学研究所
景山陽一, Tel: (018)889-2786, Fax: (018)837-5771, Email: kageyama@ie.akita-u.ac.jp

1. はじめに

近年, ヒューマンマシン間における感情のコミュニケーションの実現を目的とした顔表情認識に関する研究が注目を集めている¹⁾⁵⁾. ロボットなどが人間の行動から意図・感情を理解することができれば, 状況に応じた判断・行動が可能になるため, 知的なヒューマンマシンインターフェイスの実現が期待される.

人間同士のコミュニケーションは音声, 表情, 身振りなどの情報を用いて行われる. 中でも人間の顔表情は, 顔の構成要素である眉, 目, 鼻, 口などの形状および動きにより表出され, 多くの情報を有することから, 意思疎通において重要な役割を担っていると考える. このため, ヒューマンマシン間における

感情のコミュニケーションの実現には, 顔表情に着目することが有用であると考えられる.

表情認識に関する従来研究では, 短期間で撮影した被験者の表情画像を用いて表情認識モデルを生成し, 評価実験を行っている場合が多い¹⁾⁵⁾. しかしながら, 人間の表情は静的多様性(顔を構成する器官の形状および配置)と動的多様性(情動の発露および意思伝達時における顔面筋肉の動き)を有しているため, 短期間の撮影で個人の全ての表情パターンを取得することは困難である. また, 時々刻々と変化する顔表情は未学習の表情パターンを多数含むため, 表情認識モデルには未学習の表情パターンに適応して進化する機能が必要である.

そこで本稿では, 時間軸に頑強性を有する表情認識モデルの実現を目的とし, 適応学習

機能を有する表情認識モデルに関して基礎検討を行う。

2. 実験データ取得方法

2.1 実験データ取得環境

表情認識の従来研究では、Ekman 氏らによって提唱された顔表情記述法である FACS(Facial Action Coding System)が特徴として広く用いられている⁶⁾。FACS では、人種、文化、性別などによらない普遍的な感情の分類として「怒り」、「悲しみ」、「嫌悪」、「喜び」、「驚き」、「恐れ」の 6 種類の感情が定義されている。

本稿では、6 基本表情の中でも、顔表情の物理的変化が最も豊富かつ顕著に表れると考えられる「喜び」を対象とし、1 名の被験者(日本人 20 代男性)が意図的に表出した喜び表情と無表情の 2 表情に関して基礎検討を行う。

表情画像の取得は、日常一般的な蛍光灯下(約 500~700lx)で行った。被験者の顔は正面向きとし、頭部がフレーム内に含まれる状態で USB カメラ(Logicool 社製 QcamOrbit)を用いて撮影した。また、取得した動画データ(320×480pixel, 15fps, 60sec)を静止画像に変換し、正規化処理を施した上で表情画像データセットを作成した。なお、本研究は「秋田大学手形地区における人を対象とした研究に関する倫理規定第 6 条第 2 項」に基づき、倫理審査の申請を行い、承認を得た研究計画の下、被験者の了承を得てデータを取得している。

2.2 表情画像の正規化

学習に使用する表情画像には、顔の構成要素の形状が保持されていることが必要である。このため、顔の構成要素を含む領域を処理の対象とし、以下の手順に従って顔領域の抽出および正規化を行った。

- 1) 照明条件の変動およびノイズ、ならびに画像の解像度(顔領域のサイズ)に対して頑強な Haar-like 特徴 η を用いて顔の検出を行い、80×96pixel のサイズに正規化した顔領域画像を抽出した。
- 2) 表情認識モデルの学習処理を高速化す

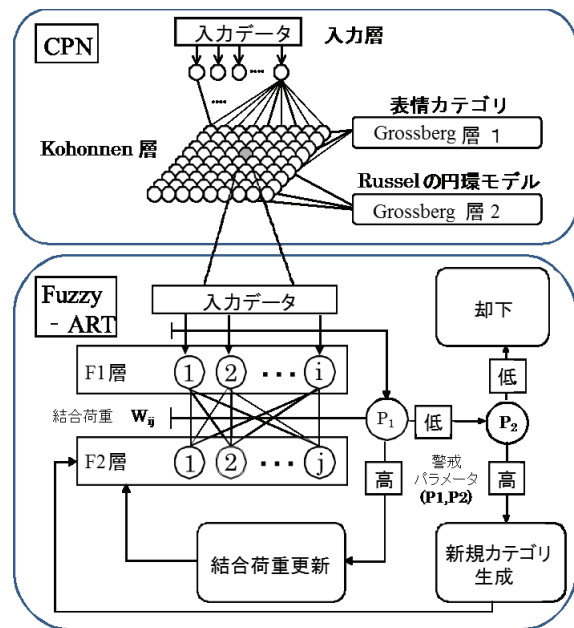


Fig.1 提案手法の概要

Outline of proposed method

るため、粗子化处理(2×2pixel のブロック単位)により画像の次元数(画像サイズ)を削減した後、平滑化处理(40×48pixel)を施した。

- 3) 画像上の(0, 20)–(20, 48)および(40, 20)–(20, 48)の座標間を結ぶ 3 次曲線を引き、その内部を顔領域として抽出した。
- 4) ヒストグラム線形変換により濃度値補正を実施した。

3. 提案手法

提案手法は、対向伝播ネットワーク(CPN)⁸⁾を用いて生成される表情特徴空間(CPN の Kohonen 層)と、特徴空間の各ユニットに連結した適応共鳴理論(Fuzzy-ART)⁹⁾から構成される。提案手法の概要を図 1 に示す。なお、本稿では既存の Fuzzy-ART にもう一つ警戒パラメータを追加し、既存カテゴリと著しく異なるデータ(例えば、喜びのカテゴリに怒りの表情画像が入力された場合)を排除する機能を追加した。以下に提案手法の処理手順を示す。

- 1) CPN を用いて初期学習データを学習し、表情特徴空間を生成する。
- 2) 表情特徴空間の各ユニットに Fuzzy-ART を連結し、各ユニットの結

合荷重を Fuzzy-ART の初期カテゴリとして設定する。

- 3) 追加学習データを表情特徴空間に入力し，勝者ユニットを探索(表情認識処理)する。
- 4) 勝者ユニット上の Fuzzy-ART が有する既存カテゴリと，入力した追加学習データの整合度を算出し，値に応じた警戒パラメータのしきい値判定の結果から，追加学習処理(結合荷重更新，新規カテゴリ生成，却下)を実施する。
- 5) 追加学習を一定期間実施した後，表情特徴空間が有する既存の表情カテゴリおよび Fuzzy-ART が有する新規の表情カテゴリを学習データとし，再学習により表情特徴空間を更新する。

4. 実験手法

はじめに，提案手法における適応学習機能の効果を検証するための基礎実験を行った(実験 1)。実験で用いた CPN および Fuzzy-ART の各パラメータを表 1 に示す。

追加学習の効果を検証しやすくするため，初期学習データを無表情と喜びの 2 表情で構成し，追加学習データには徐々に口を開いていく喜び表情を使用した。初期学習データを用いて初期表情特徴空間を生成し，追加学習データを用いて再学習を 30 回実施した。生成された表情特徴空間および再学習ごとの追加学習処理発生回数に着目し，追加学習機能の検証を行った。

次に，時間経過に伴って出現する多様なパターンを考慮し，長期間にわたって取得した表情画像を用いて検証を行った(実験 2)。実験で用いた各パラメータを表 2 に示す。

表情画像の取得日数は 22 日間であり，初日の画像を初期学習データ，以降の 21 日分の画像を追加学習データとして時間経過を考慮したデータセットを作成した。表情特徴空間を生成する際は，喜び表情と無表情の学習回数が同程度となるように学習を行った。初期学習データを用いて初期表情特徴空間を生成し，以降は 1 日分ごとのデータを再学習することにより表情特徴空間を生成した。その後，生成された表情特徴空間(識別器)に対して全ての追加学習データ(900 枚×21 日)を評価データとして入力し，追加学習処

Table.1 実験 1 で用いた各パラメータ
Parameters used for experiment 1.

	画像データサイズ(pixel)	40×48
初期学習 (CPN)	初期学習データ数(枚)	100(喜び 50, 無表情 50)
	学習回数	20000
	学習率係数	0.5
	表情特徴空間サイズ	1 次元 30,40,50
追加学習 (Fuzzy-ART)	追加学習データ数(枚)	200(喜び)
	警戒パラメータ(P ₁ ,P ₂)	(0.98,0.96)

Table.2 実験 2 で用いた各パラメータ
Parameters used for experiment 2.

	画像データサイズ(pixel)	40×48
初期学習 (CPN)	初期学習データ数(枚)	845(喜び 593,無表情 252)
	学習回数	20000
	学習率係数	0.5
	表情特徴空間サイズ	1 次元 30,40,50
追加学習 (Fuzzy-ART)	追加学習データ数(枚)	18900(900 ×21 日)
	警戒パラメータ(P ₁ ,P ₂)	(0.98,0.96)

理発生回数に着目して追加学習機能の検証を行った。

実験 2 の追加学習では，個々の追加学習データの学習回数が 1 回のみであるため，個々のデータセットの特徴を十分に学習しきれないことが予想される。そこで，1 回の追加学習において追加学習データを複数回学習した条件下において，同様の検討を行った(実験 3)。

なお，上記の実験 2 および実験 3 において 21 日分の追加学習データを入力したところ，21 日目と 22 日目の追加学習データはすべて却下された。したがって，初期学習および再学習によって生成された表情特徴空間(識別器)は 19 個である。

5. 実験結果および検討

5.1 追加学習機能の基礎検証

実験 1 において，表情特徴空間のサイズを 1 次元 30 ユニットとした場合の特徴空間の可視化結果を図 2 に示す。同図は初期学習，再学習 1 回目，12 回目，30 回目の表情特徴空間を示しており，図中の緑色のユニットは喜び表情，桃色のユニットは無表情をそれぞれ

れ意味している。追加学習を重ねるごとに、徐々に口を大きく開いた喜び表情が現われていることがわかる。また、初期表情特徴空間に表現されている表情パターンは、再学習を経た後も同様に表情特徴空間内に保持されていることがわかる。

表情特徴空間のサイズを1次元50ユニットとした場合の再学習回数ごとの追加学習処理発生回数を図3に示す。結合荷重更新処理、新規カテゴリ生成処理の発生回数は、再学習を重ねるごとに緩やかに増加し、収束している。一方、却下処理の発生回数は、再学習を重ねるごとに緩やかに減少し、収束していることがわかる。

以上の結果は、提案手法が初期に学習した表情パターンを保持しつつ、新たな表情パターンを追加的に学習可能であることを示唆している。

5.2 長期的データによる追加学習の検証

実験2における再学習回数ごとの追加学習処理の発生回数を図4に示す。表情特徴空間のサイズは1次元50ユニットである。各識別器における追加学習処理発生回数には、ばらつきが認められた。これは、時間経過による表情パターンを考慮し、追加学習データとして異なる21日分のデータを用いたためと考える。同図の結果からは追加学習の検証が困難であるため、本節では個々の追加学習データに着目した検証を行った。再学習を重ねた19個の表情特徴空間に、1日目、7日目、19日目の追加学習データを入力した場合の追加学習処理発生回数を図5に示す。

同図(a), (b), (c)を比較すると、いずれも追加学習データに対応する識別器において結合荷重更新および新規カテゴリ生成処理が増加し、却下処理が減少していることがわかる。例えば、7日目の追加学習データに対しては、7日目の再学習後の表情特徴空間において結合荷重更新および新規カテゴリ生成処理が増加している。このことは、追加学習データは対応した日にちの追加学習・再学習において正しく学習されていることを示唆している。

一方、7日目の追加学習データを入力した場合(図5(b))に着目すると、追加学習の日に



Fig.2 表情特徴空間
Facial expression feature space

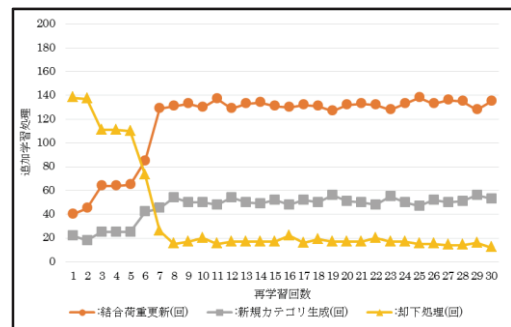


Fig.3 追加学習処理発生回数(実験1)
Frequency of incremental learning processing (experiment 1)

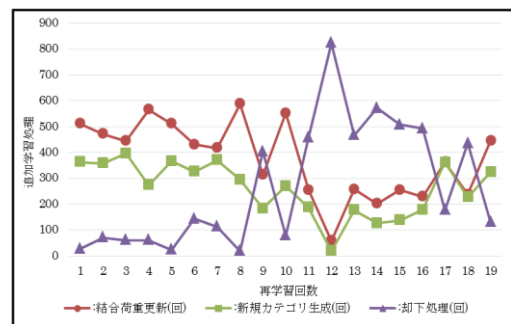
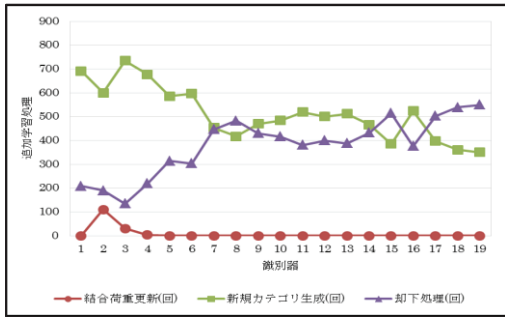


Fig.4 追加学習処理発生回数(実験2)
Frequency of incremental learning processing (experiment 2)

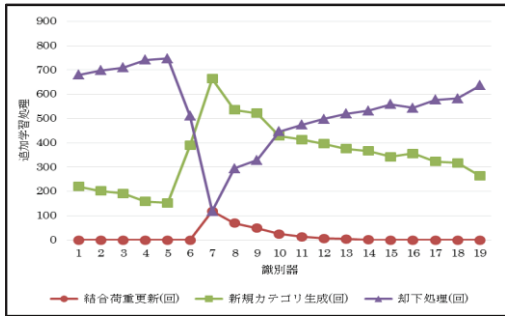
ち(ここでは再学習回数7回目)以降に、結合荷重更新および新規カテゴリ生成処理の回数が徐々に減少していることがわかる。このことは、学習した知識が再学習を重ねるごとに徐々に忘却されていることを示唆している。

5.3 複数回の追加学習による検証

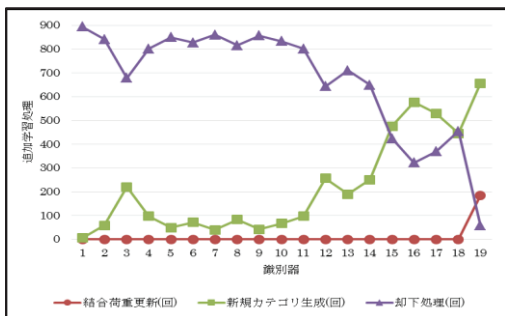
前節では、長期的なデータを用いて追加学習を行った場合、学習した知識が忘却されていることを明らかにした。これは、実験2の追加学習では個々の追加学習データの学



(a) 追加学習データ 1 日目



(b) 追加学習データ 7 日目

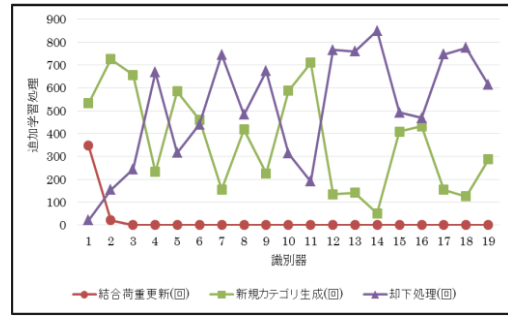


(c) 追加学習データ 19 日目

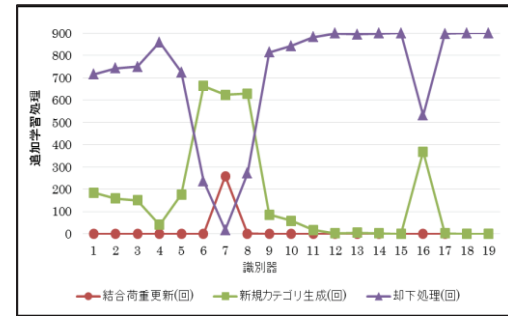
Fig.5 個々の追加学習データに着目した処理発生回数(実験 2)
Frequency of processing in individual incremental learning data (experiment 2)

習回数が 1 回のみであり、個々のデータセットの特徴を十分に学習しきれないためと考える。そこで 1 回の追加学習において追加学習データを複数回学習した場合について、前節と同様の検討を加えた。

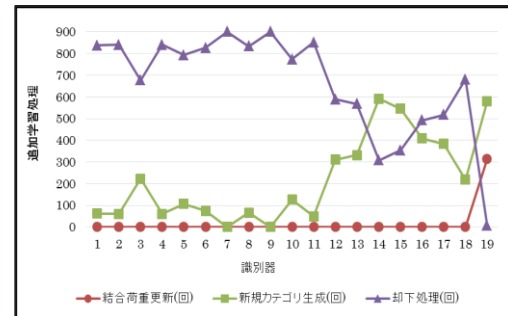
再学習を重ねた 19 個の表情特徴空間に、1 日目、7 日目、19 日目の追加学習データを入力した場合の追加学習処理発生回数を図 6 に示す。1 日分の追加学習データを複数回追加学習させた結果、対応する識別器の却下処理がさらに減少する傾向が認められる。このことは、追加学習が 1 回の場合と比較し、個々のデータセットの特徴が強く学習されていることを示唆している。しかしながら、



(a) 追加学習データ 1 日目



(b) 追加学習データ 7 日目



(c) 追加学習データ 19 日目

Fig.6 個々の追加学習データに着目した処理発生回数(実験 3)
Frequency of processing in individual incremental learning data (experiment 3)

図 5(実験 2)では却下処理が徐々に増加しているのに対し、図 6(実験 3)では急激に増加する様子が認められた。このことは、個々の追加学習データの学習が強く反映されたため、他の追加学習データに対する汎化能力が低下したことを示唆している。

6. おわりに

本稿では、時間軸に頑強性を有する表情認識モデルの実現を目的とし、Fuzzy-ART を用いた適応学習機能を有する表情認識モデルに関して基礎検討を行った。

はじめに、無表情と喜びの 2 表情を用いて

表情特徴空間を生成し、再学習ごとの追加学習処理発生回数に着目して追加学習機能の検証を行った。また、時間経過に伴って出現する多様なパターンを考慮し、長期間にわたって取得した表情画像を用いて追加学習機能の検証を行った。その結果、提案手法は初期に学習した表情パターンを保持しつつ、新たな表情パターンを追加的に学習可能であることを明らかにした。さらに、学習した知識は長期的な再学習を重ねるごとに徐々に忘却されていくことを明らかにした。

次に、追加学習で獲得した知識の長期的な保持を目的とし、複数回の追加学習を行った場合の追加学習機能について検証を行った。その結果、追加学習が1回の場合と比較して個々のデータセットの特徴が強く学習されたものの、表情特徴空間の汎化能力は低下することが明らかになった。

今後は、追加学習によって得た知識の長期的な保持を目的とし、表情特徴空間の最適なサイズの自動決定手法および追加学習に用いる警戒パラメータの自動設定手法に関して検討を行う予定である。

参考文献

- 1) 赤松茂, “人間とコンピュータによる顔表情の認識[I]-コミュニケーションにおける表情とコンピュータによるその自動解析-”, 信学誌, Vol. 85, No. 9, pp. 680-685, 2002.
- 2) 赤松茂, “人間とコンピュータによる顔表情の認識[II]-コンピュータによる顔表情認識技術(1):表情による感情の認識-”, 信学誌, Vol. 85, No. 10, pp. 766-771, 2002.
- 3) 赤松茂, “人間とコンピュータによる顔表情の認識[III]-コンピュータによる顔表情認識技術(2):顔面の表情動作の認識-”, 信学誌, Vol. 85, No. 12, pp. 936-941, 2002.
- 4) 赤松茂, “人間とコンピュータによる顔表情の認識[IV・完]-コンピュータによる顔表情認識の展望:人間による表情認知に学ぶ-”, 信学誌, Vol. 86, No. 1, pp. 54-61, 2002.
- 5) M. Pantic and L. J. M. Rothkrantz, “Expert system for automatic analysis of facial expression,” Image Vis. Comput., vol. 8, no. 9, pp. 881-905, 2000.
- 6) P. Ekman and W. V. Friesen, Facial action coding system, Consulting Psychologist Press, 1978.
- 7) R. Lienhart and J. Maydt, “An extended

set of haar-like features for rapid object detection,” Proc.IEEE Int. Conf. Image Processing, Vol.1, pp. 900-903, 2002.

- 8) R. H. Nielsen, “Counterpropagation networks,” Applied Optics, Vol. 26, No. 23, pp. 4979-4984, 1987.
- 9) G. A. Carpenter, A. Grossberg, and D. B. Rosen, “Fuzzy-ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system,” Neural Networks, Vol. 4, pp. 565-588, 1991.