

並列砂時計ニューラルネットワークと情緒生起手法を用いた感情指向型インタフェースの応用

Approach to emotion oriented intelligent system by parallel sand glass type neural networks and emotion generating calculations

市村 匠¹⁾, 石田 与志²⁾, 目良 和也¹⁾, 大枝 真一³⁾, 杉原 亮宏⁴⁾, 山下 利之³⁾
Takumi ICHIMURA¹⁾, Hitoshi ISHIDA²⁾, Kazuya MERA¹⁾, Shinichi OEDA³⁾, Akihiro SUGIHARA⁴⁾,
Toshiyuki YAMASHITA³⁾

1)広島市立大学情報科学部, 2)東芝情報システム株式会社, 3)東京都立科学技術大学工学研究科,
4)広島市立大学大学院情報科学研究科

1) Faculty of Information Sciences, Hiroshima City University, 2) Toshiba Information Systems,

3) Graduate School of Engineering, Tokyo Metropolitan Institute of Technology,

4) Graduate School of Information Sciences, Hiroshima City University

Abstract: We developed an application soft with emotional facial expression that can analyze emotions of the user and represent emotions with facial expressions. The application has two outstanding functions. One function is displaying emotional faces and the other is analysis of emotion from some sentences. The face displaying part is due to a sand glass type neural network by training the real face images. The emotional analyzing part is due to Emotion Generating Calculations (EGC) method based on the Elliot's Emotion Eliciting Condition Theory. The proposed method can judge whether an event is pleasant or not and can obtain each emotion value for 20 various emotions under the situation. Successfully, these attributes are set to the trained neural network as input signals. In this paper, our proposed system we call FaceMail is introduced as a kind of emotional oriented intelligent systems.

1. はじめに

高度情報化社会の発展に伴い、一般家庭にもコンピュータが広く普及している。特に、インターネットは新しいコミュニケーションの手段として、Webやメールなどと、幅広く使われている。中でも、メールは手紙の代わりになるもので、口頭では伝えられなかった気持ちを手紙に託して相手に伝えるという手紙の特色を、そのまま活かすことが可能であると言えよう。メールに動画ファイルを添付するなど、メールを利用した対面的コミュニケーションも実現されようとしている。これらは、通信技術の向上のみならず、人間とコンピュータとのインタラクションにおいて、ヒューマンインタフェースの開発が進められていることが、多くの利用者に賛同を得ている結果となっている。

このようなインタフェースが利用されるようになってきた理由として、コンピュータとのインタラクションにおいて、人間同士の対面的コミュニケーションと同様に“楽しさ”が求められるようになってきており、システムを使用することで、“楽しさ”を感じられるようになってきている点がある。人間同士の対面的コミュニケーションの場合、相手の表情から感情を推測し、その感情に応じて話しの内容、言い回し、表現を変えたりしながら円滑なコミュニケーションが行われている。人間とコンピュータとのインタラクションにおいても、コンピュータに人間らしさを与えるために、顔の画像をインタフェースとして用いる試みがされている^[1,2]。

山下らは、ファジィ推論による表情選択モデルを

提唱し、これを用いたインタフェースを開発した^[3]。これは、文章に表されている感情を解析し、比較的単純な線画による画像で表現するシステムであった。本研究では、このシステムをさらに発展させ、人間の顔の表情画像（以下、顔画像）を表示するシステムを開発した^[4]。この顔画像表示においては、並列砂時計型ニューラルネットワークによる学習を用いた^[5]。この手法は、まず複数の顔画像を学習することで、ニューラルネットワーク内に、感情空間を構築し、学習後のネットワークに感情を示す値を与えることで、その感情を示す顔画像を出力するものである。山下らのシステムで用いられている文章の解析部、すなわち文章を解析することにより、そこに含まれる感情を抽出する試みに対し、本研究では、情緒生起式^[6]を適用した。

本論文では、並列砂時計型ニューラルネットワークおよび情緒生起式について、簡単な説明を与え、開発したシステムについて、詳細を述べる。

2. 並列砂時計型ニューラルネットワーク^[5]

2.1 顔画像からの感情の分離

顔画像を分析することで、そこに表出する特徴から類推される感情の分類を行い、その構造を解明する研究が多く行われている。特に、顔画像と感情を関連付けるために、ニューラルネットワークにより顔画像を学習する研究が行われている。これらの中で、砂時計型ニューラルネットワークは、顔画像を感情に基づいて分類する研究であり、大変有効な手法であると考えられる。しかし、複数種類の情報を

同時に扱うことが困難であるため、1人分の顔画像を学習するのに用いられていた。このため、複数人の顔画像を処理するためには、異種の情報を取り扱うことを可能としたネットワークが必要である。これには、拡張砂時計型ニューラルネットワークが提案されている。このモデルを用いて、男女各1名、計2名の顔画像を学習し、感情空間を構築した。しかし、このモデルでも3人以上の顔画像を分類することができなかった。そこで、我々はN個の5層砂時計型ニューラルネットワークを、それぞれの第3層で連結した、並列砂時計型ニューラルネットワークを提案した(図1)。

2.2 顔画像の切り出し

表情認識において、空間周波数成分の低周波部分の認識が重要であるとされる。そこで、顔画像を2次元離散コサイン変換し、その低周波成分を用いて学習を行うことにより、パラメタの取り出しを必要としない手法を導入する。

さらに、教師信号として用いるため、撮影した写真を8ビットグレースケールに変換し、その後、画像の回転による影響を避けるため、回転補正を行った。さらに髪の毛などの影響を避けるため、顔の中心部分128×128ピクセルを切り出し、画像の正規化を行った(図2)。

2.3 基本感情

Ekmanらによると、顔の表情と感情について詳細な検討を行っており、ポーズをとった顔の写真から判断される基本的な感情カテゴリとして、“喜び、悲しみ、嫌悪、怒り、恐れ、驚き”の6種類がある^[12]。また、Schlosbergによると、表情には“愛-驚き-恐れ-怒り-嫌悪-軽蔑”の順に並んだ、円環状の順序が存在するという。この2つの提案をもとに、Ekmanの6種類の基本感情を、“喜び-悲しみ-嫌悪-怒り-恐れ-驚き”の順に並べた円環状の順序が、表情と感情の間に存在するとした研究が多く見られる。この円環状に並んでいる基本感情について、相互に近い感情どうしは混同されることがあるが、離れている感情どうしが混同することはないとされている。

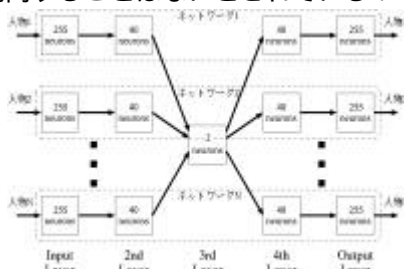


図1 並列砂時計型ニューラルネットワーク

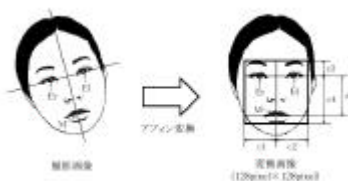


図2 顔画像の切り出し

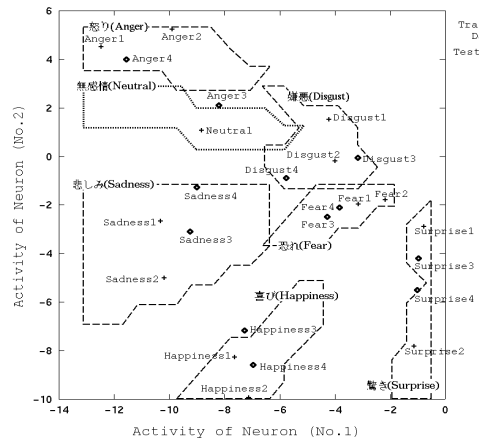


図3：学習後の第3層の活性化値(N=4)

2.4 実験結果

ここで、4人分の顔画像を学習した場合の実験結果を示す。図3は、学習後の第3層における活性化値を示している。2個のニューロンの活性化値を、2次元平面上に分布したものであり、このマップを「感情空間」と呼ぶ。結果として、環状空間にはEkmanらが提案した顔画像を円環状に並べた順序が表れている。

3. 情緒生起式^[6]

3.1 情緒生起のプロセス

本研究では、メールの文に含まれる、ある事象に対して急激に引き起こされる一時的な感情を情緒、さまざまな情緒によって変化する持続的な感情を気分として、感情を情緒と気分の2つに大別している。相手発話の内容を自分の価値観に基づいて判断し、その結果生じた快/不快を情緒とする。また、過去に生じた情緒の影響を累積したものを気分としている。

本研究では、メールに書かれている文章から情緒を生起するために、メールに含まれる語の好感度と、事象の格フレーム表現の意味構造に注目している。岡田らが定義した全12種類の事象のタイプに対して、目良らは8種類の情緒生起式を定義した。情緒生起式に、予め設定している語の好感度を代入することで、メールに書かれた内容に含まれると考えられる情緒を求める。情緒生起は、快/不快の生起判別と、強度計算の2段階に分けられる。まず、事象がメール送信者にとって快なものか、不快なものかを判別する。判別には事象中の各要素に対して好感度データが使用される。この好感度データは、その事象に対する好き/嫌いの度合を示しているもので、予め用意されている。これらの語の好感度を12種類の事象タイプに応じた情緒生起式に代入することで、快/不快の生起を判別する。情緒生起式によって生じた快/不快は、Elliottの感情誘発条件理論に基づき、20種類の情緒に分類される。情緒の強度は、情緒生起式によって快/不快を判別した後、情緒生起において、生じられた情緒の距離を計算することにより求められる。

3.2 好感度

好感度とは、ある対象について好き / 嫌いの度合を区間[-1.0, 1.0]の実数値で表現したもので、好きな事物に対し正、嫌いな事物に対し負、何とも思っていないならば0を与える。また、好感度には、あらかじめ与えられるものと、相手との対話内容や発話時の表出した感情から推測することで、状況に応じた好感度を学習により更新することができるものがある。好感度の値の更新は、相手自身による好き / 嫌いの表明の検出、好感度変化条件の使用、不快な事象内に表れるものの好感度を下げる、情緒生起式と現在の感情から逆算する、といった4つの手法により履行される。

3.3 情緒生起式

情緒生起式は、事象概念の格フレームのタイプごとに作成されている。ここで、事象概念には、格要素の結び付きによって表される単純事象概念と、単純事象概念の結び付きによって表される連結合成概念がある。この単純事象概念を表現するための、最小限必要不可欠な格要素のことを必須要素と呼ぶ。必須要素は、S(主体)、O(客体)、OF(出発点または源)、OT(目標)、OM(相互作用の相手)、OS(投げ所)、OC(属性の補足)およびI(道具)の8種類である。

事象概念を、表現するために必要となる格要素の組み合わせのタイプに分類する。このタイプは、岡田が、国立国語研究所で編纂された「分類語彙表」に収録されている、約4700の概念を12種類の単純事象概念と連結合成概念に分類したものである。本研究では、この分類に基づき、同じ格フレームタイプの事象概念は、人間が意味を捉える上でも、同じ意味構造をもつと考えている。これらの意味構造的に同じ形式の事象、つまり同じ格フレームタイプごとに、情緒生起式を与えている。情緒生起式は、表1のように、格フレームのタイプ別に定義されている。

3.4 情緒タイプの分類

情緒生起式によって求められた事象の情緒値に基づいて、各情緒タイプの快 / 不快の強度が求められる。本論文では、Elliottの提唱している24種類の情緒のうち、20種類を対象とする。

さらに、複合情緒として「幸福 / 帰属」から「感謝」、「怒り」、「自己満足」、「自責の念」、「魅力 / 帰属」から「愛」と「憎しみ」が生起する。このうち「好き」と「嫌い」は、現在の情緒生起式では生起できないので、今回は考慮していない。「愛」と「憎しみ」もこの「好き」「嫌い」を用いるため、同様な理由で考慮していない。図4は各情緒タイプの依存関係を示している。

3.5 情緒強度計算手法

語の好感度の積から、快 / 不快 / どちらでもないを判別していたが、好感度が[-1.0, 1.0]の実数値をとるため、情緒生起式から得られる情緒値も[-1.0, 1.0]の実数値となる。この値は、情緒生起式の各項に用いられる格要素の好感度の強度に比例する。

表1 事象タイプと情緒生起式

タイプ	情緒生起式
$V(S)$	$f_S \times f_P$
$V(S, OF)$	$f_S \times (f_{OT} - f_{OF}) \times f_P$
$V(S, OT)$	$f_S \times (f_{OT} - f_{OF}) \times f_P$
$V(S, OM)$	$f_S \times f_{OM} \times f_P$
$V(S, OS)$	$(f_S - f_{OS}) \times f_P$
$V(S, O)$	$f_S \times (f_O \times f_P)$
	$f_O \times f_P$
$V(S, O, OF)$	$f_O \times (f_{OT} - f_{OF}) \times f_P$
$V(S, O, OT)$	$f_O \times (f_{OT} - f_{OF}) \times f_P$
$V(S, O, OM)$	定義なし
$V(S, O, I)$	$f_O \times f_P$
$V(S, O, OC)$	$f_O \times f_{OC}$
その他	定義なし

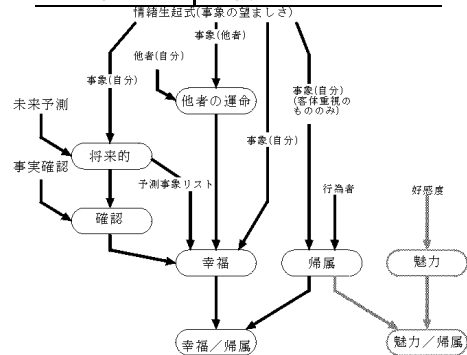


図4 感情生起の依存関係

したがって、事象に含まれる語に対して、より強い「好き嫌い」のイメージをもっているほど、その事象はより強い快 / 不快を喚起する、という人間の感覚に一致している。これより、この情緒式から生起情緒の強さを求める。しかし、情緒生起式では、好感度の絶対値が1以下であるので、積を取ると値が小さくなってしまふ。そこで、情緒生起式の各項の直交ベクトルと考え、それらが成す直方体の対角線ベクトルの大きさを、生起された情緒の強度とする。2項以下しかない情緒生起式が適用される場合には、3番目項目に一般的な値として0.5を与えている。

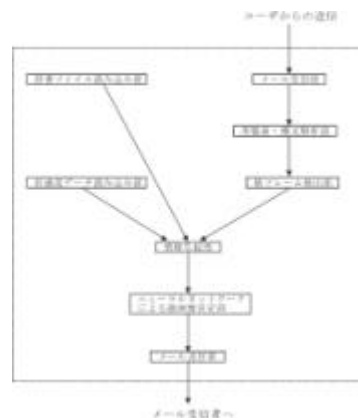


図5 システムの構成

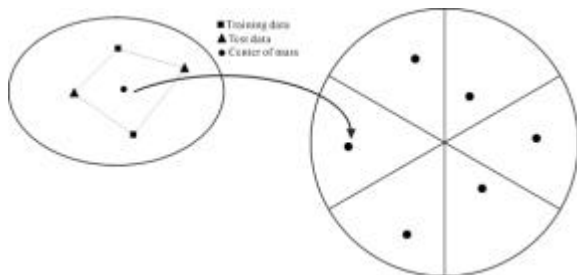
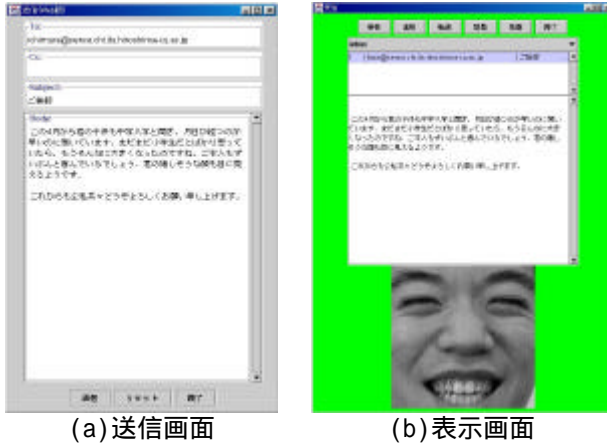


図6 感情空間での各感情に対する入力点



(a)送信画面 (b)表示画面

図7 FaceMailの実行画面

4. 感性インタフェースへの応用^[4]

本研究で提案した並列砂時計型ニューラルネットワークと情緒生起手法による、メールツールを開発した。このメールツールは、文章中に含まれるであろう感情を分析・表現可能である。本システムは、サーバ側とクライアント側の2つのプログラムが動作する。

図5は、本システムのサーバ側の構成である。つまり、メールは、クライアントで動作するインタフェースで入力され、ネットワークを経由してサーバに送られる。サーバでは、形態素・構文解析部、格フレーム抽出部、情緒生起部、顔画像決定部、メール送信部に分けられる。クライアントでは、通常のメール機能の他に、メールヘッダーに付加された解析結果より、顔画像を表示する機能をもつ。

形態素にはJUMANを、構文解析にはKNPを用いた。格フレーム抽出部は、構文解析結果を格フレーム表現に変換する。KNPの出力では、述部とそれに係る格要素が得られる。顔画像決定部では、情緒生起部で決定された情緒は20種類であり、並列砂時計型ニューラルネットワークで分類した顔画像の表情は6種類であるため、20種類の情緒を6種類に変換する必要がある。本研究では、喜び、悲しみ、困惑、怒り、嫌悪、驚きの語の意味を考慮することで、情緒を分類している。

これらの6種類の感情を表す語に対し、情緒生起式により得られた、各情緒の強さを与えることで、6種類の感情に対する強さが得られる。本論文では、メール全体で表現されている感情を総合的に判断するため、6種類の感情を1つにまとめる必要がある。このため、生起された感情の個数を数え上げ、最も

多く出現した感情の表情を提示するようにした。

さらに、求められた6種類の感情とその強度を、並列砂時計型ニューラルネットワークで構築された感情空間へ入力し、出力として顔画像を得る。ここで問題となるのが、顔画像の決定は、感情空間での入力位置に依存している。本論文では、各感情に対して、学習に2種類の顔画像、検証のために2種類の顔画像を用いているので、各感情の入力点は、図6のように、4点で結ばれた四角形の重心とした。

一方、クライアント側で動作するプログラムは、顔画像の表示機能をもった専用のメールソフトである。これはFaceMailと呼び、Java Swingを利用して開発されたものである。図7はFaceMailの送信画面と受信画面を表示したものである。

5. まとめ

本システムでは、特定の状況下に限定されているとはいえ、人間の感情を分析・表現するインタフェースを開発した。本研究で開発した技術は、コンピュータを楽しんで使うためだけではなく、高齢者の在宅健康診断システムなどにも、応用されている。このような技術は、コンピュータを使うために、我々が特殊な技術を習得するのではなく、人間同士の対面的コミュニケーションで見られる場面を再現し、コンピュータとのインタラクションにより、長時間の仕事・作業を進める際に生じる、コンピュータに伴うストレスを軽減しようとする試みを実現するものである。

参考文献

- [1] 原島 博, "顔, 表情, そして感情 - 映像ロボット技術からのアプローチ -", 日本機会学会誌, Vol.95, No.883, pp.532-540(1996)
- [2] 岩田 満, 鬼沢 武久, "顔の表情表現と結びつけた経路決定システム", 日本ファジィ学会誌, Vol.8, No.3, pp.532-540(1996)
- [3] 山下 利之, 高橋 雅博, 酒井 秀昭, 武田 利浩, 市村 匠, "ファジィ推論による表情選択モデルのヒューマンインタフェースへの応用", 日本ファジィ学会誌, Vol.12, No.2, pp.313-320(2000)
- [4] T.Ichimura, K.Mera, H.Ishida, S.Oeda, A.Sugihara, T.Yamashita, "An Emotional Interface with Facial Expression by Sand Glass Type Neural Network and Emotion Generating Calculations Method", submitted to Proc. of ISHF2001
- [5] 石田 与志, 市村 匠, 寺内 睦博, 高濱 徹行, 磯道 義典, "砂時計型ニューラルネットワークを用いた表情画像の分類", 第10回FAN講演論文集, pp.201-204(2000)
- [6] 目良 和也, 市村 匠, 相沢 輝昭, "語の好感度に基づく自然言語発話からの情緒生起手法", 人工知能学会誌投稿中

連絡先

〒731-3194 広島市安佐南区大塚東 3-4-1
 広島市立大学情報科学部 市村 匠
 Phone&Fax: 082-830-1690
 E-mail: ichimura@its.hiroshima-cu.ac.jp