

ニューラルネットワークによる高齢者向け表情推論システム

Construction of a dialog system with emotions for elderly persons by neural networks

市村匠

Takumi ICHIMURA
広島市立大学情報科学部
Faculty of Information Sciences,
Hiroshima City University

目良和也

Kazuya MERA
広島市立大学情報科学部
Faculty of Information Sciences,
Hiroshima City University

山下利之

Toshiyuki YAMASHITA
東京都立科学技術大学工学部
Faculty of Engineering,
Tokyo Metropolitan Institute of
Technology

Abstract: We developed an interactive system with dialog function that can analyze emotions of the user and response to him/her with facial expressions. The dialog system is due to our Emotion Invoking Calculations (EIC) method based on the Elliot's Emotion Eliciting Condition Theory. The proposed method can judge whether an event is pleasant or not and obtain each attribute value for 20 various emotions under that situation. Successfully, we consider adjective relation and adverbial phrases and we input their attributes into the EIC, such that we can evaluate strength of each emotion.

This system can depict the face according to the context of conversation. Although each attribute value in EIC shows one aspect of feeling we must integrate those attributes to express human-like emotions. Then, we applied a fuzzy reasoning model for selecting the face proposed by T.Yamashita. In this method, nine faces with three levels of brow and eye deflection and three levels of mouth deflection were used to construct the fuzzy reasoning model. In this study, we extend Yamashita's system with eight emotions to a sophisticated system with 20 emotions by neural network.

1. はじめに

高度情報化社会の発展に伴い、情報端末を用いた高齢者の健康支援システム^[1]が開発されているが、コンピュータの操作方法の複雑さ故に、利用しにくいという問題点がある。そのため、人間とコンピュータとのインタラクションにおいても、人間同士の対面的コミュニケーションと同様なシステムが求められるようになってきた^[2]。

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る^[3]。このように、対面的コミュニケーションにおける表情の役割が大きいため、心理学においても表情からの感情認知に関する研究は多い^[4]。そこで、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（以下、顔と略記）をインタフェースとして用いる試みが多く行われている^[5,6,7]。

一方、特定の状況における感情を顔で伝えるためには、相手の発話内容を解析する必要が生じる。ここでは、発話内容から快/不快を生起するための情緒生起式を与え、生起される情緒の種類と感情の強さを示す度合いを求めることにより、20種類の情緒を生起する手法^[8]を適用し、相手の感情を分析する。生起された情緒からそれに伴う顔の表情を選択するため、ニューラルネットワークによるファジィ推論モデルを提唱し、人間とコンピュータとのインタラクションに感情を考慮した対話システムの実現を試み、これを対話レベルでの問診を実現可能とする高齢者健康診断システムに適用することを目的とする。

2. 情緒生起手法と属性値

相手の発話内容から快/不快を生起するための計算式（情緒生起式）とElliottの感情誘発条件理論^[9]で定義された感情型を組み合わせることにより、次の20種類の情緒を生起する手法がある。つまり、「喜び」、「苦しみ」、「嬉しい」、「ほくそえむ」、「憤慨」、「気の毒な」、「望み」、「恐れ」、「満足」、「安堵」、「恐れていた通り」、「失望」、「誇り」、「賞賛」、「羞恥」、「叱責」、「感謝」、「怒り」、「自己満足」、「自責の念」である。さらに、生起されるこれらの情緒に対し、その強さの程度を表現可能とする手法が目良により提案されている。ここでは、この手法を簡単に説明する。

2.1 情緒生起式

格フレーム形式で表現された生起事象があり、主体、客体など格要素に対する語の好感度を事象のタイプごとに予め用意された情緒生起式に代入することで、事象の快/不快を求める。事象のタイプとそれに対応する情緒生起式は表1で示すものとする。また表中で使われている記号を次のように定義する。

f_S : 主体(Subject)の好感度

f_O : 客体(Object)の好感度

f_{OF} : 出発点(Object-From)の好感度

f_{OT} : 目標(Object-To)の好感度

f_{OM} : 相互作用の相手(Object-Mutual)の好感度

f_{OS} : 抛り所(Object-Source)の好感度

f_{OC} : 属性の補足(Object-Content)の好感度

f_P : 述語(Predicate)の好感度

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

この情緒生起式では、1文を主体、客体、出発点、目標、相互作用の相手、抛り所、属性の補足、述語に分割したそれぞれの好感度の内積から求められ、情緒値としてその正負の符号のみで与えられていた。ところが、快/不快にも程度があり、情緒生起における好感度以外の要素から情緒の強さを表す測度が必要となっている。そこで、[8]では「強度にのみ影響する格要素を情緒生起式に追加」、「好感度(好き/嫌い)以外の属性の使用」、「修飾構造への対応」を考慮する。これには、各要素に対して注目すべき属性を事象ごとに指定し、さらに各要素となる対象物にも好感度以外の様々な属性に対する値を設定する手法が考えられている。(詳細は、[8])

3. ファジィ推論による表情選択モデル

次に、ファジィ推論による表情選択モデルの構築について述べる。

3.1 顔の画像と基準感情

本研究では、図1に示すような、眉と目の傾きに関する3条件、口の形の3条件の組み合わせによる計9種類の顔を用いた。

3.2 ファジィ推論規則の構成

[7]では、質問紙調査の結果より、各基準感情に対して各顔が選択された比率を求めた。本研究では、2節で述べた情緒生起により求められた属性値を、各顔がその基準感情を表す表情に属する程度を表すメンバーシップ度とした。

この20個の情緒をもとに、図2に示すような、20個のファジィ規則からなるファジィ推論モデルを仮定した。すなわち、20個の各々のファジィ規則は、前件部が生起された情緒の属性値を表し、後件部はその情緒を表す表情に各顔が所属する程度を表す。

ある状況における各々の情緒の属性値が入力として与えられたとき、Mamdaniのファジィ推論法に従って、それらの複合された感情を表す表情に各顔が所属する程度を表すメンバーシップ関数が推定される。

すなわち、各情緒を表す集合を A_1, A_2, \dots, A_{20} とし、各々の基準感情を表す表情に各顔が所属する程度を表すファジィ集合を P_1, P_2, \dots, P_{20} とし、ある状況における各基準感情の程度を a_1, a_2, \dots, a_{20} とすると、ファジィ推論モデルは以下のように表される。

$$\text{規則 1 : } A_1^1 \text{ and } A_2^1 \text{ and } \dots \text{ and } A_{20}^1 \quad P_1$$

$$\text{規則 2 : } A_1^2 \text{ and } A_2^2 \text{ and } \dots \text{ and } A_{20}^2 \quad P_2$$

.....

$$\text{規則 n : } A_1^n \text{ and } A_2^n \text{ and } \dots \text{ and } A_{20}^n \quad P_n$$

$$\text{入 力 : } a_1 \text{ and } a_2 \text{ and } \dots \text{ and } a_{20}$$

結 論 : P'

ここで、入力 a_i とファジィ規則 n「 A_1 and A_2 and... A_{20} P 」から得られる推論結果 P_n' は、

$$m_{P_n'}(z) = a_i \quad m_{P_n}(z)$$

で与えられ、最終的な結論 P' は、

$$\mu_{P'}(z) =$$

$$\mu_{P_1'}(z) \quad \mu_{P_2'}(z) \quad \dots \quad \mu_{P_n'}(z)$$

で与えられる。

そして、その状況にふさわしい感情を顔によって表現する場合、最大メンバーシップ度を持つ顔、あるいは2番目に大きなメンバーシップ度を持つ顔などを選択して表示する。

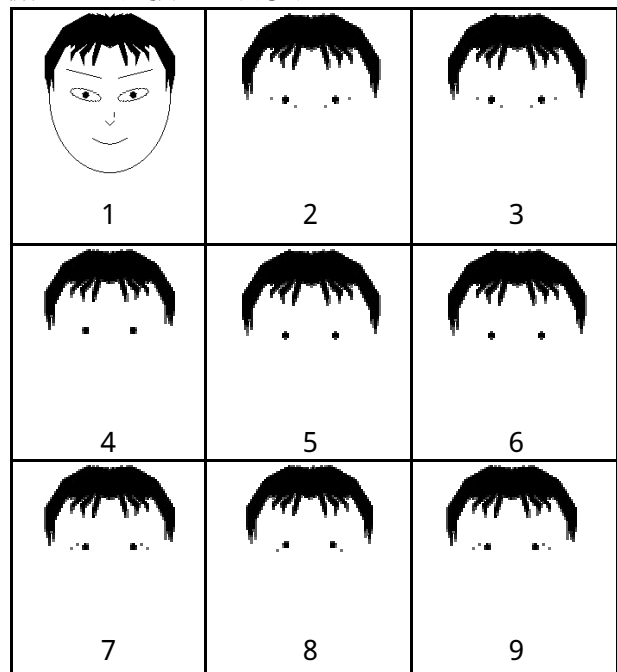


図1 本研究で用いた9種類の顔

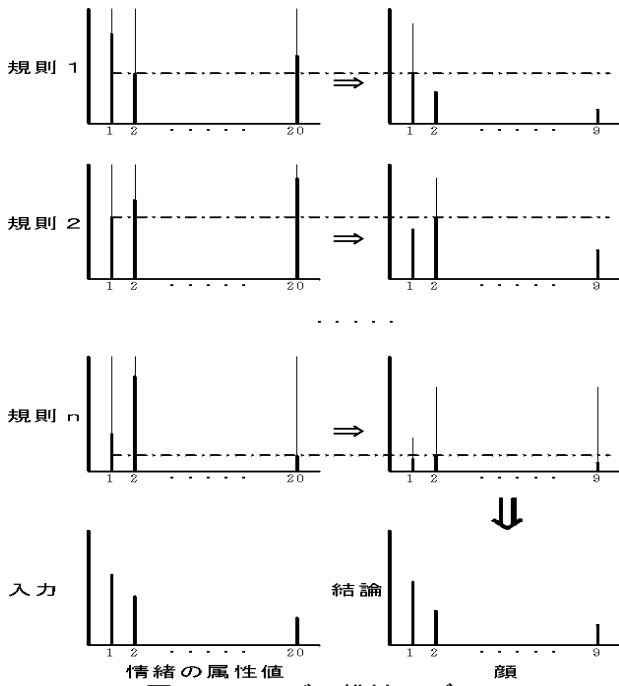


図2 ファジィ推論モデル

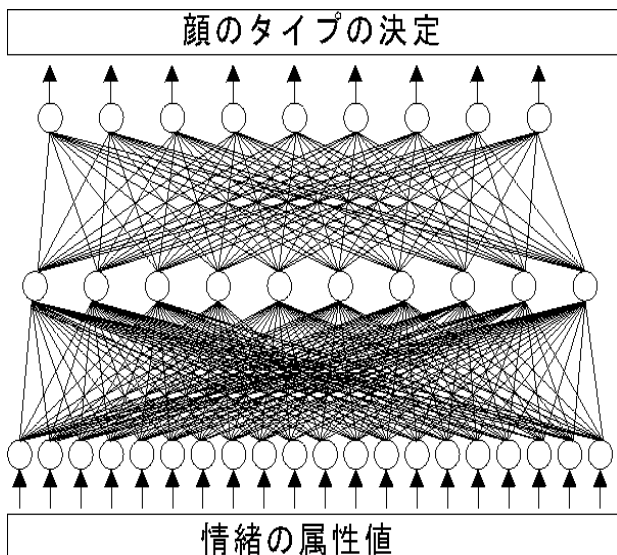


図3 顔決定のためのニューラルネットワーク

4. ニューラルネットワークによる表情選択モデル

3. 1節で述べたファジィ推論規則をニューラルネットワークで実現する。図3は表情選択を行うファジィ推論を実行するニューラルネットワークで、入力層、隠れ層、出力層の3層から構成される。入力層には $[-1,1]$ に変換された情緒の属性値 $a'=(a_1, a_2, \dots, a_n)$ を与える。出力層では、最大メンバーシップ度を持つ顔、あるいは2番目に大きなメンバーシップ度を持つ顔が選択されるように学習を行う。教師データは、予め与えた情緒の属性値とそれに対し決定された顔のタイプのペアを元に行われる。

いま、入力値 a' 、その出力活性の期待値 b' が与えられたときの各ニューロンからの出力活性は次のようになる。

入力層活性値：

$$O_i^I = a_i', \quad i=1,2,\dots,n \quad (1)$$

隠れ層活性値：

$$O_j^H = f(\text{net}_j^H), \quad j=1,2,\dots,m \quad (2)$$

$$\text{net}_j^H = \sum_{i=1}^n w_{ji} O_i^I + \mathbf{q}_j^H, \quad j=1,2,\dots,m \quad (3)$$

出力層活性値：

$$O_k^{\text{out}} = f(\text{net}_k^{\text{out}}), \quad k=1,2,\dots,l \quad (4)$$

$$\text{net}_k^{\text{out}} = \sum_{j=1}^m u_{kj} O_j^H + \mathbf{q}_k^{\text{out}}, \quad k=1,2,\dots,l \quad (5)$$

入出力関数 $f(\text{net})$ は次のようなシグモイド関数とする。ただし定数 $K>0$ は十分に小さい数とする。

$$f(\text{net}) = \frac{1}{1 + \exp^{K(-\text{net})}} \quad (6)$$

また、学習は次の評価関数が最小になるように行われる。

$$E_k = \frac{(b_k' - O_k^{\text{out}})^2}{2}, \quad k=1,2,\dots,l \quad (7)$$

重みの修正は、次のようになる。

$$\Delta_p w_{ji} = \mathbf{H} \left(-\frac{\partial E_i}{\partial w_{ji}} \right), \quad p=1,2,\dots,m \quad (8)$$

$$\Delta_p u_{kj} = \mathbf{H} \left(-\frac{\partial E_i}{\partial u_{kj}} \right), \quad p=1,2,\dots,l \quad (9)$$

属性値の値から顔のタイプを決定する際、出力活性値が最も大きい、2番目に大きいニューロンに一致する顔のタイプを表示することになる。

5. 高齢者健康診断システムへの適用

高齢化社会を迎えるにつれ、介護を要する者に対し、介護保険法が施行された。また厚生省が「21世紀における国民健康づくり運動（健康日本21）」を策定し、健在者に対する健康管理が求められている。一方、高度情報化社会に伴い、コンピュータが身近な存在になってきており、図4のような情報端末を用いた高齢者の健康診断システムが開発されている^[1]。

ところが、コンピュータを操作するために、まずその操作方法について学習をしなければならないが、その操作方法の複雑さ故に、コンピュータの使用を拒否する人も多く存在している。そのため、高齢者がコンピュータを利用するには、ど

のような機能性が必要であるかを調査した結果、感情を考慮した対話を可能なコンピュータがあれば利用するという結果が得られた。

そこで本論文では、図4で示されているような問診に対する回答を YES/NO の2種類に限定することなく、通常のコミュニケーションと同様に回答可能とし、その内容を分析することで人間同士の対面的コミュニケーションの場合と同じように、コンピュータに人間らしさを与える。

まず広島市内の老人ホームで、高齢者 15 人に対し、予め用意した質問に対し口頭で回答を求めた。これらの音声データをテキスト化し、回答された文章を分析した結果により、2 節で述べた方法により各情緒の属性値を求める。例えば次のように属性値が求まる。

質問「道に迷うことがありますか。」
 回答「広島市内では迷うことはないですね。」
 場所：広島市内
 主体：相手・・・属性値 +0.8
 事象：迷う・・・属性値 -0.5
 様相：否定・・・属性値 -1.0
 となりこれらの積により属性値が求まる。つまり、
 $+0.8 \times \{(-0.5) \times (-1.0)\} = +0.4$

さらに、図5をもとに情緒の種類が求められ、この例では「喜び」、「嬉しい」、「賞賛」の3種類の情緒が生起される。このときの教師信号は、次のようになる。

{0.4, 0.0, 0.4, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.4, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0}

これらの教師信号を与え、ニューラルネットワークの学習を行い、実際の対話システムに適用した。



図4 Web 上の高齢者健康診断システム

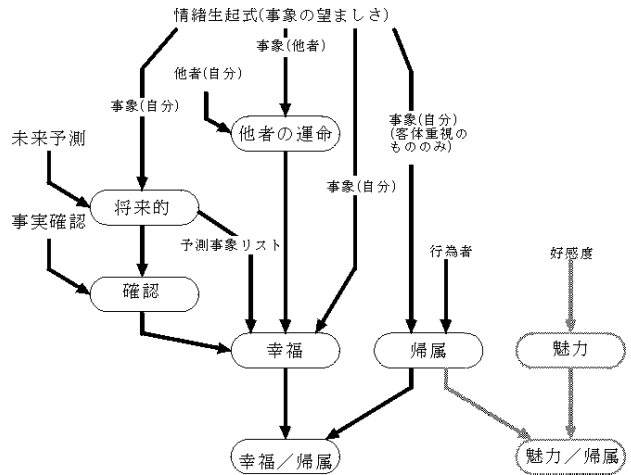


図5 感情生起の依存関係

6. 結論

本論文では、高齢者の返答した文章における情緒を分析し、それらを統合することにより、顔のタイプを表示した。人間同士の対面的なコミュニケーションには、返答内容を分析し、相手の発話意図を理解する処理する手法が必要となり、開発が急がれるところである。

参考文献

- [1]K.Yoshida, T.Ichimura et al., "Analytical System of Health Service needs among Healthy Elderly by using Internet", Proc. of Geontechnology Third Intl. Conf.(1999)
- [2] 濱保久,「コンピュータからのメッセージが作業者に与える効果」, 心理学研究, 61(1), 40-46 (1990)
- [3]山本吉伸, 松井孝雄, 開一夫, 梅田聡, 安西祐一郎,「計算システムとのインタラクション 楽しさを促進する要因に関する考察」, 認知科学, 1(1), 107-120 (1994)
- [4]山田寛,「顔面表情認識の心理学モデル」, 計測と制御, 33(12), 1063-1069 (1994)
- [5] 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」, 日本機械学会誌, 95(No.883), 503-507 (1992)
- [6]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, 8(3), 532-540 (1996)
- [7]高橋雅博, 山下利之, 酒井秀昭, 武田利治, 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掲載予定(2000)
- [8]目良和也,「事象の格フレーム表現からの情緒生起における強度計算手法」, 信学技報, TL99-32, pp.47-54(1999)
- [9]Clark Elliott,高砂美樹訳,「人間とコンピュータの間の双方向感情的コミュニケーションの構成要素 - 感情と人格に関する広範な基本モデルを用いて」, 認知科学, Vol.1, No.2, pp.16-30(1994)

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

