

# 話し言葉における感情を考慮した 知的インタラクションシステムの構築

## Emotion Oriented Intelligent Interaction System for Spoken Language

目良 和也

Kazuya Mera

広島市立大学

Hiroshima City University

黒澤 義明

Yoshiaki Kurosawa

広島市立大学

Hiroshima City University

市村 匠

Takumi Ichimura

広島市立大学

Hiroshima City University

**Abstract:** We propose the “emotion oriented intelligent interface for spoken language” to be able to access computers easily. We applied three methods about natural language dialogue and emotion, analyzing the user’s utterances, estimating and expressing the user’s emotions, and analyzing the user’s intention from his/her utterances. By using these three methods, the user would be able to communicate with the system naturally. We constructed an interface system based on the methods, and the interface system has been applied into the “web-based health service system for elderly people.”

### 1. はじめに

現在、高齢者の健康管理のための地域職域コミュニケーションネットワークが構築されている。しかし、高齢者人口の急増や金銭的成本などの問題により、それらのサービスは重度の要介護者を主に対象とすることが多く、軽度の要介護者や身体的には健康な高齢者については何のケアもされないことがある。とはいえ、精神面でのケアを考えた場合、日々のカウンセリングやコミュニケーションは非常に重要な役割を持つ。そこでこのようなコミュニケーション機能を持つインタラクションシステムが一時的にでも忙しい介護者の代わりに話し相手になることが出来れば、これらの被介護者に対する精神面でのケアを支援することができる。そこで、どのようなインタラクションシステムを望んでいるかについて高齢者に質問紙調査を行ったところ、人間と同じように対話することができ、表情を表出するものが好まれていることが分かった。

我々は現在、挨拶やあいづちなどのような簡単な会話が可能なインタフェースシステムについて研究を進めており、現在、高齢者健康診断システムへの適用を行っている。このシステムは QOL(Quality of Life)に関する 50 個の質問に対する回答をもとに、解析結果やカウンセリングコメントを返すというものである[1]。

本研究では、自然言語による対面的コミュニケーションを通じて健康診断を行い、同時にユーザの現在の感情を解析することで、ユーザを心身両面からケアするインタラクションシステムを提案する。このシステムでは、自然言語発話を入力とし、その内容からユーザの発話意図および生起感情を解析する。そして、解析結果に応じてシステムの反応を決定し、それに合った顔画像を出力する。

### 2. 感情指向型インタラクションシステム

#### 2.1. 処理の流れ

図 1 に本システムの全体構造を示す。まず、QOLに関する質問をユーザに尋ね、マイクを通じてユーザの返答発話音声を取得する。次に、この発話を解析し発話意図を抽出する。ここでの発話意図とは、各質問に対する返答 (Yes/No) である。抽出した発話意図はサーバに送られる。サーバでは、これらの返答に基づいたカウンセリングコメントを生成し、返信する。一方、インタフェースシステムでは、ユーザ発話の内容に対してユーザの視点からの感情を計算する。そして、この感情とサーバから送られてきたカウンセリングコメントをもとにインタフェースシステムの反応を決定し、表情画像を作成したのち、ユーザに提示する。

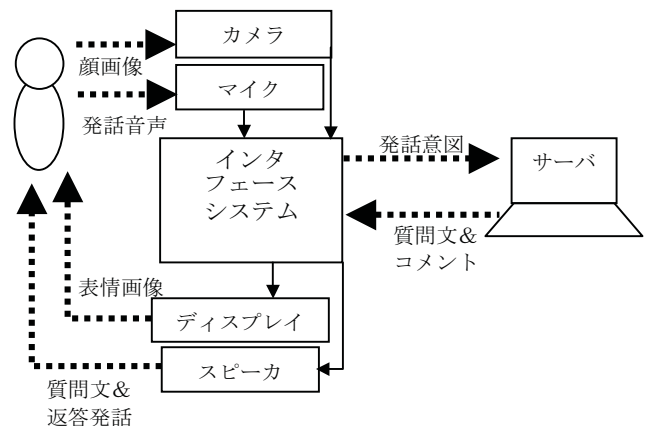


図 1 インタフェースシステムの構成

図 2 は本システムにおけるインタラクション部の処理の手順である。本システムでは自然な対話処理を目指すため、返答は「はい」、「いいえ」のような決まった形式だけではなく、自由な表現も処理できるようにする。また、言いよどみや言い間違いにも対応できるようにする。

そのため、黒澤が提案している正規表現を導入した構文解析手法[2]を用いる。そして、解析結果をもとに、肯定／否定の発話意図を計算する。この手法では、発話中に含まれる語や表現を抽出し、それぞれが持つ肯定／否定の度合をもとに発話全体の意図を計算する[3]。

また、構文解析の結果から、ユーザが生起しているであろう感情を計算する。我々は、発話が表示する事象から20種類の感情を並行して生起する情緒計算手法(Emotion Generating Calculations method: EGC)を提案している。この手法では、ユーザの嗜好情報に基づいて快／不快を計算し、その事象の状況や他者の心理状態などから、20種類の感情を生起する[4]。

生起した感情からは、ユーザに対してシステムが返すべき適切な反応が計算される。基本的には、算出したユーザの感情と同じものを表出するが、相手の状態に応じて「なぐさめる」、「元気づける」といった異なる感情を表出する。

20種類の感情やシステムの反応は、表情に関連する6種類の感情に分類され、それらの強度に基づいて1つの合成表情を生成する。表情画像のもととなる顔画像は、あらかじめシステム側で用意したものやユーザがカメラを通じて入力したものを学習させたニューラルネットワークを用いて生成する[5]。

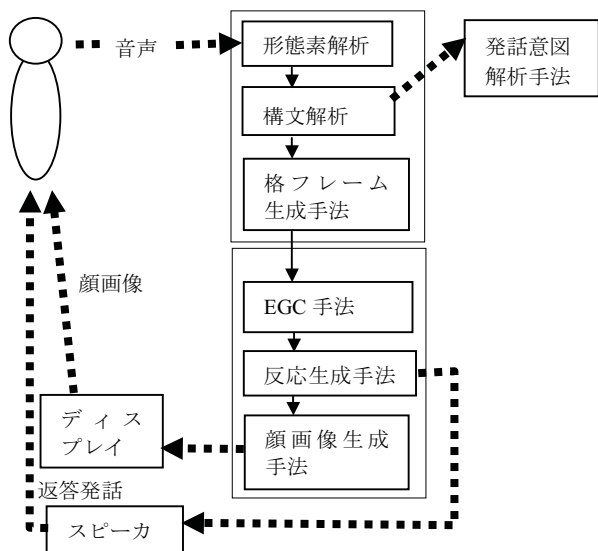


図2 インタラクションシステムの処理の流れ

## 2.2. 正規表現を用いた構文解析規則記述法

黒澤らは、表現の自由度が高く、言いよどみや言い直しを含む話し言葉を解析するため、正規表現を用いた構文解析規則記述法を提案している[2]。この手法では、音声認識装置の出力に対して形態素解析を行った結果を入力とし、それらを構文解析した結果を出力とする。そして、形態素の関係を表す規則の記述に正規表現を用

いることによって、ロバストな構文解析を実現している。図3に規則の例を示す。この規則は、「自立動詞の後に1つ以上の助動詞または終助詞の並びがあれば、それを1つのまとまりとみなす」という意味を持つ。

動詞・自立/(助動詞 助詞・終助詞)+	¥
¥	
0/1 ->{* , * , * , 0, 0, 0, 0, 0, 0}	

図3 構文解析規則

## 2.3. 格フレーム表現生成手法

構文解析処理によって得られた意味構造は、格フレーム表現に変換され、情緒計算部に渡される。格フレーム表現には、ガ格やヲ格といった格要素を持つ表層格フレーム表現と、主格や目的格といった格要素を持つ深層格フレーム表現があるが、本研究では情緒計算処理の都合上、深層格フレーム表現を作成する。しかし構文解析で得られる構造は表層表現であるため、これを深層表現に変換する処理が必要となる。

そこで我々は、表層表現の格要素を深層表現に変換するため、助詞の特性にもとづいて表1のような変換規則を作成した。この表では、ニ格は“目標”格と“抛り所”格の2つに変換されるが、これは主となる述語のタイプによって識別される。例えば「Xに行く」ならば“目標”、「Xに勝つ」ならば“抛り所”とみなされる。また、ハ格も“主体”と“対象”に変換されるが、この場合“主体”が存在しなければ“主体”、“主体”が存在するが“客体”が存在しなければ“客体”と判断する。

格フレーム表現には時間や否定に関する記述も含まれるが、これらは文中の文末表現や副詞表現から求められる。

表1 表層格と深層格の対応

表層格	深層格
ガ	主体
カラ	出発点または源
ニ, ヘ, マデ	目標
ニ	抛り所
ト	相互作用の相手
ヲ	対象
デ	道具
ハ, モ	主体, 客体

## 2.4. 情緒計算手法

続いて情緒計算手法の概要を説明する。まず、格フレーム表現された事象に対して事象タイプに応じた情緒計算式を適用することで快／不快を計算する。計算には事象中の格要素に対して

ユーザがあらかじめ持っている好感度を用いる。好感度は、自分の好きな事物に対して正、嫌いな事物に対して負の値を持つ。各事象タイプの高々3種類の必須要素の好感度を三次元空間(情緒空間)の直交軸にベクトルとして配置し、その合成ベクトルの存在する象限から快/不快を、長さから快/不快の強度を計算する。

次に、算出した快/不快を、20種類の情緒に分類する。この手法では、“他者の視点”、“予測と確認”、“行為の承認/非難”という条件に基づいて情緒の分類を行っている。20種類の情緒は、“幸福”群に「喜び、悲しみ」，“他者の運命”群に「嬉しい、ほくそえむ、憤慨、気の毒」，“将来的”群に「望み、恐れ」，“確認”群に「満足、安堵、恐れていた通り、失望」，“帰属”群に「誇り、賞賛、羞恥、叱責」，“幸福/帰属”群に「感謝、怒り、自己満足、自責の念」と分類される。

## 2.5. 反応生成手法

以前提案したインタフェースシステム[5]は、メールソフトやチャットシステムに応用されていたため、ユーザの感情をユーザ自身の顔画像によって表出していた。しかし今回目指すシステムではユーザの話し相手として反応を返すため、必要に応じてユーザと異なる反応をしなければならない。

本手法では基本的に2種類の反応を行う。1つめは従来と同様に、ユーザが生起した感情と同じ感情を表出するものである。これにより、システムがユーザに共感しているような感覚を与えることができる。もう1つはユーザの感情状態があまり良くないと判断したときに、その負の感情を緩和するような感情を表出するものである。この場合、必要があれば健康管理コンピュータネットワークを通じて、専門医やカウンセラーに通知する。現在は、感情状態を与える身体面への悪影響と、感情状態から推察される心理面での要注意状態から、表2の状態を対処すべき感情状態としている。これらに対する反応感情は、サーバや健康管理コンピュータネットワークから得られるカウンセリング知識やユーザの個人情報に基づいてサーバで決定され、本インタラクションシステムに送信される。

## 2.6. 表情画像生成手法

我々は、実際の顔画像を並列砂時計型ニューラルネットワークに学習させ、その中間層に情緒計算手法から得られた20種類の感情の値を与えることで、複数の感情が交じった1つの表情画像を生成する手法を提案している[5]。この手法では、まず“喜び、悲しみ、嫌悪、怒り、恐れ、驚き”の6種類の表情と通常の顔の画像を

表2 要注意感情状態

要注意感情状態	不安視される要因
長期間ネガティブな感情が続いている	うつ病
興奮性の感情(怒り、叱責など)が強く生起している	血圧の過度の上昇(高血圧のユーザ)
抑制性の感情(悲しみ、失望など)が強く生起している	胃潰瘍、自己免疫力の低下
強い感情同士の同時生起による葛藤が起きている(少なくとも片方がネガティブな感情)	胃潰瘍、悩み状態
全体的な感情生起が減少している	痴呆、自暴自棄、無関心、うつ病
感情生起が過度にアンバランスである	躁うつ状態

並列砂時計型ニューラルネットワークに学習させる。本手法で扱うニューラルネットワークでは、中間の第3層のニューロン数を2個にしており、この2つのニューロンの活性値を直交軸として平面に配置することで、各感情の位置関係を決定する(図4)。

次に、情緒計算手法あるいは反応生成手法によって得られた感情を6種類に分類し、それぞれの強度から平面における感情全体の重心を求める。そしてその場所に該当する活性値を第3層のニューロンに与えることで、最終的な顔画像を獲得する。

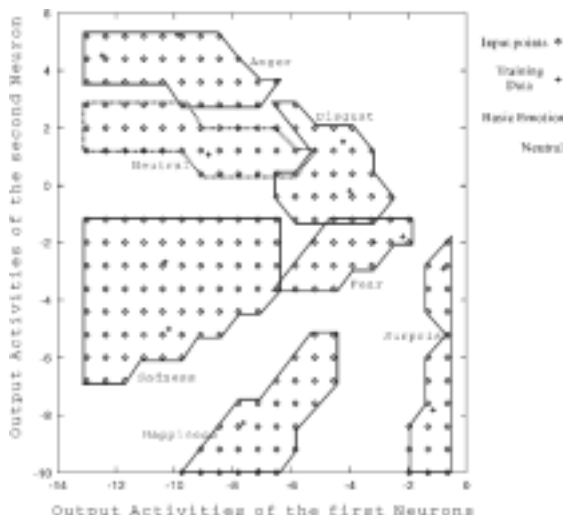


図4 第3層に構築された感情平面

## 2.7. 発話意図解析手法

我々は、真偽疑問文に対する返答発話を解析し、その発話全体が表す肯定及び否定の意図とその度合を計算する手法を提案している[3]。この手法では、まず発話の構文解析結果から肯定あるいは否定の意図を表す語(肯定否定要素)を抽出する。肯定否定要素には、「はい」、「いい

え」のような直接的なものから、「書ける」のように質問表現に依存するもの、「難しい」のように間接的な表現のものまで含まれる。それらにはそれぞれ肯定値と呼ばれる肯定／否定の度合を表す値が与えられている。さらに、「とても」や「あまり」のように、程度副詞が付いている場合には、その副詞ごとの変化倍率をもとに肯定値を変化させる。このようにして収集された肯定値をもとに、最終的に発話全体の肯定あるいは否定の度合を算出する。

### 3. 評価実験

図5に本インタラクシオンシステムの実行画面を示す。このシステムはLinux OS上で構築され、音声認識装置にJulius[6]、形態素解析装置にChaSen[7]を用いている。

本システムを実際に使用してもらったところ、80%のユーザから良い評価が得られた。しかし、「表情があいまいである」、「実画像よりアニメのような顔のほうが良い」、「やはり機械より人間のほうが良い」といった意見も聞かれた。



図5 インタラクシオンシステムの動作画面

### 4. まとめ

本論文では、ユーザの発話音声から発話意図と生起感情を解析し、その内容に応じた反応を返すインタラクシオンシステムを提案した。特に、構文解析結果の格フレーム表現への変換手法と、ユーザへの反応の生成手法について新たに提案を行い、それらを組み込んだインタラクシオンシステムを構築した。

要注意の心理状態の検知およびその対応についての評価実験も今後行っていく予定であるが、評価には専門的知識が必要となるため、注意が必要である。また、本システムを実際に現場で使用するためには、初期設備や通信費などについても考慮していく必要がある。

### 参考文献

- [1] K. Yoshida, T. Ichimura, H. Sugimori, T. Izuno, and H. Inada, "Analytical System of Health Service needs among Healthy Elderly by using Internet," *Proceedings of Gerontechnology Third International Conference*, 1999
- [2] 黒澤義明, 市村匠, 相澤輝昭: シナリオを対象とした構文解析手法, 第19回ファジィシステムシンポジウム講演論文集, pp.617-620, 2003
- [3] 吉江誠, 目良和也, 市村匠, 山下利之, 相澤輝昭, 吉田勝美: 真偽疑問文における返答発話の肯定／否定意図の解析とWebインタフェースへの応用, 日本ファジィ学会誌, Vol.14, No.4, pp.393-403, 2002
- [4] K. Mera, T. Ichimura, T. Yamashita, "Complicated Emotion Allocating Method based on Emotional Eliciting Condition Theory," *Journal of the Biomedical Fuzzy Systems and Human Sciences*, Vol.9, No.1, pp.1-10, 2003
- [5] 市村匠, 石田与志, 目良和也, 大枝真一, 杉原亮宏, 山下利之: 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, ヒューマンインタフェース学会誌, Vol.3, No.4, pp.225-238, 2001
- [6] A. Lee, T. Kawahara, and K. Shikano, "Julius --- an open source real-time large vocabulary recognition engine," *Proc. of the 7th European Conference on Speech Communication and Technology*, pp. 1691-1694, 2001
- [7] 松本祐治, 北内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸: 日本語形態素解析システム『茶筌』version 2.2.7 使用説明書, 奈良先端科学技術大学院大学, 2001

### 連絡先:

目良 和也  
 広島市立大学情報科学部知能情報システム工学科  
 〒731-3194 広島市安佐南区大塚東 3-4-1  
 Phone + Fax: 082-830-1677  
 E-mail: mera@its.Hiroshima-cu.ac.jp