
感性コミュニケーションロボットのための感情制御と感情生成

An Emotion Control and Generation Method for Emotional Communication Robots

1071005

研究代表者 名古屋工業大学大学院工学研究科 准教授 加藤 昇平

[研究の目的]

本研究では、感情を伴うコミュニケーションが可能な対個人サービスロボットを開発する。本研究の目標は、ロボットの高度な擬人化を実現する要素技術として、会話や画像認識・通信の基本処理の中に、感情認識、感情制御、感情表出、の3つが整合された感性情報処理を実現することである。これにより、本研究で開発するロボットは、生活支援や情報提供などユーザーに利便性を提供するのみならず、会話を通じたロボットとのコミュニケーションにおいてユーザーに心的作用をもたらす。

[研究の内容、成果]

近年、人とのより豊かなコミュニケーションを目指した会話ロボットの研究が盛んに行われている。我々は、先行研究において開発した感性会話ロボット ifbot (図1) を用いて以下の感情認識・制御に関する研究を行った。

1 ベイジアンネット混合モデルを用いた感性ロボットのための話者音声からの感情推定

1.1 背景と概要

ロボットと人間とのより豊かなコミュニケーションのためには、お互いの感情や情動を把握する必要がある。その実現には、人間が後天的に学習し獲得している「対話者感情を理解する知能」の推論モデルをロボットに持たせることが必要であると考えられる。そこで本研究では、感性会話ロボット Ifbot の対話者感情推定手法と

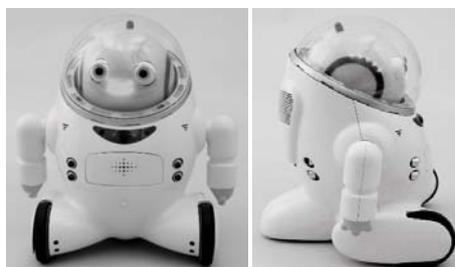


図1 感性会話ロボット ifbot

して発話音声を用いたベイジアンネットモデルを提案した。

1.2 Bayesian Network

ベイジアンネット (BN) は、複数の確率変数の間の定性的な依存関係を非循環有向グラフ (DAG) により表現し、個々の変数の間の定量的な関係を条件付確率で表した確率モデルである。

ベイジアンネットを用いて知識をモデル化することで、感情推定の知識記述量・計算量が大幅に削減される。また部分的な証拠からでも確率的に推論できる長所を持つ。このため本研究では、ロボットに搭載する感情推定のための知識モデルとして効率とロバスト性を得ることが可能なベイジアンネットを応用する。

1.3 感情推論器の学習

本研究では、対話者が発話した音声から対話者感情を推定するための感情推定知識をベイジアンネットワークとしてモデル化した。本節では、モデル構築の流れについて概説する。

音声資料：使用する音声資料は、感情表現がなされている必要がある。本研究では TV ドラマ、

映画、それに準ずるものより女優、俳優が感情を込めて発話したフレーズを抽出し「怒り」「嫌悪」「悲しみ」「恐怖」「驚き」「喜び」の6種類（以降、6感情）に分類した。それらの中から、聴取実験により感情が適切に表現されていると判断された音声資料をサンプルデータとする。

特徴量の抽出：音声は、3つの要素（韻律、音質、音韻）から成り立っている。この中で、韻律的特徴が人間の感情表現に最も関連することが過去の様々な研究から明らかになっている。そこで本研究では、音声資料から振幅構造を反映する「短時間パワー」(PW)、ピッチ構造を反映する「基本周波数」($F0$)及び時間構造を反映する「1モーラあたりの発話継続時間」(Tm)をそれぞれ計測する。 PW 及び $F0$ に関しては、平均、最大、最小、標準偏差を抽出した。このとき、短時間分析におけるフレーム長を23ms (250 samples)、フレーム周期を11msとし、窓関数としてHamming窓を使用した。以上9個の音声韻律特徴量と話者の性別情報(SE)をベイジアンネットワークの確率変数とする。

モデルの構造決定：本研究では、音声資料より抽出した各音声特徴量を適当な量子化数で量子化し学習データとする。学習データに含まれる目標属性（6感情）と属性（音声韻律特徴）との間の依存関係を表現するために属性間の結合とその強さ(CPT)を学習することでベイジアンネットワークの構造を決定する。学習方法として、本研究ではBIC (Bayesian Information Criterion) が最大となるモデルを求めこれを対話者の感情推定のための知識としてロボットに与える。BICを最大にするモデルの探索にはK2アルゴリズムを用いた。

1.4 感情推定実験

本論文で提案した手法の有効性を確認するため、感情推定実験を行った。まず、6感情のいずれかにラベル付けされた音声資料を1600事例用意し、話者の性別ラベルと音声の韻律特徴を抽出し属性を付与した後、任意に1400の学習事例と200のテスト事例を作成した。図2に学習事例から作成されたベイジアンネットワークモデルを示す。

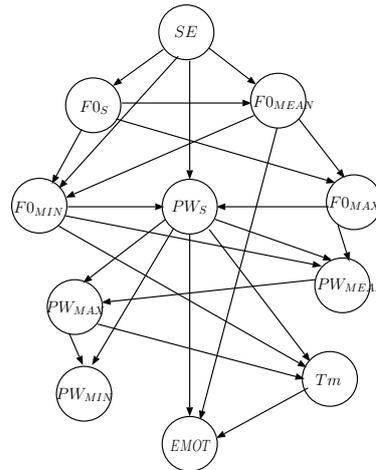


図2 生成されたベイジアンネットワークモデル

表1 感情推定実験結果

感情	正答率 (%)			
	提案 (BN) 手法			PCA
	全 (10) 証拠	6 証拠	4 証拠	
怒り	72.1	54.1	58.1	63.6
悲しみ	64.9	24.3	27.0	27.0
嫌悪	64.1	38.5	35.9	61.5
恐怖	55.6	37.0	37.0	48.1
驚き	59.3	48.1	44.4	44.4
喜び	63.0	40.7	44.4	22.2

評価実験はテスト事例から10属性すべての証拠が与えられた場合と、6証拠 ($SE, F0s, F0MAX, PWs, PWMEAN, Tm$)のみ、ならびに、4証拠 ($F0MAX, PWs, PWMEAN, Tm$)のみが与えられた場合の3つについて行った。表1(左)に感情推定の正答率を示す。全証拠を用いた実験では6感情すべてにおいて55%以上の正答率で認識された。特に「怒り」の感情について高い認識性能が確認された。一方で、6証拠および4証拠のみを用いた実験では「悲しみ」「嫌悪」の正答率が大きく低下したものの、6感情を無作為に回答した場合(16.7%)を下回ることとはなく、6感情すべてが認識されたといえる。このことから提案手法の頑健性が確認された。また、主成分分析(PCA)を用いた非線形判別分析との比較実験の結果(表1(右))からも「嫌悪」をのぞく5感情について大幅な正答率の改善が確認されており、提案手法の有効性が確認できた。

2 対話者好感度に基づいた感性ロボットの感情制御

2.1 背景と概要

まず、感情決定の主要素として快-不快に着目し、ロボットが会話中の相手（対話者）に対して抱く好感度を会話ロボットの快-不快と捉え、対話者好感度に基づく感情生成モデルを提案した。一方心理学において、感情の性質や強さの違いによる分類の一つに気分が挙げられており、感情決定において気分は重要な要素と考えられる。そこで、より豊かな感情生成を行うために、快-不快の計算モデルに加えて気分値と覚醒度を新たに考慮した感情生成モデルを提案し、会話ロボットにおけるこれらの計算方法を考案した。

2.2 感情生成モデル

本稿ではロボットが発話した質問に対して対話者 P が返答する会話形式を考え、会話内容から計算された対話者好感度に基づいてロボットの快-不快を決定し、その累積である気分値とその変化量の影響を受ける覚醒度を求め、これにより感情を制御する感情生成モデルを提案する。図3に本稿で提案する感性会話ロボットの感情生成モデルを示す。

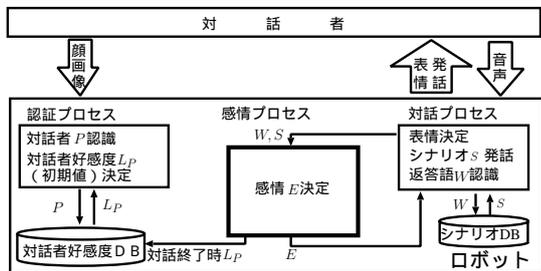


図3 感情生成モデル

認証プロセスでは対話者 P の認識や対話者好感度 L_P の初期値決定を行う。また対話プロセスでは1対話（本稿では質問と返答の一组を1対話と呼ぶ）ごとの感情 E を用いた表情決定と発話及び、返答語 W の音声認識を行う。そして感情プロセス（図4）では、まず、ロボットの発話文と対話者の返答語 W 、語の好感度 DB、質問文 DB を用いて対話者好感度 L_P を求める。次に、この L_P に基づき気分値 M と覚醒度 A を決定する。そして L_P 、 M 、 A の3要素から

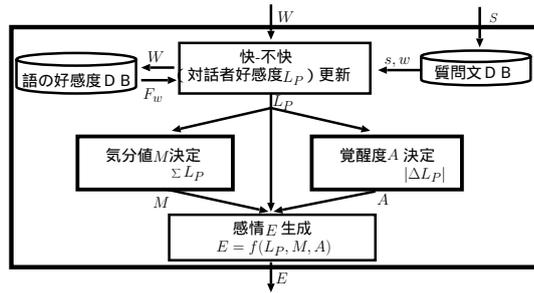


図4 感情プロセス

感情生成関数 $f(L_P, M, A)$ によって感情 E を決定する。

2.3 気分値 M の計算方法

気分とは比較的長時間にわたり持続的に生じるものである。快が連続すれば気分も良くなり、不快が連続すれば気分も悪くなる。そこで本稿では、気分値を過去から現在までの快-不快すなわち対話者好感度の蓄積と捉えた。ある1対話 n における気分値 M_n を以下の式で定義する。

$$M_n = \sum_{m=0}^{\kappa} Att(m)L_{P(n-m)} \quad (1)$$

ここで n は対話の順序を示すインデックスである。対話者好感度 L_P が気分値に与える影響は過去に遡るほど徐々に弱まると考え、 m 対話前の対話者好感度 $L_{P(n-m)}$ と減衰関数 $Att(m)$ の積の和を気分値とする。気分値は κ 回前までの対話者好感度によって決定される。 $Att(m)$ は m に対し単調減少し、 $0 \leq Att(m) \leq 1$ の値をとる。

2.4 覚醒度 A の計算方法

本稿では、生理的な一日の覚醒睡眠リズム（生理覚醒度）をベースとし、これに心理的要素（心理覚醒度）を加えたものを覚醒度と定義する。ifbotの年齢設定は5歳のため、生理覚醒度として幼児の覚醒睡眠リズムを参考に作成し、心理覚醒度として対話者好感度の変化による刺激と対話者との会話時間のストレスによる疲労を採用する。ある1対話 n における覚醒度 A_n を以下の式で定義する。

$$A_n = \alpha Stim_n - \beta Tired(t_n - ts_p) + \gamma ARhythm(t_n) \quad (2)$$

$$Stim_n = \frac{|L_{Pn} - L_{P(n-1)}|}{2} \quad (3)$$

ここで $Stim_n$ は、現在の対話者好感度 L_{P_n} と 1 対話前の対話者好感度 $L_{P(n-1)}$ の差分の絶対値によって与えられる心理的な刺激である。 t_n , ts_P はそれぞれ対話 n の行なわれた時刻, 対話者 P との会話開始時刻を示し, $Tired(t_n - ts_P)$ は対話者 P との会話時間 $t_n - ts_P$ によって生じるストレスによる疲れを示す単調増加関数である。 $ARhythm(t_n)$ は時刻 t_n における生理覚醒度を示す。 α, β, γ はそれぞれ刺激敏感度, 疲労係数, 生理覚醒係数と呼ばれ, $Stim_n, Tired(t_n - ts_P)$ 及び $ARhythm(t_n)$ がロボットの覚醒度にどの程度影響するかを示す係数である。

提案モデルをロボットに実装し種々の実験を行い有効性を確認した。ここでは、紙面の都合上割愛させていただく。詳しくは、本抄録末尾の発表論文を参照されたい。

[今後の研究の方向, 課題]

まず、音声・言語混合ベイジアンネットワークを用いて感性ロボットのための対話者感情の推定法を提案した。本手法により、音声情報から対話者の感情を推定することが可能となった。今後の課題としては、学習データを増加させることにより感情推定の性能を改善すること、および、対話者の顔表情やジェスチャを考慮した総合的な感情推論器を構築し、ロボットへの実装を行なう予定である。

つぎに、感性会話ロボットのための快-不快、気分値、覚醒度を用いた感情生成モデルを考案し、対話者好感度による感情制御の計算方法を提案した。今後は、気分の概念をより発展させた感情生成モデルを考案し、ロボットの性格づけや個性化の実現について取り組む計画である。

謝辞

本研究を進めるにあたり多大な支援をいただきました、立石科学技術振興財団に深く感謝致します。

[成果の発表, 論文等]

[学会誌論文]

1. 趙章植, 加藤昇平, 加納政芳, 伊藤英則: ベイジアンネットワークを用いた感性会話ロボットのた

めの対話者感情の推定法, 情報科学技術レターズ, Vol. 6, pp. 327-330, 2007.9.

2. 柴田寛, 加納政芳, 加藤昇平, 中村剛士, 伊藤英則: 感性ロボットの表情変化による性格付け, 情報科学技術レターズ, Vol. 6, pp. 323-326, 2007.9.
3. 竹内将吾, 酒井あゆみ, 加藤昇平, 伊藤英則: 対話者好感度に基づく感性会話ロボットの感情生成モデル, 日本ロボット学会誌, Vol. 25, No. 7, pp. 1125-1133, 2007.11.

[国際会議論文]

1. Jangsik Cho, Shohei Kato and Hidenori Itoh: Bayesian-Based Inference of Dialogist's Emotion for Sensitivity Robots, IEEE International Symposium on Robot & Human Interactive Communication (RO-MAN2007), pp. 792-797, 2007.8.
2. Minori Gotoh, Masayoshi Kanoh, Shohei Kato and Hidenori Itoh: A Neural-based Approach to Facial Expression Mapping between Human and Robot, Lecture Notes in Artificial Intelligence, Vol. 4964 (Knowledge-Based Intelligent Information and Engineering Systems), Springer-Verlag, pp. 194-201, 2007.9.
3. Yuki Matsui, Hiroshi Shibata, Masayoshi Kanoh, Shohei Kato and Hidenori Itoh: Interaction Effects in Facial Expressions of Emotional Space-using Kansei Robot "Ibot", International Conference on Kansei Engineering and Emotion Research, CD-ROM, 2007.10.
4. Jangsik Cho, Shohei Kato and Hidenori Itoh: A Biphase-Bayesian-based Method of Emotion Detection from Talking Voice Lecture Notes in Artificial Intelligence, Springer-Verlag, 2008.9. (to appear).

[学術講演会予稿]

1. 趙章植, 加藤昇平, 伊藤英則: 感性会話ロボットのためのベイジアンネットワークを用いた対話者感情推定法, 情報処理学会第 69 回全国大会, 分冊 4, pp. 235-236, 2007.3.
2. 松井裕紀, 柴田寛, 加納政芳, 加藤昇平, 中村剛士, 伊藤英則: 感情空間を用いた感性ロボットの表情表出効果, 第 17 回インテリジェント・システム・シンポジウム, pp. 447-450, 2007.8.
3. 松井裕紀, 柴田寛, 加納政芳, 加藤昇平, 中村剛士, 伊藤英則: 感性ロボットの感情空間を用いた表情変化によるインタラクション効果, 第 6 回情報科学技術フォーラム, pp. 497-488, 2007.9.
4. 加納政芳, 後藤みの理, 加藤昇平, 伊藤英則: 遠隔コミュニケーションのための感性ロボットと人の表情対応付け, 第 6 回情報科学技術フォーラム, pp. 489-492, 2007.9.