

人間とロボットの感情的インタラクション

The emotional interaction between humans and robots

田島敬士
Takashi Tajima

西田豊明
Toyoaki Nishida

東京大学大学院情報理工学系研究科電子情報学専攻
Department of Information and Communication Engineering,
Graduate School of Information Science and Technology, The University of Tokyo

The purpose of this study is smoothing communication between humans and computers by mutual adapting. This study aims at spontaneous behavior such as "Sowasowa". We explored a human and a robot recognize partner's state of mind by spontaneous behavior and behave to adapt the partner in communication. And we made an interactive model for spontaneous behavior and implemented to a robot. The model has analyzing and generating spontaneous behavior system. The analyzing system use Hidden Markov Model to analyze spontaneous behavior from motion capture input. The generating system output spontaneous behavior according to program from user's spontaneous behavior and robot's state. As a result, the interaction model can't use in real world because real world has various situation and high dimensions in sensor-motor. Future work is searching a new interactive model for mutual adapting in cleaning robot world by using simulator.

1. はじめに

1.1 本研究の目的

本研究の最終的な目的は、人と計算機の間のある目標を達成するためのコミュニケーションの円滑化である。コミュニケーションは通常、当事者達の持つ何らかの目的を達成するために行われる。当事者達はコミュニケーション相手の表情、動作等を観察し、相手の状態を類推し、自分の行動を相手に適応させることで目的達成に近づいていく。動作者が自分の状態を知らせようと意識的に行う動作だけではなく、無意識的に行う仕草からも動作者の心理状態、意図といったものが類推可能である。本研究では、まずこの仕草というものを人間とロボットとのコミュニケーションにも適用させることにした。

一般的に仕草というものにはスピーチと共に意識的に行われる例示動作というものもあるが、本研究では無意識的に行われ、心理状態を表現する仕草について取り扱う。本研究では、心理状態を表す仕草によるインタラクションを行うためのシステムを考案し、実際にロボットに実装した。

1.2 仕草を用いたインタラクションの意義

ロボットが仕草を用いたインタラクションを行うためには、ロボットがユーザの仕草を認識、理解することと、仕草を用いてロボット自身の心理状態を表現することが必要である。ロボットがユーザの仕草を認識できると、感情といったユーザの心理状態を認知できるため、それに応じた適切な行動が可能になる。また、人間とのコミュニケーションをロボットからの働きかけによって開始することができる。従来のロボットは、近寄る、話しかけるなど人間から働きかけてやらないとコミュニケーションを開始できないものや、人間がロボットの相手をしたくないときでも強制的にコミュニケーションを働きかけるものであった。しかし、仕草を感知することで、人間にとってコミュニケーションが必要な時に、それを開始することができる。例えば、イベント会場を案内するロ

ボットについて考える。会場にいるユーザの中には、会場の配置を熟知し、誰にも邪魔されず自分の好みのままに会場を回りたい者もいれば、道に迷い、誰かに会場を案内して欲しいと思っている者もいるだろう。もし道に迷っている者が案内ロボットを発見したとしても、それが案内ロボットだとは気づかないかもしれないし、ロボットにわざわざこちらから話しかけるのは億劫と感じるかもしれない。ロボットに仕草を認識する機能が備わっていれば、道に迷い「そわそわ」している人だけに、「案内しましょうか？」とロボットの方から声をかけてやることができる。

次に、ロボットが仕草によって自分の心理状態を表現できることの利点について述べる。本研究ではロボットがコミュニケーションを円滑に進めるために心理状態を持つことを仮定している。感情といった心理状態を持つことが計算機にとって有益だという考え方[Picard 1997 等]は現在認知されてきている。ロボットが仕草によって自分の心理状態を表現できることの利点の一つは、ロボットの感情表現が豊かになり、人間性が増し、人間にとって親しみやすくなることである。従来のロボットの動作は目的を効率的に遂行するための単純な動きが主であり、ユーザに無機質な印象を与えることが多かった。ロボットが仕草を用いることでより生命的な印象を与えることができる。もう一つの利点は、ロボットの心理状態を暗黙的に表示することにより、人間によるロボットの状態理解を助けることができることである。加えて人間は仕草によって類推される心理状態は言語によって表現される心理状態より、確実性が高いと判断することが多い。

具体的に例をあげて利点を説明する。目的を達成するためにユーザとロボットが協力して事を進めていたとき、ロボットが何か失敗をするとユーザは怒っていららするであろう。失敗を行ったことをロボットが検知できればよいのだが、ロボットが想定できないようなミスは現実世界には多々ある。そしてユーザがわざわざ失敗であると教示を行うことはユーザにとって面倒である。もし失敗を検知できていたとしても、そのことはユーザにはわからない。ロボットがユーザの仕草を認識できると、ユーザのいららを検知することで、自らの失敗を認知することができる。また、失敗を認知し、次からは失敗を犯さないように反省をしているという心理状態をしゅんと萎縮するという仕草で表すことができる。するとユーザはロボットが真に反省していると感じ、ユーザのいららも少しはおさまるかもしれない。

連絡先: 〒113-8656 東京都 文京区 本郷 7-3-1
東京大学大学院情報理工学系研究科電子情報学専攻
TEL 03-5841-6689 FAX 03-5841-6689
tajima@kc.t.u-tokyo.ac.jp

1.3 本研究の成果

本研究では、ロボットに仕草を扱う能力を実装した。そのために仕草認識機構と仕草出力機構、ロボットにおける仕草の動作データを作成した。そして、それらを用いて仕草によって人間とインタラクションを行うシステムを作成した。

2. 仕草によるインタラクションシステム

2.1 インタラクションシステムの概要

ロボットに実装した仕草を用いたインタラクションシステムの概要を図1に示す。システムは仕草認識機構とロボットにおける仕草の動作データ、ロボット自身の心理状態と入力された仕草と出力する仕草の対応プログラムを持つ。本来ならロボットの心理状態と出力される仕草は別物であると思われるが、簡略化のため同一のものとして実装を行った。

ユーザがある仕草を行うと、その動作情報がロボットの持つシステムに入力される。入力には今回モーションキャプチャを用いた。動作情報は隠れマルコフモデル(HMM)による仕草認識機構に入力され、具体的な仕草のシンボル、例えば図1の場合「そわそわ」に変換される。ユーザの仕草がわかると、システムを持つ目的とロボット自身の心理状態によって、次のステップの心理状態及び出力される仕草が決定される。図1の場合、ロボットは勉強を教えたいのにユーザは焦ってそわそわしてばかりいるため、ロボットは「いらいら」し、「いらいら」した仕草をユーザに示す。ユーザもロボットの仕草に対して心理状態を変える。例えば「そわそわ」するのをやめるかもしれない。システムの理想としては、再びユーザが仕草で応答することでコミュニケーションの循環が起こり、目的に向かって近づいていく。

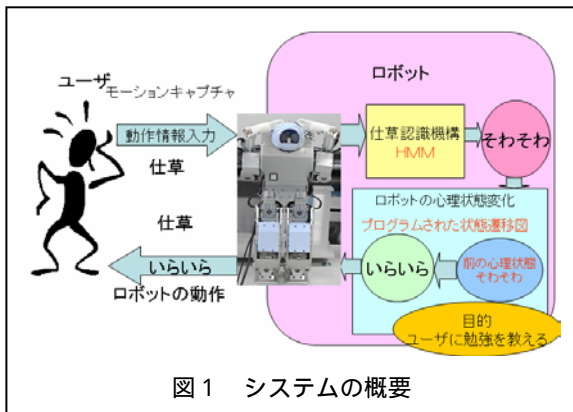


図1 システムの概要

2.2 仕草認識機構

(1) 仕草認識機構の概要

仕草認識機構は、筆者らが動態に関わる副詞を認識するために作成した認識機構[田島 2002]を簡略化して用いた。機構は128個の動態に関わる副詞を想定し、22個の副詞について実際に実装している。心理状態を表わす仕草と動態に関わる副詞には共通するものが多く、流用が可能である。

仕草認識機構について説明する。仕草認識機構はモーションキャプチャで入力された動作情報に対して、それを表す仕草シンボルを出力する。今回は情報を高い精度で得るためにモーションキャプチャを用いたが、将来的にはロボットの持つカメラで動作情報を取り入れることも考えている。

動作情報から仕草へのマッピングには一度特徴ベクトルに動作情報を変換した後 HMM を用いた。HMM は音声認識や文字分析、生物配列解析、そして動作認識[大和 1993]に用いられるマルコフモデルの一種である。状態遷移が外から見えず、現在どの状態にいるかを観測できないことがマルコフモデルとの違いである。各状態は確率的にシンボルを出力し、このシンボルのみが外から観測可能である。HMM を用いた学習及び認識とは、この出力シンボルのパターンを観測することで、モデルを推定することを基本としている。

仕草を HMM で判別する理由を説明する。仕草は本来なら判別に用いるべき特徴ベクトル、判別条件が仕草の種類により異なると考えられるが、莫大なセンサ情報と判別条件が必要になるため、システム作成は不可能である。よって仕草は動作状態の遷移によって表した。また、心理状態を表わす仕草は定義が曖昧であり、個々人により異なる。よって明確な定義をつけるのは難しく、学習により共通概念を探る必要がある。仕草は動作状態の遷移によって表されると考えるが、仕草の示す動作には揺らぎがあり、人間が仕草を認識する際にも、仕草に対応する状態遷移モデルのうち、観察対象の動作は現在どの状態にあるかをはっきり理解しているわけではない。以上のことから仕草を HMM により記述することにした。

(2) 仕草認識機構が外部から取り込む動作情報

腰、頭、左肩、右肩、左肘、右肘、左手、右手、左足、右足にモーションキャプチャのセンサを取り付け、0.25 秒間隔でセンサの3次元位置情報(X,Y,Z 情報)及びセンサ自身のもつ XYZ 軸の傾きの情報(Azimuth, Elevation, Roll)を取り込む。

(3) 特徴ベクトル抽出

入力情報から「頭、左肩、右肩、左肘、右肘、左手、右手、左足、右足の腰を原点としたときの相対的な位置ベクトル(X,Y,Z)情報」、「腰の速度ベクトル(r, ϕ, θ)情報」、「腰の位置情報(Z)」、「頭のオイラー角変化(Azimuth, Elevation, Roll)」の34次元の要素を持つ特徴ベクトルを抽出する。そしてベクトルの要素のとりうる値の幅が同程度となるようにパラメータ正規化を行う。学習データ量を少なくするために、腰センサからモーションキャプチャ本体への位置ベクトルの向きが常に一定方向になるように回転移動を行った。具体的には両肩を結んだ線分の垂線を体の向きと考え、その向きが常に一定方向になるように特徴ベクトルの各要素(x,y成分)を回転移動した。また、人間の体格の差に対処するために測定した身長によりパラメータに補正をかけた。

(4) シンボル化

特徴空間中の代表点コードワード $g_i \in \mathbb{R}^{34}$ の集合であるコードブック g を用いて、特徴ベクトルをユークリッド距離が一番近いコードワード $g_j (j=1,2,\dots,40)$ で表すことによりシンボル化は行われる。コードブック g は、コードワード g_j で表される特徴ベクトル集合の平均値を新たなコードワード g_j としてコードブックの更新を繰り返すことによって作成される。

(5) 仕草の学習

仕草の学習はそれぞれの仕草に対応する HMM を作成することで行う。外部から取り込んだセンサ情報から特徴ベクトル抽出し、シンボル化を行い、仕草学習シンボル列 V を得る。そ

して HMM 推定の基本である Baum-Welch アルゴリズムを用いて V を出力しやすいような HMM のパラメータを推定する。

図2の「そわそわ」を例として説明する。仕草一つに対し、10フレーム間のモーションキャプチャから得た学習データ列 L を 40 個用意する。この学習データ列 L をそれぞれ特徴ベクトル列に変換し、コードワード g に基づき 1~40 の値にシンボル化してシンボル列 V (11,11,13, ...,23[図2])にする。Baum-Welch アルゴリズムを用いて 40 個のシンボルデータ列 V を出力しやすいような HMM のパラメータを推定し、これを「そわそわ」の HMM とする。HMM の初期パラメータについて状態数 N は 4 で固定し、時間 T は 10 フレームとした。HMM 初期パラメータについてはランダム値とした。

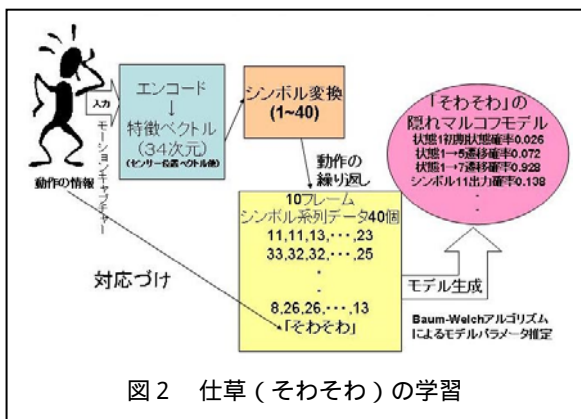


図2 仕草(そわそわ)の学習

(6) 仕草の認識

入力されたセンサデータ列を特徴ベクトル抽出、シンボル変換し、それぞれの仕草に対応する HMM がシンボル系列を出力する確率を前向きアルゴリズムによって計算し、出力確率が最も高い仕草を出力する。だが、出力確率がこれより低い HMM に対応する仕草の方が適切である可能性は多々ある。文脈情報などを考慮する機構にすると認識率は上がると思われる。

具体的に図3を利用して説明する。モーションキャプチャから入力されたデータから特徴ベクトル f_i を抽出し、コードブック g を用いてシンボル化する。10 フレームごとにデータを区切り、シンボル系列 $A (=11, 8, 9, 25, 37, 21, 21, 28, 28, 30)$ [図3]とする。シンボル系列 A の出力確率 $P(A|M)$ を全ての HMM について前向きアルゴリズムを使って計算する。そして出力確率が最も高い仕草を認識する。

学習データ量の少なさを補うため、本システムでは連続した 4 回の認識中、最も高い出力確率をその仕草の出力確率とする。よって 10 秒(40 フレーム)ごとに動作を認識することになる。

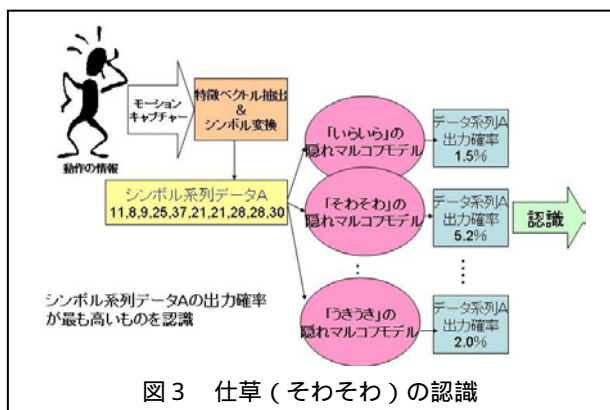


図3 仕草(そわそわ)の認識

2.3 仕草の出力機構

仕草によるインタラクションを行うためには、入力された仕草に対して、システムの持つ心理状態や目的に応じて適切な仕草を出力する必要がある。今回は、システム簡略化のため、システムの心理状態とシステムの出力する仕草を同じものとし、心理状態を入力された仕草によって状態遷移させる。遷移図については、システム制作者がシステムの持つ目的を考慮しながら予めプログラムしておく。

2.4 仕草の動作データ生成

ロボット自身の仕草動作データ生成については、幾つかの手法が考えられる。初めは「右手を上げる」「右肘を90度曲げる」等個々のモータへの指令の組み合わせにより仕草を実現しようとしたが、動作が機械的になるという欠点があった。よって、ロボットを人間が手動で操作しながら、ロボットの持つセンサによりモータ情報を計測することで仕草の動作データを生成した。

2.5 システムの実装

2.2、2.3、2.4 節で述べた仕草認識、出力機構や動作データを既存のヒューノイドロボット HOAP-1[富士通 2001]に実装した。HOAP-01 は2足歩行可能なロボットであり、関節数は20である。

(1) システムの持つ目的

考えられる全ての状況を想定するとシステムは莫大なものになる。今回は、人間、ロボット共に「うきうきした楽しい気分になる」ということを目的として設定し、目的に関わる仕草のうち幾つかをシステムに実装した。

(2) ロボットが認識する人間の仕草

ロボットには表1のような人間の仕草を認識できるように学習させた。システム簡略化のため、ほとんど学習させていないが、認識システム自体は20個程度の仕草は認識可能である。心理状態を表わす仕草だけでは単純なインタラクションしかできないため、心理状態を表わす仕草の定義からは外れたものも含んでいる。

表1 ロボットが認識する人間の仕草

A	まあまあとロボットの肩を抱く
B	ぼんぼんとロボットをたたく
C	うきうきする
D	そわそわする
E	いらいらしながら怒る
F	何もしない

(3) ロボットが行う仕草

今回は「うきうきした気分になる」という目的を達成するというタスクに合わせてロボットに、表2のような動作を学習させた。

表2 ロボットが行う仕草

① うじうじする(しゃがみ込んで指で「の」の字を描く)
② しゃがみ込んだ状態でユーザを見上げる
③ 普通に立つ
④ うきうきする
⑤ うとうとする
⑥ ふあへとあくびのような仕草をして起きる
⑦ 怒りにわなわなと震える
⑧ うがーと両手を上げて怒る
⑨ そわそわする
⑩ 焦っているようにする
⑪ もうダメだーと頭を抱え込む
⑫ へこへこ土下座して謝る

(4) 仕草の状態遷移図

ロボットは人間の仕草を認識すると現在の仕草に応じて次の仕草を出力する。表3にロボットにおける仕草の出力例の一部を記す。

表3 ロボットにおける仕草の出力例の一部

現在の仕草	認識した仕草	出力する仕草
①うじうじ	Bぼんぼんと叩く	②ユーザを見上げる
②ユーザを見上げる	Aまあまあと肩を抱く	③立ちあがる
③立っている	Cうきうき	④うきうき

2.6 実装したシステムの考察

仕草によるインタラクションシステムをロボットに実装してみたが、目的に応じた心理状態遷移を設計者自身がプログラムするという方法では、扱う仕草を増やすなどのシステム拡張が困難であり、実世界には適用できない。よって、実世界に適用できるようにシステムを改良する必要がある。しかし実世界でのロボット問題を扱おうとするとセンサーモータ空間が非常に大きく、次元数も高いため扱いにくい。まずはセンサーモータ空間の次元を下げて、新たなインタラクションモデルを探っていく必要がある。

2.7 システムの相互適応モデルとしての考察

作成したシステムによって、人間とロボットが相手の状態に適応しながら目的に向かって近づいていくという相互適応が行われているか考察を行う。

本研究で作成したシステムは図4の左半分である。本システムでは、人間とロボットは自発的に行われる仕草を用いた認識-応答のループにより、目的に近づき、到達することを目指している。しかし、このモデルは目的に向かう状態遷移の道筋のバリエーションが少なく、結局制作者自身が予めプログラムした道筋しか存在しないという問題がある。そして目的に向かって人間、ロボットが自然に近づいていくことを表現するのは困難である。

対して、図4右側のモデルを考察する。ユーザの出力などの状態を (x_t, y_t, \dots) 、ロボットの状態を (X_t, Y_t, \dots) とすると、ユーザ、ロボット相互をあわせた状態はそれぞれの次元の数を加算した多次元空間の中で、 $(x_t, y_t, \dots, X_t, Y_t, \dots)$ というベクトルで表わされる。そして、ユーザ、ロボットの目的とする状態はその多次元空間の中で、ある領域、例えば図4ではサイクルとして表現される。相

互適応とは $(x_t, y_t, \dots, X_t, Y_t, \dots)$ というベクトルが時間の経過を経ながら段々目的とする状態に引き込まれていくことであるかもしれない。このように、ベクトルを目的とする状態まで引き込んでいくようなベクトル場をデザインすることである種の相互適応モデルが作成できると考えられる。力学系においてベクトルがある時間を経て引き込まれる定常状態はアトラクタと呼ばれ、このアトラクタをシステムの持つ目的とする。今後はこのようなモデルを用いて研究を行う。アトラクタを用いたロボット研究にはアトラクタを用いたロボットの行動生成の研究[岡田 2002]があるが、本研究ではアトラクタを人間にも適用し、相互適応モデルとして取り扱う予定である。

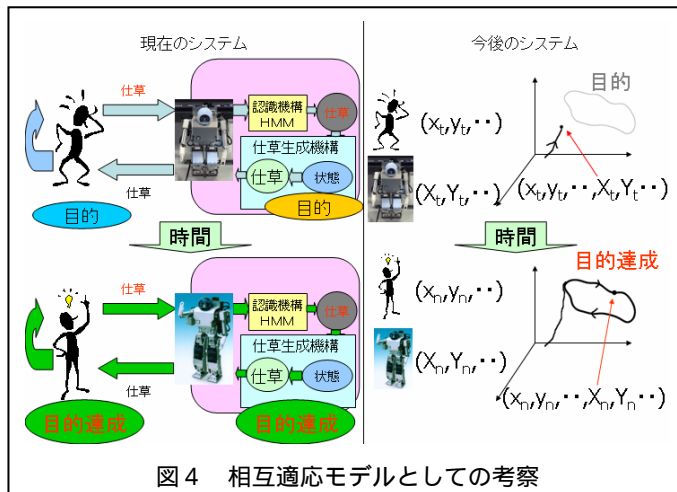


図4 相互適応モデルとしての考察

3. 本研究の成果及び今後の研究

本研究では、心理状態を表す仕草によるインタラクションシステムを作成し、ロボットに実装した。結果、センサーモータ空間の次元が大きく、少ない状態数ならまだしも、実世界を想定したアプリケーションを作るためには新たなモデルの作成が必要であった。今後の研究ではシミュレータを用いて新たな相互適応モデルを探っていく。具体的には 2.7 節で述べたアトラクタを用いたモデルを考えている。また、状態空間を狭めるために、ロボットが実現すべき目的を明確に定める必要がある。これについては大掃除、引っ越しなどを想定し、床掃除と家具運搬に目標を定め、掃除ロボットと家具を運搬する人間の相互適応について調査していこうと考えている。

参考文献

[Picard 1997] Rosalind W.Picard: Affective Computing, The MIT Press, 1997.
 [田島 2002] 田島敬士, 黒橋禎夫, 西田豊明: 動態に関わる副詞のオントロジー, 人工知能学会全国大会(第16回)論文集, 2002.
 [大和 1993] 大和淳司, 大谷淳, 石井健一郎: 隠れマルコフモデルを用いた動画画像からの人物の行動認識, 電子情報通信学会論文誌 D-II Vol.J76-D-II No.12 pp.2556-2563, 1993.
 [岡田 2002] Polynomial Design of the Nonlinear Dynamics for the Brain-Like Information Processing of Whole Body Motion, Proc. of IEEE International Conference on Robotics and Automation(ICRA2002), Vol.2, pp.1410-1415, Washington D.C., U.S.A., 2002.