

感情動因学習モデルを用いたペットロボット

Pet Robot Using Emotion Triggered Learning Model

鈴木 薫
SUZUKI Kaoru金澤 博史
KANAZAWA Hiroshi

顔認識、音声認識、ジェスチャ認識などのヒューマンインタフェース技術を用いて、利用者と感情や欲求のレベルで対話するペットロボット“コモコモ”を開発した。このロボットは、“感情動因学習モデル”に基づく感情機能を持ち、人物識別能力と対人関係能力を自動的に学習発展させられる。学習のトリガーは利用者の動きかけに応じてロボットに発生する強い情緒であり、そのときの情緒と利用者の顔パターンとが関連付けられて記憶され、その利用者に対する好き嫌いを決めるのに使われる。この結果、好感情を与える利用者にロボットは懐(なつ)くようになり、ペットロボットが特定の人物と親密になるメカニズムが実現される。

We have developed a pet robot called ComoComo which converses with a user on feelings and desires by using human interface technologies such as face, speech, and gesture recognition. This robot has a function of emotion based on an emotion triggered learning model, and can automatically evolve its person discernment capability and interpersonal skills.

The trigger of learning is strong feeling generated in the robot according to the approach of the user. The feeling and the user's facial images are associated, memorized, and used to determine whether the user is good or bad for the robot. Consequently, the robot can detect and attach to a user who gives a good feeling, realizing a mechanism that makes a pet robot become on intimate terms with a specific person.

1 まえがき

近年、ロボットと人間を取り巻く状況は新しい段階に踏み込んだ感がある。それは、人間と活動空間を共有するロボットの登場であり、“ロボットと暮らす”時代が本格的に始まるようとしている。

ペットロボットも、人と暮らすことを前提にしたロボットの一つである。その主な用途は、作業の代行や補助という実用的なものではなく、むしろ利用者に活力や癒(いや)しを与えるというメンタルなものである。もちろん、これに実用的な機能を加えることも可能であり、そのときのペットロボットの使い方は、愛犬にお使いを頼むのに似てくるであろう。しかし、十分な技術を投入して作られたペットロボットは、生きた犬よりもずっと意思疎通しやすく、他の機器と連動するなどの高度な役だち方ができるはずである。機械の利点を生かしつつご主人さまをメンタルな面までサポートする、それがわれわれの目指すペットロボットである。そして、このようなペットロボットは極めてパーソナルな存在であるため、特定の人物との結び付きを強化できるように作られていることが重要である。

2 ペットロボット

2.1 従来のロボット観

これまで、ロボットの用途はパワーアシストや組立加工作

業の自動化に重点が置かれてきた。そのニーズは、今後もさほど変わるものではないだろう。その一方で、ソフトロボティクス¹⁾のように、人に危害を及ぼさない、人に怖がられないロボットの研究も進められ、ロボットの作業領域が工場などの整備された環境から、家庭や公共空間などの、より一般的な環境に拡大されようとしている。昨今話題の人型二足歩行ロボット²⁾も、人間のために用意された施設を共用でき、親しみやすい形態を持つことからこの流れに沿っている。

ロボットが家庭や公共空間で利用されるシーンは様々に想定できる。例えば、施設内を巡回する清掃ロボットや警備ロボット、家庭に置かれるヘルパーロボット、ネットワーク家電のインタフェースロボット、施設内の案内誘導ロボットなどである。これらは、人間の活動空間で作業するロボットであり、そのいくつかは人間と直接対話して指示を受け、場合によっては個人を識別する必要にも迫られる。したがって、音声対話や個人識別などのヒューマンインタフェース技術が、ロボットのための要素技術として重要度を増す。そして、これらのロボットは、原則的に利用者の希望どおりに動作することが期待されている。

2.2 ペットロボットの特殊性

一方、ペットロボットは、人間の心に作用して利用者に活力や癒しを与えるロボットである。その手本は生き物のペットであり、指示どおりに動く単なる道具ではなく、独立した存在として利用者 と接するものでなければ意味がない。

そのため、ペットロボットは生き物を模した独自の欲求と感情を持ち、その発露である自律的な行動や感情表現を通して人間と対話できるようにデザインされる。行動の動機となる欲求を持つということは、ロボットが独自にその行動を決定するということであり、必ずしも利用者の思わくどおりにならないところが、他のロボットと決定的に異なっている。この意のままにならないロボットから利用者が得られるものは、ペット飼育によってもたらされる効果³⁾とほぼ同じであり、次に示すようなものである。

- (1) なでる気持ち良さなどの触覚的喜びを得る。
- (2) 相手を意識し、意思を通じ合える喜びを得る。
- (3) 困難を乗り越えて育てる喜びを得る。
- (4) 家族の一員を得て、孤独感が癒される。
- (5) 表情や会話が豊かになる。
- (6) 感覚や作業の代行をしてもらえる。

生きたペットの役割を人工物で置き換えることには賛否両論があろう。われわれも、ペットロボットが生き物に成り代われるとは考えていない。しかしながら、ペットロボットはペット飼育のリスク(人畜共通感染症、ペットロス:死別による精神的ショック)とデメリット(飼育疲れ、鳴き声・におい、留守にできない)を完全に排除しつつ、上記効果を程度の差こそあれ達成可能である。特に、ペットを導入したくてもできない人達にとって有望な選択肢となる。実際、病院などで利用できる小型のペットロボットは、ロボットセラピーという目的で試されつつある。

ところで、結果を最大効率で提供することが目的の実用ロボットと異なり、ペットロボットでは利用者との交流するプロセス自体に価値がある。そして、交流の結果得られる報酬が高いほど、そのプロセス自体がおもしろくなる。ペットロボットでは、ロボットが利用者に懐くことが報酬となる。このとき、懐かせるための特別な操作が要るようではプロセスが台なしになるので、実際のペットに接するような自然なやり取りの中で、良い関係性を発揮した相手にだけロボットが懐くというのが好ましい。そして、懐かれた人物は、ロボットの自発的な協力を得て様々な便利さを引き出すことができる。これがペットロボットを使うだいたいご味であり、懐かれることに成功した人物は、そのロボットに特別な愛着を持つようになる。これは上記の効果(2)、(3)、(6)の達成である。

3 感情動因学習モデル

提案する感情動因学習モデルは、自然なやり取りの中でロボットが利用者に懐いていくことを可能にする。その原理を以下に述べる(図1)。

3.1 欲求とその充足行動

生き物と同様、ペットロボットの行動も、欲求を充足するために行われるというのが基本原理である。欲求には大き

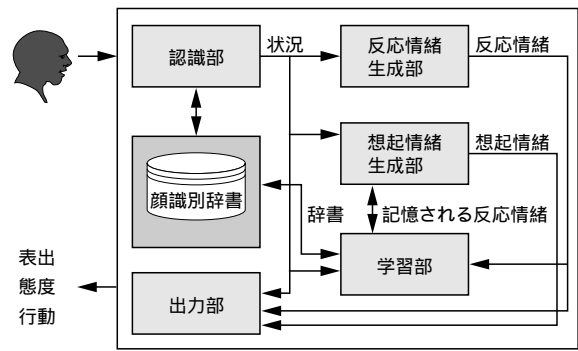


図1. 感情動因学習モデルの原理 認識結果に応じて情緒(反応情緒と想起情緒の和)が生成される。強い反応情緒が発生すると、その原因となった人物がその反応情緒とともに記憶される。再び同じ人物に会うと、記憶された反応情緒が想起されて、この人物への好悪を与える。

Principle of emotion triggered learning model

く生存欲求と受容欲求があり、生存欲求の充足は危機的状況に対する自衛行動(回避など)で、受容欲求の充足は人間とのかわりを欲する愛着行動(接近など)で達成される。特に、ペットのようにふるまうロボットでは、相手が好きなら愛着行動、嫌いなら自衛行動をすなおに表すことが許される。このとき、相手に対する好き嫌いの判断のよりどころになるのが感情機能である。

3.2 感情機能の学習性

感情には、喜びや怒りなどの一時的で急激な感情(情緒)と、好き嫌いなどの持続的な感情(感情傾向)とがある⁴⁾。感情動因学習モデルでは、情緒は反応情緒生成部による反応情緒と、想起情緒生成部による想起情緒の和で計算され、感情傾向は想起情緒で与えられる。ペットロボットがなでられると快感情を生じ、たたかれると不快感情を生じるとき、この情緒を生成しているのは反応情緒生成部である。一方、人物に対してロボットが抱く好き嫌いは、想起情緒生成部で決定される。このとき、ロボットが人物に懐くプロセスは次のようになる。

例えば、ロボットが人物Aになでられたとき、反応情緒が強い快状態に至ったとする。すると、学習部がこの反応情緒を人物Aに対応付けて記憶する。そして、次に人物Aが現れたとき、想起情緒生成部がこの記憶された反応情緒を想起情緒として生成する。人物Aの登場に伴って発生するこの想起情緒は、以前なでてもらったときに覚えた快感情である。ロボットの情緒が反応情緒と想起情緒の和であるとすると、特に刺激がないかぎりロボットはこの想起情緒に支配されて、人物Aのいる状況でなでられる前から快感情を覚える。これは人物Aに対してロボットが懐いた状態と言える。この例からわかるように、反応情緒は情緒における生得的・固定的成分であり、想起情緒は経験的・可变的成分である。

なお、愛着行動は人物に対する快い想起情緒によって、自衛行動は不快な想起情緒によって各々出力部を通じて発現する。また、出力部からは、行動のほかに表出(喜びや悲しみの表情など)と態度(指示に対する同意/非同意など)も出力される。

3.3 認識系の学習性

人物に対する想起情緒は、所定閾(しきい)値を超える反応情緒の発生をトリガーとして感情動因的に学習される。同時に、例えば、顔画像パターンのようなその人物を識別するための特徴も学習され、ロボットは強い反応情緒を発生させた相手の顔画像パターンを識別できるようになる。学習された顔画像パターンには、そのときトリガーとなった反応情緒値が関連付けられる。想起情緒生成部は、学習された顔画像パターンで識別できる人物に関連付けられたこの反応情緒を、想起情緒として出力する。この結果、ロボットは特定人物を識別する能力と、その人物への好悪感情(想起情緒)を運用中に無手順で学習することができ、ペットロボットが特定の人物に懐くメカニズムが実現される。

4 ぬいぐるみ型ロボット“コモコモ”

4.1 コモコモの概要

以上の議論を踏まえてわれわれが試作したペットロボット“コモコモ”(図2)は、本体(高さ20cm)と台座(高さ15cm)に各種センサ類を装備した9自由度のぬいぐるみ型ロボットである。コモコモは焦電センサ、感圧センサ、マイクロホン、カラーCCD(電荷結合素子)カメラ、モーションプロセッサ™を備えている(図3)。コモコモは、これらのセンサ情報を外部のホストコンピュータで処理して、利用者の接近、利用者によるなでやたたきの接触刺激、利用者の音声語彙(ごい)、顔の有無と人物の別、ハンドジェスチャーを認識しつつ、かわいらしい動作と音声で応答する。



図2. コモコモの外観 ハムスターをデフォルメしたフォルムのぬいぐるみ型ロボットである。
Appearance of ComoComo

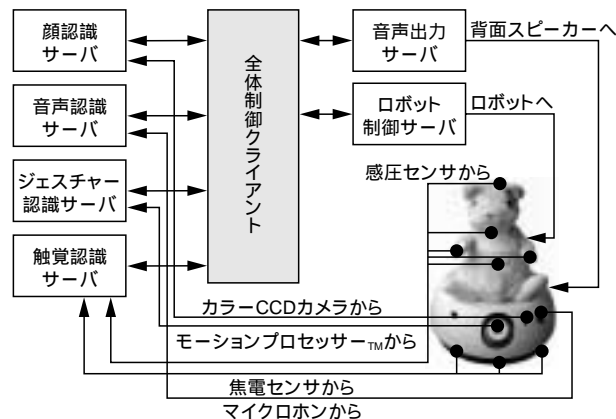


図3. コモコモのソフトウェア構成 四つの認識サーバ、二つの出力サーバ、及びペットロボットとして機能するための全体制御クライアントから成る。
Software configuration of ComoComo

4.2 コモコモの動作

- (1) 反応情緒を操作する報酬/罰系刺激 コモコモは利用者の顔を検出すると、甘え声を出しながらおねだりする(愛着行動)。このとき、名前を呼んでもらえたり、なでてもらえたりすると、反応情緒が快状態に傾けられ、コモコモは喜びを動作と声で表出する。呼びかけや褒めことばやなで刺激は報酬系刺激であり、入力ごとに反応情緒を所定値ずつ快状態に偏らせる。一方、しかりことばやたたき刺激は罰系刺激であり、入力ごとに反応情緒を所定値ずつ不快状態に偏らせる。また、報酬/罰系の刺激がなければ、反応情緒は中間状態に収束する。
- (2) 想起情緒の生成 既知の利用者の顔は識別されて想起情緒を発生させる。顔が識別されるごとに、その顔と結び付けられて記憶されている反応情緒値が想起情緒に代入される。識別可能な顔がなければ想起情緒も中間状態に収束する。顔は所定閾値以上類似している顔を一つのクラスとして記憶されており、各クラスには学習時の反応情緒値が関連付けられている。顔を識別した結果、所定閾値以上で最大類似度を獲得したクラスの反応情緒値が想起情緒に設定されることになる。なお、コモコモの情緒は快/不快の別を表す一次元のスカラー値を使っている。その推移のグラフを図4に示す。
- (3) 態度表明 情緒(反応+想起)が所定閾値を超えて不快状態にないとき、コモコモは利用者による遊びへの誘いかけに同意する(態度表明)。遊びは利用者によるジェスチャーと発声を認識して行われる物まねであり、モーションプロセッサ™で得られる近距離物体の配置と動きから、いくつかのハンドジェスチャーを識別してコモコモがまねをする。なお、遊びはより実用的なタス

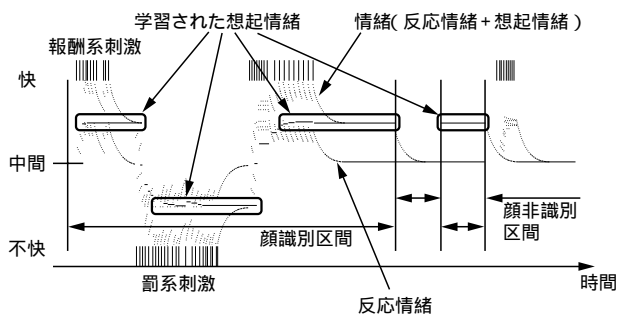


図4. 反応情緒と想起情緒の推移グラフ 刺激を与えることにより, 反応情緒が想起情緒に学習される。顔が識別できている区間だけ想起情緒が発生する。

Transition graph of reaction feeling and remembrance feeling

くに置き換えてもよい。一方, 情緒が所定閾値を超えて不快状態にあるときは, コモコモは利用者からの誘いかけを拒否してしまう。

(4) 顔パターンと想起情緒の学習 反応情緒が所定閾値以上の快状態あるいは不快状態に達し, かつ利用者の顔が検出されていれば, その顔を識別するための辞書が生成される。このとき, その顔がまだ識別可能でないなら, この顔のクラスを新規に生成して新しい辞書を構築し, 現在の反応情緒をこれに対応付けて記憶する。また, 顔が既知のクラスとして識別可能なら, このクラスの辞書を更新するとともに, このクラスに結び付けられている反応情緒と現在の反応情緒の加重和(実験では1:1)を記憶する。この結果, 既に嫌われている人物でも, その後の接し方で好かれることができるようになる。なお, 識別辞書の生成と更新は顔認識サーバで行われる。このサーバは, 顔向き, 表情変化などにロバスト(頑健)な相互部分空間法⁵⁾で顔を識別する。

4.3 耐雑音性のある音声認識

ところで, ペットロボットが家庭やモバイル環境で使われる可能性は非常に高い。そのため, 人間とロボットの意思疎通のかぎとなる音声認識は, 環境雑音の混入や発話様式の違いに対処し, 不特定利用者が自由に利用できるものである必要がある。そこで, 雑音キャンセル処理と雑音適応処理を組み合わせ, 耐雑音性の向上した音声認識方式を開発して実装した(図5)。

この方式は, 最初に音声から雑音成分を除去するために, 定常的な背景雑音の除去に効果のある“スペクトルサブトラクション法(Spectral Subtraction法:SS法)”を実行する(雑音抑圧処理部)。しかし, SS法は非定常雑音に対しては十分な抑圧ができないため, 非定常雑音及びSS法によりキャンセルできなかった定常雑音の残差成分に対しては, 入力音声を音声特徴ベクトルに変換する特徴抽出処理と, 実際に音声特徴ベクトルを認識する処理の二つの段階で環境適応辞書を使って対処する。特徴抽出部では, 雑音による時

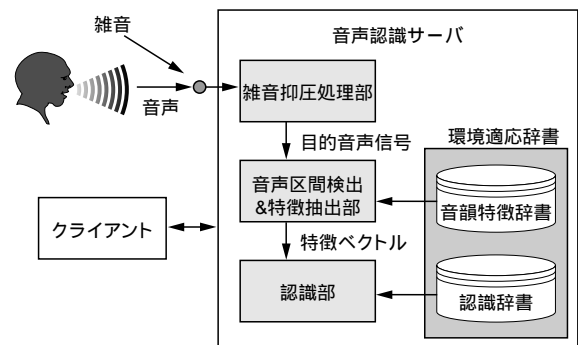


図5. 音声認識サーバのブロック構成 雑音抑圧処理と環境適応辞書で雑音の混入に対処している。

Configuration of speech-recognition server

間・周波数成分の様々な変動を表現したセグメント単位のマッチングにより, 雑音に影響されない安定した特徴ベクトルを求める。更に, 認識部で用いる辞書を雑音重畳音声から生成し, 特徴抽出部で吸収しきれない変形に対処する(雑音免疫法)。この結果, 高い耐雑音性を実現することができる。

5 あとがき

ここでは, ペットロボットの効果と使われ方について考察し, ロボットが自然なやりとりのなかで利用者に懐くための, 感情動因学習モデルを提案した。このモデルは, ロボットが自己の体験を通じて, 物の意味とそれを見分ける能力とを自律的に学習する基礎を与えるものである。また, これを支える技術として, パターン変動に強い顔認識方式と音声認識方式を実装した。

文献

- (1) [特集] ソフトロボティクス. 日本ロボット学会誌. 17, 6, 1999, p.1 - 55.
- (2) 日経メカニカル・日経デザイン共同編集. RoBolution. 東京, 日経BP社, 2001, 127p.
- (3) 林 良博. 検証アニマルセラピー. 東京, 講談社, 1999, 173p.
- (4) 斎藤 勇. 感情と人間関係の心理. 東京, 川島書店, 1986, 288p.
- (5) 福井和広, ほか. 制約相互部分空間法を用いた環境変動にロバストな顔画像認識 - 照明変動を抑える制約部分空間の学習. 電子情報通信学会論文誌. J82-D-II, 4, 1999, p.613 - 620.



鈴木 薫 SUZUKI Kaoru
研究開発センター マルチメディアラボラトリー研究主務。
コンピュータグラフィックス, 文字・画像認識, ヒューマンインタフェース, ロボットの研究・開発に従事。情報処理学会, システム制御情報学会会員。
Multimedia Lab.



金澤 博史 KANAZAWA Hiroshi
研究開発センター マルチメディアラボラトリー研究主務。
音声インタフェース, Bluetooth™ 応用システムの研究・開発に従事。電子情報通信学会, 日本音響学会会員。
Multimedia Lab.