

人とのインタラクションに基づくロボットの親和性向上に関する研究

— 人とロボットが同居するためのロボット要素技術に関する研究 —

稲葉 昭夫 張 勤 小川 行宏 光井 輝彰 田中泰斗 山本 和彦*

Development of Human Interface based on Interaction between Human beings and Robots

Development of Basic Technologies for Coexistence of Human Being and Robots

Akio INABA Qin ZHANG Yukihiro OGAWA Teruaki MITSUI
Taito TANAKA Kazuhiko YAMAMOTO*

あらまし わが国において世界的に例を見ない高齢化が進行する中で、労働人口の減少、障害者や寝たきり老人の増加、それに伴う介護者等の不足などから、福祉をはじめとする多くの分野において、各種の支援にロボットを利用することが強く期待されている。このような分野に使用するロボットに求められる基本コンセプトは、産業用ロボットとは異なり、「ロボットが人に合わせる」ということである。本研究では、これを実現する要素技術の1つとして、ロボットが人に対してリアクションを返すことにより、人とロボットの親和性を向上させる手法について検討する。

キーワード ヒューマノイドロボット、2足歩行、人とロボットの共存、画像処理

1. はじめに

世界的に例を見ない高齢化が進行する我が国において、高齢者支援ロボットや福祉関連ロボットは、今後、市場の拡大が期待できる民生用ロボットの有力な候補である。特に、超高齢化社会においては、高齢者の自立を支援するロボット技術に大きな期待が寄せられている。

この分野に使用するロボットに求められる基本コンセプトは、産業用ロボットとは異なり、「ロボットが人に合わせる」ということである。これを実現する技術の1つとして、人とロボットの親和性向上技術がある。この技術は、人のロボットに対する違和感や恐怖感を和らげるための技術である。本研究では、ロボットが人に対してリアクションを返すことにより、人とロボットの親和性を向上させる手法について検討する。

する人のリアクションを認識することによって(③、④)、ロボットの行動に対する評価・学習を行う。状況の認識と行動(決定・学習)の操作を繰り返すことにより人とロボットの間に徐々に親近感が創出されることを期待する。

図2に本研究で取り扱うインタラクションを実現するシステム構成例を示す。ロボットにはカメラとマイクを装着する。カメラからの画像を用いて人物認識及び顔ボ

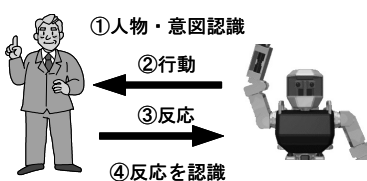


図1 インタラクションの概要

2. インタラクションの概要

本研究で対象とする人とロボットとの間のインタラクションの概要を図1に示す。

ロボットが人と自然に振舞うためには、相手が誰であるかを特定し、どのような意図を持っているのかを知ることが重要である。最初に、ロボットは人物とその意図の認識を行い(①)、この結果に基づいて行動を切り替える(②)。次に、ロボットが行動した後に、その行動に対

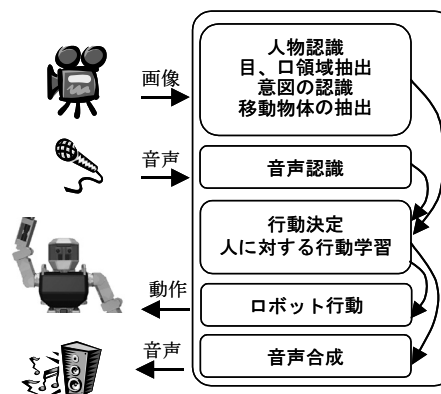


図2 システムの構成例

* 岐阜大学 工学部 応用情報学科

ーズ認識（物理量に基づく表情（意図）認識）を行う．マイクからの音声を用いて音声認識を行う．画像認識，音声認識ともに常に入力を受け付けることが可能で，双方の認識結果に基づいて行動を決定し，実際に行動を生成する．ロボットの行動にはロボット動作及び音声を使用する．ロボットの動作に対する人のリアクションがロボットに対する次の入力となる．さらに，ロボットは顔ポーズの認識を行うことにより直前の行動を評価し学習する．

3. ロボットシステムの概要

3. 1 機構系

本研究で開発したヒューマノイドロボットの機構系の概要は次のとおりである．人とロボットのインタラクションに関する研究のプラットフォームとして使用することを考慮し，上肢の自由度を多くしている．

- 1) 身長 約80cm
- 2) 体重 約15kg
- 3) 関節自由度 30自由度（図3）
- 4) 電源 バッテリーと外部電源の併用

3. 2 制御系

ヒューマノイドロボットは，非常に多くの関節を同時に制御するとともに画像処理や行動計画の生成等多くの情報を処理する必要がある．さらに，関節自由度が多いため，可能な限り配線量を少なくする必要がある．このため，処理内容のレベルを3つの階層（アプリケーション，運動制御，モータ制御）に分類し，独立のコントローラで処理することとした（図4）．また，各コントローラのCPUには，アプリケーションと運動制御に関しては，SH4（（株）ルネサステクノロジ）を，モータに関してはSH2（（株）ルネサステクノロジ）を使用した．体内LANには，1Mbps以上の高速通信が可能，マルチキャストな通信が可能，実装が容易である等の理由から，RS-485を使用した．

3. 3 歩行制御ソフトウェア

歩容の生成に関してはこれまでいろいろな手法が提案されている^{[1]-[5]}．これらの多くがオフラインであらかじめ歩容を生成する手法である．しかし，ロボットが環境に適応して歩行したり，人の行動に対して歩行を伴ったリアクションを返すためには，リアルタイムに歩容を生成する必要がある．本研究では問題を簡単にするため，歩容をSagittal Plane（XZ平面），Lateral Plane（YZ平面），XY平面に関して独立に定める．各平面に関する歩容は，倒立振子の倒れ込みの運動を基本とし，各フェーズのパターンが滑らかにつながることを考慮して定める（図5）^[6]．各平面における具体的な歩容を以下に示す．ここで，支持脚の運動は，足首からみた腰の運動として，誘客の運動は腰からみた足首の運動として記述する．

○Sagittal Plane の歩行パターン

Sagittal Plane に関して，線形倒立振子モード^[3]の考え方に基づいて，支持脚の片足支持期の腰の運動を次の様に定める．線形倒立振子モードにおいては，床からの腰の高さは変化しないので，x座標における運動のみを定めればよい．

$$x = -X' \cosh\left(\frac{t}{T_C}\right) + T_C V_{0x} \sinh\left(\frac{t}{T_C}\right)$$

となる．ここで，

Z_C : 足首から見た重心の高さ

T_S : 片足支持期の時間

$$T_C = \sqrt{Z_C/g}$$

g : 重力加速度

X_S : 歩幅

X_W : 両足支持期における重心のX軸方向の移動量

$$X' = (X_S - X_W)/2$$

$$V_{0x} = \left\{ X' + X' \cdot \cosh\left(\frac{T_S}{T_C}\right) \right\} / \left\{ T_C \cdot \sinh\left(\frac{T_S}{T_C}\right) \right\}$$

である．

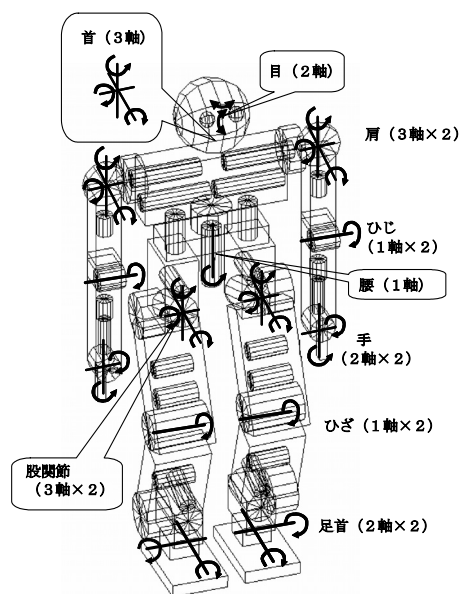


図3 ロボットの関節自由度

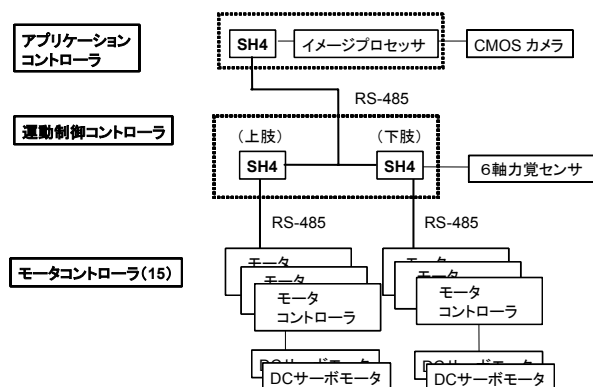


図4 制御ハードウェアの構成

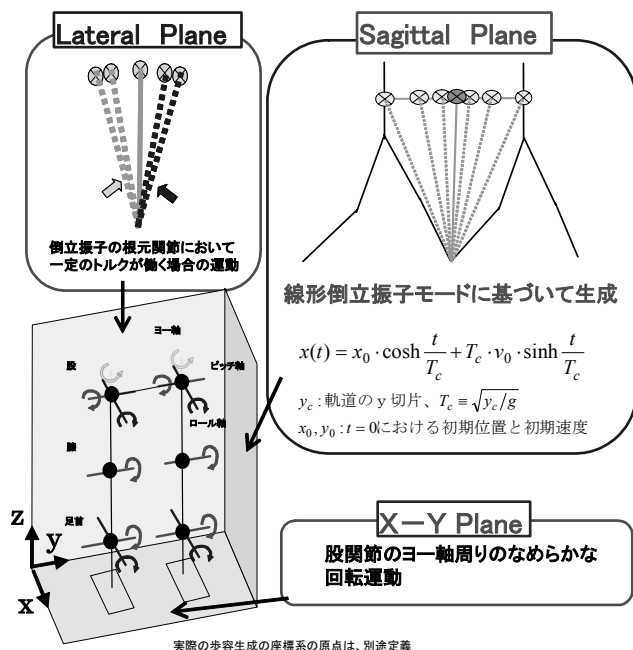


図5 歩行アルゴリズム

遊脚の運動は着地時に足裏が地面に対し垂直／水平方向に速度が0となるような4次関数で記述する．4次で運動を定義した理由は、各フェーズの位置、速度を連続的に結合させるためである．

○Lateral Planeでの歩行パターン

Lateral Plane に関する運動は、重心に質点が集中していると仮定した倒立振子の根元関節に一定のトルクが働く場合の運動に基づいて次のように定める．支持脚に関する片足支持期の腰の軌跡は、

$$y = Z_c \cdot \sin(A \cdot \exp(\alpha \cdot t) + B \cdot \exp(-\alpha \cdot t) + C)$$

となる．ここで、

θ_s : 倒立振子の最大振れ角

θ_w : 両足支持期へ移行時の倒立振子の振れ角

$$\alpha = \sqrt{g/Z_c}, \beta = \exp\{\alpha \cdot (T_s/2)\}$$

$$A = \frac{\theta_s - C}{2\beta}, B = \frac{\theta_s - C}{2\beta^{-1}}, C = \frac{\theta_s + \beta^2 \theta_s - 2\beta \theta_w}{(\beta - 1)^2}$$

である．

また、遊脚の運動は、支持脚に平行になるように定める．

○XY平面の運動

XY平面の運動は、片足支持期における支持脚の股関節のヨー軸周りの運動として3次関数で次のように定義する．

$$\theta = 2 \cdot \frac{\alpha - \beta}{T_s^3} \cdot t^3 - 3 \cdot \frac{\alpha - \beta}{T_s^2} \cdot t^2 + \alpha$$

ここで、

α : 片足支持期開始時の支持脚のヨー軸周りの回転移動量

β : 片足支持期終了時の支持脚のヨー軸周りの回転移動量

T_s : 片足支持期の時間

である．

実際のロボットの歩容は、これらの各平面の運動を合成することにより得られる(図5)．例えば、直線歩行に関しては、Sagittal Plane とLateral Planeの運動を、旋回歩行に関しては、直線歩行の運動にXY平面の運動を合成すればよい．

本研究では、上述したアルゴリズムに基づいて、アプリケーションコントローラからの指令により、前進、後退、旋回、旋回歩行をリアルタイムに実行する制御ソフトウェアを試作した．

3. 4 上肢制御ソフトウェア

本研究においては、上肢は人に対するリアクションのジェスチャを表現するために使用する．このため、試作したソフトウェアには、アプリケーションコントローラの指令に基づき次のような動作を実行する機能を組み込んだ．

○腕

- ・あらかじめ決められた動作(万歳、バイバイ 等)
- ・指定した物体(位置情報はアプリケーションコントローラが獲得)を指す．
- ・指先を与えられた目標座標(位置と指の姿勢)に移動する．

○顔

- ・指定した物体(位置情報はアプリケーションコントローラが獲得)の方向を向く．

腕に関して、目標座標を用いて指示するほかに、特殊コマンドを設定した理由は、人に対するリアクションには定型的なものが多く、これらを容易に設定できるようにするためである．

また、試作したロボットの腕においては、手先の3自由度の回転軸が1点で交わらないため、解析的に逆運動学を解くことが困難であることや腰の自由度を含めると冗長になることを考慮して、図6に示すような制御系を

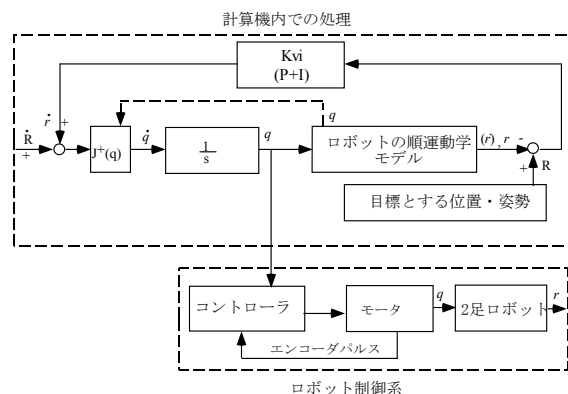


図6 制御のブロック線図

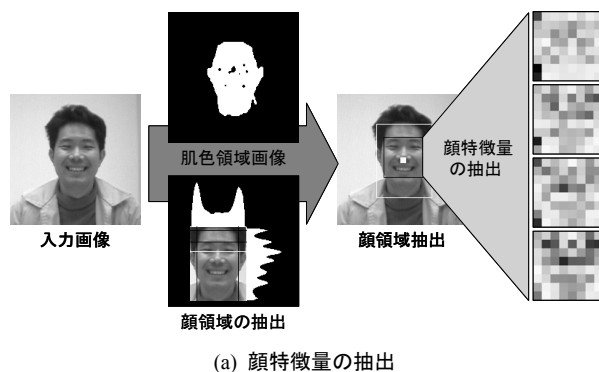
用いて制御することとした。この制御系においては、目標とする位置・姿勢 (R) とそれらの移動速度 ($\dot{R}$) を指令値として与え、ロボットの順運動学モデルにより算出した現在の位置・姿勢 (r) との偏差に対して、比例成分と積分成分を速度指令値にフィードバックする。これにより得られた移動速度に逆ヤコビ行列を乗じて各関節の速度に変換し、それを積分して目標の関節位置を求める。これを各モータコントローラに指令値として与える。

4. インタラクションシステムの開発

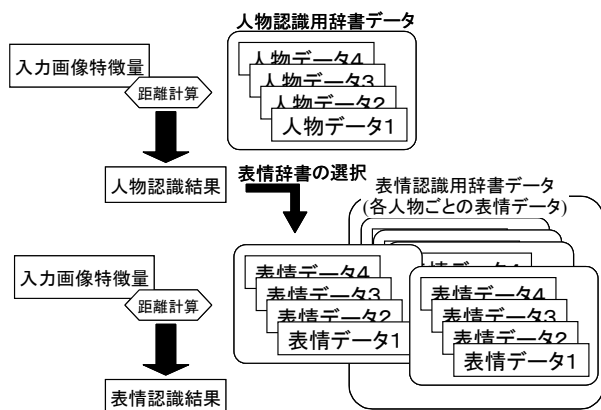
4. 1 人物・表情認識の実装

実装するソフトウェアの認識処理の流れを図7に示す。まず、入力画像から顔の特徴量を抽出する(図7(a))。色情報を用いて人物の肌色領域を抽出し、抽出した肌色領域のうち最大の肌色領域を顔領域候補とする。顔候補領域の水平・垂直方向エッジを利用して顔候補領域を決定する。得られた顔領域からエッジ特徴である4方向面特徴量を抽出する。

人物認識用の辞書及び各個人ごとの表情認識用の辞書は予め学習用データから抽出した特徴から線形判別分析により作成し使用する。抽出した顔特徴量と辞書のユークリッド距離を計算し、辞書ごとの距離を比較することにより人物・表情認識を行う^[7,8]。線形判別分析はより少ない次元でクラス間の分散を最大限強調する線形写像を構成する多変量解析手法である。



(a) 顔特徴量の抽出



(b) 人物・表情の認識

図7 人物・表情認識手法

図7(b)に人物・表情認識の流れを示す。まず、入力画像から得られた特徴量を人物認識用の辞書データと比較し人物認識を行う。次に得られた人物認識結果を用いて認識した個人の表情辞書を選択する。選択した辞書と入力された特徴量を比較することによって表情認識結果を得る。

図8に人物・表情認識の様子を示す。特定環境における認識実験の結果、人物認識・表情認識ともに高い認識率を達成した。

画像処理を行うためのビジョンセンサとしてVision sensor IP7500EB(株日立ハイコス)をロボットに搭載し認識を実現した。

4. 2 音声認識・合成の実装

音声認識・合成はその処理を外部PCにおいて実行し、TCP通信を介してロボットとデータ交換をすることにより実現した。音声認識には単語ベースの認識ソフトであるメディアドライブ社製“Wspeech”を利用した。音声合成には沖電気工業“Smart Talk”ライブラリを使用した。各ソフト間の通信用モジュールを開発しこれにより音声認識・合成による対話を可能にした。この様子を図9に示す。

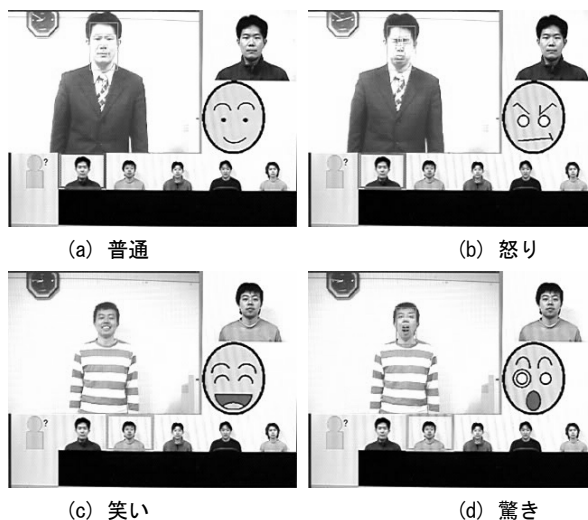


図8 人物・表情認識の様子



図9 音声認識・合成による対話の様子

5. インタラクシオンシステムの開発

5. 1 インタラクシオンによる表情の学習

5. 1. 1 表情自動学習・認識システム

実環境においてロボットが表情認識を利用する場合に、いつ、どのように表情の学習を行うのが重要な問題となる。そのため、自然な流れの中で人間とロボットがふれあいながら顔の表情を自動的に学習し、表情を認識できるようになるロボットシステム^[9, 10, 11]を構築した。

人の表情は、感情の自然表出としての表情と意図的なメッセージを伝えるための表情の二面性を有すると言われる^[12]。例えば、意図的なメッセージを伝えるための表情には、母親が赤ちゃんを叱る時に見せる表情があげられる。本研究で対象とする表情は、この意図を伝えるために表出した表情とする。しかし日常の生活で表出される表情はこの2つのうちのいずれか一方のみというのではなく、これらを複合した意味を持ち、これらを識別することは困難であると考えられる。このため、本研究では、学習・認識モデルにおいて表情を触覚刺激による感情と自然な形で結びつけて学習させることとした。

5. 1. 2 表情学習モデル

人が自然なコミュニケーションと感じられるようにするため、本研究ではロボットが表情を学習する手法として、赤ちゃんが母親の表情を学習する過程を模倣したモデルを提案する。一般に、赤ちゃんは、頭を撫でる、頭を叩くなどの母親の行動を触覚によって感じ、人間が本質的に持つ感情を変化させる。この時に同時に入ってくる視覚情報(母親の顔の表情)を結びつけることによって母親の表情を学習する。このモデルをロボットシステムに適用することで、人にとって自然に感じられる形の学習を実現する。本システムにより、あらかじめ表情の意味を教示しなくても表情を学習・認識することが可能になる。さらに、これにより個人の顔に応じた表情を学習することも可能となる。

図10に赤ちゃんの表情の学習・認識モデルを示す。学習の過程は次のとおりである。

- (1) 母親が特定の行動によって赤ちゃんを刺激する。この時、母親は行動に対応した表情をする。
- (2) 赤ちゃんは与えた刺激に対する感情を持ち、刺激に対して反応し行動する。
- (3) 赤ちゃんが感情に対応した表情を覚える。
- (4) (1)から(3)の繰り返しにより、赤ちゃんは母親の表情を見るだけで、刺激された状態と同じ感情を持ち、表情に反応するようになる。

怒りの表情を学習する過程を例に挙げる。はじめに母親は怒った顔をしながら、赤ちゃんを叩く。叩くという行動は怒った時に自然と出てくる行動の一つである。

叩くという行動が行われた時に、赤ちゃん側では触覚情報と感情と視覚情報の統合が行われる。赤ちゃんは触覚によって痛いと感じると、悲しいという感情を持つ。

同時に視覚から得られた情報として母親の怒った顔の情報を得て、この時に赤ちゃんは母親の怒った顔を学習している。悲しい感情と母親の叩く行動と母親の表情の結びつきが学習されている。繰り返し学習することにより、その結びつきは強くなり入力として叩くという行動により発生する触覚の情報がなくても、表情を認識し、悲しいという感情を持つようになる。同様に喜びの表情を学習する過程では撫でる動作と喜びの表情を嬉しいという感情に結びつけて学習する。また何も動作が無い時の表情を通常の表情として学習する。

5. 1. 3 表情自動学習・認識システム

図11に自動学習・認識システムのフローチャートを示す。システムにおける処理は学習フェーズ、認識フェーズの2つの部分に大別できる。

学習フェーズは表情を自動的に学習し、認識するために必要な情報を抽出するフェーズである。最初にセンサ情報及び画像情報から動作を検出する。行動を検出すると同時に、ロボットは検出した行動に関連する反応を示す。さらに、行動を検出するとそれがトリガとなり、画像処理を行う。入力画像の顔領域から特徴量を抽出し、行動検出の結果を用いて顔の特徴量を蓄積し学習する。行動と結びつけて特徴を蓄積することは、赤ちゃんが感情と顔の表情の関係を学習することに由来する。

認識フェーズは学習した表情の情報を使用して、表情を認識し、反応を返すフェーズである。学習フェーズと同様の手法を用いて入力画像から顔の特徴を抽出する。抽出した特徴を、学習フェーズで作成した表情の情報と比較することにより表情を認識する。これによりロボットは認識した表情に対応した反応を示す。

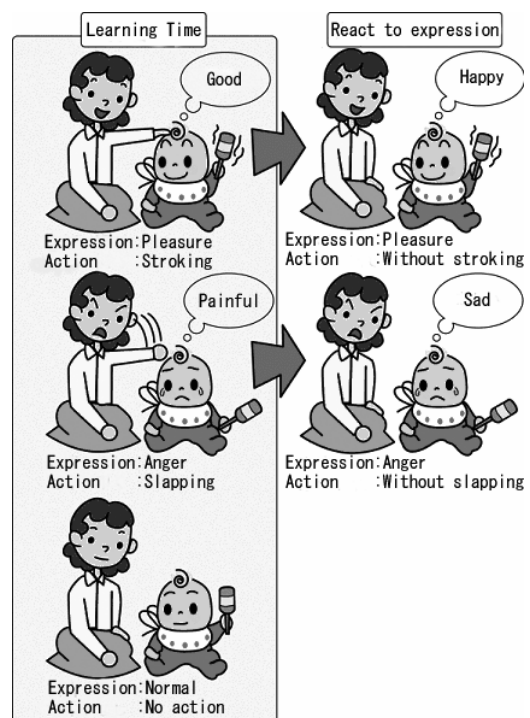


図10 表情学習モデル

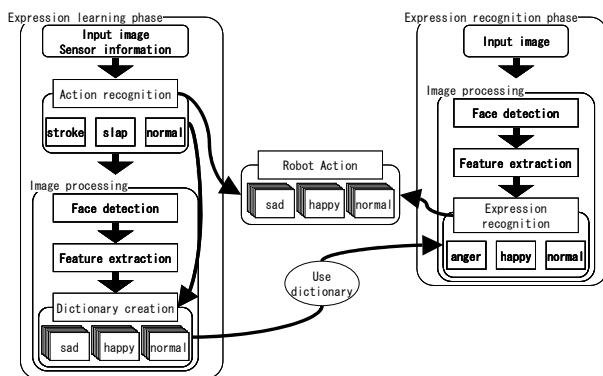


図 1 1 表情自動学習・認識システムの流れ

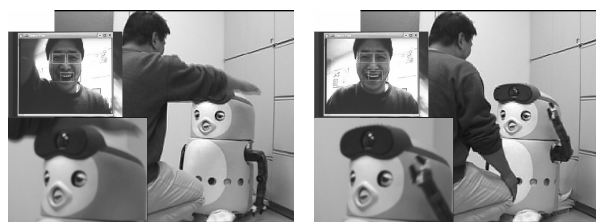


図 1 2 表情学習・認識システムのロボットへの実装
(左) 学習フェーズ (右) 認識中フェーズ

本システムでは人間の触覚の代わりにCdSセンサアレイを用いて人間の動作の検出を行う。さらに、カメラ画像を用いて人間がロボットの前にいるかどうかを検出する。

5. 1. 4 ロボットへの実装

表情自動・学習認識システムのロボットへの実装例を図 1 2 に示す。左上はロボットが見ている顔画像、左下はロボットの反応を拡大表示したものである。学習フェーズにおいて笑った顔を教示している様子を図 1 2 左に示す。CdSセンサを用いて人間の行動(撫でる)を検出し表情を学習している。学習後の認識フェーズの様子を図 1 2 右に示す。カメラにより表情(笑った顔)を検出し、笑った顔に反応してロボットが動作している。

6. まとめ

本研究では、ロボットが人に対してリアクションを返すことにより、人とロボットの親和性を向上させる手法について検討した。具体的には、ヒューマノイドロボットの機構を制御するための基本ソフトウェアを開発するとともに、人物・表情認識システムをロボットに実装し、人とロボットのインタラクションシステムを構築した。さらに、人とのインタラクションをより自然に行えるようにするため、各人に対する行動の学習の枠組みについて基礎的な検討を行い、ロボットへの実装を試みた。

文 献

- [1] 山口仁一, 高西淳夫, 加藤一郎 “上体運動により 3 軸モーメントを補償する 2 足歩行ロボットの開発”, 日本ロボット学会誌, Vol.11, No.4, pp.581-586, 1993
- [2] 西脇光一, 北川知伸, 杉原知道, 加賀美聡, 稲葉雅幸, 井上博允, “ZMP 導出の線形・非干渉化, 離散化によるヒューマノイドの動力学安定軌道的高速生成”, 第 1 8 回ロボット学会学術講演会予稿集, pp.721-722, 2000
- [3] 梶田秀司 “線形倒立振子モードを規範とする動的 2 足歩行ロボットの実時間制御”, 機械技術研究所報告, 第171号
- [4] 遠藤謙, 山崎文敬, 北野宏明, 前野隆司, “進化的計算法を用いた 2 足歩行ロボットの形態と歩行パターンの生成”, 第 1 9 回日本ロボット学会学術講演会予稿集, pp.519-520, 2001
- [5] Jiang Shan, Fumio Nagashima, “Biologically Inspired Spinal locomotion Controller for Humanoid Robot”, 第 1 9 回日本ロボット学会学術講演会予稿集, pp.517-518, 2001
- [6] 稲葉昭夫, 張勤, 光井輝彰, 千原健司, 小川行宏, 田中等幸 “ヒューマノイドロボット基本ハードウェアの開発 (第 2 報)”, 岐阜県生産情報技術研究所報告 No4, pp.1-6, 2003
- [7] K. Yamamoto : “Present State of Recognition Method on Consideration of Neighbor Points and Its Ability in Common Database”, Trans IECE, Vol.E79-D, No.5, pp.417-422, 1996.
- [8] 栗山聖・山本和彦・本郷仁志・加藤邦彦: 「顔認識システムにおける高次特徴面の効果と性能評価」, 画像の認識・理解シンポジウムMIRU2000講演論文集 I , pp.475-480, 2000.
- [9] Y. Ogawa, K. Kato, and K. Yamamoto, "Self-learning System of Facial Expression for Intelligent Robot through Interaction with Human", Proceedings of The 5th IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp.1233-1238 (2003)
- [10] T. Kobayashi, Y. Ogawa, K. Kato, K. Yamamoto and G. Wang, "Proposal of an Expression Self-Learning System for a Family Robot", Proceedings of 6th Japan-France Congress on Mechatronics and 4th Asia-Europe Congress on Mechatronics, pp.380-385, 2003.
- [11] 小川行宏, 加藤邦人, 山本和彦, “表情自動学習・認識システムの提案”, 電気学会論文誌, Vol.124-C, No.3, pp.699-705, 2004.
- [12] 吉川左紀子, “表情は何をつたえるのか-メディアとしての表情”, 岡田, 三嶋, 佐々木共編, “身体性とコンピュータ”, 共立出版bit別冊, pp.376-388, 2000