

# ヒューマノイドインターフェースに関する研究

～ 表情認識ソフトウェアの実装と精度検証 ～

小川 行宏

山本 和彦\*

稲葉 昭夫

杉山 正晴

## Research on Humanoid Interface

- Implementation and Accuracy Verification about Expression Recognition -

Yukihiro OGAWA Kazuhiko YAMAMOTO\* Akio INABA Masaharu SUGIYAMA

あらまし 著者らはこれまでに人とロボットが共存するための要素技術として人とロボットのインタラクションの中で人の意図に着目したシステムを提案してきた。本システムは人の意図の中で顔の表情に着目し、顔の表情を検出することにより、それに応じて行動するシステムである。本報告ではその中で重要な要素となる人物認識及び表情認識について、ヒューマノイドロボットに実装した。また、ヒューマノイドロボットにより人物・表情認識を行い、システムの有効性を確認した。表情の認識手法に人物認識の認識結果により表情認識辞書を選択する手法を用い実験により手法の有効性を確認した。

キーワード インタラクション、画像処理、人物認識、表情認識

### 1. はじめに

現在、成長が期待されている産業の一つとしてロボット産業がある。これまでのロボットは産業用ロボットに見られるように人の代わりに重労働や過酷な環境での作業に用いられ、生産の高度化、効率化、省力化のために活用されてきた。しかし、今後、超高齢化社会を迎える中で、ロボットがさまざまな形で人間社会に関わり、人の日常生活を支援することが予想される。この人と共生するロボットにおいては信頼感や親近感を感じられるロボットであることが重要になる。

しかしながら、人がロボットに対して信頼感や親近感を持つためには様々な問題が存在する。鉄の塊であるロボットが動作することに対して恐怖を感じるといったハードウェア的な問題や、人の生活環境をいかにセンシングするか、人等に対してどのように動作したらよいかというソフトウェア的な問題などが存在する。

これまでに人と共存するロボットの研究開発において様々な手法が検討されてきている<sup>[1-5]</sup>。個人を識別し個人ごとにアクションする手法や、人とロボットのインタラクションにおけるインターフェースのモデルの提案などが挙げられる。しかし、より親密なインタラクションを実現するためには相手の意図に基づいてリアクションすることが大切であり、このような取り組みを行っているものは少ない。

著者らはこれまでに人がロボットに親しみを感じるための手法として、人とロボットのインタラクションの中

で人の意図に着目するシステムについて検討してきた<sup>[6]</sup>。本システムは人の意図の中で顔の表情に着目し、顔の表情を検出することにより、それに応じて行動するシステムである。本報告ではこのシステムの枠組みをヒューマノイドロボットに適用するために、人物認識及び人の表情認識をヒューマノイドロボットに実装したので報告する。図1に実装したヒューマノイドロボット及び人とヒューマノイドロボットのインタラクションの様子を示す。

本報告は次のような構成になっている。2章ではこれまでに提案してきたシステムについて述べる。3章ではロボットへの人物・表情認識の手法について述べ、4章で認識手法の評価実験を行う。5章でまとめる。

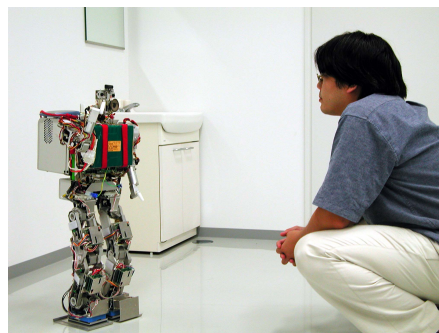


図1 ヒューマノイドロボット「ながら-2」と人間とのインタラクションの様子

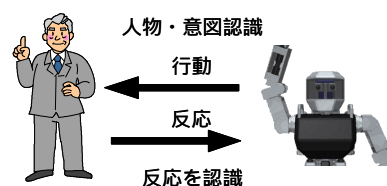


図2 システムの概要

\* 岐阜大学 工学部 応用情報学科

## 2. 親近感創出のための戦略

### 2.1 基本戦略

著者らはこれまで人とロボットの親和性を向上させる手法として、人とロボットのインタラクションの中で人の意図に着目したシステムを検討してきた。本章ではこのシステムの概略について述べる。

図2にシステムの概略を示す。本システムにおいて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。まず、ロボットは人物認識と人の意図認識を行い( )、得られた認識結果に応じて行動を切り替える( )。さらに、ロボットが行動した後、ロボットの行動に対する人のリアクションを認識することによって( )、ロボットの行動に対する評価・学習をする。環境の認識と行動(決定・学習)の操作を繰り返し行う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。

本システムの中で最も重要となるのが意図の認識であり、著者らは意図の中でも顔の表情に着目する。顔の表情は意図の中で目に見える形で認識可能なものの一つであり、意図の表現の中で重要な役割を果たしていると考えられる。本研究では予め顔のポーズを表情として定義しそれを認識する。例えば、「顔をしかめている」場合は「怒っている表情」とする。この定義は全ての場合で成立するわけではないが一般的な表情においては多くの場合に成立すると考えられる。

### 2.2 システム構成

図3に提案したシステムの構成を示す。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ポーズ認識を行う。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、実際に行動に移る。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。

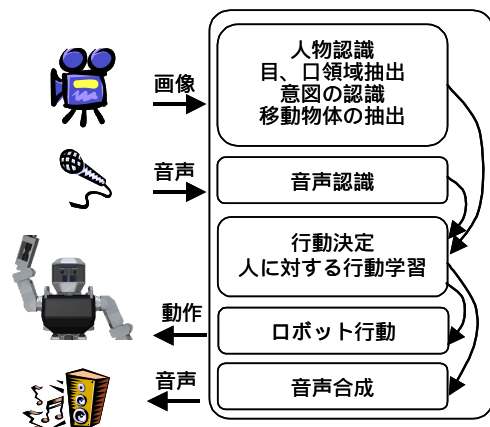


図3 システムの構成

## 3. 人物・表情認識ソフトウェアの実装

本システムではカメラから得られる画像を用いて人物・表情認識を同時に行う。本章ではロボットの画像処理ハードウェア及びそのハードウェアを使用した認識手法について述べる。

### 3.1 画像処理ハードウェア

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)を使用した。図4にIP7500EBの外観を示す。IP7500EBはCPUにSH-4<sub>0S</sub>にLinuxを採用した組込み型のビジョンセンサである。専用画像処理プロセッサSuperVchip(166MHz)を搭載しており高速に画像処理を行うことが可能である。

### 3.2 顔領域の抽出

入力画像から顔の特徴量を抽出するまでの流れを図5、抽出過程を図6に示す。まず、入力画像から色情報を用いて人物の肌色らしき領域を抽出し<sup>[7]</sup>(図6(b))、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。肌色領域の抽出には人の色に対する感覚に近い均等知覚色空間の一つであるLUV表色系を用いることにより環境の変化や個人差に対応した肌色抽出を行う<sup>[8]</sup>。

次に顔候補領域の水平・垂直方向エッジを抽出し、水平方向エッジをy軸方向(縦)に、垂直方向エッジをx軸方向(横)に射影したヒストグラムを作成する(図6(c))。水平方向エッジのヒストグラムのピーク値を求めることにより顔候補領域中の目の位置を推定する。図6(c)では水平方向の白いラインで示したのが検出した目の位置である。また、垂直方向エッジのヒストグラムの左右からそれぞれのピーク値を求めることにより顔のサイドラインの位置を求めることができる。サイドラインにより顔の幅を決定する。図6(c)中の垂直方向の黒ラインで示したのが検出したサイドラインである。求めた目の位置と顔幅から顔領域の上部、下部を決定する。これらは顔幅に予備実験で求めた一定の比率をかけた値を目の位置の上下に取ることで求める。図6(d)の黒のラインが抽出した顔領域である。



図4 IP7500EB外観

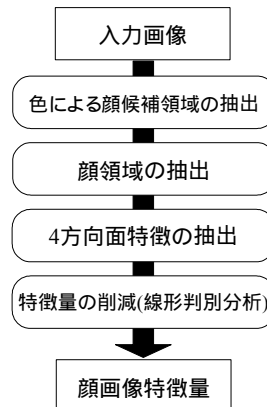


図5 顔特徴量抽出の流れ

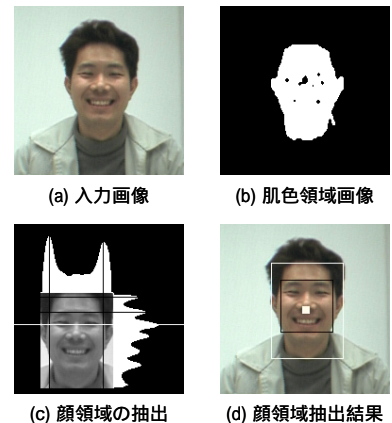


図6 顔領域の抽出

### 3.3 特徴量の抽出

顔領域からエッジ特徴である4方向面特徴<sup>[9]</sup>を抽出し、線形判別分析により特徴量を算出する。一般的にエッジ特徴は文字認識や人物認識の分野で良く用いられている手法である。その中でも4方向面特徴は濃淡の入力画像に対し方向検出フィルタを用いて、方向毎にエッジ強度を濃淡特徴とする手法である。ここで4方向面とは水平方向面、右上がり方向面、垂直方向面、右下がり方向面のことである。4方向面特徴は高次特徴面に分けた後、低解像度化しているため、1つの画像を4方向面に分ける前にエッジを抽出し解像度を落とす場合と比べて、高解像度でのエッジ情報が残る。また、4方向面特徴では非常に低い解像度まで落とすため、顔自体や顔部品の位置のわずかなずれが特徴量に及ぼす影響が少なくなるという性質を持っている。図7に方向検出フィルタを用いて4方向面特徴を抽出した例を示す。

### 3.4 人物・表情判別

本システムでは、人物認識用の辞書及び各個人ごとの表情認識用の辞書を予め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<sup>[10]</sup>。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。

図8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。

## 4. 人物・表情認識実験

本章ではロボットに実装した画像認識システムの認識実験を行う。

### 4.1 実験手法

各認識実験は6人の4つの表情(無表情・笑い・怒り・驚き)を対象とする。実験に用いる画像は「(無表情 笑い) × 3 (無表情 怒り) × 3 (無表情 驚き) × 3」と連続して顔の表情を変化した映像を約4フレーム/秒で取得した画像を用いる。1人につき約1000枚の画像を取得し、取得した画像を無表情・笑い・怒り・驚き・その他(表情変化の途中)と分別し、1人1表情につき54枚、合計1296枚の画像を使用した。そのうち半数を学習用データとし、残りの半数を未学習用データとして使用した。図9に人物・表情認識実験に使用した画像の一部(顔部分を切抜いたもの)を表示する。学習用データを用いて人物認識用辞書、各個人の表情認識用辞書を作成した。

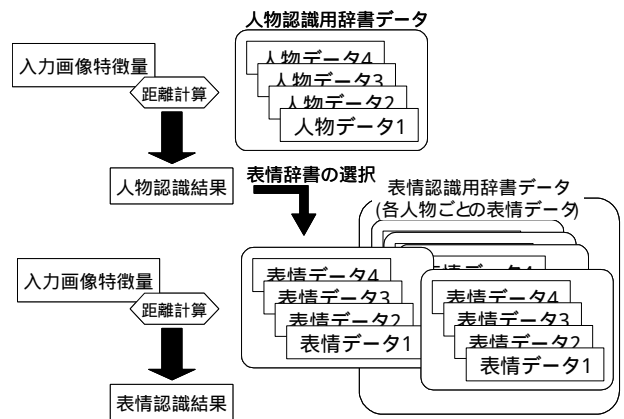


図8 人物・表情認識の流れ

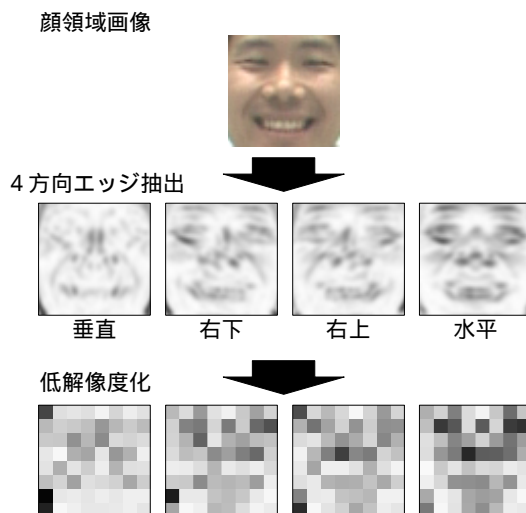


図7 4方向面特徴の抽出



(A) 驚き (B) 喜び (C) 怒り (D) 通常

図9 人物・表情認識の実験画像の一部(切抜き)

#### 4.2 人物認識結果

未学習用データに対する人物認識の結果を図10に示す。縦軸は認識率である。未知データの認識率は平均で99.54%と各人物ともに高い認識率となり、本システムにおいて人物の顔特徴の空間をうまく分割しており信頼性の高いものといえる。

#### 4.3 表情認識結果

本システムでは人物認識結果に基づいて表情認識の個人用辞書を選択する。この手法の有効性を確認するため全員分の表情を分類する1つの共通辞書を用いた場合と本システムの個人辞書を用いた場合の比較を行った。

未学習用データに対する表情認識の結果を図11に示す。共通辞書を用いた認識結果と本システムで用いた個人毎の認識結果を示す。縦軸は認識率である。

各表情ともに個人用辞書を用いた手法の認識率が高くなっている。共通辞書の認識率の平均は91.15%，個人毎辞書の認識率の平均は平均94.6%であり、本システムで用いた手法の有効性が確認できた。

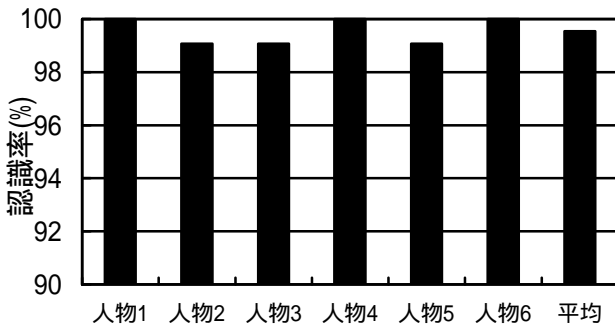


図10 人物認識の認識率

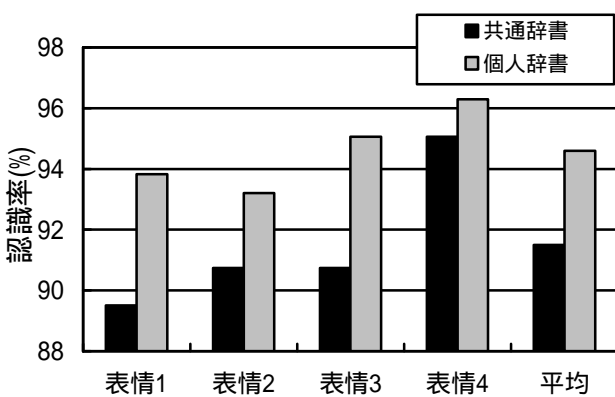


図11 表情認識の認識率

#### 5. まとめ

著者らは人とロボットの親和性の創出を目的として、インタラクションに基づくシステムを提案してきた。本報告ではその中で重要な要素となる人物認識及び表情認識について、ヒューマノイドロボットに実装した。実装したソフトウェアを用いて人物・表情認識の評価実験を行い高い認識率が得られることを確認した。また、表情の認識手法に人物認識の結果に基づいて表情認識辞書を選択する手法を導入することにより、認識率の向上を図った。

実際に人とロボットのインタラクションの中でシステム全体を動作させることが今後の課題である。

#### 文献

- [1] 日本電気株式会社, “ パーソナルロボット「Papero」 ”, <http://www.incx.nec.co.jp/robot/>
- [2] 藤田雅博, “ ロボットエンタテインメント: 小型4脚自律ロボット ”, 日本ロボット学会誌 Vol.16 No.3, pp.313-314, 1998.
- [3] 株式会社国際電気通信基礎技術研究所, “ 日常活動型ロボットプロジェクト ”, <http://www.mic.atr.co.jp/~michita/everyday/>
- [4] 小野哲雄, 今井倫太, 江谷為之, 中津良平, “ ヒューマンロボットインタラクションにおける関係性の創出 ”, 情報処理学会論文誌 Vol.41 No.1, pp.158-166, 2000.
- [5] 佐部浩太郎, 藤田雅博, “ エンターテインメントロボットの商品化 ”, 日本ロボット学会誌 Vol.18 No.2, pp.185-187, 2000.
- [6] Y. Ogawa, G. Wang, K. Yamamoto, A. Inaba, "Improving Human Friendly Interface based on Interaction with Robot -Development of Basic Technologies for Human Friendly Robot-", In Proc. of the Eighth International Conference on Virtual Systems and MultiMedia, pp.500-505, September, 2002.
- [7] 加島英樹, 本郷仁志, 加藤邦人, 山本和彦: “ マン・マシンインタフェースのための高速な虹彩検出法の考察 ”, 電気学会論文誌, Vol.123-A, No.4, pp.346-352
- [8] K. Yamamoto, "Present State of Recognition Method on Consideration of Neighbor Points and Its Ability in Common Database", IECE Transactions, Vol.E79-D, No.5, pp.417-422, 1996.
- [9] 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の効果と性能 ”, MIRU2000, pp.475-480, 2000.