

社会的共創知能に向けたアンドロイドの動作獲得

港隆史 (JST ERATO) 石黒浩 (大阪大学)

Generating Natural Motion in an Android towards Socio-Synergistic Intelligence

*Takashi MINATO (JST ERATO), Hiroshi ISHIGURO (Osaka University)

Abstract— In order to develop a robot working in daily situations, it is necessary to discover the principles to establish and keep social interaction between human and robot. One of the important issues is how to animate a robot to keep social interaction. This study tackles the issue through implementing natural motions in the android which closely resembles human beings. This paper proposes a method to implement postures that look human by mapping their three-dimensional appearance from a human performer to the android.

Key Words: Social interaction, Motion mapping, Posture mapping, Android

1. はじめに

近年、人間とのコミュニケーションを目的とした日常生活型ロボットが注目され、その開発が盛んになってきた。しかしながら、これまでに人間とロボットの相互作用、コミュニケーションの設計方法はあまり考慮されてこなかった。人間との相互作用を表層的なロボットの動作の作り込みにより実現したとしても、長期的な社会的相互作用を実現するのは困難である。したがって、社会的相互作用を維持するための原理を明らかにしなければならない。すなわち、人間の日常生活におけるロボットの動作の基礎的な表現系が必要である。

人間とロボットのコミュニケーションの原理を探る上で問題となるのは、ロボットの見かけと振る舞いの関係である。これまでにロボットの振る舞いが人間とのコミュニケーションに与える影響について研究されているが、ロボットの振る舞いだけでなくその見かけもまたコミュニケーションに影響を与えている。これまでに種々のヒューマノイドロボットが開発されているが、ロボットの見かけが機械的でよいのか、人間にできるだけ近い方がよいのかという問題は明らかにされていない。したがって、ある特定のロボットの振る舞いを調べただけでは、そのロボットに固有の機械的外見がコミュニケーションに与える影響を見落としている可能性がある。このように、人間とロボットのコミュニケーションを調べる上では、ロボットの見かけと振る舞いが互いにどのように影響を与えているかを区別することが難しいという問題がある。このような問題に対して、まず人間に酷似したロボットを開発し、そこから見かけと振る舞いを区別しながらコミュニケーションに与える影響を探るトップダウン的アプローチが考えられる。また、人間が人間の体や顔を認識するための脳機能を有している（たとえば [1]）ことから、人間に酷似した見かけがコミュニケーションにおいて重要であることが示唆される。そこで、我々は人間に見かけも動作も酷似したロボット-アンドロイド-の開発を通して、人間とロボットのコミュニケーション原理の解明に取り組んでいる。すなわち、アンドロイドを人間に近づける過程で、自然なコミュニケーション

を実現する要素を明らかにすることが目的である。

アンドロイドは人間らしい見かけと多自由度の運動機構を有しており、姿勢や動作による高い表現能力を持つ。すなわち、人間の細かな運動や、わずかな姿勢の変化が表現可能である。バイオリジカルモーションの研究では、人間は他者のバイオリジカルモーションから性別、感情、押している物体の重さの違いなどが判別できることが明らかにされている [2, 3, 4, 5]。すなわち、わずかな姿勢や運動の違いから様々な情報を読み取っていることがわかる。アンドロイドにおいても細かな姿勢や運動の違いを表現することにより、人間とのコミュニケーションにおいて多くの情報を伝えることができると考えられる。また細かな姿勢や運動の違いが、コミュニケーションにどのように影響を与えるかを調べるのが可能となる。

アンドロイドに人間らしい動作を行わせるためには、人間の動作を模倣させることが考えられる。ヒューマノイドロボットの動作生成に関する従来研究では、人間の動作をモーションキャプチャシステムで計測し、それをロボットに写像する方法が用いられる。Riley et al. [6] や Nakaoka et al. [7] はモーションキャプチャで計測された 3 次元データから、逆運動学を解いて人間の関節角度を計算し、それをロボットの関節角度に写像することにより、人間の動作をロボットに写像している。これらの研究では、人間がロボットと同じ関節構造を持つと仮定している。しかしながら、実際には関節構造は異なるため、関節角度を写像しても見かけの姿勢や動作が異なる可能性がある。また、人間の動作をロボットに写像する際に、類似度をどのように評価すべきかが問題となるが、Harada et al. [8] は、CG で描かれた 2 つのヒューマンフィギュアの姿勢の類似性を人間が評価するとき、人間の主観的な類似度はフィギュアの関節角度の類似度ではなく、手先等の位置の類似度から説明できることを明らかにしている。一般に、写像元の人間とヒューマノイドロボットの形や大きさが異なるために拡張性のある関節角度に基づいた写像が行われるが、アンドロイドは人間と酷似した体型をしているため、体の部位の位置を人間とアンドロイド間で比較することができる。これらのことから、ア

ンドロイドに人間らしい動作を行わせるためには、関節角の類似度ではなく体の部位の位置の類似度に基づいて人間の動作の写像を行うべきであると考えられる。

人間の動作をロボットに写像するためには、目標姿勢系列を作成し、さらにロボットのダイナミクスを考慮して目標姿勢系列を再生するコントローラを作成しなければならない。本研究ではこの問題を姿勢の写像と、姿勢系列を実現するコントローラの構築の2つの問題に分割する。そして本報告では、前者の姿勢の写像の問題を扱い、体の部位の位置の類似度に基づいて人間の姿勢をアンドロイドに写像する手法を提案する。

2. 人間からアンドロイドへの姿勢写像

2.1 開発したアンドロイド

Fig. 1 に開発したアンドロイドを示す。見かけ（頭部）は日本人成人女性をモデルとし、人間に酷似させてある。皮膚はシリコン製で人間の皮膚のように柔らかい。全身をシリコン製皮膚で覆うと可動範囲が制約されるため、関節部分は皮膚で覆われていない。アクチュエータは空気圧シリンダおよびエアモータで、上半身に42自由度配置されている。その内、13自由度が顔部、4自由度が指に使われている。自由度配置をTable 1に示す。肩には上腕の前後、脇の開閉だけでなく、肩全体の前後、上下運動を行うための自由度が設けてあり、多くのヒューマノイドロボットが持つ肩構造より複雑である。下半身にはアクチュエータがないため、図にあるように椅子に腰をかけた状態から移動できない。指定したシリンダ位置、ロータ位置（関節角度に相当）を実現するコントローラを体外に持っており、コントローラに指令値を送信することで動作させることができる。空気圧アクチュエータは高コンプライアンスを持つため人に対して安全である。その反面、複雑な動特性のため単純な制御で軌道追従を行うのは困難である。そのため本報告では、まず姿勢の写像のみを扱う。

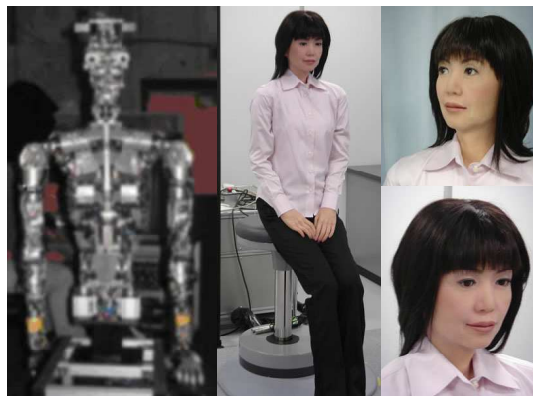


Fig.1 The developed android Repliee Q2.

2.2 基本的考え

人間の見かけの姿勢をアンドロイドに写像するため、アンドロイドと体型の似た人間を姿勢提示者とし、体の各部位（頭頂、肘、手首、肩など）の相対的位置を一致させるように、人間の姿勢をアンドロイドに写像

Table 1 The DoF configuration of Repliee Q2.

	Degree of freedom
Eyes	pan×2 + tilt×1
Face	eyebrows×1 + eyelids×1 + cheeks×1
Mouth	7 (including the upper and lower lips)
Neck	3
Shoulder	5×2
Elbow	2×2
Wrist	2×2
Fingers	2×2
Torso	4

する。人間からヒューマノイドロボットへの動作写像においては、ロボットの関節の可動範囲が人間と対応していない問題が生ずることが多い。同様に、見かけの姿勢の写像においても、体の部位の可動範囲が人間と対応していない問題が生ずる。

関節角に基づいた写像手法では、人間の動作に応じて適応的に関節角を可動範囲に収まるように拡大縮小する手法が提案されている[9]。このような手法では、人間とロボットの関節角のずれを、各関節ごとの評価を用いて補償している。しかし、全身の姿勢を人間に近づけることが目的であれば、姿勢の類似度を全身で評価する必要がある。たとえば、手先の位置のずれを、肘や肩の関節だけでなく、腰の関節を使って補償可能な状況も考えられる。

そこで本手法では、人間とアンドロイドの互いに対応する全身の体部位の位置のずれを同時に評価し、アンドロイドの姿勢を人間の姿勢に近づける。アンドロイドの関節構造は複雑であり、かつ弾性のあるシリコン製皮膚上にマーカーを取り付けるため、逆運動学を計算するのは困難である。そこで全身の体部位の位置のずれを評価関数として、探索的手法により人間の姿勢にアンドロイドの姿勢を近づけるアクチュエータ指令値を求める。

2.3 写像手法

アンドロイドと体型の似た人間を姿勢提示者とし、姿勢提示者の上半身の体表面にモーションキャプチャで位置を計測するためのマーカーを取り付ける。取り付ける位置は、アンドロイドの自由度を考慮して決定する。また姿勢提示者はアンドロイドが座っている椅子と同様の椅子に腰をかけた状態で姿勢を提示するものとする。さらにアンドロイドの体表面にも、姿勢提示者に取り付けたマーカー位置に対応する位置にマーカーを取り付ける。アンドロイドと姿勢提示者のそれぞれに固有のマーカー位置座標系を用意し、それぞれ対応する位置に原点を設ける。座標系のスケージングは行わない。

取り付けたマーカー数を n とし、姿勢提示者、アンドロイドのそれぞれの座標系におけるマーカーの3次元位置を $x_{hi}, x_{ai} (i = 1, \dots, n)$ とする。制御するアンドロイドのアクチュエータ数を m とし、アクチュエータ指令値を $u \in \mathcal{R}^m$ とする。まず姿勢提示者の姿勢を計測し、 x_{h1}, \dots, x_{hn} を得る。そして以下の評価関数を

最小にする u を山登り法により求める。

$$E_p = \sum_i^n \delta_i, \quad \delta_i = \|x_{ai} - x_{hi}\|$$

ここで、 δ_i は各マーカー位置誤差である。

山登り法は常に評価値が改善される方向に探索するため、焼き鈍し法などと比較して局所解に陥りやすい。この問題に対しては、評価関数を多目的化すると局所解を回避しやすくなる。三木ら [10] は単目的の最適化問題をダミー目的関数を用いて多目的化することで、局所解から抜け出しやすい山登り法を提案している。本手法でも局所解を回避しやすくするために、 E_p 以外の評価関数を加える。ここでは位置誤差 $\{\delta_1, \dots, \delta_n\}$ の標準偏差 σ_s を別の評価関数として加える。すなわち位置誤差の偏りを減らす評価を加える。よって評価関数は以下ようになる。

$$E = E_p + \alpha \sigma_s$$

ここで α は位置誤差に対する標準偏差の評価の重みである。

3. 実験

3.1 実験設定

人間の上半身の姿勢を写像するために、頭部に 4 個、首に 2 個、胸に 2 個、肩に 2 個 $\times 2$ 、肘に 1 個 $\times 2$ 、手首に 2 個 $\times 2$ の合計 18 個 ($n = 18$) のマーカーを取り付けた。取り付けた状態を Fig. 2 に示す (図では評価に使用していないマーカーも取り付けられている)。姿勢提示者、アンドロイドとも腰は動かないと仮定し、腰部に取り付けた 2 個のマーカーの midpoint をそれぞれの座標系の原点に定めた。制御するアンドロイドのアクチュエータは顔部の 13 カ所、両手首から先の 8 カ所を除いた 21 カ所 ($m = 21$) である。モーションキャプチャは Motion Analysis 製 Hawk Digital System を用いた。計測誤差は 1mm 以下である。

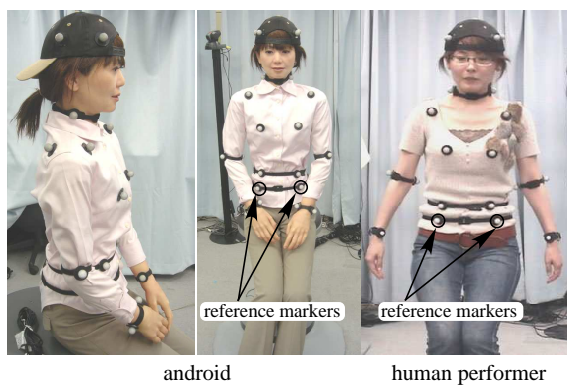


Fig.2 The marker positions.

わずかな姿勢の変化で意図や状態が伝わるような姿勢として、重いものを片手で持っている姿勢および肩を落として落胆している姿勢を写像した。評価関数中の標準偏差の重みは、位置誤差と標準偏差が同じオーダーになるように $\alpha = 18.0$ とした。探索の終了条件は、

一定ステップ (20 ステップ) 評価値が変化しないこととした。また探索時のアンドロイドの初期姿勢は、手動である程度目標姿勢に近づけることにより与えた。

3.2 実験結果

基準姿勢として座っている姿勢および上述した 2 つの姿勢を写像した結果を Fig. 3 から Fig. 5 に示す。図中の $\bar{\delta}$ はマーカー位置の平均誤差、 σ_s は標準偏差を表す。姿勢提示者とアンドロイドの体型は完全に等しいわけではないため、探索収束時点で 1 マーカー当たり平均 30mm 程度の誤差が残っている。またアンドロイドのアクチュエータの可動範囲が人間と対応していないことも誤差が残っている原因である。たとえば Repliee Q2 は肘をまっすぐに伸ばすことができないため、Fig. 3 の姿勢でも手首、肘部には誤差が生ずる。Fig. 6 には、Fig. 3 の姿勢探索時の評価関数の変化を示した。手動で与えた初期姿勢から目標姿勢に近づくように探索が行われていることがわかる。



Fig.3 The performer's and mapped posture (sitting, $\bar{\delta} = 26.4\text{mm}$, $\sigma_s = 13.8\text{mm}$).



Fig.4 The performer's and mapped postures (having a heavy baggage, $\bar{\delta} = 37.5\text{mm}$, $\sigma_s = 13.2\text{mm}$).

Fig. 4 では、アンドロイドのアクチュエータの可動範囲が足りずに誤差が大きくなっているが、腰を曲げて体を傾けて重いものを持っている状態が写像されている。Fig. 5 は肩を落として落胆している姿勢を写像した結果である。この姿勢では、姿勢提示者の肩が下

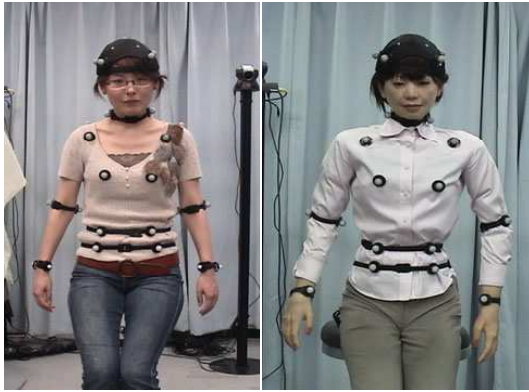


Fig.5 The performer's and mapped postures (discouraged, $\bar{\delta} = 27.6\text{mm}$, $\sigma_s = 11.1\text{mm}$).

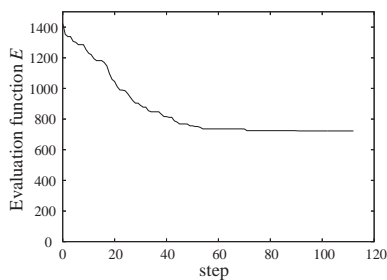


Fig.6 The searching process (sitting posture).

がってすばまった状態になっている。実際に、右肩前のマーカー位置は Fig. 3 より鉛直方向に 9.5mm 下がっている。また両肩前のマーカー間の距離は Fig. 3 より 2.7mm 狭まっている。これに対してアンドロイドの対応する部位もそれぞれ 2.6mm 下がり、1.0mm 狭まっている。アクチュエータの可動範囲の制限のため、姿勢提示者の肩の位置の変化量と同程度の変化は実現できていないが、肩が下がってすばまった状態が写像されている。

本報告では姿勢の写像に焦点を置いた。これを動作の写像に拡張するためには、提示者の動作からキーフレームごとの姿勢、あるいは一定時間間隔ごとの姿勢を計測し、その姿勢系列を実現するコントローラを設計する必要がある。本手法で各姿勢をアンドロイドに写像することにより、目標となる関節角（シリンダ位置、ロータ位置）系列を得ることができる、すなわち逆運動学を解くことができるため、関節角系列を実現するコントローラを設計すればよい。現在、ニューラルネットワークを用いて、アンドロイドのアクチュエータで目標関節角系列を実現するコントローラを構築する手法を開発している [11]。この手法と組み合わせることにより、人間の動作をアンドロイドに写像することができると考えている。

本実験の山登り法による探索では、多次元空間を実際のアンドロイドを用いて探索するため、探索時間が大きくなる。初期姿勢にも依存するが、本報告の実験では 1 姿勢の探索時間は数時間のオーダーであった。この問題に関しては、目標動作から抽出した姿勢系列を

時系列順に探索することにより、探索時間を減らすことができると考えている。

4. おわりに

本報告では、人間に酷似した見かけと多自由度の運動機構を持つアンドロイドの人間らしい動作の実現に向けて、人間の姿勢をアンドロイドに写像する手法を提案した。従来の関節角を介した手法ではなく、全身の体部位の位置の類似度を評価とすることにより、全身のアクチュエータを用いて目標姿勢に類似の姿勢を実現した。今後は動作の写像へ拡張するとともに、提示姿勢と写像された姿勢の類似度、写像された姿勢の自然さの主観的評価を行う。

- [1] D. I. Perrett, M. W. Oram, and E. Ashbridge. Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. *Cognition*, Vol. 67, pp. 111–145, 1998.
- [2] G. Mather and L. Murdoch. Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London Series B-Biological Sciences*, Vol. 258, No. 1353, pp. 273–279, 1994.
- [3] L. T. Kozlowski and J. E. Cutting. Recognizing the gender of walkers from dynamic point-light displays. *Perception and Psychophysics*, Vol. 21, pp. 575–580, 1977.
- [4] W. H. Dittrich, T. Troscianko, S. E. Lea, and D. Morgan. Perception of emotion from dynamic point-light displays represented in dance. *Perception*, Vol. 25, No. 6, pp. 727–738, 1996.
- [5] H. M. Paterson, F. E. Pollick, and A. Ude. Shaping biological motion: Adding realistic form cues to biological motion displays. *Journal of Vision*, Vol. 2, No. 7, p. 336a, 2002.
- [6] M. Riley, A. Ude, and C. G. Atkeson. Methods for motion generation and interaction with a humanoid robot: Case studies of dancing and catching. In *Proceedings of AAAI and CMU Workshop on Interactive Robotics and Entertainment*, 2000.
- [7] S. Nakaoka, A. Nakazawa, K. Yokoi, H. Hirukawa, and K. Ikeuchi. Generating whole body motions for a biped humanoid robot from captured human dances. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, 2003.
- [8] T. Harada, S. Taoka, T. Mori, and T. Sato. Quantitative evaluation method for pose and motion similarity based on human perception. In *Proceedings of IEEE-RAS/RSJ International Conference on Humanoid Robots*, pp. 1–19, 2004.
- [9] N. Pollard, J. K. Hodgins, M. J. Riley, and C. G. Atkeson. Adapting human motion for the control of a humanoid robot. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2002.
- [10] 三木光範, 廣安知之, 日和悟, 實田健. ダミー目的関数をもつ多目的確率的山登り法. 第 3 回情報科学技術フォーラム講演論文集, 2004.
- [11] H. Ben-Amor, S. Ikemoto, T. Minato, and H. Ishiguro. Learning android robot control using growing neural network. 日本機械学会ロボティクス・メカトロニクス講演会講演論文集, pp. 1P1–E07, 2006.