

Jerk 最小モデルに基づく指差し動作予測の印象評価

鈴木 優太[†] 大村 廉[†]

我々は、人とロボットとの自然なインタラクションを目的とし、Jerk 最小モデルに基づく人の指差し行動予測手法を提案している。本稿では、この手法を実際にロボット上に実装し、16 人の被験者を対象に印象評価を行った結果について述べる。予測の間違いなどから、「正確さ」については提案手法は追従手法に対しては評価が悪化するものの、「自分の行動を予測されている」という感覚や、自然さやなめらかさでは、提案手法は追従手法に対して良い印象を与えることが分かった。また、従来手法に対しても、提案手法は人や生物としてよりよい好印象を与える手法であることが示唆された。

Impression Evaluation of Pointing Prediction Based on Minimum-Jerk Model

YUTA SUZUKI[†] and REN OHMURA[†]

For establishing a model of a natural human-robot interaction, we are proposing a method that predicts human's pointing gesture using minimum jerk model. This paper reports an inquiry result obtaining from sixteen subjects with implementation on a real robot. While prediction error decreases the impression of "correctness", the feeling of "the robot predicts our behavior", "naturalness" and "smoothness" are increased by our proposed method. Moreover, the result implies our proposed method can give us better impression of a robot as a living thing more than existing method.

1. はじめに

近年、様々な生活支援ロボットの研究・開発が多く行われており、近い将来、生活支援ロボットは日常生活の中に溶け込んでいくことが予想されている。そのような社会において、ロボットが人間と自然にインタラクションを行えるようになることは極めて重要である。そのような研究の一環として、我々は人間同士の対話において頻繁に生じる「指さし（指示）行動」を対象とし、人が何か対象物を指した時のロボットの自然な振る舞いに関する研究を行っている。

対話において相手が何かを指示する場合、我々は話し相手が指を指し終わる前に大体の方向や対象を無意識に推定し、その方向に注意を向けている場合が多い。つまり、人は相手が指示行動を行う場合、その行動が完了する前に指示対象を「予測」しており、ロボットの自然な振る舞いを実現するには、この『相手の行動を予測する』ことが必要不可欠となる。そこで我々は、Jerk 最少モデルに基づく指さし方法を提案している²⁾。しかし、画面上のエージェントでの実装にと

どまり、実際にロボットを用いた評価実験は行われていなかった。

そこで本稿では、2) において提案した手法を実際にロボット上に実装し、印象評価を行った結果を報告する。なお、ロボット上に提案手法を実装し、ロボットを動作させるに当たり、提案手法ではどうしても「初動の動作が遅れる」という問題が生じた。このため、提案手法の改良を行い、Sugiyama らの研究で用いられた、0.3 秒後先の手の位置を予測する手法¹⁾を、手の初動時に取り入れてレスポンスの向上を行った。本稿では、予測を全く行わない場合、Sugiyama らの手法、提案手法それぞれについて、ユーザの印象評価を行った結果を示す。

以下、第 2 章では 2) において提案した手法を説明し、第 3 章においてロボットへの実装、および、ロボットへの実装に伴う問題点、改良点を述べる。そして、第 4 章では、行った印象評価実験について述べる。第 5 章で実験の考察を述べる。最後に、第 6 章で本論文のまとめを述べる。

2. 提案手法

指さし時に手が止まる位置を予測するためには、人間の手の動きがわからなければならない。Flash ら⁴⁾

[†] 豊橋技術科学大学
Toyohashi University of Technology

は平面上の2点間を移動する人間の手の運動を表すモデルとして、Jerk 最小モデルを提案した。Jerk 最小モデルは、完全ではないものの比較的良く人間の運動と一致し、かつ、その軌跡や運動を解析的に解くことができる、という特徴を持つ。本研究では人間の行動モデルとしてこの Jerk 最小モデルを用いる。このモデルは人間の手が移動するとき、躍度の積分(式(1))が最小となるというものである。

$$\int_t^T x'''(t)^2 dt \quad (1)$$

そして、このモデルにおいて、人間の手の動きの開始点、最終点で速度と加速度が0であり、開始点を \mathbf{x}_0 、最終点を \mathbf{x}_f とすると、ある時刻 τ における手先の座標 $\mathbf{x}(\tau)$ は以下のように表すことができる。

$$\mathbf{x}(\tau) = \mathbf{x}_0 + (\mathbf{x}_f - \mathbf{x}_0)(6\tau^5 - 15\tau^4 + 10\tau^3) \quad (2)$$

$$\tau = \frac{t}{T} \quad (3)$$

ここで、 t は動作開始時からの時間であり、 T は動作開始から終了までにかかる時間である。このとき $\mathbf{x}(\tau)$ に対する速度は図1となり、また、加速度は図2のようになる。

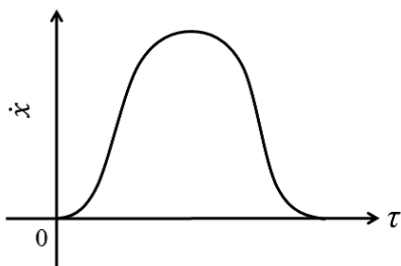


図1 速度の概形

Fig. 1 Approximate shape of velocity.

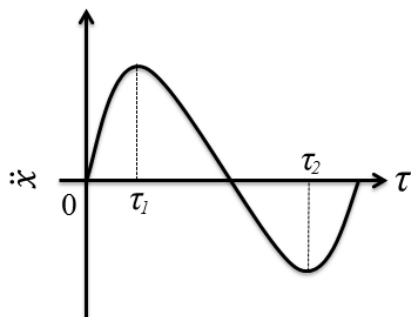


図2 加速度の概形

Fig. 2 Approximate shape of acceleration.

ここで、加速度の傾きが変化する点 τ_1 、 τ_2 は、式(2)を3回微分した式

$$\frac{d^3 \mathbf{x}(\tau)}{d\tau^3} = (\mathbf{x}_f - \mathbf{x}_0)(360\tau^2 - 360\tau + 60) \quad (4)$$

から、 $\frac{d^3 \mathbf{x}(\tau)}{d\tau^3} = 0$ を満たす τ を求めると

$$\begin{cases} \tau_1 = \frac{3 - \sqrt{3}}{6} \approx 0.21 \\ \tau_2 = \frac{3 + \sqrt{3}}{6} \approx 0.79 \end{cases} \quad (5)$$

となる。よって、動作開始から加速度が極大になるまでの時間 τ_1 は動作終了までの時間の約21%であることがわかる。つまり、指さし行動時に、手先の加速度が増加から減少に変わる(加速度が極大になる瞬間)が分かれば、その指さし行動にかかるおおまかな時間および最終的な手先の座標を求めることができる。終点 \mathbf{x}_f を求めるためには式(2)を

$$\mathbf{x}_f = \frac{\mathbf{x}(\tau) - \mathbf{x}_0}{6\tau^5 - 15\tau^4 + 10\tau^3} + \mathbf{x}_0 \quad (6)$$

と変形し、その時の手先の座標 $\mathbf{x}(\tau_1)$ 、動作開始時の座標 \mathbf{x}_0 、 τ_1 (式(5))を代入し、

$$\mathbf{x}_f = \frac{4\mathbf{x}(\tau_1) - (2 + \sqrt{3})\mathbf{x}_0}{2 - \sqrt{3}} \quad (7)$$

を求めることで手先の最終点の座標 \mathbf{x}_f を求めることができる。

3. 実装

3.1 ロボットへの実装

本研究では、手先の軌跡の認識に3次元深度画像センサ(Microsoft社のKinect)を使用した。そして、Kinectのデバイスコントロールおよび画像処理からのジェスチャ認識を行うためにOpenNI⁷⁾を使用し、手先の座標を得た。また、Kinectから得られた予測された値を、卓上型ロボット(ATR知能ロボティクス研究所のRobovie-W⁵⁾)に無線でデータを送信し、ユーザと同じ場所を指すような指さし動作を行わせた。

提案手法を実現するためのアルゴリズムを図3に示す。指さし行動の開始の認識は以下のように行った。まず、人間の手先の位置をフレームごとに取得し、1フレーム前の手先の位置との差を求めることにより、フレームごとの速度を求めた。このとき、閾値を決め、その値以上の速度となった時、指さし行動開始時点としてその時刻および手先の座標を記録した。

以降は、手先の位置のフレーム間差分により速度、並びに速度の差分により加速度、加速度の差分により躍度をフレーム毎にそれぞれ求めた。式(5)の τ_1 に相当する点を検出するために、加速度が極大となる時間を探すが、実際に0になることは稀であるため、時

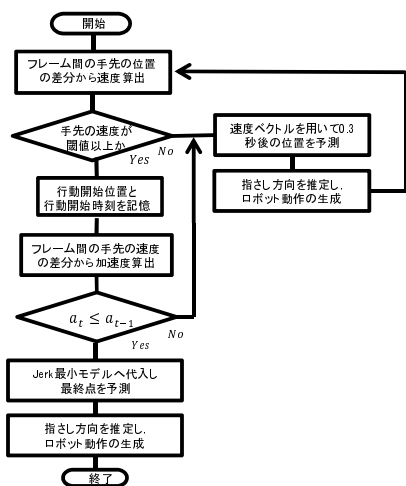


図 3 提案手法を実現するためのアルゴリズム

Fig. 3 The flowchart to achieve proposed method.

刻 t において、指さし行動開始後はじめて

$$a_t \leq a_{t-1} \quad (8)$$

となるとき、 $t-1$ を τ_1 の時刻とし、この時の手先の座標を $\mathbf{x}(\tau)$ として、式 (7) に代入した。その後、予測された座標の値から、指さしの方向を予測し、その方向に対応したデータをロボットに無線で送信し、ロボットに指差しを動作させた。

なお、本研究で用いるロボットは、あらかじめ決められた動作を再生させる形で制御を行うロボットである。このため、Kinect より得られる画像の x 座標と y 座標から空間を縦 $9 \times$ 横 9 の 81 マスに区切った後、それぞれマスの位置に対応するロボットのポーズを作成した。そして、指さし行動からロボットを動かす際は、指さし動作により動作プログラム上で出力された座標の値を使用して 81 マスのどこに当てはまるかを判断して、そのマスに対応したポーズをロボットに取らせた。

3.2 問題点とその対応

図 3 にあるように、提案手法では、手先の速度を計算し、あるスレッシュホールド値以上になったら 2 章で述べた計算式を用いて手先の最終点を予測する。このとき、このスレッシュホールド値の決定は非常に困難であり、値が小さければ、少しの手の動きで動作が予測されてしまい、思わぬ時にロボットが思わぬ方法に向いてしまう、ということが起こる。一方、スレッシュホールド値の値を高く設定すると、ロボットの反応が悪くなり、あまり意識せずに気なく手を動かしたとしても最終点を予測するアルゴリズムが動作しないか、反応したとしても手先の動きがかなり速くなってから予測アルゴリズムに動作が移り、結果的に反応が遅れてしまう、

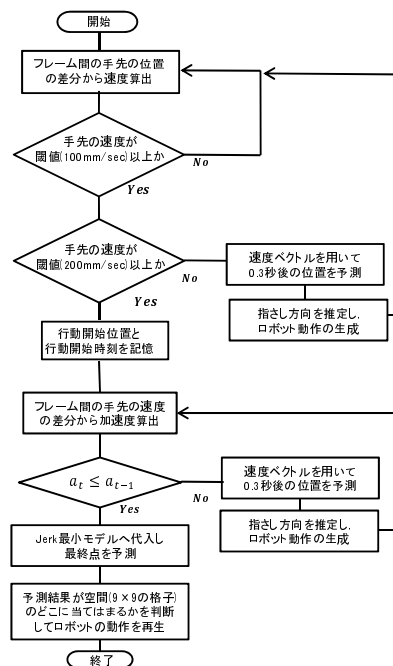


図 4 提案手法を実現するためのアルゴリズム

Fig. 4 The flowchart to achieve proposed method with prediction after 0.3 sec.

ということが起こる。

当初、我々は「手先をぶらぶらさせている時」と「指さし行動を行う時」それぞれの速度を何度か測定し、その時の速度の分布から、2つのケースが良く分離される速度を決定してこれをスレッシュホールド値として用いた。しかし、この値を用いた場合、反応が遅く成る傾向にあり、手がかなり動かないとアルゴリズムが動かない傾向となった。

そこで、我々は、予測アルゴリズムを開始する前のロボットの動作として、Sugiyama らの手法¹⁾を導入した。Sugiyama らの手法では、指の速度ベクトルを用いて、今ある指の位置から 0.3 秒後の位置を予測し、ロボットにその方向を向かせる。本研究でも、スレッシュホールド値を超える前はロボットをそのように動作させ、スレッシュホールドを超えた時に 2 章で述べた予測アルゴリズムを動作させるようにした。このような改良を加えた提案システムのアルゴリズムを図 4 に示す。

4. 印象評価実験

4.1 実験方法

実験では、被験者から 1.60m のところに Kinect を設置し、そのすぐ後方にロボットを設置した。また、被験者の左右 0.75m、前方 0.80 の位置にオブジェクトを設置した (図 5)。被験者には、左右前に置いたオブジェクトに指差しを行ってもらい、このときロボッ

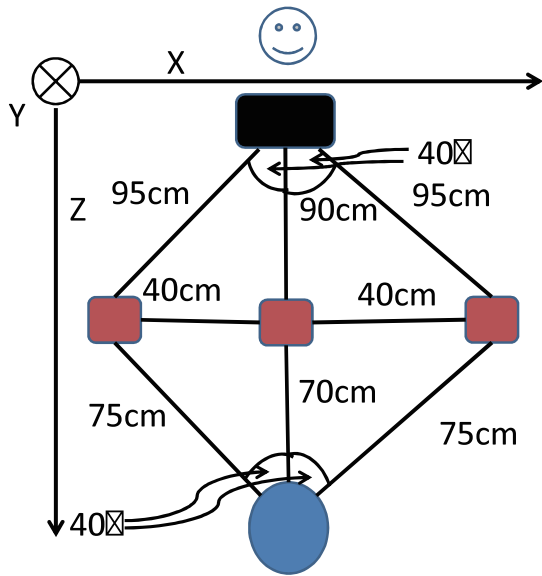


図 5 実験図
Fig. 5 Experiment figure

トを、予測なし動作 (追従), 従来手法 (0.3 秒後を予測¹⁾), 提案手法, の 3 パターンで動作させた。

予測なし動作 (追従) 人の指の動きを予測はせずに人の指を追う動き (追従) させる。

0.3 秒後予測 指の速度ベクトルを用いて, 今ある指の位置から 0.3 秒後の位置を予測し続けることで, 指差しの予測を行う。

提案手法 Jerk 最小モデルという人の数理的行動モデルを元に指さし先を予測することで, 初動のわずかな動作だけで手の指す最終点を予測する (2 章のアルゴリズム)。また, Jerk による指差しの認識が行われるまでは, Sugiyama らの手法を用いて予測を行わせる。

そして, 被験者にはそれぞれに対する印象をアンケートに回答してもらった。この時, 被験者にはどのロボットの動きがどのような手法を用いているのかは説明せずに, 動作パターンの順番についても被験者によってランダムに行った。被験者は, 理工系大学生および大学院生の学生, 計 16 名とした。実装環境を表 2 に示す。また, 実験の様子を図 6 に, 同じモーションデータを与えた時の, 各手法におけるロボットの動きの違いを図 7 に示す。

アンケートについては SD 法と 5 段階評価を採用し, SD 法による項目は Robovie の印象評価を行っている神田ら³⁾ を参考にした。アンケートの項目を表 1 に示す。また, アンケート結果からそれぞれに差があるかどうか t 検定を行って有意差の有無を判定した。

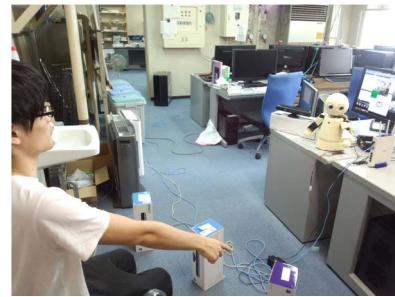


図 6 システムを用いた指さし動作
Fig. 6 Pointing action using the system.

表 1 アンケート項目
Table 1 Question Item.

SD 法	
	速い-遅い
	自然な-不自然な
	正確な-不正確な
	スムーズ-ギクシャク
	人間的な-機械的な
	優しい-怖い
	すばやい-のろい
	タイミングの良い-タイミングの悪い
	敏感な-鈍感な
	なめらかな-ぎこちない
	賢い-頭の悪い
	感じのよい-感じの悪い
5 段階評価	
	自分が指した方向にロボットが向いたか 自分の行動を予測されているように感じたか

表 2 実験環境
Table 2 Experiment environment.

CPU	Intel Core i7-2600K 3.40GHz
メモリ	16GB
OS	Windows7 64bit
速度の閾値	200mm/s

5. 考 察

図 9 および, 表 3 より, 「自分の指した方向にロボットが向いたかどうか」という点について, 提案手法は追従方式よりも評価が低い形で有意差が存在した。提案手法は予測が外れ, 被験者が意図した場所とは異なる方向を見てしまった場合などがあり, このような差が生じていると考えられる。また, 「自分の行動を予測されているかどうか」という点については, 提案手法は追従方式に対して良い評価では優位差が存在した。これより, 提案手法は追従方式にくらべ, 時々予測が外れるものの, 概ね「自分の行動を予測している感」は与えられていると言える。なお, これらの内容について, 提案手法と従来方式 (0.3 秒後方式) は, 平均値は提案手法と追従方式の間であるものの, 提案方式

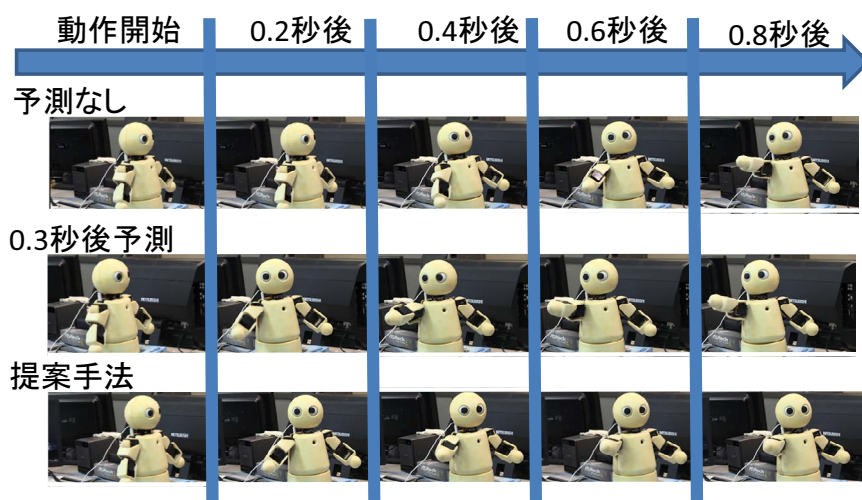


図 7 各手法におけるロボットの動作
Fig. 7 Robot motions of each method

表 3 有意検定 (5 段階評価)(○ : 有意差あり ($p < 0.5$), △ : 有意傾向 ($p < 0.10$), × : 有意差なし)
Table 3 Significance test.(5 levels quistion)(○ : ($p < 0.5$), △ : ($p < 0.10$), × : ($p > 0.10$))

	自分が指した方向 にロボットが向いたか	自分の行動を予測 されているように感じたか
提案-追従	○	○
追従 0.3 秒	△	×
提案 0.3 秒	×	×

表 4 有意検定 (SD 法)(○ : 有意差あり ($p < 0.5$), △ : 有意傾向 ($p < 0.10$), × : 有意差なし)
Table 4 Significance test.(SD method)(○ : ($p < 0.5$), △ : ($p < 0.10$), × : ($p > 0.10$))

	速い	自然な	正確な	スムーズ	人間的な	優しい
提案-追従	○	○	○	○	×	×
追従 0.3 秒	○	×	×	×	×	△
提案 0.3 秒	×	○	○	○	○	×
	すばやい	タイミング	敏感な	なめらか	賢い	感じのよい
提案-追従	○	×	△	△	×	×
追従 0.3 秒	○	×	△	×	×	×
提案 0.3 秒	×	×	×	○	△	×

との優位差は見られなかった。

また、図 8 および表 4 より、提案手法と追従手法との間に良い方向で優位差および優位傾向があったのは、「速い」「自然な」「スムーズ」「すばやい」「敏感な」「なめらか」といった項目であった。また、悪い方方稿で「正確な」にも優位差が存在した。一方、「人間的」「優しい」「タイミングのよい」「賢い」「感じのよい」といった項目には優位差が存在しなかった。前者の形容詞群は、その傾向として機能やその性能を評価する形容詞と考えることができる。一方、後者の形容詞群の傾向は、人や生物を相手にした時に用いられる形容詞である。このことから、提案手法は、その機能や性能としては優位に追従手法より優れるものの、まだ対象であるロボットについて、人や生物としてその

評価を向上させるには至っていない、と考察できる。ただし、「自然な」という項目については優位性が存在しており、動く物体として、あるいは、コミュニケーション対象としての「違和感」は、提案手法によってある程度緩和されているものと推察できる。

一方で、提案手法と従来手法 (0.3 秒後) については、「自然な」「スムーズ」「人間的」「なめらか」「賢い」と言った形容詞について、提案手法に良い方向で優位差あるいは優位傾向が存在した。従来手法より、「人間的」や「賢い」と言った点で優位差があることは非常に興味深く、他にも「優しい」「感じのよい」といった項目ではほとんど差が無いことから一概にはいえないものの、従来手法より提案手法の方が、「人や静物としての評価を向上させる」手法である、ということが示

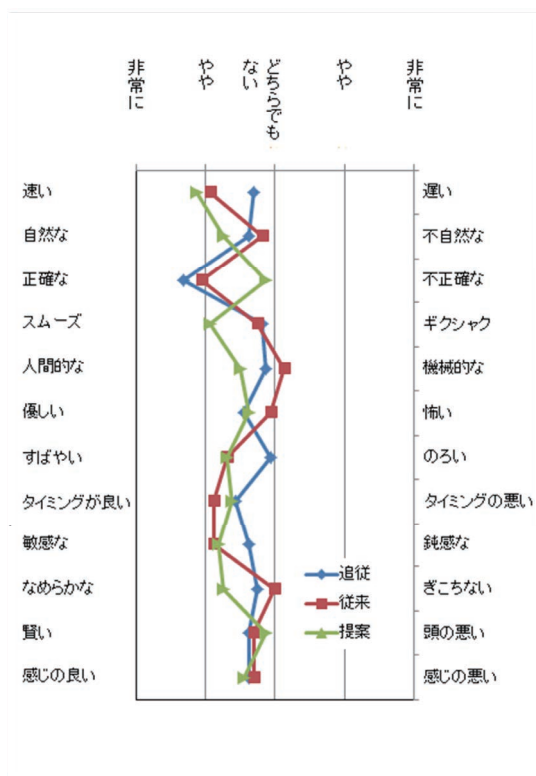


図 8 SD 法によるグラフ
Fig. 8 The graph by a SD method.

峻されているとみてとれる。なお、「正確な」については、追従手法同様、悪い方稿に優位差が存在した。これは、図 9 および、表 3 からの考察同様、やはり、提案手法では、間違う場合がある、ということが影響しているものと考えられる。

以上をまとめると、今回の印象評価実験では、提案手法と追従手法では、「自分の指した方向を見ている」という印象や「正確さ」については追従手法の方が勝るものの、「自分の行動を予測されている」という感覚や、自然さ、すばやさ、なめらかさと言った項目では優れていると言える評価が得られた。また、従来手法と比較した場合、提案手法は人や静物としてよりよい好印象を与える手法であることが示唆された。

6. ま と め

我々は、自然な人とロボットのインタラクションの実現を目的とし、人の指さし行動の予測を行うために、Jerk 最小モデルに基づく指さし予測手法の提案を行っている。本稿では、提案手法を実際にロボット上に実装し、16 人の被験者にアンケート調査を行った結果について報告した。ただし、提案手法は、Jerk 最小モデルを用いた予測に加え、初動の応答を改善するため、指さし動作の初動では 0.3 秒後の手先の位置を予

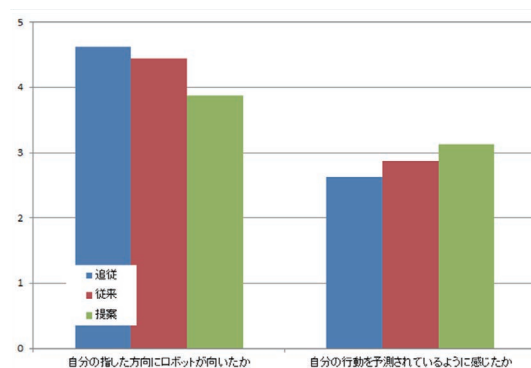


図 9 5 段階評価によるグラフ
Fig. 9 The graph by five-step evaluation.

測するという、従来の制御手法に基づく改良を加えたものとした。

理工系学生 16 人に対してアンケート調査を行った所、予測による間違いなどから、「正確さ」については提案手法は追従手法に対して評価が悪化するものの、「自分の行動を予測されている」という感覚や、自然さやなめらかさでは優れた印象を与えられることが分かった。また、従来手法に対しても、提案手法は人や生物としてよりよい好印象を与える手法であることが示唆された。

参 考 文 献

- 1) O. Sugiyama, T. Kanda, M. Imai, H. Ishiguro, N. Hagita: *Natural Deictic Communication with Humanoid Robots*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2007), pp.1441-1448 (2007).
- 2) 草野 祐樹, 大村 廉 *Jerk 最小モデルを用いた人の指さし行動予測手法の検討*, 情報処理学会インタラクション 2012 (IPSJ Interaction 2012), pp.929-934(2012).
- 3) 神田 崇行, 石黒 浩, 小野 哲雄, 今井 倫太, 中津 良平 *人間と相互作用する自律型ロボット Robovie の評価*, 日本ロボット学会誌 Vol. 20 No. 3, pp.1-9, (2002)
- 4) T. Flash, N. Hogan: *The Coordination of Arm Movements*, J. Neuroscience, vol.5, pp. 1688-1703 (1985).
- 5) <http://www.irc.atr.jp/>
- 6) <http://www.xbox.com/ja-JP/kinect> .
- 7) <http://www.openni.org/> .