

## 異種のヒューマノイドロボットにおける動作獲得と知識化に関する研究

熊谷 賢弥

指導教員：小林 邦和

## 1 はじめに

近年ロボットにおけるクラウドネットワーク技術の利用が注目を浴びてきている。よく目にする事例として、ロボットに搭載されているカメラの映像をスマートフォンやパソコンを使用して見られるものや、ロボットが一方的な対話によって得た情報を他のロボットが用いるものがある。また、研究分野としてクラウドロボティクスが発足し、クラウド技術を利用した物体認識や、各センサーの情報をクラウドで統合しロボットの行動制御を行う研究がされている。しかし、ロボットの動作を知識化し、クラウドネットワークを用いて他のロボットがその知識を利用する研究はあまりみられない [1]。

ロボットの動作はロボットの個体に依存するため、人が動作を作り込むか、学習によって動作を獲得するかである。作り込みでは、人が想定した動きにしか対応できないため、学習により獲得する方法が適用分野が広い。動作を獲得する学習法はたくさんあり、人の動作学習に似たモデルである模倣学習が特に注目を浴びている。そのため、本研究では動作を知識化することを想定し模倣学習を行う。また、他のロボットが動作の知識を利用できるように動作情報を変換し、知識化する。

## 2 運動学習法

ロボットの動作獲得では、人間の動作学習に似たモデルである模倣学習が注目を浴びている。模倣学習は教師あり学習のひとつで、ロボットに動作を学習させるのに向いている手法であり、多くの研究が行われている [2, 3, 4]。

模倣学習の手法の一つとして、中西ら [4] は運動学習プリミティブによって運動学習を行っている。運動学習には、到達運動学習と周期運動学習の二種類を定義しており様々な動作に対して運動学習プリミティブが適応できることを示している。また、学習は線形回帰モデルである LWR (Locally Weighted Regression) を用いるため、高速な学習が可能である。

本研究では周期運動のみを取り扱うため、周期運動学習についてのみ述べる。

## 2.1 周期運動学習

周期運動は、基本的な周期運動パターン生成器を極座標系  $(\phi, r)$  で表した次の位相振動子を考える。

$$\tau \dot{\phi} = 1, \quad \tau \dot{r} = -\mu(r - r_0) \quad (1)$$

ここで、 $(\phi, r)$  はそれぞれ振動子の位相、振幅を表しており、 $\tau$  は時定数、 $r_0$  は目標振幅、 $\mu$  は正の定数である。運動出力  $y$  は位相振動子の状態  $(\phi, r)$  と座標変換した  $\tilde{\mathbf{v}} = [r \cos \phi, r \sin \phi]^T$  を入力とする次の 2 次のダイナミクスで表現される。

$$\tau \dot{\mathbf{y}} = \mathbf{z} + \mathbf{f}(\tilde{\mathbf{v}}, \phi), \quad \tau \dot{\mathbf{z}} = \alpha_z (\beta_z (\mathbf{y}_m - \mathbf{y}) - \mathbf{z}) \quad (2)$$

ここで、 $\alpha_z, \beta_z$  は正の定数、 $\mathbf{y}_m$  は出力角度のオフセットを表している。 $(y, \dot{y})$  はそれぞれ出力の目標関節軌道と速度を表している。また、式 (2) における  $\mathbf{f}(\tilde{\mathbf{v}}, \phi)$  は次式で表される非線形関数近似器である。

$$\mathbf{f}(\tilde{\mathbf{v}}, \phi) = \frac{\sum_{i=1}^N \Psi_i \mathbf{w}_i^T \tilde{\mathbf{v}}}{\sum_{i=1}^N \Psi_i} \quad (3)$$

式 (3) における  $\Psi_i$  は局所モデルのガウシアンカーネルであり、次式のように定義される。

$$\Psi_i = \exp\left(-\frac{1}{2\sigma_i^2}(\text{mod}(\phi, 2\pi) - c_i)^2\right) \quad (4)$$

ここで、 $\text{mod}(\phi, 2\pi)$  は入力  $\phi$  を  $0 \sim 2\pi$  で正規化することを表している。式 (4) によって局所モデルの出力は重み付けされる。式 (3) における  $\mathbf{w}_i$  は局所線形モデルのパラメータベクトルであり、LWR (図 1) [5] によって決定される。

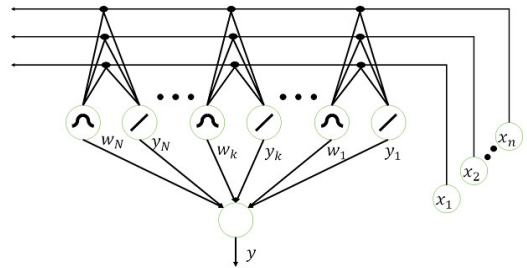


図 1 局所加重回帰 (LWR)

## 3 提案システム

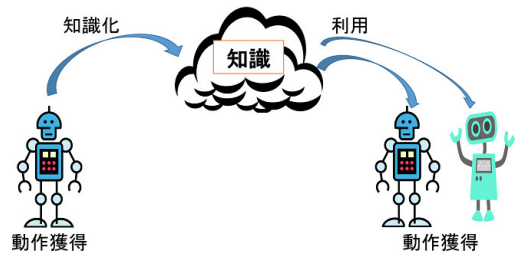


図 2 システム概要

本研究のシステム概要を図 2 に示す。提案するシステムでは、「動作獲得」と「知識化と利用」の 2 つで構成されている。

## 3.1 動作獲得

動作獲得では、周期運動学習を用いて動作を獲得する。また動作は各自由度で動くことはなく、動作に関係している自由度が相互に影響を与えている。そのため本研究では、動作に必要な複数自由度をまとめて学習できるようにモデルを拡張する。

式 (2) と局所回帰パラメータを求める式に改良を加える。

$$\tau \dot{\mathbf{y}} = \mathbf{z} + \mathbf{f}(\tilde{\mathbf{v}}, \phi), \quad \tau \dot{\mathbf{z}} = \alpha_z (\beta_z (\mathbf{y}_m - \mathbf{y}) - \mathbf{z}) \quad (5)$$

このとき出力と出力オフセットが複数自由度に合わせてスカラー  $y, y_m$  からベクトル  $\mathbf{y}, \mathbf{y}_m$  になる。そのため、パラメータ  $z$  と非関数近似器  $\mathbf{f}(\tilde{\mathbf{v}}, \phi)$  もそれぞれベクトル  $\mathbf{z}, \mathbf{f}(\tilde{\mathbf{v}}, \phi)$  となる。また、局所回帰パラメータを求める式は、LWR を基にして周期運動学習に適応させる。これらの拡張により、一度に複数自由度の学習が行えるようになる。また、拡張したことにより動作を 1 つのモデルで表現できるようになり動作を表現する知識として意味を持つようになる。

### 3.2 知識化と利用

ロボットが獲得した動作は当該ロボットのみが使用できる。しかしながら、本研究ではロボット間で共通の認識を持たせることで、他のロボットでも獲得した動作を活用できるようにする。本研究ではヒューマノイドロボットを用いるため、共通の認識として人間を基準とする。ロボットが動作するには、角度情報が必要である。そのため、ロボットの角度情報を人間の角度情報へと変換することで他のロボットでも扱える情報にする。このとき、共通の認識である人間の情報を用いて知識化と利用を行う。

#### 3.2.1 知識化

変換された情報を知識化するために、Left-to-Right 型の隠れマルコフモデル (HMM) を用いる。HMM は時系列データを入力としてパラメータ推定を行い、確率モデルで表現される。そのため動作時系列、つまり動作知識を確率モデルで表現できる。また、パラメータ推定が出来ている HMM に対して時系列データを入力することで、以前獲得した動作との尤度が分かる。そのため、HMM を用いて知識化を行う。

#### 3.2.2 利用

知識となった HMM から動作時系列を取得するために、状態遷移確率と出力確率を用いる。Left-to-Right 型であるため、開始状態と終了状態が決まっている。そのため、状態遷移確率によって開始状態から終了状態までの状態系列を保持し、出力確率によって各状態からの出力を得る。これにより時系列データを取得する。

## 4 計算機シミュレーション

異種のヒューマノイドロボットとしてソフトバンクロボティクス社の Pepper と NAO を用いて、動作獲得と動作の知識化と利用を行う。

### 4.1 動作獲得

片手を振る動作を学習し獲得させる。手を振る動作は周期的な動作であり、ヒューマノイドロボットの個体差も少ないため評価用の動作として用いる。手を振る動作は肩と肘を用いて動作し、実験で使用する Pepper と NAO どちらも同じ 4 自由度 (肩 2, 肘 2) である。そのため、学習に必要なパラメータはどちらも共通の値で設定することができ、 $r_0 = 1, \mu = 1, \alpha_z = 8, \beta_z = 1$  とする。また、局所モデルのガウシアンカーネルのパラメータ  $c_i$  を  $0 \sim 2\pi$  の間で均等に配置し、 $\sigma_i$  を  $2\sigma_i$  と隣接する局所モデルのガウス分布が  $2\sigma_{i+1}$  で重なるように設定した。

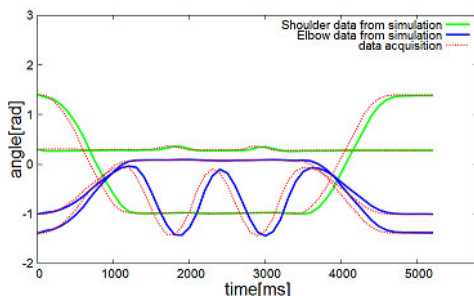


図3 動作獲得の結果 (赤:教師データ, 青と緑:獲得したデータ)

まず、NAO が動作獲得を行う。動作獲得の結果が図3であり、赤の点線が教師データ、青 (肘) と緑 (肩) の実線が獲得した

時系列データである。この結果から動作獲得において運動学習における一括学習が有効であると分かる。

### 4.2 知識化と利用

NAO が獲得した動作を共通の認識である人間の情報に変換する。変換後の時系列データを HMM に入力することでパラメータを推定し、知識化を行う。このとき、HMM の状態数は 28、更新回数は 2 回に設定した。知識を利用するために、パラメータ推定した HMM から時系列データを獲得する。その結果が図4となる。

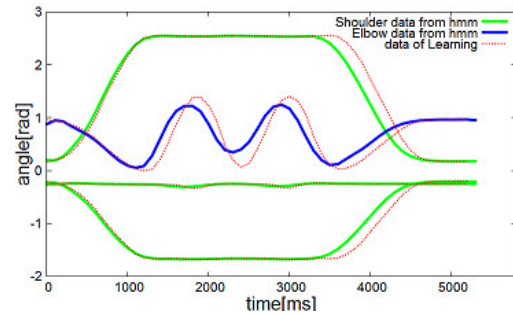


図4 知識化と利用の結果 (赤:教師データ, 青と緑:獲得したデータ)

知識を利用して得た時系列データ (図4: 赤の点線) は、HMM に入力した時系列データ (図4: 青と緑の実線) と最大最小値の誤差は多少見られるが、近い形状をしていることが分かる。そのため HMM のパラメータを知識とし、知識から時系列データが得ることが分かる。Pepper は得られた時系列データを自身の角度情報へと変換することで動作を行う。動作を確認したが、NAO と同様に手を振る動作を行っていた。

## 5 おわりに

本研究では、複数自由度の一括学習による動作獲得、HMM における知識化と利用を行った。計算機シミュレーションによって、ヒューマノイドロボットにおいて動作の獲得と知識化、知識利用が可能であることを確認した [6]。

今後の課題として、HMM によるパラメータ推定がオーバーフロー、アンダーフローにより不安定であるため、安定してパラメータ推定できるようにする必要がある。また、異種ロボットに適用するために共通認識の情報として人間の角度基準値を用いたが、人間の角度基準では汎用性に限界がある。ヒューマノイドロボットの自由度は人間の自由度と同じ数とは限らなく、人間の自由度数では対応できないことがある。そのため、より汎用性の高い情報を共通認識として扱う必要がある。

## 参考文献

- [1] 熊谷, 小林: "クラウドネットワークロボットにおける RNN を用いた知識の獲得と利用", 電気学会研究会資料, ST-14-027, pp.7-12, (2014)
- [2] 宇田川, 内田, 堀内: "模倣学習を用いた二脚駆動ロボットの行動獲得", ファジィシステムシンポジウム講演論文集, WG1-2, pp.1312-1315, (2011)
- [3] 稲邑, 戸嶋, 江崎, 中村: "ミメシス理論に基づく見まね学習とシンボル創発の統合モデル", 日本ロボット学会誌, vol.22, No.2, pp.256~263(2004)
- [4] 中西, Auke, Stefan, Gordon: "運動学習プリミティブを用いたロボットの見まね学習", 日本ロボット学会誌, Vol.22, No.2, pp.165-170, (2004)
- [5] S.Schaal, C.G.Atkeson: "Constructive Incremental Learning From Only Local Information", Neural Computation, 10, 8, pp.2047-2084, (1998)
- [6] 熊谷, 鈴木, 小林: "異種のヒューマノイドロボットにおける動作の知識化と利用", 第 17 回システムインテグレーション部門講演会 (SI2016), 資料, pp.0223-0228, (2016)