

参照点に依存したトラジェクトリ HMM による動作生成

杉浦孔明, 岩橋直人, 柏岡秀紀 (情報通信研究機構)

1. はじめに

生活支援ロボットにとって、日用品等のオブジェクトを操作する機能は必要不可欠であるが、各種の日用品や棚に対応する動作を事前にプログラムするコストは非常に大きい。加えて、事前にプログラムされた動作がユーザにとってイメージしにくいものであった場合、安心して動作指示できないという問題もある。本研究は、これらの問題を解決するため、模倣学習の枠組みにより動作を学習する手法の構築を目的とする。このような学習手法を構築することで、プログラミングスキルが必要とされないユーザフレンドリーな動作教示方法を実現できる。

物体操作の模倣学習における主な問題のひとつは、観測された軌道そのものはモデルとしての利用性が低いということである [3]。これは同じ「載せる」動作であっても、オブジェクトの配置の違いによって軌道が全く異なるためである。この問題に取り組んだ先行研究では、物体操作を 2 つのオブジェクト間の相対軌道として表現するものが多い (例えば [5, 6])。これらの手法においては、2 つのオブジェクトが与えられているという前提があるため、関連物体をユーザが指定しなくてはならないという制約があった。この制約に対し、我々は、関連物体の推定と軌道学習を同時に行う方法を提案している [7, 8]。

一方、物体操作に限らず、模倣学習における軌道のモデル化には、隠れマルコフモデル (HMM) が用いられることが多い。ただし、HMM を用いて単純なサンプリングにより軌道を生成すると、生成された軌道は不連続なものとなる。これに対し、本論文では、トラジェクトリ HMM [9] を利用して連続的な軌道を生成する手法について述べる。提案手法の独自性は、以下の 2 点である。

- 学習時に動作の表現に適切な座標系 (固有座標系) を推定する。軌道生成時に固有座標系上の HMM を変換する。
- トラジェクトリ HMM を用いて連続的な物体操作軌道を生成する

2. タスク

2.1 参照点に依存した動作

空間的移動の概念には、参照点に依存しているものがある。「A を B に載せる」動作や「A を回す」動作がこれにあたる。いま、参照点に依存する動作の概念を、模倣学習の枠組みによりロボットに学習させる問題を考える。このとき、動かされるオブジェクトをトラジェクタ、トラジェクタの基準になるオブジェクトをランドマークと呼ぶことにする [4]。上記の 2 種類の動作のように、ランドマークを必要とする動作と必要としない動作を同じ枠組みで学習させることは簡単ではない。

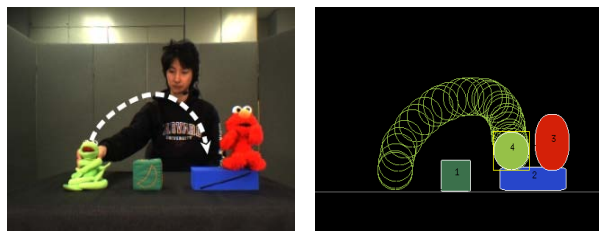


図 1 左:カメラ画像の例。右:動画画像から抽出された観測情報。



図 2 実験に用いたロボット

このような情報は、通常訓練データ中に明示されていないためである。

2.2 ハードウェア

実験に用いたロボットシステムを図 2 に示す。ロボットシステムは、7 自由度のロボットアーム (三菱重工製 PA-10)、4 自由度のロボットハンド (Barrett Technology 製 BarrettHand)、マイクロフォン、ステレオカメラ (Point Grey Research 製 Bumblebee 2)、視線表出ユニット (Directed Perception 製 PTU-46-70 にロボットヘッドを取付) からなる。

オブジェクトの画像特徴および 2 次元座標 (カメラ座標系) は、固定されたステレオカメラを用いて得る。なお、オブジェクトの抽出およびトラッキングは、色およびステレオカメラから得られる距離に基づくヒューリスティックな手法により行う。カメラのフレームレートは 30[frame/sec] であり、解像度は 320×240 である。図 1 にカメラ画像の例とその内部表現を示す。新規動作の学習や動作認識を行う際には、オブジェクトの座標時系列が利用される。

3. 参照点に依存した HMM による動作学習

3.1 デルタパラメータを用いた軌道のモデル化

移動するオブジェクトが一つで複数の静止オブジェクトが存在する動画画像が与えられたとする。この動画画像から、移動する物体 (トラジェクタ) の重心軌道 O_t と、静止オブジェクトの重心位置の集合 O_s を抽出し、学

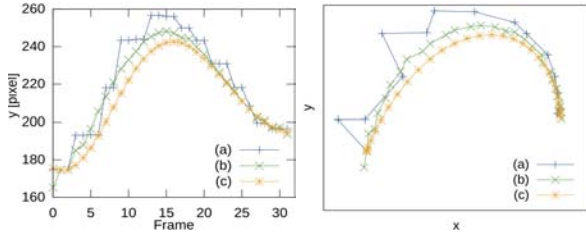


図3 デルタパラメータを用いた軌道生成の例。(a)位置のみ、(b)位置および1次のデルタパラメータ、(c)位置および1,2次のデルタパラメータを用いた軌道生成の結果。

習時の訓練サンプル \mathcal{V} とする。

$$\mathcal{V} = (\Xi, O_S) \quad (1)$$

$$\Xi = [\xi_1^\top, \xi_2^\top, \dots, \xi_T^\top]^\top \quad (2)$$

$$\xi_t = [x_t^\top, \Delta^{(1)}x_t^\top, \Delta^{(2)}x_t^\top]^\top \quad (3)$$

ここに、 x_t は時刻 t におけるトラジェクタの位置ベクトルであり、 x_t の次元を m とする。本研究では、 $m=2$ を仮定するものの、本手法は3次元空間($m=3$)にも適用可能である。 $\Delta^{(k)}x_t$ ($k=0,1,2$)は、音声処理分野ではデルタパラメータと呼ばれ、HMMから軌道を生成する際に用いられている。デルタパラメータは次のように定義される。

$$\Delta^{(k)}x_t = \sum_{\tau=-L^{(k)}}^{L^{(k)}} w^{(k)}(\tau)x_{t+\tau} \quad (k=0,1,2) \quad (4)$$

ここに、 $L^{(0)}=0$ 、 $w^{(0)}(0)=1$ 、 $L^{(1)}=L^{(2)}=1$ とする。また、 $w^{(1)}(\tau)=\frac{\tau}{2}$ 、 $w^{(2)}(\tau)=\frac{3}{4}\tau^2 - \frac{1}{2}$ と定義する。よって、

$$\Delta^{(1)}x_t = -\frac{1}{2}x_{t-1} + \frac{1}{2}x_{t+1} \quad (5)$$

$$\Delta^{(2)}x_t = \frac{1}{4}x_{t-1} - \frac{1}{2}x_t + \frac{1}{4}x_{t+1} \quad (6)$$

となる。便宜的に、位置を0次のデルタパラメータと呼ぶ。定義より、1次および2次のデルタパラメータは速度および加速度の定数倍とみなすことができる。

デルタパラメータによる軌道生成の効果を図3に示す。図3は、動作「載せる」に関して、軌道生成を行った結果である。図において、左図にt-yプロット、右図にx-yプロットを示す。図より、位置のみを用いた場合には不連続な軌道が生成されているのに対し、1,2次のデルタパラメータを用いることで滑らかな軌道が生成できることがわかる。

HMMとして、状態数 N のleft-to-right型HMMを用いる。以下では、HMMの全パラメータを λ と表すこととする。 λ は、初期状態分布 $\pi = \{\pi_i\}$ 、状態遷移確率 A 、出力確率分布 $b = \{b_i(\cdot)\}$ からなる。出力確率分布として、混合数1のガウス分布を用いる。

3.2 固有座標系タイプ

本研究では、固有座標系は有限個のタイプに分類できるものとする。固有座標系タイプを k とし、その総

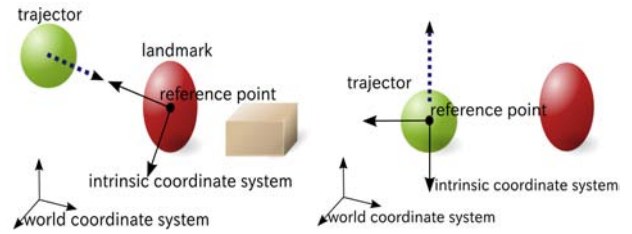


図4 動作と参照点・固有座標系の関係。左:「近づける」、右:「上げる」

数を K とする。固有座標系として、以下の C_1 から C_4 を定義する。

- C_1 : ランドマークを原点とする、カメラ座標系を平行移動した座標系。ただし、変換後の座標系において x_1 の x 座標が負になる場合には、 x 軸を反転させる。 y 軸は重力方向とする。動作の例としては、「載せる」などが挙げられる。
- C_2 : ランドマークを原点とし、 x_1 に向かう軸を x 軸とする直交座標系(図4左図参照)。動作の例としては、「近づける」などが挙げられる。
- C_3 : x_1 を原点とする、カメラ座標系を平行移動した座標系(図4右図参照)。動作の例としては、「上げる」などが挙げられる。
- C_4 : 画面中心 x_{center} を原点とする、カメラ座標系を平行移動した座標系。固有座標系に関係ない動作、例えばアルファベットを描くなどが C_4 に分類される。

C_3, C_4 では参照点が唯一に定まるので、参照点の探索を行なわない。参照点の候補集合 O_R を以下のように定義する。

$$O_R = \{O_S, x_1, x_{center}\} \triangleq \{x^r | r = 1, 2, \dots, |O_R|\}$$

3.3 参照点に依存したHMMの学習

固有座標系タイプ k と参照点 r によって決定される固有座標系 $C_k(x^r)$ 上での軌道を ${}^{C_k(x^r)}\Xi$ と表記することにする。以下の尤度最大化基準により、固有座標系タイプ k 、参照点インデックス列 r 、HMMパラメータ λ を探索する。

$$(\hat{\lambda}, \hat{k}, \hat{r}) = \operatorname{argmax}_{\lambda, k, r} \sum_{l=1}^L \log P({}^{C_k(x^r)}\Xi_l; {}^{C_k(x^r)}\lambda) \quad (7)$$

ここに Ξ_l は l 番目の訓練サンプルにおける軌道を表す。式(7)の解は[8]を参照されたい。

軌道生成時に用いる状態系列において、状態継続長は任意のものを設定可能である。ここでは合理的な状態継続長として、訓練集合における平均の状態継続長を用いる。訓練集合中の平均状態系列 \bar{q} は以下で表される。

$$\begin{aligned} \bar{q} &= (q_{\text{init}}, q_1, q_2, \dots, q_{\bar{T}}, q_{\text{final}}) \quad (8) \\ &= (S_{\text{init}}, \underbrace{1, \dots, 1}_{\bar{d}_1}, \underbrace{2, \dots, 2}_{\bar{d}_2}, \dots, \underbrace{N, \dots, N}_{\bar{d}_N}, S_{\text{final}}) \end{aligned}$$

ここに \bar{d}_i は状態 i の平均継続長を表し、 $\bar{T} = \sum_i \bar{d}_i$ である。

4. 参照点に依存したトラジェクトリ HMM による動作生成

学習された HMM から軌道を生成することを考える．状態系列 \bar{q} と HMM パラメータ ${}^W\lambda$ が与えられたうえで，生成軌道の確率を最大化する最適軌道 ${}^W\hat{\Xi}$ を得ることを考える．世界座標系 W における最適軌道 ${}^W\hat{\Xi}$ は以下により与えられる．

$${}^W\hat{\Xi} = \operatorname{argmax}_{W\Xi} P(W\Xi | \bar{q}, {}^W\lambda) \quad (9)$$

$$= \operatorname{argmax}_{W\Xi} P(W\Xi | \bar{q}, {}^C\lambda, \mathbf{x}^{r\text{traj}}, \mathbf{x}^r) \quad (10)$$

ここに，固有座標系を C ，トラジェクタ初期位置を $\mathbf{x}^{r\text{traj}}$ ，参照点を \mathbf{x}^r とする．

HMM パラメータ λ は固有座標系上で学習されているため，軌道生成時には座標変換を行う必要がある．まず，状態 s における $k(=0, 1, 2)$ 次のデルタパラメータの平均ベクトル ${}^C\boldsymbol{\mu}^{(k)}(s)$ と共分散行列 ${}^C\Sigma^{(k)}(s)$ に関して， C から W への座標変換を行う．位置に関する平均ベクトル ${}^C\boldsymbol{\mu}^{(0)}(s)$ は，以下の同次変換行列により変換される．

$$\begin{bmatrix} {}^W\boldsymbol{\mu}^{(0)}(s) \\ 1 \end{bmatrix} = \begin{bmatrix} R & \mathbf{x}^{r\text{traj}} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} {}^C\boldsymbol{\mu}^{(0)}(s) - {}^C\boldsymbol{\mu}^{(0)}(1) \\ 1 \end{bmatrix} \quad (s = 1, 2, \dots, N)$$

ここに， R は回転行列を表す．同様に， R を用いて位置以外の平均ベクトルと共分散ベクトルを変換する．

$${}^W\boldsymbol{\mu}^{(k)}(s) = R {}^C\boldsymbol{\mu}^{(k)}(s) \quad (k = 1, 2) \quad (11)$$

$${}^W\Sigma^{(k)}(s) = R {}^C\Sigma^{(k)}(s) R^T \quad (k = 0, 1, 2) \quad (12)$$

最後に，トラジェクタ初期位置 $\mathbf{x}^{r\text{traj}}$ を起点とする軌道を生成するために，正の数 $\varepsilon \ll 1$ を，第 1 状態の位置の分散に設定する．

$${}^W\Sigma^{(0)}(1) = \varepsilon I \quad (13)$$

以上より， ${}^C\lambda$ から ${}^W\lambda$ が得られた．式 (9) の解を付録に示す．

5. 実験設定

本実験の目的は，生成軌道の学習が収束するために必要な訓練サンプル数を示すことである．

5.1 訓練データおよびテストデータ

動作ラベルを自然言語で被験者に呈示し，机の上に置かれたオブジェクトを操作させた．用いられた動作ラベルは，以下の 7 種類である．

「上げる」、「近づける」、「離す」、「回す」、「載せる」、「下げる」、「飛び越えさせる」

表 1 に訓練集合名とサンプル数 L を示す．TA-3 から TA-9 は，TA-2 に訓練サンプルを一つずつ追加したものである．学習時に用いられた配置が，動作生成時に用いられないようにするため，訓練集合とは異なる別のテスト集合 RA-5 を用意した．RA-5 は，TA- n と同様に 7 種類の動作からなり，各動作に対し軌道を 5 個含む．

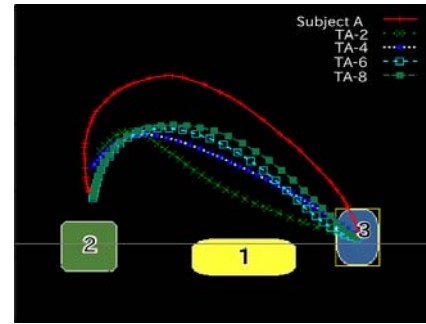


図 5 訓練サンプル数と生成された軌道の関係

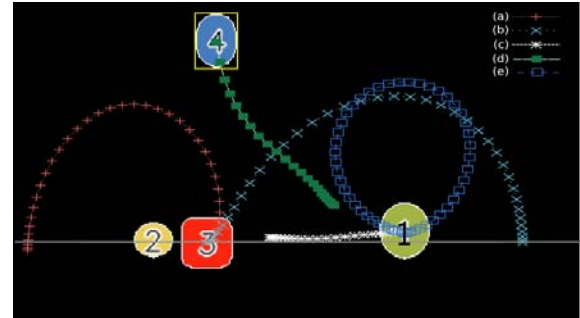


図 6 生成軌道の例

6. 結果

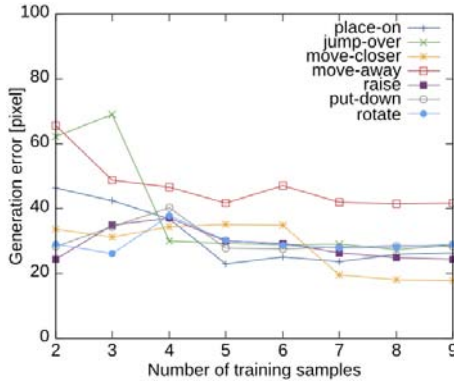
6.1 定性的結果

図 5 に，訓練サンプル数と生成された軌道の関係を示す．図における軌道は，オブジェクト 3 をオブジェクト 2 に「載せる」を軌道を示す．図において，Subject A は，同じ物体配置において被験者 A が実行した軌道，TA- n は n 個のサンプルを用いて学習された HMM から生成された軌道を示す．図より，TA-2 の最終位置は被験者のものと離れているが，TA-8 の最終位置は被験者のものとほぼ等しいことがわかる．このことから，TA-8 の最終状態の分散は適切に学習されていたと考えられる．実際に，TA-2，TA-4，TA-6，TA-8 の最終状態の分散(縦方向)は，それぞれ 26.6，23.5，5.41，7.44 であった．

図 6 に，様々な物体配置における動作生成の結果を示す．各動作の[動作ラベル，トラジェクタ ID，ランドマーク ID]はそれぞれ，(a) [飛び越えさせる，3，2]，(b) [飛び越えさせる，3，1]，(c) [近づける，4，1]，(d) [近づける，1，3]，(e) [回す，1] である．図より，関連物体の組み合わせに応じて，適切な軌道が生成されていることがわかる．例えば，軌道 (a) と (b) は同じ HMM から生成されているものの，配置に応じて逆向きの軌道が生成されている．

表 1 訓練集合およびテスト集合

Data set	Subject	# of motion labels	# of trajectories per motion label
TA-2	A	7	2
TA-3	A	7	3
⋮	⋮	⋮	⋮
TA-9	A	7	9
RA-5	A	7	5

図7 訓練サンプル数と生成誤差 $D(\Xi, \hat{\Xi})$ の関係

6.2 定量的結果

提案手法により生成された軌道 $\hat{\Xi}$ を、被験者が実行した軌道 Ξ と比較し、生成誤差を評価する。評価尺度として、時系列の比較のために一般的に用いられているユークリッド距離を用いる [2]。生成誤差 $D(\Xi, \hat{\Xi})$ は、フレーム長 T で正規化されたユークリッド距離で与えられる。

$$D(\Xi, \hat{\Xi}) = \frac{1}{T} \sum_{t=1}^T \sqrt{|x_t - \hat{x}_t|^2} \quad (14)$$

ここに、 x_t および \hat{x}_t は、それぞれ Ξ と $\hat{\Xi}$ のフレーム t における位置ベクトルを表す。ただし、 \hat{x}_t は、 x_t と同数のフレームを含むようにリサンプリングを行うものとする。

図7に定量的結果を示す。図は、 $D(\Xi, \hat{\Xi})$ を訓練サンプル数で平均化したものをプロットしたものである。図の縦軸は生成誤差、横軸は訓練サンプル数を表す。図より、サンプル数が7以上の場合において、生成誤差が収束していることがわかる。このことから、提案手法による軌道学習のために、7個程度のサンプルが必要であることがわかる。

7. おわりに

本論文では、「載せる」「回す」などの参照点に依存した動作を模倣学習する手法を述べた。提案手法は、デルタパラメータを利用して連続的な軌道を生成することができるという特徴を持つ。提案手法の応用タスクとしては、音声対話による物体操作 [10] や、RoboCup@Home [1] における日用品ハンドリングの模倣学習、などが挙げられる。

A トラジェクトリ HMM を用いた軌道生成

本付録では、[9] で提案された軌道生成について説明する。以下では、簡単のため、座標系の表記を省略する。つまり、 $W\lambda$ を λ と表記することにする。

HMM λ の出力確率分布はガウス分布であることから、式 (9) を変形すると、

$$\hat{\Xi} = \underset{\Xi}{\operatorname{argmax}} \log P(\Xi|q, \lambda) \quad (15)$$

$$= \underset{\Xi}{\operatorname{argmax}} \left\{ -\frac{1}{2} \Xi^\top \Sigma^{-1} \Xi + \Xi^\top \Sigma^{-1} \mu + K \right\} \quad (16)$$

ここに、

$$\Sigma^{-1} = \operatorname{diag} \left[\Sigma_{q_{L'+1}}^{-1}, \Sigma_{q_{L'+2}}^{-1}, \dots, \Sigma_{q_{T-L'}}^{-1} \right] \quad (17)$$

$$\mu = \left[\mu_{q_{L'+1}}^\top, \mu_{q_{L'+2}}^\top, \dots, \mu_{q_{T-L'}}^\top \right]^\top \quad (18)$$

$$L' = \max_k L^{(k)} \quad (19)$$

μ_{q_i} は $3m \times 1$ の平均ベクトル、 Σ_{q_i} は $3m \times 3m$ の共分散行列、 K は定数である。 q は状態系列であり、 $q = \bar{q}$ とする。

式 (4) を行列表現すると、

$$\Xi = Wx \quad (20)$$

となる。ここに、

$$x = [x_1^\top, x_2^\top, \dots, x_T^\top]^\top$$

$$W = [w_{L'+1}, w_{L'+2}, \dots, w_{T-L'}]^\top$$

$$w_t = [w_t^{(0)}, w_t^{(1)}, w_t^{(2)}]^\top$$

$$w_t^{(k)} = \left[\mathbf{0}_{m \times (t-L^{(k)}-1)m}, w^{(k)}(-L^{(k)}) I_{m \times m}, \dots, w^{(k)}(0) I_{m \times m}, \dots, w^{(k)}(L^{(k)}) I_{m \times m}, \mathbf{0}_{m \times (T-t-L^{(k)})m} \right]^\top$$

制約 (20) の下で、式 (15) を解くために、以下を考える。

$$\frac{\partial \log P(Wx|q, \lambda)}{\partial x} = 0 \quad (21)$$

式 (16)、(20)、(21) より、

$$W^\top \Sigma^{-1} Wx = W^\top \Sigma^{-1} \mu \quad (22)$$

を得る。式 (22) は Cholesky 分解を用いて、 $O(Tm^3)$ で解ける。

参考文献

- [1] <http://www.ai.rug.nl/robocupathome/>: RoboCup@Home Rules & Regulations.
- [2] Keogh, E. and Kasetty, S.: On the need for time series data mining benchmarks: A survey and empirical demonstration, *Data Mining and Knowledge Discovery*, Vol. 7, No. 4, pp. 349–371 (2003).
- [3] Krüger, V., Kragic, D., Ude, A. and Geib, C.: The meaning of action: a review on action recognition and mapping, *Advanced Robotics*, Vol. 21, No. 13, pp. 1473–1501 (2007).
- [4] Langacker, R. W.: *Foundations of Cognitive Grammar: Theoretical Prerequisites*, Stanford Univ Pr (1987).
- [5] Ogawara, K., Takamatsu, J., Kimura, H. and Ikeuchi, K.: Modeling manipulation interactions by hidden Markov models, *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1096–1101 (2002).
- [6] Regier, T.: *The Human Semantic Potential: Spatial Language and Constrained Connectionism*, Bradford Books (1996).
- [7] Sugiura, K. and Iwahashi, N.: Learning object-manipulation verbs for human-robot communication, *Proceedings of the 2007 workshop on Multimodal interfaces in semantic interaction*, pp. 32–38 (2007).
- [8] Sugiura, K., Iwahashi, N., Kashioka, H. and Nakamura, S.: Learning, Generation, and Recognition of Motions by Reference-Point-Dependent Probabilistic Models, *Advanced Robotics*, Vol. 25, No. 6-7, pp. 825–848 (2011).
- [9] Tokuda, K., Yoshimura, T., Masuko, T., Kobayashi, T. and Kitamura, T.: Speech parameter generation algorithms for HMM-based speech synthesis, *Proceedings of ICASSP*, pp. 1315–1318 (2000).
- [10] 杉浦孔明, 岩橋直人, 柏岡秀紀, 中村哲: 言語獲得ロボットによる発話理解確率の推定に基づく物体操作対話, *日本ロボット学会誌*, Vol. 28, No. 8, pp. 978–988 (2010).