「『不確実性に挑むロボティクス』特集号」　　解　説

模倣学習における確率ロボティクスの新展開

杉浦　孔明*

1.　はじめに

模倣学習研究は，ロボティクスおよび人工知能分野における中心課題のひとつであり，機械学習，ヒューマンロボットインタラクション，画像処理，動作計画などにおける基盤技術と関連が深い [1]．模倣学習は単に「人間が見せた動作をロボットが模倣する」問題を指すだけでなく，後述するようにロボット以外の入出力デバイス（CG など）を扱う研究や，人間の動作解析に比重を置いた研究を含む．このように模倣学習研究は広い意味を持ち得るので，本稿では

- 教示入力の一部に軌道を含み，その入力から確率モデルの学習を行うとともに，動作実行時には学習済みの確率モデルから状況に応じた軌道を生成する問題

に限定する．紙面の都合上割愛するが，模倣学習の対象タスクとしては，動作レベルの模倣以外にも，行動の連鎖の学習や目的レベルの模倣などがある．

模倣学習の問題を大きく分類すると，動作検出，動作認識，動作生成の側面が存在する．ただし，動作検出および動作認識については含まれない問題設定も存在する．多次元の時系列を扱う点で，模倣学習と音声処理研究が相似形であることは非常に興味深い．音声処理分野では，上記の三者は発話区間検出，狭義の音声認識，音声合成，に対応しており，極めて類似した手法を適用可能である．一方，状況に応じた座標変換や，外乱に応じた逐次動作生成などは模倣学習特有の問題設定である．以降では，動作レベルの模倣学習を単に模倣学習と呼ぶこととする．また，動作認識・動作生成についても，混乱がなければ認識・生成と呼称する．

模倣学習のメリットとしては，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な動作教示方法を実現できることが挙げられる．例えば，「食器棚からコップを取り出す」動作をユーザの教示から学習し，他の状況において「グラスを取り出す」等の動作をロボットに実行させるような状況が想定される．第 1 図に物体操作の模倣学習を行うロボットの例を示す．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．本稿では，このような課題に対する我々の取り組みを紹介する．

第 1 図　学習した「載せる」動作を実行するロボットの例

模倣学習研究のもうひとつの側面としては，人間の行動の科学的理解がある．この研究目的のために認識と生成に同じ確率モデルを用いることが多い点は，模倣学習研究を特徴付けるひとつの観点として興味深い．

本稿の構成は以下の通りである．まず，2 章で，模倣学習と確率ロボティクスに関連する文献を中心に，模倣学習研究の関連研究を概観する．3 章では，模倣学習で一般的に用いられる確率モデルである Hidden Markov Model (HMM) の基本要素について述べ，4 章では HMM に基づく模倣学習手法について述べる．5 章では物体操作の模倣学習に焦点を当て，我々が提案した Reference-Point-Dependent HMM (RPD-HMM) の適用について述べる．最後に，6 章で結論と今後の展望について述べる．

2.　模倣学習と確率ロボティクス

本章では，模倣学習と確率ロボティクスに関連する文献を中心に，広い視点から模倣学習研究の関連研究を紹介する．なお，模倣学習は “Imitation Learning” の訳であり，“Learning from Demonstration (LfD)”，“Programming by Demonstration (PbD)” などの語もほぼ同じ意味で使われることが多いので検索の際には注意されたい．

模倣学習分野では，確率モデル，強化学習，リカレントニューラルネット，Dynamic Movement Primitives (DMP) によるアプローチが代表的である．本稿では特に確率モデルによる模倣学習に焦点を当て，それ以外のアプローチに関しては本段落に解説論文と代表的文献を

* 国立研究開発法人　情報通信研究機構
Key Words: imitation learning, HMM, reference point, trajectory

挙げるにとどめる．模倣学習分野における強化学習の適用に関しては, [2] が詳しい．計算論的神経科学における模倣学習の代表的解説論文としては [3] がある．初期の模倣学習研究は計算論的神経科学と関連が深く，組み立てタスク [4] や剣玉タスク [5] が対象であった．ニューラルネットに基づく模倣学習手法として代表的な手法としては，Recurrent Neural Network with Parametric Bias (RNNPB) が挙げられる [6]．また，認知科学の面からの模倣学習の用語定義および解説としては, [7] が挙げられる．DMP による模倣学習手法の代表的事例としては, [8,9] などがある．確率モデルに基づく模倣学習手法のチュートリアルとしては [10] が詳しい．画像処理，ロボティクス，人工知能における模倣学習に適用可能な機械学習手法の解説としては [11] がある．

確率モデルによるアプローチを分類すると，HMM などの Markov Switching Process[12]，Gaussian Mixture Regression (GMR)[13]，ガウス過程 [14] を用いるものが多い．HMM を用いた模倣学習手法については，次章以降で述べる．GMR を用いる手法では，関節角軌道のモデル化のために，姿勢（位置）だけでなく時間も特徴量としてガウス分布を学習させるアプローチが多い [13,15]．ガウス過程の適用例としては，モーションキャプチャデータを扱った事例 [16] や，物体操作 [17] を扱った事例がある．

模倣学習研究では，入力デバイスとして，カメラ，モーションキャプチャデバイス，力覚センサ，教示用のコントローラ，などが用いられることが多い．これらのデバイスから得られる情報として，関節角の位置・姿勢，オブジェクトの位置・姿勢，力覚センサ情報，教示入力，などがある．出力デバイスまたは対象としては，ヒューマノイド，ロボットアーム，CG エージェント，などが考えられる．また，模倣学習の評価尺度としては，動作認識の側面では認識精度が一般的に用いられる．動作生成の側面では，軌道の Root Mean Squared Error (RMSE)，軌道の自然性に対する評価（例：Mean Opinion Score)，強化学習における報酬，などがあり得る．

3. HMM による軌道のモデル化

前章で述べたように，確率モデルを用いる模倣学習手法のなかで，HMM には学習・認識・生成に高速なアルゴリズムが存在する，というメリットがある．模倣学習分野において HMM を利用した初期の事例としては，Ogawara らによる物体操作への適用 [18] や，ヒューマノイドの全身動作への適用 [19] などがある．HMM の適用事例としては，全身運動に対する Factorial HMM の適用 [20]，物体操作に対する RPD-HMM の適用 [21]，動作の教師なし分節化に対する Sticky HDP-HMM の適用 [22] などがある．

HMM を用いた時系列のモデル化は，音声認識分野において 1980 年代から活発に研究が行われている．特徴量としては Mel Frequency Cepstral Coefficient (MFCC) が用いられ，その後，デルタパラメータと呼ばれる動的特徴量（動作時系列においては速度にあたると考えてよい）を加えることで精度が上がることが報告された [23]．また，動的特徴量を用いた HMM（以下，HMM-d と略記）による音声合成手法が Tokuda らによって提案された [24]．HMM-d は，主に音声合成において広く適用されてきたが，模倣学習を含む多次元の時系列生成にも適用可能である．動作への適用例は，歩行や物体操作などから開始された [25,21]．HMM に基づく音声合成のソフトウェアとしては HTS が広く利用されている [26]．

HMM は，非観測の（内部）状態がマルコフ的に遷移し，状態に応じて定まる出力確率分布から観測値が生成されるものと考えるモデルである [27]．基本的な HMM は，初期状態分布，状態遷移確率，出力確率密度分布で表現される．これらの全パラメータをまとめて λ と表す．HMM を状態遷移の面から分類すると，エルゴディック HMM と left-to-right HMM に大別できる．left-to-right HMM では，状態は一方向的に遷移するため，次の状態に遷移すると前の状態には戻ることができない．left-to-right HMM は，主にターゲットに向かうような軌道を表現しやすく，高速な学習手法が存在するなど音声認識・合成分野において十分に実績がある．一方，エルゴディック HMM では，ある状態から任

第 1 表　本稿で用いる主な記号表記

次元数（およびそのインデックス）	
$T(t)$	軌道の総フレーム数（軌道ごとに異なる）
$N(i)$	各動作の訓練集合サイズ
$K(k)$	固有座標系タイプの数
$M(j)$	HMM の状態数
$R(r)$	参照点候補の数
ベクトルおよび行列	
$\boldsymbol{x}$	直交座標系における位置
$\boldsymbol{\xi}$	任意の多次元ベクトル（例:$\boldsymbol{\xi}=\left[\boldsymbol{x}^\top,\dot{\boldsymbol{x}}^\top\right]^\top$）
$\dot{\boldsymbol{\xi}},\ddot{\boldsymbol{\xi}}$	$\boldsymbol{\xi}$ の速度，加速度
$\hat{\boldsymbol{\xi}}$	$\boldsymbol{\xi}$ の推定値
Ξ	$\left[\boldsymbol{\xi}_1^\top,\boldsymbol{\xi}_2^\top,...,\boldsymbol{\xi}_T^\top\right]^\top$ で表される軌道
RPD-HMM	
$\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$	出力確率分布（Gaussian）
$\boldsymbol{\mu},\boldsymbol{\Sigma}$	上記の平均ベクトルと共分散行列
λ	初期状態分布・出力確率分布・遷移確率からなる HMM の全パラメータ
$\boldsymbol{S_j}$	HMM の状態
$C_k(\boldsymbol{x})$	参照点 $\boldsymbol{x}$ に依存する座標系（タイプ k）
$\boldsymbol{O}$	参照点候補集合 $\{\boldsymbol{x}^r\}$

意の状態に遷移し得る．エルゴディック HMM は，他の手法においてリミットサイクルで表現されるような軌道や，(軌道より上位レベルの）行動間の遷移を表現しやすい．本稿では，主にターゲットに向かう軌道を扱うため，left-to-right HMM を用いることとする．

出力確率密度分布としては，混合ガウス分布が用いられることが多い．音声認識分野においては，混合数が非常に大きい混合ガウス分布が用いられることが一般的であるが，音声合成分野においてはパラメータ数削減のため，混合数を 1 とすることが多い．また，共分散行列についても，パラメータ数削減のため対角行列を用いることが多い．このように，模倣学習分野と比べ大きいデータセットを用いる音声合成においても，単純なガウス分布や対角共分散行列の使用は一般的である．よって，模倣学習分野において，同様のパラメータ削減を行うことに著しい不合理性はない．

4. HMM-d による模倣学習

4.1 軌道の学習

単純な HMM に基づく模倣学習手法と異なり，HMM-d は最尤軌道の生成が保証できるというメリットを持つ．最尤軌道の生成結果について，単純な HMM（Simple HMM）と HMM-d を比較した結果を第 2 図に示す．第 2 図は，動作「載せる」に関して，軌道生成を行った結果である．図より，Simple HMM では不連続な軌道が生成されているのに対し，HMM-d 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-d では，学習対象の軌道を位置・速度・加速度を用いてモデル化する．フレーム t における位置 $\boldsymbol{x}_t$ に対して，速度と加速度を以下のように定義する．第 1 表に本稿で用いる記号表記を示す．

$$\dot{\boldsymbol{x}}_t = -\frac{1}{2}\boldsymbol{x}_{t-1} + \frac{1}{2}\boldsymbol{x}_{t+1} \tag{1}$$

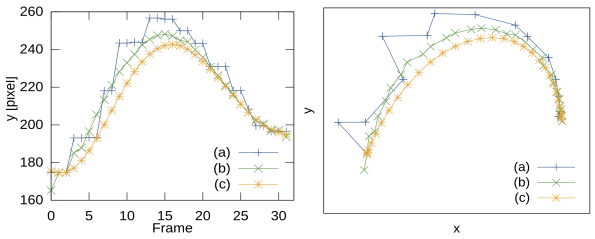
$$\ddot{\boldsymbol{x}}_t = \frac{1}{4}\boldsymbol{x}_{t-1} - \frac{1}{2}\boldsymbol{x}_t + \frac{1}{4}\boldsymbol{x}_{t+1} \tag{2}$$

上記は 2 次精度の中心差分近似として係数を決定した場合である．2 次以外の精度の差分近似式の係数や，Jerk を導入し 3 階微分の差分近似係数を用いてもよい．また，本稿では $\boldsymbol{x}_t$ を 2 次元と仮定するが，3 次元空間にも容易に拡張可能である．差分近似係数の行列を用いて，軌道に含まれるデータ点を次のように表現する．

$$\boldsymbol{\xi}_t = \begin{bmatrix} \boldsymbol{x}_t \\ \dot{\boldsymbol{x}}_t \\ \ddot{\boldsymbol{x}}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} \\ -\frac{1}{2}\boldsymbol{I} & \boldsymbol{0} & \frac{1}{2}\boldsymbol{I} \\ \frac{1}{4}\boldsymbol{I} & -\frac{1}{2}\boldsymbol{I} & \frac{1}{4}\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{t-1} \\ \boldsymbol{x}_t \\ \boldsymbol{x}_{t+1} \end{bmatrix} \tag{3}$$

ここに，$\boldsymbol{I}$ および $\boldsymbol{0}$ はそれぞれ，2×2 の単位行列および零行列である．

上記を用いて，ある動作の軌道を以下の 2 通りで表現する．



第 2 図　Simple HMM と HMM-d の最尤軌道の比較．左図：t-y プロット．右図：x-y プロット．(a)Simple HMM（位置のみ），(b)HMM-d（位置・速度），(c)HMM-d（位置・速度・加速度）．

$$\boldsymbol{x} = \left[\boldsymbol{x}_1^\top, \boldsymbol{x}_2^\top, ..., \boldsymbol{x}_T^\top\right]^\top \tag{4}$$

$$\Xi = \left[\boldsymbol{\xi}_1^\top, \boldsymbol{\xi}_2^\top, ..., \boldsymbol{\xi}_T^\top\right]^\top := \boldsymbol{\Phi}\boldsymbol{x} \tag{5}$$

ここに，$\boldsymbol{\Phi}$ は差分近似係数の行列を並べた行列であり，具体的には以下の右辺第 1 項である．

$$\begin{bmatrix} \vdots \\ \boldsymbol{\xi}_t \\ \boldsymbol{\xi}_{t+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} & \vdots & \vdots & \vdots & \\ \cdots & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \cdots \\ \cdots & -\frac{1}{2}\boldsymbol{I} & \boldsymbol{0} & \frac{1}{2}\boldsymbol{I} & \cdots \\ \cdots & \frac{1}{4}\boldsymbol{I} & -\frac{1}{2}\boldsymbol{I} & \frac{1}{4}\boldsymbol{I} & \cdots \\ & \cdots & \boldsymbol{0} & \boldsymbol{I} & \boldsymbol{0} & \cdots \\ & \cdots & -\frac{1}{2}\boldsymbol{I} & \boldsymbol{0} & \frac{1}{2}\boldsymbol{I} & \cdots \\ & \cdots & \frac{1}{4}\boldsymbol{I} & -\frac{1}{2}\boldsymbol{I} & \frac{1}{4}\boldsymbol{I} & \cdots \\ & & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \vdots \\ \boldsymbol{x}_{t-1} \\ \boldsymbol{x}_t \\ \boldsymbol{x}_{t+1} \\ \boldsymbol{x}_{t+2} \\ \vdots \end{bmatrix}$$

上記では簡単のためインデックス i を省略したが，実際には各動作に対して得られる訓練集合は $\{\Xi_i | i = 1,...,N\}$ で表される．ここで，煩雑さを避けるために記号上は区別しないものの，Ξ_i ごとに T が異なることに注意されたい．この訓練集合に対し，HMM-d の学習を行う．HMM-d のパラメータ学習は，通常の HMM と同様に行えばよい．学習手法については，多くの解説（例えば [27]）がなされているため割愛する．Calinon の解説 [10] には，Matlab による模倣学習のサンプルコードが示されるなど，模倣学習における HMM の利用を考える読者に一読を薦める．

4.2 HMM-d を用いた軌道生成

学習済みの HMM-d から最尤軌道を生成する手法 [24] について述べる．本稿では，状態系列は入力として与えられることを前提とする．合理的な状態系列を得るために，訓練集合における状態継続長の平均について考える．HMM の学習では，サンプル Ξ のフレーム t においてどの状態にあったかの推定結果が得られる．この推定結果から，訓練集合中の平均状態系列 $\bar{\boldsymbol{s}}$ が以下のように表される．

$$\bar{\boldsymbol{s}} = (S_{\text{init}}, \underbrace{\underbrace{S_1,\ldots,S_1}_{\bar{d}_1}, \underbrace{S_2,\ldots,S_2}_{\bar{d}_2}, \ldots, \underbrace{S_M,\ldots,S_M}_{\bar{d}_M}}_{T}, S_{\text{final}})$$

ここに $\bar{d}_j$ は状態 S_j の平均継続長（ただし離散とする）を表す．よって，生成される軌道のフレーム数 T は $\bar{d}_j$ の和となる．なお，S_{init} および S_{final} は初期状態および終了状態であり，出力確率分布を持たず，総状態数に含めて考えないことが多い．本稿では割愛するが，状態系列を含めて最尤軌道生成を議論することもできる．

上述のように，状態系列 $\bar{s}$ が与えられるので t に対応する S_j が定まる．ここで，λ には全状態における出力確率分布の学習済みパラメータが含まれていたので，S_j が定まれば平均ベクトルと共分散行列は定まる．いま，t に対応する平均ベクトルおよび共分散行列を，$\boldsymbol{\mu}_t$ および $\boldsymbol{\Sigma}_t$ と表記することとする．$\bar{\boldsymbol{s}}$ が与えられたうえで，軌道 Ξ の尤度は以下で与えられる．

$$p(\Xi|\bar{\boldsymbol{s}},\lambda) = \prod_{t=1}^{T} p(\boldsymbol{\xi}_t|\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t) \tag{6}$$

これは以下のように書き直すことができる．

$$p(\Xi|\bar{\boldsymbol{s}},\lambda) = p(\Xi|\boldsymbol{\mu},\boldsymbol{\Sigma}) \tag{7}$$

ここに，

$$\boldsymbol{\mu} = \left[\boldsymbol{\mu}_1^\top, \boldsymbol{\mu}_2^\top, \ldots, \boldsymbol{\mu}_T^\top\right]^\top \tag{8}$$

$$\boldsymbol{\Sigma} = \text{diag}[\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \ldots, \boldsymbol{\Sigma}_T] \tag{9}$$

である．いま，(7) 式を最大化する軌道 $\hat{\Xi}$ は，以下のようにして得られる．

$$\hat{\Xi} = \underset{\Xi}{\text{argmax}}\ \log\, p(\Xi|\boldsymbol{\mu},\boldsymbol{\Sigma}) \tag{10}$$

$$= \underset{\Xi}{\text{argmax}} \left\{ -\frac{1}{2}\Xi^\top \boldsymbol{\Sigma}^{-1}\Xi + \Xi^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + const. \right\} \tag{11}$$

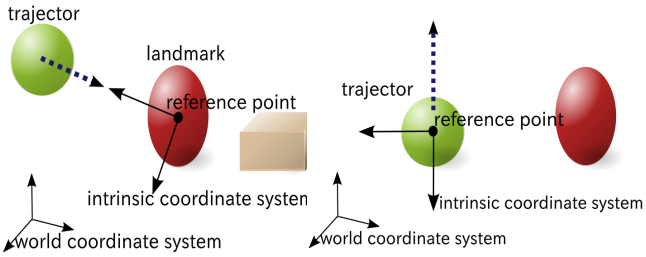
ここで，出力確率密度分布がガウス分布であることを用いた．また，(5) 式では $\Xi = \boldsymbol{\Phi}\boldsymbol{x}$ という関係があったので，この関係のもとで (10) 式を解くために，(10) 式の対数尤度の偏微分を $\boldsymbol{0}$ とおく．

$$\frac{\partial \log p(\boldsymbol{\Phi}\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial \boldsymbol{x}} = \boldsymbol{0} \tag{12}$$

(10)〜(12) 式より，最尤軌道 $\hat{\boldsymbol{x}}$ が以下のように得られる．

$$\hat{\boldsymbol{x}} = (W^\top \Sigma^{-1} W)^{-1} W^\top \Sigma^{-1} \boldsymbol{\mu} \tag{13}$$

(13) 式は Cholesky 分解を用いて，高速に解ける．外乱に応じた逐次動作生成手法については [28] を参照されたい．



第 3 図　動作と参照点・固有座標系の関係．左:「近づける」，右:「上げる」

5. 物体操作の模倣学習

5.1 問題設定

「A を B に載せる」や「A を回す」など，物体を操作する動作を模倣学習の枠組みでロボットに獲得させることを考える．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[1]．本稿では，以下のような問題設定に分類できるものと考える．

(1) ランドマークを必要としない動作（例：「A を回す」）のみに限定できる場合は，前節の HMM-d を用いてトラジェクタの軌道を生成することができる．また，ランドマークを必要とする動作（例：「A を B に載せる」）に限定し，さらにランドマーク候補が 1 つしかない状況に限定できるのであれば，オブジェクト同士の相対座標系上で軌道を学習すればよい [18]．

(2) 上記 (1) において，ランドマーク候補が複数ある場合は，ランドマークの選択と軌道を同時に学習する必要がある．また，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する場合は，ランドマークの選択を参照点の選択として一般化し，参照点候補にランドマーク候補座標に加えて世界座標系の中心等を含める．

(3) 上記 (1)(2) において，参照点の選択に加えて座標系の方向（またはアフィン変換の種類）について K 通りの可能性が考えられる場合，動作に応じた K 個の固有座標系タイプが存在すると考え，{固有座標系タイプ，参照点，HMM パラメータ} の三者について推定を行う問題として定式化する．

上記 (2)(3) 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．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，軌道の学習と同時に推定する必要がある．

事前知識として，固有座標系は有限個のタイプに分類できるものと仮定する．固有座標系タイプを k とし，そ

[1] 認知言語学では，外部世界を解釈する主体のプロセスにおいて焦点化される存在のうち，相対的に際立って認知される対象をトラジェクタ，これを背景的に位置付けるオブジェクトをランドマークと呼ぶ [29]．

の総数を 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．

C_2: ランドマークを原点とし，$\boldsymbol{x}_1$ に向かう軸を x 軸とする直交座標系 (第 3 図左図参照)．

C_3: $\boldsymbol{x}_1$ を原点とする，カメラ座標系を平行移動した座標系 (第 3 図右図参照)．

C_4: 画面中心 $\boldsymbol{x}_{\mathrm{center}}$ を原点とする，カメラ座標系を平行移動した座標系．

教示データは動画像として与えられるものとし，動画像に 1 個のトラジェクタと 0 個以上の静止オブジェクトが存在するものとする．画像から物体の重心位置は抽出可能であるものとし，得られた位置集合を参照点候補集合 $\boldsymbol{O}=\{\boldsymbol{x}^r|r=1,2,...,R\}$ とする．物体検出およびトラッキングについては，一般的な手法を利用することを想定する．C_1,C_2 では，$\boldsymbol{O}$ の全て要素に対して探索を行う．C_3,C_4 では参照点が唯一に定まるので，参照点の探索を行う必要はない．

5.2 RPD-HMM の学習

以下では，前節の問題 (2)(3) を同じ枠組みで解くために我々が提案した，Reference-Point-Dependent HMM (RPD-HMM) を紹介する．教示入力 $\mathcal{V}_i$ は以下のように構成されるものとする．

$$\mathcal{V}_i=(\Xi_i,\boldsymbol{O}_i) \tag{14}$$

すなわち，前章で説明した軌道に加え，参照点候補集合が学習サンプルに含まれるものとする．

RPD-HMM では，Ξ_i を各固有座標系上の軌道に変換し，複数の軌道として捉える．ここで，固有座標系タイプ k と参照点 $\boldsymbol{x}^r$ によって決定される固有座標系 $C_k(\boldsymbol{x}^r)$ 上での軌道を ${}^{C_k(\boldsymbol{x}^r)}\Xi$ と表記することとする．以下の尤度最大化により，固有座標系タイプ k，参照点インデックス r，HMM パラメータ λ を探索する．

$$(\hat{k},\hat{r},\hat{\lambda})=\underset{k,r,\lambda}{\mathrm{argmax}}\sum_{i=1}^{N}\log\ p({}^{C_k(\boldsymbol{x}^r)}\Xi_i;\lambda) \tag{15}$$

EM アルゴリズムを用いた上式の解法は [21] の付録を参照されたい．RPD-HMM とは，このようにして得られた固有座標系タイプ k および HMM パラメータ λ の組を指す．なお，音声指示と v を対応付けておけば，「A を B に載せて」などの指示を理解し動作を生成することが可能になる．

5.3 RPD-HMM による動作認識

RPD-HMM を用いて動作認識を行うことを考える．(15) 式の学習の結果，固有座標系タイプおよび HMM パラメータの組が，動作 v_m と対応付けられる．すなわち，$v_m=(k_m,\lambda_m)$ が得られる．動作認識は，軌道 Ξ を最も高い確率で出力する RPD-HMM λ_m を求める問題として定式化することができる．

いま，前節と同様に，Ξ と $\boldsymbol{O}$ が得られたとする．k について探索する必要がないので，固有座標系を $C(\boldsymbol{x}^r)$ と書くことにする．このとき，認識結果である動作インデックス $\hat{m}$ は以下のようにして求められる．

$$(\hat{m},\hat{r})=\underset{m,r}{\mathrm{argmax}}\ p({}^{C(\boldsymbol{x}^r)}\Xi|\lambda_m) \tag{16}$$

なお，付加的に参照点の推定結果 $\hat{r}$ も同時に得られる．連続動作の認識への拡張については [30] を参照されたい．

5.4 RPD-HMM による動作生成

さらに，ある動作について，学習済みの RPD-HMM から軌道を生成することを考える．RPD-HMM に単純に軌道生成手法を適用すると固有座標系上の軌道が生成されるため，生成時に RPD-HMM の座標変換を行う必要がある．本節では，混乱を避けるため座標系を明示することとする．例えば，固有座標系上および世界座標系の HMM パラメータを，それぞれ ${}^C\lambda$ および ${}^W\lambda$ と表記する．

まず，位置の平均ベクトル ${}^C\boldsymbol{\mu}_{\boldsymbol{x}}$ を，以下の同次変換行列により変換する．

$$\left[\begin{array}{c}{}^W\boldsymbol{\mu}_{\boldsymbol{x}}\\ \hline 1\end{array}\right]=\left[\begin{array}{c|c}A & {}^W\boldsymbol{x}_1\\ \hline \boldsymbol{0} & 1\end{array}\right]\left[\begin{array}{c}{}^C\boldsymbol{\mu}_{\boldsymbol{x}}-{}^C\boldsymbol{\mu}_{\boldsymbol{x}}(S_1)\\ \hline 1\end{array}\right] \tag{17}$$

ここに，A は回転行列，$\boldsymbol{\mu}_{\boldsymbol{x}}(S_1)$ は S_1 の位置の平均ベクトルを表す．なお，平均や共分散行列は状態ごとに定義されているが，煩雑さを避けるために省略した [1]．また，A を用いて速度の平均ベクトルと，位置の共分散行列 $\boldsymbol{\Sigma}_{\boldsymbol{xx}}$ を変換する．

$${}^W\boldsymbol{\mu}_{\dot{\boldsymbol{x}}}=A\ {}^C\boldsymbol{\mu}_{\dot{x}} \tag{18}$$

$${}^W\boldsymbol{\Sigma}_{\boldsymbol{xx}}=A\ {}^C\boldsymbol{\Sigma}_{\boldsymbol{xx}}\ A^\top \tag{19}$$

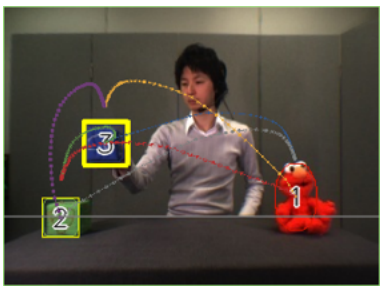
加速度の平均ベクトル，および位置・加速度の共分散行列についても同様である．以上より世界座標系に変換された HMM について，(13) 式を適用して最尤軌道を生成する．第 4 図に同一の RPD-HMM から生成された軌道を示す．各状況ごとに滑らかな軌道が生成できるという RPD-HMM の特徴を示すために，6 通りの組み合わせについて最尤軌道を生成させた．

6. おわりに

模倣学習研究は，機械学習，ヒューマンロボットインタラクション，画像処理，動作計画などにおける基盤技術と関連が深く，人間の行動の科学的理解と工学的応用の両側面から活発に研究が行われている．どちらの側面においても，人間の多様な行動を観察し実世界の事物とどう関連するかを推定することが不可欠である．

本稿では，確率ロボティクスに分類されるアプローチ

[1] 前章で述べたように状態系列は既知であるので，各フレームに対応する出力確率密度分布は容易に得られる．



第 4 図 RPD-HMM から生成された各対象ごとの最尤軌道. 同一の RPD-HMM から状況に応じた軌道が生成可能であることを示す.

を中心に模倣学習研究を概説するとともに，HMM-d による模倣学習手法と，物体操作に対する RPD-HMM の適用例を紹介した．本稿で紹介した RPD-HMM に関する動画は http://komeisugiura.jp を参照されたい.

謝　　辞

本研究の一部は，JST CREST および JSPS 科研費 15K16074 の助成を受けて実施されたものである.

参 考 文 献

[1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, “Robot Programming by Demonstration,” Springer Handbook of Robotics, pp.1371–1394, Springer, 2008.

[2] B.D. Argall, S. Chernova, M. Veloso, and B. Browning, “A Survey of Robot Learning from Demonstration,” Robotics and Autonomous Systems, vol.57, no.5, pp.469–483, 2009.

[3] S. Schaal, A. Ijspeert, and A. Billard, “Computational approaches to motor learning by imitation,” Philosophical Transaction of the Royal Society of London: Series B, Biological Sciences, vol.358, no.1431, pp.537–547, 2003.

[4] Y. Kuniyoshi, M. Inaba, and H. Inoue, “Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance,” IEEE Transactions on Robotics and Automation, vol.10, no.6, pp.799–822, 1994.

[5] H. Miyamoto, S. Schaal, F. Gandolfo, H. Gomi, Y. Koike, R. Osu, E. Nakano, Y. Wada, and M. Kawato, “A Kendama Learning Robot Based on Bi-directional Theory,” Neural Networks, vol.9, no.8, pp.1281–1302, 1996.

[6] J. Tani and M. Ito, “Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment,” IEEE Trans. on Systems, Man and Cybernetics, Part A, vol.33, no.4, pp.481–488, 2003.

[7] C. Breazeal and B. Scassellati, “Robots That Imitate Humans,” Trends in Cognitive Science, vol.6, pp.481–487, 2002.

[8] A.J. Ijspeert, J. Nakanishi, and S. Schaal, “Learning Attractor Landscapes for Learning Motor Primitives,” Advances in Neural Information Processing Systems, vol.15, pp.1523–1530, 2002.

[9] T. Matsubara, S. Hyon, and J. Morimoto, “Learning Stylistic Dynamic Movement Primitives from Multiple Demonstrations,” Proc. IEEE/RSJ IROS, pp.1277–1283, 2010.

[10] S. Calinon, “A Tutorial on Task-Parameterized Movement Learning and Retrieval,” Intelligent Service Robotics, vol.9, no.1, pp.1–29, 2016.

[11] V. Krüger, D. Kragic, A. Ude, and C. Geib, “The Meaning of Action: A Review on Action Recognition and Mapping,” Advanced Robotics, vol.21, no.13, pp.1473–1501, 2007.

[12] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky, “Bayesian Nonparametric Methods for Learning Markov Switching Processes,” Signal Processing Magazine, IEEE, vol.27, no.6, pp.43–54, 2010.

[13] S. Calinon and A. Billard, “Incremental Learning of Gestures by Imitation in a Humanoid Robot,” Proc. ACM/IEEE HRI, pp.255–262, 2007.

[14] C.E. Rasmussen and C.K.I. Williams, Gaussian Processes for Machine Learning, The MIT Press, 2005.

[15] K. Gräve and S. Behnke, “Incremental Action Recognition and Generalizing Motion Generation Based on Goal-Directed Features,” Proc. IEEE/RSJ IROS, pp.751–757, 2012.

[16] J.M. Wang, D.J. Fleet, and A. Hertzmann, “Gaussian Process Dynamical Models,” Advances in Neural Information Processing Systems, vol.18, pp.1441–1448, The MIT Press, 2006.

[17] 岩田健輔，中村友昭，長井隆行，持橋大地，小林一郎，麻生英樹，“参照点に依存したガウス過程隠れセミマルコフモデルに基づく連続動作の分節化,” 第 34 回日本ロボット学会学術講演会資料，pp.RSJ2016AC3Z2–07, 2016.

[18] K. Ogawara, J. Takamatsu, H. Kimura, and K. Ikeuchi, “Generation of a Task Model by Integrating Multiple Observations of Human Demonstrations,” Proc. IEEE ICRA, pp.1545–1550, 2002.

[19] T. Inamura, I. Toshima, H. Tanie, and Y. Nakamura, “Embodied Symbol Emergence Based on Mimesis Theory,” International Journal of Robotics Research, vol.23, no.4, pp.363–377, 2004.

[20] D. Kulic, W. Takano, and Y. Nakamura, “Representability of Human Motions by Factorial Hidden Markov Models,” Proc. IEEE/RSJ IROS, pp.2388–2393, 2007.

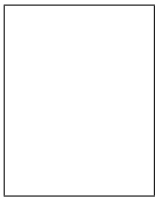
[21] K. Sugiura, N. Iwahashi, H. Kashioka, and S. Nakamura, “Learning, Generation, and Recognition of Motions by Reference-Point-Dependent Probabilistic Models,” Advanced Robotics, vol.25, no.6-7, pp.825–848, 2011.

[22] T. Tagniguchi, K. Hamahata, and N. Iwahashi, “Unsupervised Segmentation of Human Motion Data Us-

ing Sticky HDP-HMM and MDL-based Chunking Method for Imitation Learning," Advanced Robotics, vol.25, no.17, pp.2143–2172, 2011.
[23] S. Furui, "Speaker-Independent Isolated Word Recognition Using Dynamic Features of Speech Spectrum," IEEE Trans. on Acoustics, Speech, and Signal Processing, vol.34, no.1, pp.52–59, 1986.
[24] K. Tokuda, T. Kobayashi, and S. Imai, "Speech Parameter Generation from HMM using Dynamic Features," Proc. IEEE ICASSP, pp.660–663, 1995.
[25] N. Niwase, J. Yamagishi, and T. Kobayashi, "Human Walking Motion Synthesis with Desired Pace and Stride Length Based on HSMM," IEICE Trans. on Information and Systems, vol.88, no.11, pp.2492–2499, 2005.
[26] H. Zen, T. Nose, J. Yamagishi, S. Sako, T. Masuko, A.W. Black, and K. Tokuda, "The HMM-Based Speech Synthesis System (HTS) Version 2.0," Proc. Sixth ISCA Workshop on Speech Synthesis, pp.294–299, 2007.
[27] 鹿野清宏，河原達也，山本幹雄，伊藤克亘，武田一哉，IT Text 音声認識システム，オーム社，2001.
[28] 杉浦孔明，是津耕司，"RPD-HMM に基づく逐次動作生成による物体操作の模倣学習，" 第 28 回人工知能学会全国大会資料，pp.1I5–OS–09b–5，2014.
[29] R.W. Langacker, Foundations of Cognitive Grammar: Theoretical Prerequisites, Stanford Univ Pr, 1987.
[30] K. Sugiura and N. Iwahashi, "Motion Recognition and Generation by Combining Reference-Point-Dependent Probabilistic Models," Proc. IEEE/RSJ IROS, pp.852–857, 2008.

著 者 略 歴

杉浦 孔明 (非会員)



1978 年 6 月 9 日生．2002 年京都大学工学部電気電子工学科卒業．2004 年同情報学研究科修士課程修了．2007 年同博士後期課程修了．博士（情報学）．日本学術振興会特別研究員，ATR 音声言語コミュニケーション研究所研究員を経て，現在，情報通信研究機構先進的音声翻訳研究開発推進センター主任研究員．サービスロボット，機械学習，クラウドロボティクス，音声対話システム，情報検索，推薦システムなどの研究に従事．計測自動制御学会，情報処理学会，人工知能学会，日本ロボット学会，IEEE などの会員．