

# HMMでの動作認識における類似動作からの特徴部位抽出

井口 茂† 山名 早人†  
早稲田大学理工学部情報学科†

## 1 はじめに

近年ロボットは、その技術発達や低価格化に伴い、日常生活により身近になってきている。しかし、未だ十分な自律性とそのための学習機構を持つには遠く、ロボットの思考や行動は、逐一設計者やユーザが編集、作成しなければならない。行動学習における動作の獲得や認識、再生、合成を扱った研究は多いものの、動作から概念形成に役立つ情報を得ることを直接ねらいとする研究は少ない。動作データ間に存在する類似性とその理由を見出して整理することは、柔軟な認識や思考に役立つ情報と考えられる。本稿では、隠れマルコフモデルを用いた行動認識に基づき、観測した複数の類似動作群から、類似性の原因となっている特徴部位を抽出する手法を提案する。

## 2 関連研究

HMMを用いた動作認識には、データとしてカメラで撮影した2次元動画像を用いるもの[2][3]や、人間やロボットの関節角のような身体情報を用いるもの[4][5][6]がある。前者では動作をHMMに入力する際のデータのベクトル量子化に関する工夫[2]や、特徴量としてMPEGのDCT(Discrete Cosine Transform)係数を用いる方法[3]などが示されている。後者では動作の認識だけでなく再生、合成、抽象化、階層化などが提案され、またその中では動作間の類似度としてHMM間の距離を用いている。

しかし、従来の研究は動作の認識や合成の効率化にとどまり、動作から役立つ情報を取り出しロボットの思考に活かすことを中心に考えた研究は見当たらない。文献[4][5][6]では動作の類似度を算出しているが、類似性は単なるスカラー値として求まるに過ぎず、類似の理由など思考に役立つような情報は得られない。

そこで本稿では「動作に本質的な動きをしている部位」という情報を思考に役立つ情報の一つと考え、観測した動作サンプル集合から、その動作の特徴部位を特定する手法を提案する。

### 2.1 HMM間の距離 [1][5]

HMMは確率モデルであるためKullback-Leibler情報量を用いてHMM間の隔たりを次式のように表すことができる。

$$D(\lambda_1, \lambda_2) = \frac{1}{n} \sum_i \frac{1}{T_i} \left[ \log p(\mathbf{y}_1^{T_i} | \lambda_1) - \log p(\mathbf{y}_1^{T_i} | \lambda_2) \right]$$

A Characteristic-part Extraction Scheme from Similar Motions in the Motion Recognition with HMM  
†Shigeru IGUCHI, †Hayato YAMANA  
†Department of Information and Computer Science, Science and Engineering, Waseda University

ここで $\lambda_1, \lambda_2$ は対象となる2つのHMMのパラメータ、 $\mathbf{y}_1^{T_i}$ は $\lambda_1$ の学習に用いた長さ $T_i$ の時系列データ、 $n$ は観測された時系列データの数を表す。

一般に $D(\lambda_1, \lambda_2) \neq D(\lambda_2, \lambda_1)$ と非対称なので、対称にした

$$D_s(\lambda_1, \lambda_2) = \frac{1}{2} (D(\lambda_1, \lambda_2) + D(\lambda_2, \lambda_1)) \quad (1)$$

が用いられている[4][5][6]。

本稿でもHMM間距離の算出に式(1)を用いる。

## 3 特徴部位の抽出手法

人間の動作を対象として次の手順で特徴を抽出する(図1)。

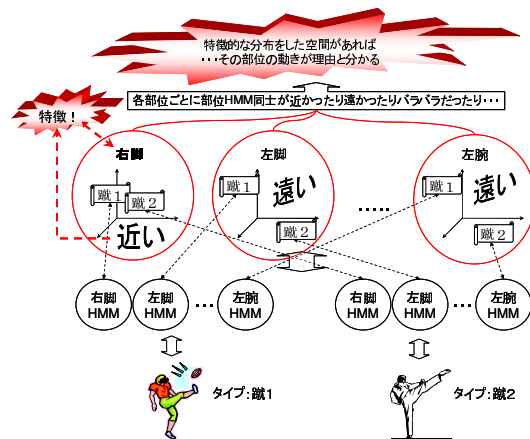


図1: 提案手法の概要

1. 人間の体をR個の部位に分割する。
2. 同一のカテゴリに属する動作をNサンプル観測する。
3. Nサンプルの動作をVタイプ( $2 \leq V \leq N$ )の動作として捉え、各タイプについて部位の数だけ部位HMM(RHMM)を用意する(計VR個)。動作 $v$ ( $2 \leq v \leq V$ )は $\text{RHMM}_{v1}, \text{RHMM}_{v2}, \dots, \text{RHMM}_{vR}$ によって表わされることになる。
4.  $v$ タイプに属する各動作ごとに、観測した各部位のデータで $\{\text{RHMM}_{vr}\}$ を学習させる。
5. 各部位 $r$ について $\text{RHMM}_{1r}, \dots, \text{RHMM}_{Nr}$ 間の近さ $close_r$ を求める。
6. R個の部位の中でRHMMが近く密集している、つまり $close_r$ が最大となる部位 $\hat{r}$ が、動作のサンプルに共通する最も特徴的な部位として抽出する。

## 4 評価

提案手法で特徴部位が抽出できるかを検証するために、2つの実験を行う。実験1ではあらかじめ特徴を設定した擬似データセットをプログラムで作成しそのデータに対して、実験2では人間の動作をモーションキャプチャして得た実際のデータに対して提案手法を用いる。

### 4.1 実験1

データは5次元ベクトルの時系列とする ( $R = 5$ )。データセット内の各データは"flat", "yama", "tani", "up", "down", "sin"の5カテゴリあり、各カテゴリに第1から第5次元のうち1つが特徴的に変化する「特徴次元」を設定する。各種類の特徴次元は順に、特徴次元無し、第1、第2、第3、第4、第5次元とする。

各カテゴリについて9サンプルずつ5タイプ(タイプごとにデータの長さや振幅の大きさが微妙に異なる)の合計45個の学習サンプルを作り ( $V = 5$ )、タイプ  $v(=1, \dots, V)$  に対して  $\{RHMM_{vr}\}(r=1, \dots, R)$  を作成する。各部位についてRHMM間距離

$$close_r = 2 / \sum_v \sum_{v'} D_s(RHMM_{vr}, RHMM_{v'r}) \quad (2)$$

を計り、 $close_r$  が最大となる第  $\hat{r}$  次元が、設定しておいた特徴次元であれば正解とする。

また時系列はsin波を元に生成するが、特徴次元以外は位相や周期をサンプルごとにランダムに変化させたものである。特徴次元以外の振幅が特徴次元の振幅に対して  $k = 0.2, 0.6, 1.0, 1.7, 5.0$  倍の時それぞれについて以上の実験を行う。 $k > 1$  の場合の実験は、特徴以外の次元の動きが大きいために特徴次元が埋もれてしまい、抽出できなくなるかを調べるためである。

### 4.2 実験2

データは20次元ベクトルの時系列であり、内訳は体の10個の部位である ( $R = 10$ )。動作のカテゴリは"蹴る", "パンチ", "投げる", "手を振る", "座る", "寝る", "歩く"の7種類で、それぞれについて特徴部位を抽出する。各カテゴリ3サンプルずつ3タイプ ( $V = 3$ 、動作の意味が変わらない範囲で各タイプごとに変化をつける)の合計9個のサンプルである。実験1と同様に各部位の  $close_r$  を求め、 $close_r$  の大きいものほど特徴的な部位であるとして、大きい順に順位をつける。

実験2では人間の自然な動きをデータとしたため、実験1のように明確な正解があるわけではない。そこで筆者の主観で考えて特徴的と思われる部位を正解として結果を検討する。動作ごとに正解とする部位は表1に示した部位とする。

### 4.3 結果

実験1では  $k$  の値に関わらず正解率100%を得た\*。これにより、提案手法は部位の動きの大きさに依らず有効であることが示された。

実験2の結果を表1にまとめる。"蹴り", "パンチ", "手を振る", "寝る"は正解部位が比較的上位にあり1位や2位となっているが、他の動作については悪い結果といえる。

\*ただし"flat"はそもそも特徴次元を設けていないためカウントしない。結果は  $\hat{r}$  は常に一意ではなく、3の場合と5の場合とがあった。

表1: 実験2の結果のまとめ

動作	蹴る		パンチ		手を振る		投げる		座る		寝る	歩く					
	右股関節	右膝	上半身	右肩	右肘	右肩	右肘	右肩	右肘	右股関節	左股関節	体全体	右股関節	左股関節	右膝	左膝	
正解																	
順位	2	5	2	1	3	6	2	5	8	3	7	2	8	10	3	7	

## 5 おわりに

本稿では観測した動作サンプル集合から、その動作の特徴部位を特定する手法を提案した。検証実験の結果、あらかじめ正解を設けておいた擬似データに対しては、部位の動きの大きさによらず100%の正解率で抽出に成功したが、実際の動作データに対しては良好な結果を示さなかった。今回作成したような、特徴やデータが比較的単純なデータに対しては提案手法は良い結果を示したが、人間の動作のような特徴が明確でなく複雑なデータに対しては有効性が示されず、改善が必要がある。

提案手法は人間の動作に限らずあらゆる多変量時系列データに対して汎用的な方法である。しかし人間の動作データに対して有効性が示されなかった(実験2)ことや、計算量が部位数の2乗のオーダーで増加するなど問題は多い。手法の改良や他方法との連携による性能の向上が今後の課題である。

## 参考文献

- [1] B.H.Juang, L, "A probabilistic distance measure for hidden Markov models", AT & T Technical Journal, Vol.64, pp.391-408, 1985.
- [2] 大和淳司, 倉掛正治, 伴野明, 石井健一郎, "カテゴリ別VQを用いたHMMによる動作認識法", 信学論D-II, Vol.J77-D-II, No.7, pp.1311-1318, 1994.
- [3] 大和淳司, 村瀬洋, "HMMを用いたMPEGデータからの動画像検索の検討", 信学技報, Vol.96, No.140, PRMU96-34, pp.65-72, 1996.
- [4] 稲邑哲也, 谷江博昭, 中村仁彦, "運動パターンの認識/生成の単一モデルとそれにもとづく離散的階層化による行動知能の集積", 人工知能学会全大(第17回)予稿集, p.3, D4-04, 2003.
- [5] 谷江博昭, 稲邑哲也, 中村仁彦, "連続分布型HMMへの位相構造導入による全身運動を埋め込んだ原始シンボル空間構成法", 人工知能学会全大(第17回)予稿集, p.3, D4-03, 2003.
- [6] 稲邑哲也, 中村仁彦, 戸嶋巖樹, 江崎英明, "ミメシ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.