

ノンパラメトリックベイズによる Trajectory HMMの自動選択と手書き文字生成

○古林 邦彬 谷口 忠大 (立命館大学)

1 はじめに

手書き文字や音声を認識する研究やロボットの動作生成を学習させる研究の多くは隠れマルコフモデル (Hidden Markov Model; HMM) が用いられている. 手書き文字や音声はヒトや環境によって揺らぎが発生する行動であり, HMMはその揺らぎに対してある程度まで柔軟に対応できるためである. しかし, 通常の HMM では隠れ状態数の指定が必要であり, 動作生成時に不連続な軌道になる問題がある.

本研究では従来の HMM で発生する問題を解決するためにノンパラメトリックベイズ理論に基づき自動的に Left-to-Right 型の HMM のモデルを生成手法に提案する.

2 実験手法

通常の Trajectory HMM では, 元の HMM の隠れ状態を指定する必要があった. 本実験ではノンパラメトリックベイズでモデルの隠れ状態数を自動推定し, 静的・動的特徴量との関係を導入することで再定式化する方法を提案する. 今回はテストケースとして動作に揺らぎが生じやすい一方で, なめらかな動きが要求される手書き文字生成を行った. 具体的には Sticky HDP-HMM の学習により隠れ状態数を自動推定し, 状態遷移と各隠れ状態の出力パラメータを学習する. Sticky HDP-HMM の学習で得たモデルを Trajectory HMM に再定式化する. この一連の流れで学習した後, ロボットに手書き文字の生成を行わせる. ロボットは Aldebaran 社の二足歩行ロボット Nao を使用し, 筆ペンを持たせて文字生成を行なう. Fig.1 で目標達成までのイメージ図を記す.

2.1 Sticky HDP-HMM ¹⁾

Sticky HDP-HMM は自己遷移確率にバイアスを付加した ergodic 型 HDP-HMM である. また, HDP-HMM は HMM に無限次元の潜在状態を仮定する事により, データに応じて潜在状態の数を決定する柔軟なモデルである. このモデルの自己遷移確率を大きくすることで潜在状態の過剰遷移を抑えることが出来るため, 動作の連続性を仮定するモデリングを効率的に行える.

2.2 Trajectory HMM ²⁾

Trajectory HMM は通常の HMM に対して静的・動的特徴量間の関係を明示的に導入して再定式化したモデルである. これにより通常の HMM で挙げられる問題点 (状態内で出力確率分布が一定である, 独立性の仮定) を避ける事ができ, 学習データや評価データに対する尤度の向上が期待できる.

3 実験設定

本実験ではペンタブレットから得た「cat」という文字の x, y 軸の座標データと速度データのセットを教示データとして5セット利用している. はじめに教示デー

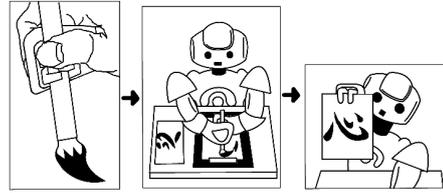


Fig. 1: 研究目標のイメージ図

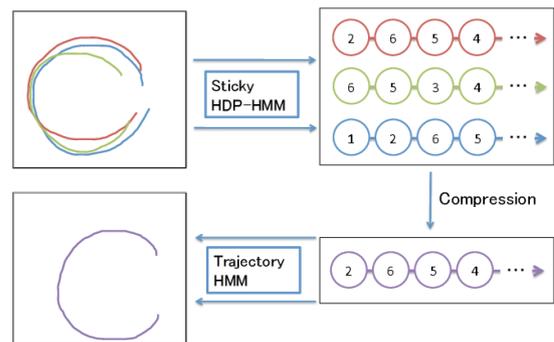


Fig. 2: x 軸に注目した時の遷移図

タを用いて Sticky HDP-HMM によるモデルの学習を行う. 次に学習で得た状態系列情報と各隠れ状態の出力パラメータを用いて Trajectory HMM に再定式化を行なう. 状態系列は Sticky HDP-HMM による学習で得た5データ分の状態系列のそれぞれの編集距離が最も小さいデータを基に決定する. また, 自己遷移回数は各状態の平均値を利用している. 今回は静的特徴量と1次動的特徴量を利用している. Fig. 2 で本実験におけるプロセスを示す.

4 今後の課題

現在の状態系列を決定する方法はヒューリスティックである. そのため, 各データの状態系列の違いが大きくなりすぎると正しい状態系列を得るのが難しくなる. この問題を解決するため, 解析的な状態系列の導出方法を模索する.

参考文献

- 1) E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky. The sticky hdp-hmm: Bayesian nonparametric hidden markov models with persistent states. Technical Report 2777, MIT Laboratory for Information and Decision Systems(2007).
- 2) Keiichi Tokuda, Heiga Zen, Tadashi Kitamura. Trajectory modeling based on HMMs with the explicit relationship between static and dynamic features. Proc. of Eurospeech(2003). pp.865-868.
- 3) Sugiura, K. and Iwahashi, N. Learning object-manipulation verbs for human-robot communication. Proceedings of the 2007 workshop on Multimodal interfaces in semantic interaction(2007). pp.32-38.