H-045

Coupled HMM を用いた複数人の行動モデリング

Modeling Multiple People's Behavior Using Coupled HMM

宋海俊† Haijun Song　平野　靖‡ Yasushi Hirano　梶田　将司‡ Shoji Kajita　間瀬　健二‡ Kenji Mase

1　はじめに

人の行動パターンを記録し，モデリングする行動モデリングの目的は，行動の分類と行動の予測，行動のシミュレーションなどが考えられる．

人の行動パターンを記録する場合，映像とセンサを利用することが一般的である．特に映像で記録する場合，画像処理のアルゴリズムを利用し，人あるいは人の一部の位置データを抽出する必要がある [1, 2]．本稿では，室内環境に設置している超音波位置センサ装置を用いて，複数人の位置データを収集した．

近年，行動パターンをモデリングする場合，確率グラフモデルを用いることが多くなってきた．青木らは隠れマルコフモデル (HMM) を用いて，人の動作と行動パターンを学習し，新たな動作と行動パターンについて分類することで HMM の有効性を検証した [1]．しかし，実世界の問題では複数のプロセスが相互に影響し合っている場合が多い．このため，モデリングを行う際，複数の変数が必要と考えられる．一般に，HMM では，隠れ変数が 1 つしかないことから，実世界の問題に適していない場合が多いと言われている [2]．複数のマルコフ過程を用いて，各マルコフ過程間の影響を考慮するモデルとして，Coupled HMM(CHMM) がある．CHMM は複数の隠れ変数を持つため，アルゴリズムを構築することが非常に困難である．Oliver らは "N-heads" dynamic programming を利用して，アルゴリズムを構築した [2]．

本稿では，変分手法を用いて，CHMM のアルゴリズムを構築する方法 [3] を採用し，文献 [3] では提示されていなかった具体的なアルゴリズムを実装する．変分手法を使用することで，複雑なモデルを単純なモデルに分解でき，個々の単純なモデルに対応するアルゴリズムを利用できるという大きな利点がある．

2　CHMM

CHMM は図 1 のようなグラフで表現することができる．図 1 では $Z_n^{(1)}, \ldots, Z_n^{(M)}$ が時刻 n における隠れ変数，$X_n^{(1)}, \ldots, X_n^{(M)}$ が時刻 n における観測値として表記する．また，各矢印を持つ線を条件付き確率分布とする．図 1 から，時刻 n における隠れ変数状態が時刻 $n-1$ の２つの隠れ変数の状態に影響されていることが分かる．各隠れ変数の関係を具体的に式で表す場合，以下のようになる．

$$p(Z_n^{(m)}|Z_{n-1}^{(1)}, \ldots, Z_{n-1}^{(M)}) = \sum_{i=1}^{M} B_{im} p(Z_n^{(m)}|Z_{n-1}^{(i)}) \quad (1)$$

†名古屋大学大学院情報科学研究科
‡名古屋大学情報連携基盤センター

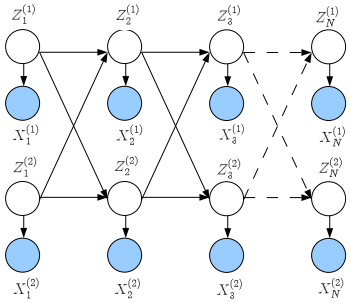

図 1　2 つのマルコフ過程を持つ CHMM

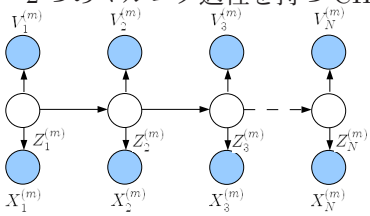

図 2　近似した HMM

式 (1) において，行列 B は M 次正方行列であり，かつ $\sum_{i=1}^{M} B_{im} = 1$ を満たしている．

CHMM のアルゴリズムを構築する場合にあたって，本稿では変分手法を用いる．変分手法の考えは CHMM に対応する隠れ変数の事後分布 $p(\boldsymbol{Z}|\boldsymbol{X})$ を事後分布 $q(\boldsymbol{Z}|\boldsymbol{X})$ に近似する．近似した後の事後分布について，簡単に計算できることが望ましい．

本稿では，CHMM を図 2 のような M 個の HMM に近似した．各 HMM が独立であるため，$q(\boldsymbol{Z}|\boldsymbol{X})$ を簡単に計算できる．しかし，近似を行う際，変分パラメータ $V_n^{(1)}, \ldots, V_n^{(M)}$ を導入する必要がある．$\boldsymbol{V}$ の計算式は近似性能を表す Lower Bound(式 (2)) を最大化することによって決めることができる．

$$\ell(q) = \sum_{\boldsymbol{Z}} q(\boldsymbol{Z}|\boldsymbol{X}) \ln \frac{p(\boldsymbol{X}, \boldsymbol{Z})}{q(\boldsymbol{Z}|\boldsymbol{X})} \quad (2)$$

CHMM を用いて学習を行う場合，図 2 のような HMM において，元の CHMM の事前確率分布，遷移確率分布，出力分布のそれぞれのパラメータを更新することができる．

3　実験結果

アルゴリズムの検証のため実験データとして古河機械金属株式会社の ZPS(Zone Positioning System)[1] 3 次元位置センサを使用して，20 日分の 13 人の位置の時系列データを収集した．センサの型番は FK04D005，タグ (発信機) の型番号は ZPS-TAG-001 である．

[1] http://www.furukawakk.jp/products/ZPS_1.html

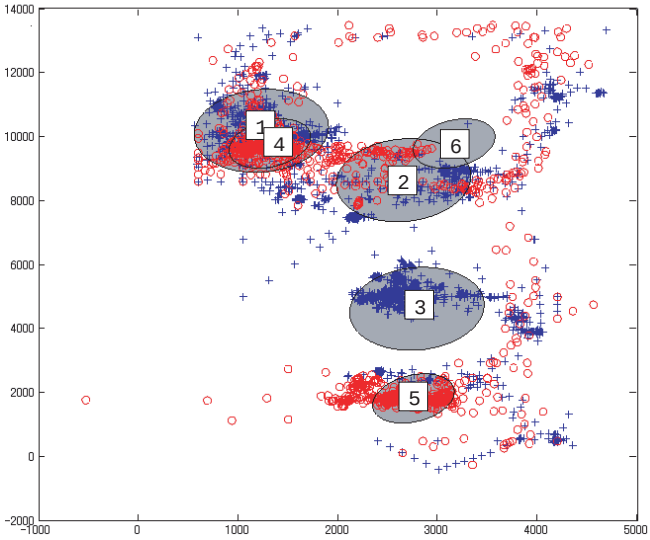

図 3　プロット図.+は被験者 1，○ は被験者 2 のそれぞれの位置座標となっている．

実験データから，比較的関係ある行動をとるある 2 人についての位置の時系列を抽出した．時系列データに関して，1 系列の長さは 289 秒，合計 10 系列となっている．

CHMM を用いて学習する場合の初期設定値は以下の通りである．

- 2 つのマルコフ過程それぞれが 3 状態を持つように設定した．
- 遷移行列 A の初期設定値をランダムで決めた．
- m 番目のマルコフ過程においての出力正規分布の共分散行列 $\Sigma^{(m)}$，出力正規分布の平均 $\mu^{(m)}$，事前確率 $\pi^{(m)}$ の初期設定値はそれぞれ以下のような方法によって決めた．
  - $\Sigma^{(m)}$ の初期設定値は m 番目のマルコフ過程に対応する人の位置の時系列データを用いて決めた．
  - $\mu^{(m)}$，$\pi^{(m)}$ の初期設定値は K-mean 法 (10 回) によって決定した．
- B の初期設定値は $\begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$ とした．

モデリングの過程において，10 回の試行を繰り返して，そのうち近似性能を表す値が一番大きい場合をモデリング結果とした．

推定された各状態に対応する正規分布の平均値と共分散を楕円形でプロットした結果を図 3 に示す．図 3 において，1，2，3 は被験者 1(+印) のマルコフ過程の状態，4，5，6 は被験者 2(○ 印) のマルコフ過程の状態である．軸の単位はミリメートルである．

図 3 から，各状態に対応する出力正規分布はほとんどデータのクラスタ周辺に推定されたことが確認できる．ただし，状態 6(図中右上) に対応する正規分布は右上に弧状に広がる点列 (○ 印) に影響され，悪い推定結果を得たことも図から簡単に確認できる．原因は正規分布が外れ値に影響されやすいためだと考えられる．

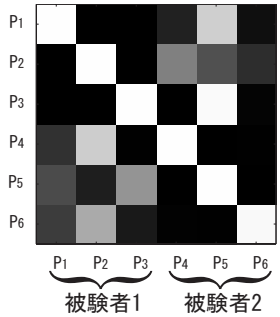

図 4　遷移行列

遷移行列の結果を図 4 に示す．図 4 において，白い領域は対応する遷移行列成分の値が 1 に近いことを意味し，黒い領域は 0 に近いことを意味する．遷移行列の構造は 4 つの 3×3 のブロックに分けられる．左上と右下のブロックは各マルコフ過程内部の遷移行列で，左下と右上は各マルコフ過程間の遷移行列となっている．

図 4 の対角線成分は 1 に近いことが確認できる．つまり各マルコフ過程内において，各状態がそのまま維持する傾向が見られた．その原因は被験者の位置があまり変化しない場合が全体のデータの中で高い割合を占めていると考えられる．そして，各マルコフ過程間の影響は非対称的であることも図 4 から確認できる．非対称の理由を簡単な例で説明できる．たとえば被験者 1 が場所 P_1 にいる時被験者 2 は場所 P_2 にいる確率が高いとしても，被験者 2 が場所 P_2 にいる時被験者 1 が場所 P_1 にいる確率は必ずしも高いとは限らないことが言える．

4　まとめ

本稿では，実世界の 2 人の位置の時系列データを CHMM を用いてモデリングした．今回収集したセンサデータをモデリングした結果から，2 人が位置を維持する傾向があることを遷移行列の結果から確認できた．他の行動パターンの例についてはさらに調査が必要である．

今後の課題として，各マルコフ過程間の影響を表す行列 B の値を学習過程から推定する方法が挙げられる．また，変分手法で構築した CHMM の性能について評価する必要があると考えられる．評価基準として，CHMM を用いて予測を行う場合の誤差，そして [2] で提案された CHMM との比較などが考えられる．また，本稿でデータ収集の手段として位置センサタグを利用したが，今後被験者への制限が少ない映像からのデータ収集を検討する必要があると考えている．

参考文献

[1] 青木茂樹, 岩井嘉男, 大西正輝, 小島篤博, 福永邦雄. 人物の位置・姿勢に注目した行動パターンの認識. 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol. 102, No. 318, pp. 13–18, September 2002.

[2] N. Oliver, B. Rosario, and A. Pentland. Graphical models for recognizing human interactions. In *Advances in Neural Information Processing Systems*, pp. 924 – 930, November 1998.

[3] Vladimir Pavlovic. *Dynamic Bayesian Networks for Information Fusion with Application to Human-Computer Interfaces*. PhD thesis, University of Illinois at Urbana-Champaign, January 1999.