

分離型2次元格子HMMに基づく顔画像認識

Face Recognition Based on Separable 2-D Lattice Hidden Markov Models

布目哲也
Tetsuya Nunome

南角吉彦
Yoshihiko Nankaku

徳田恵一
Keiichi Tokuda

北村正
Tadashi Kitamura

名古屋工業大学 情報工学専攻
Department of Computer Science and Engineering, Nagoya Institute of Technology

1 はじめに

画像認識では、入力画像における対象物の位置や大きさなどの変形による認識精度の低下が問題となる。この問題に対し、分離型2次元格子HMM[1]が提案されている。本研究では、分離型2次元格子HMMを顔画像のモデル化に適用し、顔画像の認識実験を行なった。また、EMアルゴリズムの改良であるDAEMアルゴリズム[2]を学習アルゴリズムに適用し、顔画像認識における有効性を検討した。

2 分離型2次元格子HMM

分離型2次元格子HMMは、横と縦の2つの状態遷移系列を持ち、それぞれ独立に遷移することが可能である。画像のモデル化においては矩形を維持しながら縦横の伸縮する画像の変形を表現することができる。図1に、グラフィカルモデルによる分離型2次元格子HMMの構造を示す。また、その尤度関数を次式に示す。

$$\begin{aligned}
 P(\mathbf{O}, \mathbf{S} | \Lambda) &= P(\mathbf{S}^{(1)})P(\mathbf{S}^{(2)})P(\mathbf{O} | \mathbf{S}^{(1)}, \mathbf{S}^{(2)}) \\
 &= \prod_{m=1,2} \left[P(S_1^{(m)}) \prod_{t^{(m)}=2}^{T^{(m)}} P(S_{t^{(m)}}^{(m)} | S_{t^{(m)}-1}^{(m)}) \right] \\
 &\quad \times \prod_{t^{(1)}=1}^{T^{(1)}} \prod_{t^{(2)}=1}^{T^{(2)}} P(O_{t^{(1)}t^{(2)}} | S_{t^{(1)}}^{(1)}, S_{t^{(2)}}^{(2)}) \quad (1)
 \end{aligned}$$

ここで、 Λ はモデルパラメータ、 $\mathbf{O} = \{O_{t^{(1)}t^{(2)}} | t^{(1)} = 1, \dots, T^{(1)}, t^{(2)} = 1, \dots, T^{(2)}\}$ は2次元に配置された観測ベクトル、 $\mathbf{S} = \{\mathbf{S}^{(1)}, \mathbf{S}^{(2)}\}$ は状態を表す隠れ変数であり、 $\mathbf{S}^{(1)}, \mathbf{S}^{(2)}$ は、それぞれ縦方向と横方向の状態遷移を表す。 $S_{t^{(m)}}^{(m)} \in \{1, \dots, K^{(m)}\}$ は、 m 次元方向のマルコフ連鎖 $S^{(m)}$ における座標 $t^{(m)}$ の状態を表す隠れ変数であり、 $K^{(m)}$ は状態数を表す。

分離型2次元格子HMMと似た構造を持つ2次元HMMとしてEmbedded HMM[3]がある。Embedded HMMは縦方向のHMM(Super-state)の各状態に、横方向のHMM(Embedded-state)が組み込まれた構造となっている。横方向の各HMMが独立しているため、少ない計算量で状態遷移を表現することができるが、縦方向の画像の連続性が失われるという問題がある。これに対し、分離型2次元格子HMMは、縦、横方向ともに連続性が保たれており、画像の位置ずれや大きさの変動を適切に表現できると考えられる。

3 変分EMアルゴリズムに基づく学習

分離型2次元格子HMMの学習は、一般的なHMMと同様にEM (Expectation Maximization) アルゴリズムによって実現できる。しかし、Eステップにおける事後

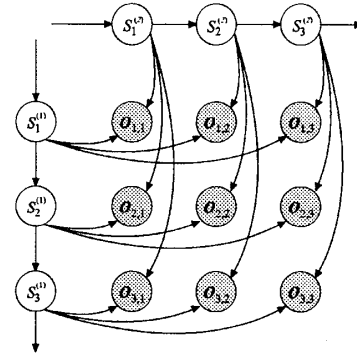


図1 分離型2次元格子HMM

確率の計算に $O(\{K^{(1)}\}^{T^{(1)}} \times \{K^{(2)}\}^{T^{(2)}})$ の演算が必要となり、直接的に計算するのは困難である。そこで、事後確率に近似を加えた変分EMアルゴリズム[1]を適用する。ここでは、状態遷移を表す隠れ変数の事後確率分布 $P(\mathbf{S} | \mathbf{O}, \Lambda)$ を以下のように近似する。

$$Q(\mathbf{S}) = Q(\mathbf{S}^{(1)})Q(\mathbf{S}^{(2)}) \quad (2)$$

式(2)は、縦方向の状態遷移と横方向の状態遷移の独立性を仮定することを意味する。変分EMアルゴリズムでは、このような構造的な近似のもとで、 $Q(\mathbf{S})$ と $P(\mathbf{S} | \mathbf{O}, \Lambda)$ のKL-divergenceが最小という意味で、最適な事後分布を推定する。導出された学習アルゴリズムは、一般的なEMアルゴリズムと同様に、EステップとMステップの繰り返しとなる。Eステップは、 $Q(\mathbf{S}^{(1)})$ と $Q(\mathbf{S}^{(2)})$ の更新を繰り返しから成り、縦方向と横方向のForward-Backwardアルゴリズムにより計算できる。変分近似を用いた場合のEステップの計算量は $O(K^{(1)}K^{(2)}T^{(1)}T^{(2)})$ となる。

4 確定的アニーリングEMアルゴリズム

変分EMアルゴリズムは、一般的なEMアルゴリズム同様に、勾配型アルゴリズムであり、局所最適性の問題を持つ。この問題を改善する手法としてDAEMアルゴリズムを適用する。DAEMアルゴリズムでは、最尤推定を“温度”に依存する自由エネルギー関数の最小化問題として再定式化し、この“温度”を用いてアニーリング過程を制御する。自由エネルギー関数は次式で定義される。

$$F_\beta = -\frac{1}{\beta} \ln \sum_{\mathbf{S}} P(\mathbf{O}, \mathbf{S} | \Lambda)^\beta \quad (3)$$

ここで、 β は逆温度を表す温度パラメータである。DAEMアルゴリズムでは、式(3)の自由エネルギー関数を最大化することにより、尤度関数 $P(\mathbf{O} | \Lambda)$ を極大に導く。自

由エネルギー関数から導出される事後確率 $Q_\beta(S)$ は、次式で表される。

$$Q_\beta(S) = \frac{P(O, S|\Lambda)^\beta}{\sum_S P(O, S|\Lambda)^\beta} \quad (4)$$

ここで、 β が小さな値のとき、 $Q_\beta(S)$ は一様分布に近い分布となり、尤度関数が局所解の少ない単純な形状となる。また、 $\beta=1$ のとき $Q_\beta(S)$ は事後確率分布 $P(S|O, \Lambda)$ と一致する。よって、 β を 0 から 1 にゆっくりと増加させながら EM ステップを繰り返すことにより、局所最適性の問題を改善することができる。

4.1 DAEM の分離型 2 次元格子 HMM への適用

DAEM アルゴリズムを分離型 2 次元格子 HMM の学習に適用する場合、変分近似による事後確率分布 $Q_\beta(S^{(m)})$ は次式のように書き換えることができる。

$$Q_\beta(S^{(m)}) = \frac{P(S^{(m)}|\Lambda)^\beta \prod_{t^{(m)}=1}^{T^{(m)}} h(t^{(m)}, S_{t^{(m)}}^{(m)})^\beta}{\sum_{S^{(m)}} P(S^{(m)}|\Lambda)^\beta \prod_{t^{(m)}=1}^{T^{(m)}} h(t^{(m)}, S_{t^{(m)}}^{(m)})^\beta} \quad (5)$$

$$\ln h(t^{(m)}, S_{t^{(m)}}^{(m)}) = \sum_{S^{(n)}} Q(S^{(n)}) \left[\sum_{t^{(n)}} \ln P(O_{t^{(1)}t^{(2)}} | S_{t^{(1)}}^{(1)}, S_{t^{(2)}}^{(2)}) \right] \quad (6)$$

ただし、 $n \neq m$ とする。ここで、式 (5) は、一般的な HMM の事後分布と同じ形をしており、 $P(S^{(m)}|\Lambda)^\beta$ と $h(t^{(m)}, S_{t^{(m)}}^{(m)})^\beta$ をそれぞれ出力確率と遷移確率に置き換えて、Forward-Backward アルゴリズムを実行することにより、計算可能である。

5 認識実験

本研究では XM2VTS データベース [4] を用いて顔画像モデルを作成し、認識実験を行った。画像は、位置・大きさを正規化した画像とガウス乱数により変動させた画像の 2 種類を用いた。使用画像データを表 1 に示す。

状態数 32×32 における実験結果を表 2 に示す。状態出力確率はガウス分布とした。ここで、“SL-EM”、“SL-DA” はそれぞれ変分 EM アルゴリズム、DAEM アルゴリズムを適用した分離型 2 次元格子 HMM を表す。“Emb-V”、“Emb-H” はそれぞれ Super-state を縦、横とした Embedded HMM を表す。Embedded HMM の学習には EM アルゴリズムを用いた。“SL-DA” における、温度パラメータ β の更新は予備実験の結果から $(i/I)^{\frac{2}{3}}$ 、 $I=1000$ とした。ただし、 i と I はそれぞれ、現在の更新回数と総更新回数を表す。また、“Gauss” は状態遷移を持たないガウス分布を表しており、状態遷移を等分割に固定した分離型 2 次元格子 HMM と等価である。

結果より、変動無しにおいて、従来法の 69% (“Emb-V”) に対し、“SL-EM” では 80%、“SL-DA” では 84% の認識率が得られた。また、変動有りにおいても従来法に比べ提案法がより高い認識率を得ており、分離型 2 次元格子 HMM の画像変動に対する有効性が確認できる。また、“SL-EM” と “SL-DA” の結果を比較すると、変動有り・無しともに、“SL-DA” が高い認識率を示しており、DAEM アルゴリズムの効果が確認できる。

顔画像モデルの平均ベクトルを図 2 に示す。“Gauss” では変動有りを用いた場合、変動が吸収されず、ぼやけた

表 1 使用画像データ

| | |
|--------|-------------------------------|
| 画像サイズ | 64 × 64 |
| 階調数 | 256, grayscale |
| 特徴量 | 画素の輝度値 |
| 学習データ | 7 枚 × 100 名分 |
| テストデータ | 1 枚 × 100 名分 |
| 位置の変動 | 中心から 40 × 20 以内 |
| 大きさの変動 | 500 × 500 ~ 600 × 600 (縦横比固定) |

表 2 実験結果 (%)

| 変動 | Gauss | Emb-V | Emb-H | SL-EM | SL-DA |
|----|-------|-------|-------|-------|-------|
| 無 | 60 | 69 | 62 | 80 | 84 |
| 有 | 17 | 56 | 58 | 70 | 76 |

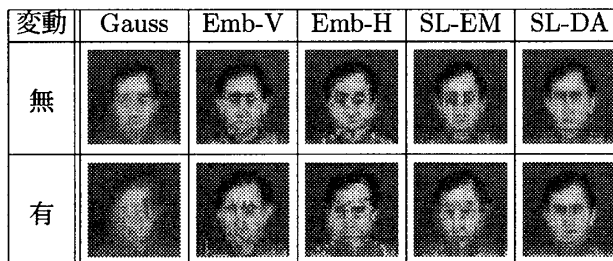


図 2 顔画像モデルの平均ベクトル

平均ベクトルが作成されている。また、“Emb-V”、“Emb-H” では、各 Embedded-state 方向の連続性が失われているのに対し、“SL-DA” は顔のパターンを保持したモデル化が行われていることが確認できる。これは、分離型 2 次元格子 HMM が画像の伸縮を適切に吸収しているためと考えられる。また、“SL-DA” では、変動有り・無しにおいて、ほぼ同じような平均ベクトルが得られており、画像の変動による初期値の影響が緩和されていることが確認できる。

6 むすび

本研究では、分離型 2 次元格子 HMM を用いて顔画像の認識実験を行った。その結果、従来法の Embedded HMM に比べ、高い認識率が得られた。また、学習において DAEM アルゴリズムを適用することにより、認識率が改善されることを確認した。今後の課題として、適切な特徴量の選択が挙げられる。

参考文献

- [1] 大野也澄志, 南角吉彦, 徳田恵一, 北村正, Zoubin Ghahramani, “変分近似に基づいた分離型 2 次元 HMM 学習アルゴリズム,” 信学技報, Vol.102, No.652, PRMU2002-211, pp.1-6, Feb. 2003.
- [2] 上田修功, 中野良平, “確定的アニーリング EM アルゴリズム,” 電子情報通信学会論文誌, J80-D-II, no.1, pp.267-276, Jan. 1997.
- [3] S. Kuo, and O.E. Agazzi, “Machine Vision for Keyword Spotting using Pseudo 2D Hidden Markov Models,” In Proc. International Conference on Acoustics, Speech and Signal Processing. Minneapolis, Minnesota, Apr. 1993.
- [4] K. Messer, J. Mates, J. Kitter, J. Luetten, and G. Maitre, “XM2VTSDB: The Extended M2VTS Database,” Audio and Video-Based Biometric Person Authentication, pp.72-77, 1999.