

分離型2次元格子HMMを用いた顔画像認識のための状態クラスタリング

State Clustering for Face Recognition Using Separable 2-D Lattice HMM

倉田大介 南角吉彦 徳田恵一 北村正
Daisuke Kurata Yoshihiko Nankaku Keiichi Tokuda 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 まえがき

画像認識では対象物の移動や伸縮、回転などによる変形に対応する必要がある。これまでに画像の位置や大きさの変動にロバストなモデルのひとつとして分離型2次元格子HMM[1]が提案されている。本稿では、分離型2次元格子HMMに基づく画像認識において、画像データの生成過程をより適切に表現するための状態クラスタリングについて述べる。従来法では、各モデルが独立に学習されていたのに対し、提案法ではモデル間で状態を共有することにより、識別に適したモデルパラメータの推定が可能となることが期待できる。また、顔画像認識実験により提案法の有効性を検討する。

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つの独立した状態遷移系列をもち、2次元データの矩形を維持した縦横の伸縮を表現することができることから、画像の位置や大きさの変動に対応することができる。分離型2次元格子HMMを図1に、その尤度関数を次式に示す。

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

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、それぞれ横方向と縦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。本研究では状態出力確率を画像の輝度値を特徴量とするガウス分布と仮定する。

2.1 変分EMアルゴリズム

分離型2次元格子HMMのパラメータ推定には一般的なHMMと同様に、EMアルゴリズム (Expectation Maximization algorithm)を用いる。しかし、分離型2次元格子HMMは2つの状態遷移系列を持ち、通常のEMアルゴリズムでは計算量が膨大となるため、変分近似を用いた変分EMアルゴリズム [1]を適用する。変分EMアルゴリズムにおける Q 関数を次式に示す。

$$Q(\Lambda, \bar{\Lambda}) = \sum_{\mathbf{S}} Q(\mathbf{S}) \log P(\mathbf{O}, \mathbf{S} | \bar{\Lambda}) \quad (2)$$

ここで、 $Q(\mathbf{S})$ は事後確率分布 $P(\mathbf{S} | \mathbf{O}, \Lambda)$ の近似分布である。変分EMアルゴリズムでは、式(2)の Q 関数を

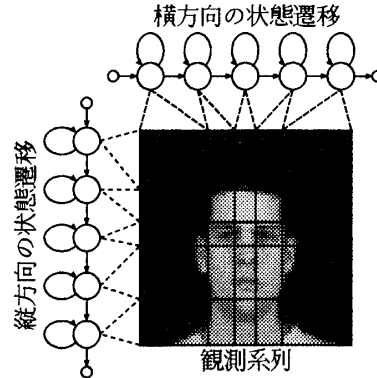


図1 分離型2次元格子HMM

最大化することで尤度を極大に導く。分離型2次元格子HMMの学習では近似分布 $Q(\mathbf{S})$ に以下の制約を与える。

$$Q(\mathbf{S}) = Q(\mathbf{S}^{(1)})Q(\mathbf{S}^{(2)}) \quad (3)$$

ただし、 $\sum_{\mathbf{S}^{(m)}} Q(\mathbf{S}^{(m)}) = 1$ とする。式(3)は変分EMアルゴリズムにおける近似分布 $Q(\mathbf{S})$ が横、縦方向の状態遷移の独立性を仮定することを意味する。

3 状態クラスタリング

統計モデルではデータの生成過程を良く表現するモデル構造を与えることが重要である。顔画像認識では背景等の領域は被験者が変わっても同一の情報源からデータが生成されたと考えられるため、このような個人識別に重要ではない状態をモデル間で共有することは妥当であると考えられる。

本研究では分離型2次元格子HMMにおける状態クラスタリングを検討する。状態クラスタリングの概要を図2に示す。状態クラスタリングでは状態を各モデル固有の状態とモデル間で共有する状態に分割する。各モデル固有の状態は被験者を識別する際に重要な個人性を多くもち、モデル間で共有する状態は個人性以外の情報をもつ。状態共有によって背景などの領域に対する尤度がモデル間で一定となり、背景などの領域の変動による影響を削減することができる。また、共有構造を考慮したモデルのパラメータ再学習を行うことで、より識別的なモデル学習が可能となることが期待できる。クラスタリング基準は後述する分散基準と尤度基準を用いる。

3.1 分散基準

全モデルの学習データを用いて学習したモデルを共通モデルと呼ぶこととする。分散基準では共通モデルの分散が小さい状態 N 個を個人性に関する情報を持たない状態として全モデルで共有する。

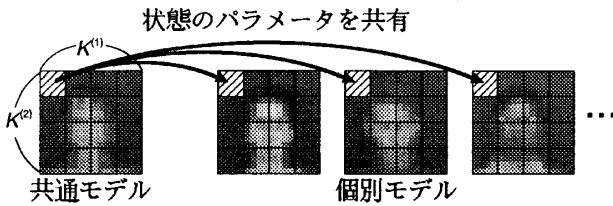


図2 状態クラスタリング

3.2 尤度基準

尤度基準では、全学習データに対する尤度を最大化する状態の共有構造を選択する。式(3)のQ関数を状態単位に分割し、尤度の増加が小さい状態N個を個人性に関する情報を持たない状態として全モデルで共有する。Q関数の状態単位での尤度の変化量は次式により各状態で独立に計算できる。

$$\Delta Q(i, j) = \left[\sum_c Q_c(i, j) \right] - Q_{all}(i, j) \quad (4)$$

$$Q_c(i, j) = - \left[\sum_{t^{(1)}} \sum_{t^{(2)}} \gamma_{t^{(1)}t^{(2)}}(i, j) \right] \log |\Sigma(i, j)| \quad (5)$$

ここで、(i, j)は横、縦方向の状態がそれぞれi, jであることを表す。Q_c(i, j)は顔画像モデルcにおけるQ関数の状態(i, j)の項であり、Q_{all}(i, j)は共通モデルにおけるQ関数の状態(i, j)の項である。また、Σ(i, j)は状態(i, j)における共分散行列であり、γ_{t⁽¹⁾t⁽²⁾}(i, j)は(S_{t⁽¹⁾}⁽¹⁾, S_{t⁽²⁾}⁽²⁾)が(i, j)である確率P(S_{t⁽¹⁾}⁽¹⁾ = i, S_{t⁽²⁾}⁽²⁾ = j | O, Λ)である。次元mにおける時間t^(m)での状態kの停留確率をγ_{t^(m)}(k) = P(S_{t^(m)}^(m) = k | O, Λ)とすると式(3)より、γ_{t⁽¹⁾t⁽²⁾}(i, j) = γ_{t⁽¹⁾}(i)γ_{t⁽²⁾}(j)と書くことができる。実験では、Q_c(i, j)を計算する際のγ_{t⁽¹⁾t⁽²⁾}(i, j)は共通モデルで求めた値を用いることとした。

4 認識実験

本研究ではXM2VTSデータベース[2]を用いた顔画像認識実験を行った。実験条件を表1に示す。被験者は100名とし、各被験者の顔画像を分離型2次元格子HMMでモデル化した。提案法として、状態クラスタリングを用いて状態の共有構造を決定した。また、クラスタリング後に共有構造を考慮したモデルパラメータの再推定を行った。

分散基準と尤度基準を用いたクラスタリングによる結果をそれぞれ図4, 5に示す。図中の共有構造は共有部分を白、非共有部分を黒で表す。図4より、分散基準において共有状態数が90, 110, 130のとき、従来法を上回る認識率が得られた。このとき、背景領域を表す状態が共有されていることが図より確認することができる。この結果から、個人識別において、適切な共有構造を選択することにより、認識性能が改善することを確認した。その一方で、図5より、尤度基準では従来法を上回る認識率を得ることはできなかった。尤度基準によるクラスタリングでは、顔領域を表す状態が共有され、逆に背景領域を表す状態が分割されている。この原因として、画像撮影時におけるカメラの自動補正等によって、背景領域に個人性が含まれてしまったことが考えられる。こ

表1 実験条件

画像サイズ	64 × 64(pixel)	
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	
HMM状態数	16 × 16	
共有状態数	従来法	提案法
	0	70, 90, 110, 130, 150

モデル	共通モデル	モデルA	モデルB	モデルC
平均				
分散				

図3 モデルの平均と分散

共有状態数	0	70	90	110	130	150
共有構造						
認識率(%)	78	78	80	85	83	72

図4 分散基準での実験結果

共有状態数	0	70	90	110	130	150
共有構造						
認識率(%)	78	57	56	49	51	47

図5 尤度基準での実験結果

のような変動を正規化することで、尤度基準により適切な状態の共有構造が得られる可能性がある。また、本研究では背景領域が一様な画像データを用いたが、複雑な背景領域を持ったデータでは、分散基準による共有構造の決定は困難であり、このような場合にも適切なクラスタリングが可能な基準の検討が今後の課題である。

5 むすび

本稿では分離型2次元格子HMMに基づく顔画像認識において状態クラスタリングを検討した。認識実験より、顔画像の背景領域を適切に共有することによって認識率が改善することを確認した。今後の課題としては、尤度基準における状態クラスタリングの改善や、識別的基準等の他のクラスタリング基準の検討が挙げられる。

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] K. Messer, J. Mates, J. Kitter, J. Luetttin, and G. Maitre, “XM2VTSDB: The Extended M2VTS Database,” Audio and Video-Based Biometric Person Authentication, pp.72-77, 1999.