

改良型 Embedded HMM を用いた顔画像認識の検討

A Study on Face Recognition using Improved Embedded HMM

菊池 洋光 甲藤 二郎
Hiromitsu Kikuchi Jiro Katto
早稲田大学 大学院 理工学研究科

1. まえがき

静止画や動画から顔画像を認識する技術は、金融機関 ATM の個人認証や、建物の入退出監視、ロボットとの対話、データベースからの人物検索システムなど様々な応用が考えられ、盛んに研究が行われている。

代表的な認識手法は、幾何学的特徴に基づく方法やテンプレートに基づく方法、固有顔による方法[1]などが挙げられる。従来研究では表情の変化、照明条件の変動などの要因に対して、撮影時に制約を課すことで安定な認識を実現してきた。しかし顔画像認識システムは顔の向きや性別・人種の違いの他、照明変化や位置変化を伴う環境下でもロバストに動作することが望まれる。

近年モデルベースによる認識方法として HMM を用いた顔画像認識手法が報告されている[2][3][5]。HMM は非線形伸縮マッチングの機能を持っており、音声認識やジェスチャ認識に多く用いられている。本稿では、局所的な顔の変化や位置のずれにロバストな認識方法として、画像を 2 次元の HMM でモデル化し、認識対象画像を部分領域ごとに分割して評価を行う方法を提案する。

2. 従来手法と問題点

2.1 A Pseudo 2-D HMM

2 次元データをモデル化するために、1 次元 HMM を拡張した Pseudo 2-D HMM と呼ばれる構造が提案されている[2]。これは画像の各ラインの終端にマーカブロックを加えて、水平方向の各 HMM の終端に end-of-line 状態を付加したものである。end-of-line 状態は、同じ行の HMM の開始状態に戻る遷移と、次行の HMM の開始状態に移る遷移の 2 通りの状態遷移が可能である。マーカブロックの挿入や状態の追加が必要になるという欠点が存在する。

2.2 An Embedded HMM

Embedded HMM の構成要素を図 1 に示す。2.1 のモデルと同様に 2 次元データをモデル化する HMM であるが、その構造が異なる。この構造を P2-D HMM と呼ぶこともあるが[4]ここでは文献[3][5]の表現を引用し Embedded HMM と呼ぶ。Embedded HMM は “super state” と “state” と呼ばれる 2 種類の状態がある。Embedded HMM そのものは “super state” から構成され、各々の内部に “state” が構成される。すなわち各 super state の中には独立の left-to-right 型 HMM が含まれる構造となり、これによって横一列の画像特徴の整合度を評価する。Embedded HMM は super state 内の状態遷移の途中で、他の super state へ遷移することを許していない。super state の内部にある HMM によって水平方向に対する尤度を求め、これを用いて super state の遷

移方向の尤度を求めるため、上下左右同時に最適化されおらず、2.1 と同様に擬似的な 2 次元構造の HMM となる。

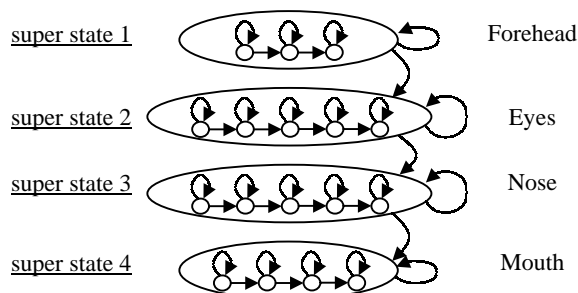


図 1. An Embedded HMM の例

3. 提案手法

3.1 提案手法の概要

Embedded HMM(以下 EHMM)は顔の向きや傾きの変化にロバストな認識が実現できる。しかし観測する特徴ベクトルが画素値もしくは空間周波数である以上、影の影響や眼鏡の着脱等、学習時の画像と比較して著しい部分変化のある画像には、十分な認識が行えない。これは認識時に、顔全体に対して評価を行っているためである。顔の一部を隠されても、人間はそれ以外の情報を元に個人を特定できる。

そこで本稿では、学習時は完全な顔画像を用いるものの、認識時はテスト画像を部分領域に分割し、各々で評価を行う方法を提案する。この改良方法によって部分的な隠れや変化に対してロバストになることが期待される。

本システムの概要は、特徴抽出部、EHMM 学習部、認識部の 3 つに分けることができる。以下に詳細を述べる。

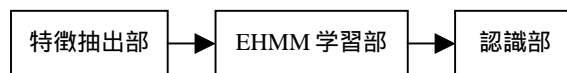


図 2. 提案手法処理の流れ

3.2 特徴抽出部について

観測可能な画像特徴として、本稿では DCT 係数を用いる。各画像に対し 8×8 のウィンドウを縦横それぞれ 2pixel オーバラップして走査する。ウィンドウ内の 64 次元のデータに DCT を施し 9 次元のデータに圧縮する。これによって、ブロック内平均階調値を特徴として用いるよりも情報量を多くでき、画素をそのまま用いるよりも次元を削減することができる。

3.3 Embedded HMM のパラメータ

1 人の人物は、1 つの EHMM を構成する。つまり EHMM パラメータが個人を特徴づけるパラメータとなる。EHMM のパラメータは以下の要素からなる。

super state に関するパラメータ

- super state 数: N_0

- super state の初期状態分布確率: $\pi_{0,i}$
 - super state の状態遷移確率: $a_{0,ij}$
- super state 内に存在する HMM に関するパラメータ
- super state k の内部 HMM の状態数: N_1^k
 - super state k の内部 HMM の初期状態分布確率: $\pi_{1,i}^k$
 - super state k の内部 HMM の状態遷移確率: $a_{1,ij}^k$
 - super state k の内部 HMM の状態 i が与えられたとき、ブロック位置 $(x, y) = (t_1, t_0)$ において特徴ベクトル O_{t_0,t_1} を観測する出力確率: $b_i^k(O_{t_0,t_1})$

ただし、 $b_i^k(O_{t_0,t_1})$ は無相関多次元正規分布と仮定する。すなわち共分散行列は対角行列となる。

3.4 Embedded HMM の学習部について

1 人物ごと、複数の画像を用いて学習を行い、EHMM パラメータを推定する。EHMM の学習部のフローチャートを図 3 に示す。

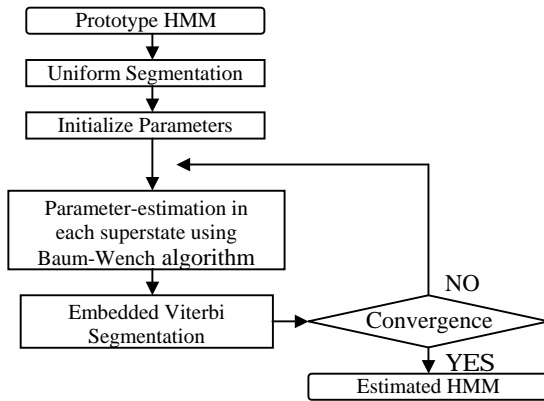


図 3 . 学習の流れ

super state 内部に存在する HMM パラメータは Baum-Welch アルゴリズムを用いて再推定され、この結果 super state 内部の HMM パラメータ $\lambda^{(1)} \sim \lambda^{(N_0)}$ が求まる。

($\lambda^{(k)} = \{ \pi_{1,i}^k, a_{1,ij}^k, b_i^k(O_{t_0,t_1}) \}$) .
 この $\lambda^{(k)}$ を用いて横方向のデータ系列尤度 $P(O_{t_0,1} \dots O_{t_0,T_1}, q_{1,1}^{(t_0)} \dots q_{1,T_1}^{(t_0)} | \lambda^{(k)})$ を全てのブロック番号 ($1 \leq t_1 \leq T_1, 1 \leq t_0 \leq T_0$)、全ての super state ($1 \leq k \leq N_0$) について求め、Embedded Viterbi アルゴリズムによって EHMM の最終尤度 $P(O_{1,1} \dots O_{1,T_1}, \dots, O_{T_0,1} \dots O_{T_0,T_1}, q_{0,1} \dots q_{0,T_0} | \lambda)$ が計算される。($\lambda = \{ \pi_{0,i}, a_{0,ij}, \lambda^{(1)}, \dots, \lambda^{(N_0)} \}$)

$a_{0,ij}$ の再推定は、Embedded Viterbi Segmentation によって最適状態パスを決定した後、super state i から遷移する回数と super state i から j に遷移する回数の比を求めることによって更新を行う。

3.5 部分画像を用いた認識部について

部分画像を用いた人物認識部のイメージを図 4 に示す。部分画像とは、テスト画像を幾つかに分割し、矩形に切り出したものである。この部分画像を逐次、学習後の人物毎 EHMM へ入力し、最終尤度の中で最大値を選ぶ。この値を人物のマッチング度合いを測る尤度とみなし、一番高い尤度を出力する EHMM のラベルを認識人物結果とする。

部分画像の評価を行う際、切り出した位置から初期状態分布 $\pi_{0,i}, \pi_{1,i}^k$ にある程度拘束条件を加える必要がある。これは例えば、画像が鼻から下を写していると思われる場合、

額や眼の super state から評価を始めさせないための拘束である。

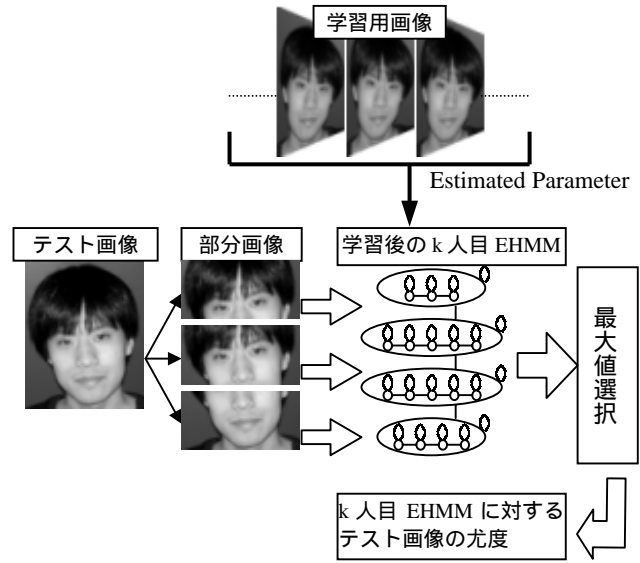


図 4 . 部分画像を用いた評価方法

4 . 実験

本提案手法による実験を縦 52 × 横 40pixel の画像に対して行う。学習用画像には照明条件の変化や位置のずれ、表情の変化を含み、背景のないほぼ正面を向いた画像とする。提案手法による実験結果は発表時に行う。

5 . まとめ

本稿では個人ごとの顔画像を Embedded HMM でモデル化し、部分領域に分割し認識を行う手法を提案した。今後は様々なデータに対して評価を行うと共に、一般背景下でも適用可能なアルゴリズムの検討を行っていく予定である。

参考文献

- [1] M.Turk and A.Pentland, "Face recognition using eigenfaces," in Proceedings of International Conference on Pattern Recognition, pp.586-591, 1991.
- [2] R.Samaria and S.Young, "HMM based architecture for face identification," Proceedings of Image and Computer Vision, vol.12, pp.537-543, October 1994.
- [3] A.V.Nefian and M.H.Hayes, "Face recognition using an embedded HMM," IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing, pp.3553-3556, 1999.
- [4] S.Kuo and O.Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-D Hidden Markov Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, pp.842-848, August 1994.
- [5] 杉江 嘉昭, 小林 哲則, "Dempster-Shafer 理論を用いた音声・画像情報の統合による個人認識システム," 信学技報 PRMU2001-122, 2001.11.