

改良型 Embedded HMM を用いた顔画像認識の検討

菊池 洋光[†] 甲藤 二郎[†]

[†]早稲田大学大学院 理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

E-mail: † {kikuchi, katto}@katto.comm.waseda.ac.jp

あらまし モデルベースによる顔画像認識手法として、複数のHMM(Hidden Markov Model)を空間的に配置し、一定方向へ連結する Embedded HMM と呼ばれる手法が報告されている。Embedded HMM を用いた手法は主成分分析を用いた代表的な認識方法と比べ、顔の向きや位置の変化にロバストな学習・認識が実現できるが、画像平面内で顔が回転した場合などにおいて、顔の構造を十分に表現できない恐れがあった。そこで本稿では、2種類の方向の異なる構造を併せもつ Embedded HMM を提案する。また、眼鏡の有無や部分的な隠れなどの問題に対応するために、認識対象として尤もらしい画像の部分領域のみに着目し評価を行う認識アルゴリズムを提案する。

キーワード 顔画像認識・隠れマルコフモデル・擬似2次元HMM・Embedded HMM

A Study on Face Recognition using improved Embedded HMM

Hirimitsu KIKUCHI[†] and Jiro KATTO[†]

[†] Graduate School of Science and Engineering, WASEDA University 3-4-1 Okubo, Shinjyuku-ku, Tokyo, 169-8555
Japan

E-mail: † {kikuchi, katto}@katto.comm.waseda.ac.jp

Abstract: As one of model-based methods for face recognition, a technique called Embedded HMM (Hidden Markov Model) was proposed, in which two or more HMMs are arranged spatially and connected in the fixed direction. This Embedded HMM can provide robust learning and recognition against changes of direction and positions of a face compared with the conventional methods based on PCA (Principle Component Analysis). However, when a face is rotated within an image plane, there is a possibility that face structure could not be expressed well by the Embedded HMM. Therefore, in this paper, we propose a novel Embedded HMM that has two kinds of structures with different direction. Furthermore, in order to improve its performance against partial hiding of a face, such as, existence of glasses, we propose a new recognition algorithm using plausible parts of face image.

Keyword Face Recognition, Hidden Markov Model, Pseudo 2-D HMM, Embedded HMM,

1. まえがき

画像や映像から顔を発見・認識する技術は、バイオメトリクス認証に代表されるセキュリティシステムへの応用の他、マン・マシンインターフェースやビデオデータベースの顔サーチ、顔インデックス等への適用など、様々な利用法が考えられ、盛んに研究が行われている。

コンピュータによる顔の自動認識は、画像中から顔を発見する検出部と、検出した顔画像を用いて本人を特定する識別部の2つのプロセスに大きく分けられる。

従来手法[1,2]の多くは顔が正確に切り出されていることを仮定しており、表情の変化や照明の変動などの要因に対しては、撮影時に制約を課すことで安定な認識を実現させてきた。つまり、顔検出時の結果が理想的で、入力される画像は学習画像に対して変動が少ないという前提の下で議論されていた。しかし顔検出自

体が一つの大きな問題となっており、画像から顔を正確に切り出すことは容易ではない。仮に顔の位置を制限できるシステムであっても、照明や表情の変化を避けて議論することは一般的ではなく、実環境下でも耐えうる性能を得るには様々な変動を吸収し、且つ安定に動作する認識器の実現が必要となる。

以上の問題を考慮して、本稿では以下に示す新たな認識アルゴリズムの実現を目指し、その有効性を確認することを目的とする。すなわち、(1)顔の位置が多少ずれていても認識可能なアルゴリズムの実現、(2)表情や髪型の変化、眼鏡の有無にもロバストな認識アルゴリズムの実現、を目的としている。

顔画像認識の代表的な手法には、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[3]、主成分分析による方法[1]などが挙げられる。また、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[3,4,5].その中でも Embedded HMM を用いた手法[5]は主成分分析を用いた代表的な認識方法と比べ、顔の向きや位置の変化にロバストな学習・認識が実現できることが報告されている。ただし、画像平面内で顔が回転した場合などにおいて、顔の構造を十分に表現できない問題も指摘されている。そこで今回は、顔の様々な変化にも対応させるため、2種類の構造をもつ Embedded HMM を考え、併せてその評価方法を検討した。また、顔画像認識に共通する課題である眼鏡の有無や顔の部分的な隠れに対し、本稿ではより尤もらしいと思われる部分領域に着目して評価を行う手法を提案する。

2. 従来手法と問題点

2.1. A One-dimensional HMM

顔画像の重要な特徴は、髪、額、両目、鼻、そして口などである。それらの特徴は顔が多少傾いても、また向きが多少変わったとしても、上から下に向かって順番に現れる。したがって、図1に示すような1つの状態に顔の領域を割り当てた1次元HMMを用いることで、顔の構造をモデル化することができる[3][4]。

各状態において観測される特徴量(特徴ベクトル)は、顔画像を水平方向に区切って得られる、複数の行から抽出されることになる。このため画像の横サイズは終始固定となり、この手法による認識システムは学習時と同じ画像サイズしか扱うことができない。

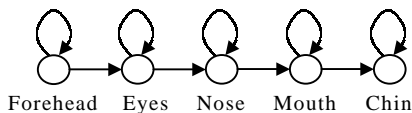


図1 顔認識のための1次元HMM

2.2. An Embedded HMM

2次元データをより柔軟にモデル化することを目的として、1次元HMMを拡張したEmbedded HMMと呼ばれる構造が提案されている[5]。Embedded HMMは1次元HMMの各状態を1つのHMMとみなして一般化され、その構成要素は図2のように示される。このような2次元構造のHMMをP2-D HMM(pseudo 2-D HMM)と呼ぶこともあるが[6]ここでは文献[5]の表現を引用しEmbedded HMM(以下EHMM)と呼ぶ。

EHMMは“super state”と“state”と呼ばれる2種類の状態から成る。“super state”が1つのleft-to-right型HMMを構成し、各super stateの内部に“state(embedded state)”が構成される。すなわちsuper stateの中にleft-to-right型HMMが独立に含まれる構造となる。EHMMはsuper state内の状態遷移の途中で他のsuper stateへ遷移することが許されていない。super

stateの内部のleft-to-right型HMMによって水平方向の画像の整合度を評価し、この整合度に基づいてsuper state遷移方向のHMMの尤度、すなわち全体の整合度を評価するため、水平方向と鉛直方向が同時に最適化されておらず、擬似的な2次元構造のHMMとなる。このため水平方向および鉛直方向の独立な歪みを補正する能力には優れるものの、それらの複合した歪みには対応できないという問題が挙げられる。

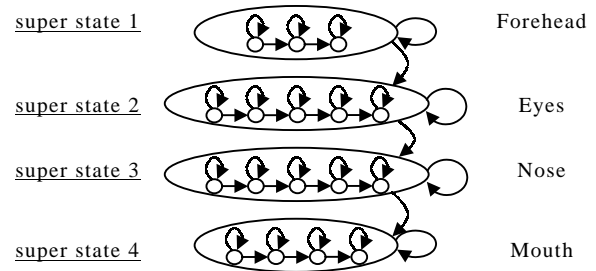


図2 顔画像とEmbedded HMMの依存関係

2.2.1. Embedded HMMのモデルパラメータ

1人の人物は1つのEHMMによって表現される。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に内包された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: $b_i^k(O_{t_0,t_1})$

2.2.2. 学習アルゴリズム

従来EHMM[5]の学習法について以下に示す。

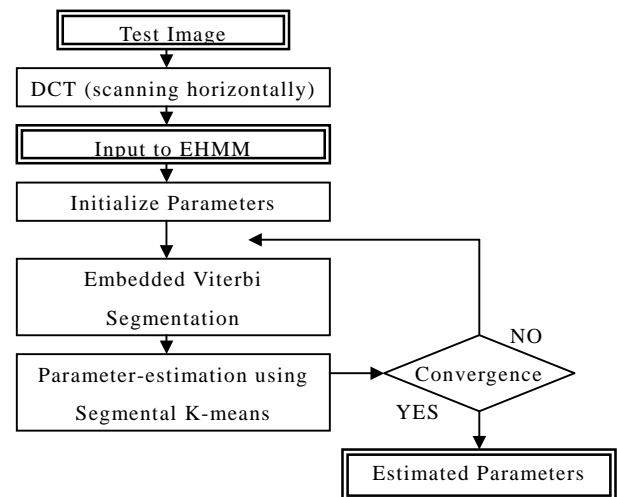


図3 EHMMの学習アルゴリズム

従来手法では ,super state パラメータ ,embedded state パラメータともに ,Segmental K-means 法に基づいて学習が行われる .

2.2.3. Segmental K-means

Segmental K-means による学習手続きは ,クラスタリングでよく知られる K-means 繰り返し手法の変形である . まず , ある初期パラメータを与え , Viterbi アルゴリズムによって顔画像を super state, embedded state に分割する (Viterbi Segmentation). 分割した結果は , 現在のパラメータにしたがったときの最尤推定になっており , この最適パスに基づいて状態遷移確率や観測確率密度関数を更新する . 全体の尤度が収束するまでこの操作を繰り返すことでパラメータ推定を行うことができる .

3. 提案手法

3.1. ハイブリッド型 Embedded HMM

3.1.1. 概要

EHMM は , まず super state に内包される HMM によって水平方向の尤度を評価し , その出力尤度に基づいて鉛直方向の尤度を評価するという 2 段階の手順を踏んでいる (図 4). このため横向きの顔など水平方向の歪みは吸収することが可能であるが , 一方で鉛直方向の歪みがある場合などを考えると , 全ての画像が水平方向の尤度に基づいて評価されることは , 必ずしも望ましいとは言えない . 初めに鉛直方向の尤度を評価し , 鉛直方向の尤度に基づいて水平方向の尤度を評価する , 図 5 のような構造を考えることも可能である .

そこで本稿では , 従来のモデルに加えて , 90 度回転させた図 5 のような EHMM を導入し , 2 つ併せて評価するハイブリッド型の方法を検討した . 従来の EHMM と同様 , 顔画像の上下左右同時に最適化することはできないが , 顔の様々な変化に対してよりロバストになることが期待できる .

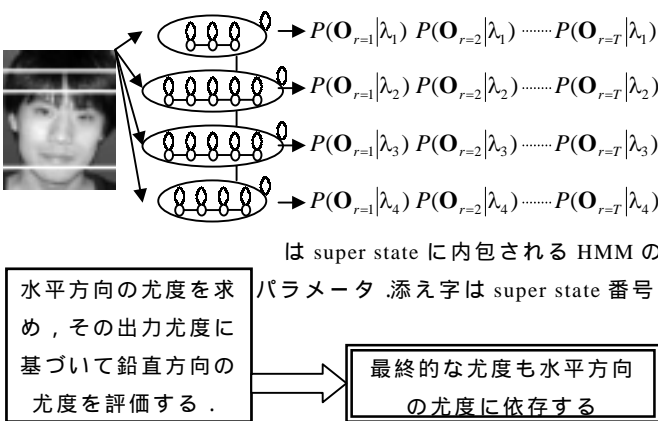


図 4 従来の Embedded HMM の尤度評価手順

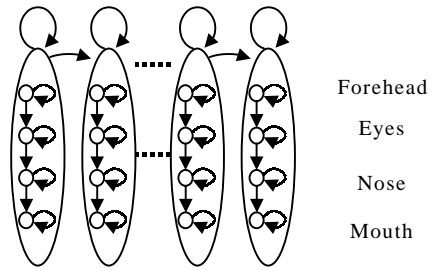


図 5 提案手法で追加する Embedded HMM

3.1.2. 特徴抽出部

顔画像の特徴量として DCT 係数を用いる . DCT 係数を用いることにより空間方向の冗長性を軽減した特徴表現が可能となる . またブロック内平均値を用いるよりも画像の性質を保存することが可能であり , 画素値を直接用いるよりも次元を削減することができる .

DCT は 8×8 pixel の正方ブロックに対して行い , 得られた 64 個の係数のうち , 本稿では直流成分を除いた左上方 35 個を特徴として採用した . ヒストグラム処理だけでは対処しきれなかったコントラストの違いに対し , 直流成分を除くことによって , 特徴量の偏りが軽減されると期待できる . さらに今回は各 DCT 係数において , レンジを一定にさせるための正規化処理を行った . 正規化後の DCT 係数は , 物理量としては意味をもたないが , 観測される特徴量を多次元正規分布で近似して表現する上で重要な処理になる .

3.1.3. 学習部

ハイブリッド型の 2 つの EHMM は共通の画像を用いて独立に学習が行われる . 学習の流れを以下に示す .

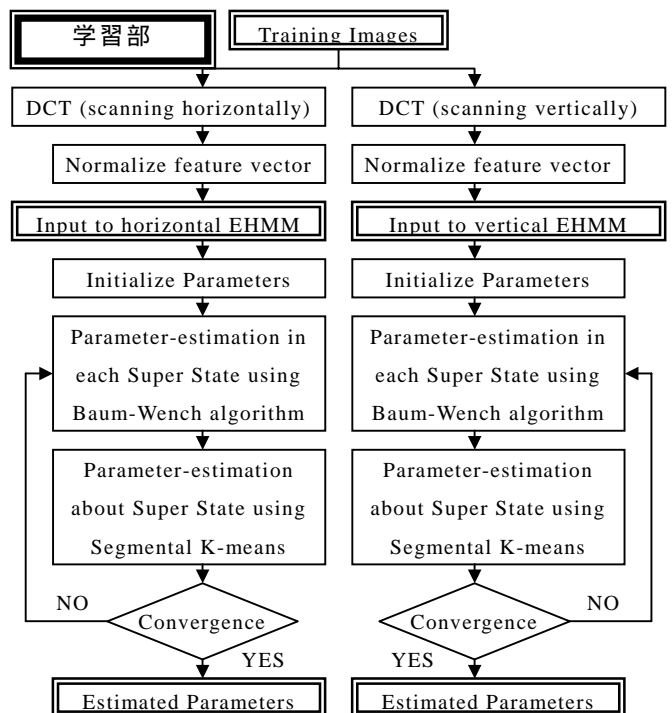


図 5 ハイブリッド型 Embedded HMM の学習の流れ

3.1.4. 認識部

学習部と同様、認識部もまた2つのEHMMによって独立に尤度を計算するが、評価時には2つの尤度の最大値を選択して行うことにした。認識の流れを以下に示す。

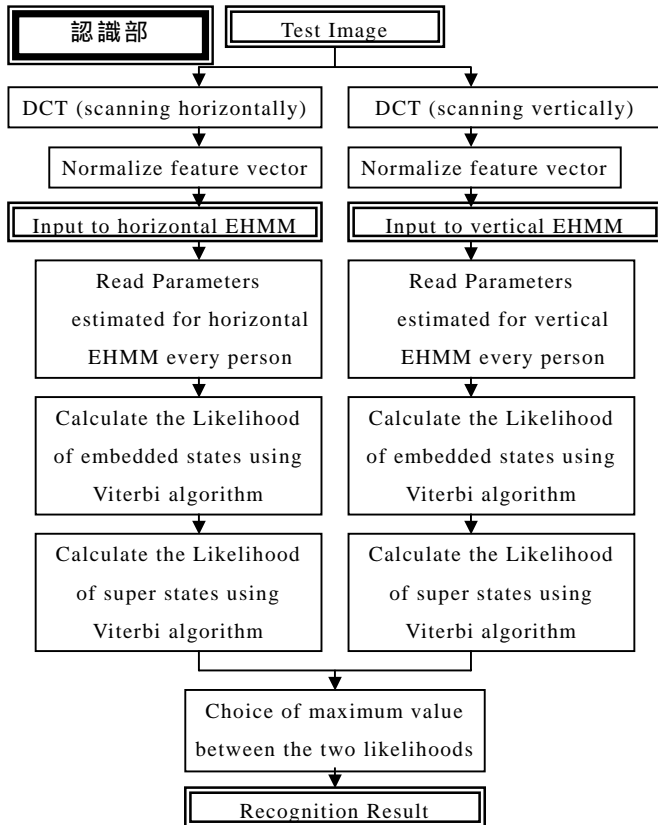


図6 ハイブリッド型 Embedded HMM の認識の流れ

3.2. 眼鏡の有無や部分的な隠れへの対応

DCT 係数を特徴として用いた EHMM は、従来手法のままでも、顔の向きや位置の変化に比較的ロバストな認識が実現できる。しかし影の影響や眼鏡の着脱等、学習画像と比べ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、入力される顔画像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基に個人を特定できる(図7)。

そこで本稿では、学習時には理想的な正面向き顔画像を用いるものの、認識時にはテスト画像全てではなく、尤もらしいと判断される部分領域のみを用いて評価を行う方法を検討した。



図7 眼鏡や隠れがある画像の例

学習画像とは異なった眼鏡や部分的な変化がある場合、その変化領域はどの super state の内部 HMM に入力しても、出力尤度は低くなると考えられる(図8)。Viterbi アルゴリズムによる評価を行う際、この局所的な尤度低下が、誤った状態選択を引き起こす可能性がある。このような問題を防ぐために、本稿では以下に示す対策を検討した。

画像から得られる T 個の各系列を、全ての super state(状態数 N)の内部 HMM に入力し、その系列ごと最大となる尤度を保存しておく。

得られた T 個の最大尤度をソートし、上位 K 個の尤度に対応する系列に関しては何も変更せず、下位(T - K)個の尤度に対応する系列に対しては、全ての super state からの確率値として 1.0(対数尤度 0)を与える。例えば、m 番目の系列からの出力尤度が低いと判断された場合、

$$P(O_{r=m}|\lambda_n)=1.0 \quad \text{for } 1 \leq n \leq N$$

のように尤度の変更を行うことになる。

以上の操作によって一部変更された尤度に基づいて Viterbi アルゴリズムによる尤度評価を行う。

以上の方法によると、眼鏡や隠れた部分は全体の尤度に影響を与えなくなるため、よりロバストになることが期待される。

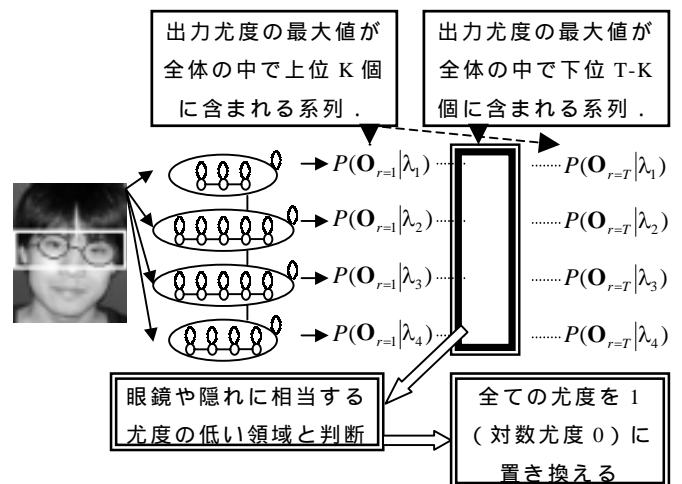


図8 眼鏡や隠れによる局所的な尤度の低下

4. 実験

4.1. 実験条件

提案手法の有効性を検証するため、140 枚の画像に対する識別実験を行った。学習は 40 名の人物に対して行う。学習画像は各 5 枚とし、学習時とは別の画像群をテスト画像とする。40 名から選んだ 20 名分の画像を 1 つのテストセットとし、7 つのテストセットに対して識別実験を行った。テストセット 1 と 2 は学習画像と同じような正面向きの画像であり位置の正規化も行われている。テストセットの番号が増えるにつれ、

顔の向きや位置が異なる画像が含まれるようになり、テストセット 6 や 7 では、顔の向きや位置、表情が大きく異なる画像がほとんどになる。尚、今回用いた顔画像は、電子・情報通信学会 第 4 回アルゴリズムコンテスト[7]のデータを使わせて頂いた。

画像の大きさは全て縦 72 × 横 54pixel であり、EHMM の状態数は、水平方向・鉛直方向 EHMM とともに super state 数 8、内包される HMM の embedded state 数は、上から 4, 4, 8, 8, 8, 8, 4, 4 とした。

DCT 係数を観測可能な特徴として採用し、正規化を行った後、その出力確率を無相関正規分布で近似した。

学習に用いた各人物 5 枚テスト画像を図 9 に示す。実際には以下の 20 名に加えてさらに 20 名分、合計 40 名分のデータベースが作られることになる。また、テストセットの画像例を図 10 示す。



図 9 学習画像の例



(a)



(b)



(c)



(d)

図 10 テスト画像の例 (a)TestSet1 (b)TestSet2 (c)TestSet6 (d)TestSet7

以上のテスト画像に対して、従来手法とハイブリッド型 EHMM による評価実験を行った。また、今回は主成分分析を用いた認識手法[1]との比較も同時に行った。表 4.1 に 7 つのテストセットごとの認識正解数を、図 11 に全テストセットから算出した識別率を示す。

表 4.1 従来手法と提案手法の認識結果
(正解数 / テスト画像数)

	PCA	EHMM	Hybrid EHMM
TestSet1	18/20	20/20	20/20
TestSet2	12/20	18/20	18/20
TestSet3	7/20	15/20	17/20
TestSet4	7/20	14/20	15/20
TestSet5	4/20	16/20	15/20
TestSet6	4/20	15/20	16/20
	5/20	12/20	11/20

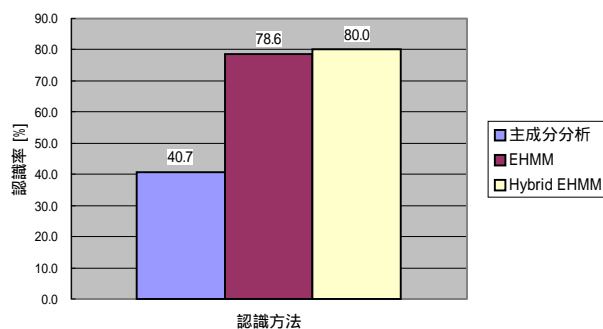


図 11 従来手法と提案手法の認識率の比較

次に、眼鏡や部分的な隠れのある画像に対する提案手法の評価実験を行った。実験に用いる画像を図 12 に示す。図 12 の画像は、幾つかのテスト画像に対し、画像処理ソフトを用いて眼鏡や帽子、マスクやヒゲを加えたものである。実験は、画像処理を加える前と加えた後の各画像に対して行い、隠れに対する効果を純粹に確認するため、ハイブリッド構造ではなく従来の

EHMM 構造のみを用いて行った。表 4.2 に、EHMM による従来の方法と、隠れに対する対策を行った場合の識別結果を示す。



図 12 眼鏡や部分的隠れのあるテスト画像

表 4.2 眼鏡やマスクのある画像に対する認識結果
(正解数 / テスト画像数)

	PCA	従来の EHMM	隠れに対する対策を行った EHMM
眼鏡やマスクを追加する前	12/20	20/20	20/20
眼鏡やマスクを追加した後	10/20	15/20	19/20

5. 考察

5.1. 主成分分析との結果比較

EHMM を用いた手法は主成分分析による手法に比べ、どのテストセットにおいても優れた結果が得られている。主成分分析による方法は、画像全体を 1 つのパターンとして捉え、グローバルな情報を用いて照合する手法である。このため多少の雑音や欠損に対しては画像全体で平均化されるという点で頑健であるが、物体が半分程度隠れている場合や、複雑な背景の中に存在したときには影響を強く受ける。その結果は本実験からも確認することができる。

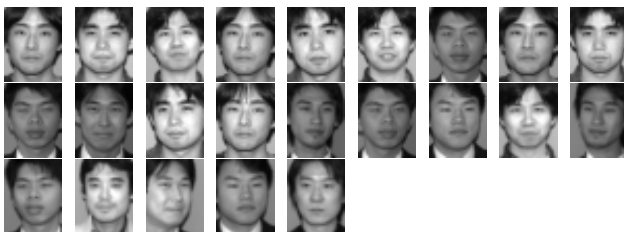
5.2. EHMM と Hybrid EHMM の結果比較

TestSet1 ~ TestSet7 までの計 140 枚の画像の中で、従来の EHMM と Hybrid EHMM による識別結果を、提案手法によって識別率が改善された画像、どちらも誤認識した画像、提案手法によって逆に誤認識を起こした画像、に分けて以下に示す。

識別率が改善された画像



識別率が改善されなかった画像(共に誤った画像)



識別率が改悪された画像



一部の画像において識別率が低下したが、これは縦方向に評価する EHMM を加えることによって、横に平行移動した特徴に対する尤度を不当に下げたためと考えられる。また、認識率の悪い画像はほぼ同じ人物であったため、学習時に必要十分なデータが得られていない可能性も考えられる。

5.3. 眼鏡や隠れのある画像に対する結果比較

眼鏡や隠れを加える前の画像に対して、EHMM を用いた手法では、隠れへの対策如何に拘らず 100% の認識率が得られている。眼鏡やマスクの処理を行った後、誤認識を起こした画像について以下に示す。

従来の EHMM で誤った画像



隠れへの対策を行った EHMM で誤った画像



従来の EHMM では、マスクをした画像のように大きな隠れのある画像に対し不得意であったが、隠れに対する対策を導入することにより、認識率が改善された。

6. むすび

本稿では 2 種類の構造をもつ Embedded HMM を考え、併せて評価する方法を提案した。また部分領域のみに着目する評価方法についても検討し、眼鏡や隠れのある画像に対し有効性を確認した。今後はさらに多くのデータベースに対して実験を行い、背景のある画像に対しても適用可能なアルゴリズムを加えていく予定である。

文 献

- [1] M.Turk and A.Pentland, "Face recognition using eigenfaces," Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp.586-591, 1991.
- [2] D.Beymer, "Face Recognition under varying pose," in Proceedings of 23rd Image Understanding Workshop, vol 2, pp.837-842, 1994.
- [3] R.Samaria and S.Young, "HMM based architecture for face identification," Image and Computer Vision, vol.12, pp.537-543, October 1994.
- [4] A.V.Nefian and M.H.Hayes, "Face detection and recognition using Hidden Markov Models," in International Conference on Image Processing, 1998.
- [5] 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.
- [6] 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.
- [7] 電子情報通信学会 PRMU 研究会第 4 回アルゴリズムコンテスト「私は誰？顔画像からの個人識別」, <http://www.toriwaki.nuie.nagoya-u.ac.jp/~sai/alcon/index.html>