社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

HMMを用いた動画像注目領域フィルタリング

加藤 浩康　　松尾　啓志　　岩田 彰

名古屋工業大学 電気情報工学科
〒466-0061 名古屋市昭和区御器所町
E-mail : fuji@mars.elcom.nitech.ac.jp

あらまし

時系列画像は、細かな時系列の区分に分割することができ、その中には、認識に必要な複数の状態と認識に必要でない状態とが含まれている場合がある。本論文では、HMMを用いることにより時系列を複数の状態に分割し不必要な状態を省く処理、フィルタリングを行う手法を提案する。また、HMMのシンボル出力確率を、状態に合わせて再学習することによってフィルタリングの性能を向上させる時系列フィルタリング設計法についても示す。本手法を評価するため、時系列顔画像による人物認識の実験を行った。

キーワード　　HMM、フィルタリング、時系列画像処理、人物認識

Image Sequences Filtering Using HMM

Hiroyasu KATO and Hiroshi MATSUO and Akira IWATA
Dept. of Electrical and Computer Eng., Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, 466-0061, Japan

Abstract

It is possible to divide image sequences into some states, where each state has same motion phase. By applying such operation before recognition processing, accuracy of the recognition will be improved. Then, new image sequence filtering by using HMM is proposed in paper, which can divide image sequences into multiple states. Performance of the filtering is improved by doing re-learning of the observation symbol probability. In addition, effectiveness of proposed method is shown by human identification using image sequences.

key words　　HMM, filtering, video image processing, human identification

1 はじめに

近年、計算機の処理速度の向上とともに特別なハードウェアを用いなくても動画像の取り込み、処理が可能となってきた。それに伴い、オプティカルフロー計算、動物体のトラッキング等の動画像処理や、ジェスチャや顔画像などの動画像認識等において様々なアルゴリズムが提案されている。

時系列画像を用いた認識手法として、隠れマルコフモデル (HMM) を用いる手法が数多く提案されている。例えば、時系列画像から得られた動き情報をKL展開することで次元圧縮を行い、ジェスチャを認識する手法である。これは、オートマトンを用いることによりジェスチャ間のつながりを考慮してジェスチャ認識をしている [1]。また、特徴ベクトルにメッシュを用いた動作認識手法もある。これは各カテゴリごとにVQコードブックを作成しHMMを用いて動作認識を行っている [2]。また、部分隠れマルコフモデルをジェスチャ認識に用いるものがある [3]。他に、HMMを顔の表情認識に用いる研究がある。これは、表情筋の動きに対応した状態を遷移するHMMを用いて表情抽出を行っている [4]。以上のように様々な適用例がある。

また、他の時系列画像を用いた認識手法として、時系列顔画像の部分空間を使う手法がある。例えば、部分空間射影により学習に最適なフレームを検出する手法 [5] がある。また、複数枚の画像を用いて部分空間を作成し、部分空間と部分空間とを比較する相互部分空間法で人物識別をする手法 [6] 等が提案されている。

ジェスチャ認識等の時系列画像の認識では、時系列そのものに意味があり、そのすべての画像群が必要な情報である。しかし、時系列画像は、細かな時系列の区分に分割することができ、その中には、認識に必要な複数の状態と認識に必要でない状態とが含まれている場合も多数存在する。このような場合には、認識処理を行う前段階として時系列画像中の意味のある複数の状態を考慮し、予め不必要な状態を省く処理、すなわち、フィルタリングを行うことにより、後に行う認識の精度を飛躍的に向上させることが期待できる。そこで本論文では、HMMを用いることにより時系列を複数の状態に分割しフィルタリングを行う手法を提案する。また、HMMのシンボル出力確率を、状態に合わせて再学習することによってフィルタリングの性能を向上させる時系列フィルタリング設計法についても示す。

以下、2章では、従来のHMMを用いた認識手法の概要について示し、3章では、HMMを用いた時系列フィルタリングと再学習を提案する。4章では、人物認識実験の結果とその考察を述べ、最後に5章でまとめと今後の課題について述べる。

2 HMMを用いた認識手法

2.1 HMM

HMM[7] は、出力シンボルによって一意に状態遷移先が決まらないという意味での非決定性確率有限状態オートマトンとして定義されている。出力シンボル系列が与えられても状態の遷移系列は唯一に決まらない、つまり、観測できるのはシンボル系列だけであることからHMMと呼ばれる。HMMは、特徴のパターンを統計的、確率的に捉えてモデル化するため、雑音や伸縮に対してロバストである。ここでは、シンボル出力確率が離散分布であり、状態で出力されるものを用いる。

HMMパラメータ$\lambda = (\pi, \mathbf{A}, \mathbf{B})$の内容を以下に示す。

$\pi = \{\pi_i\}$:初期状態確率π_iは初期状態がS_iである確率。

$\mathbf{A} = \{a_{ij}\}$:状態遷移確率a_{ij}は状態S_iから状態S_jに遷移する確率。

$\mathbf{B} = \{b_j(k)\}$:シンボル出力確率$b_j(k)$は状態S_jのときにシンボルkを出力する確率。

また、状態数を$(1 \leq i, j \leq N)$、シンボル数を$(1 \leq k \leq M)$とする。

2.2 認識と学習

HMMを用いた認識の概略を図1に示す。まず、それぞれの事象iに対応したシンボル系列Oを複数使いHMMパラメータλ_iを学習し、観測されたシンボル系列$O = o_1, o_2, \cdots, o_T$が事象$i$である確率$P(O|\lambda_i)$が最大になるHMMを選択することである。

HMMの認識アルゴリズムには、一般に前向きアルゴリズム (Forward Algorithm) が用いられる。このアルゴリズムは、直前のフレームにおける確率から現在の確率を求め、これを繰り返すことにより最終的な生成確率$P(O|\lambda)$を求める。時刻tの時に観測シンボル系列$o_1, o_2, \cdots, o_t$を出力して、状態jにいる確率を

$$\alpha_t(j) = P(o_1, o_2, \cdots, o_t, s_t = j|\lambda)$$

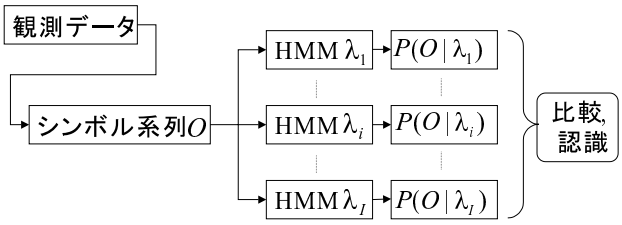

図 1: 従来の HMM

とする。これは、

$$\begin{aligned}\alpha_1(j) &= \pi_j b_j(o_1)\\ \alpha_t(j) &= \sum_{i=1}^{N}\alpha_{t-1}(i)a_{ij}b_j(o_t)\end{aligned}$$

の漸化式で求まる。生成確率 $P(O|\lambda)$ は

$$P(O|\lambda) = \sum_{j=1}^{N}\alpha_T(j)$$

となる。また、最適な状態系列から生成確率を求める Viterbi アルゴリズムがある。モデルλにおいて観測シンボル系列 $O = o_1, o_2, \cdots, o_T$に対する最適な状態系列 $S = s_1, s_2, \cdots, s_T$を求める。時刻 t で状態 jに至るまでの最適状態確率$\delta_t(j)$ を

$$\delta_t(j) = \max_{s_1,\cdots,s_{t-1}} P(s_1, \cdots, s_t = j, o_1, \cdots, o_t|\lambda)$$

とする。時刻 t における最適状態確率$\delta_t(j)$ は

$$\begin{aligned}\delta_1(j) &= \pi_j b_j(o_1)\\ \delta_t(j) &= \max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}b_j(o_t)]\end{aligned}$$

の漸化式で求まる。最適な状態系列の生成確率 $P^*(O|\lambda)$ は

$$P^*(O|\lambda) = \max_{1\le j\le N}[\delta_T(j)]$$

となる。最適な状態系列 $S = s_1, s_2, \cdots, s_T$は、時刻 t で状態 jにおいて生成確率を最大にする状態遷移を$\psi_t(j)$ とすると、

$$\begin{aligned}\psi_t(j) &= \underset{1\le i\le N}{argmax}[\delta_{t-1}(i)a_{ij}b_j(o_t)]\\ s_T &= \underset{1\le i\le N}{argmax}[\delta_T(i)]\\ s_t &= \psi_{t+1}(s_{t+1})\end{aligned}$$

で求まる。

HMM パラメータλの学習法として、Baum-Welch アルゴリズムがある [7]。このアルゴリズムは、多量の学習データを用いて繰り返しパラメータλを更新することにより、観測系列の生成確率を最大にするパラメータλの局所的最適値を求めるものである。パラメータ a_{ij}、$b_j(k)$ の更新は

$$\begin{aligned}a_{ij} &\leftarrow \frac{\sum_{t=1}^{T-1}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{t=1}^{T-1}\alpha_t(i)\beta_t(i)}\\ b_j(k) &\leftarrow \frac{\sum_{t\in(o_t=k)}\alpha_t(j)\beta_t(j)}{\sum_{t=1}^{T}\alpha_t(j)\beta_t(j)}\end{aligned}$$

で行われる。ここで$\beta_t(j)$ は時刻 t に状態 jにおいて、 以後 $o_{t+1}, o_{t+2}, \cdots, o_T$を出力する確率

$$\beta_t(j) = P(o_{t+1}, o_{t+2}, \cdots, o_T, s_t = j|\lambda)$$

である。

3 フィルタリングと再学習

3.1 時系列フィルタリング

HMM を用いて時系列を複数の区分に分割し、フィルタリングする方法について述べる。HMM を認識に用いる場合には、一般に観測された事象を 1 次元の観測シンボル系列に変換するため、情報量の減少が生じる。HMM を認識に用いるのではなく、時系列の状態の分割に用いることにより、図 2に示すように、観測された事象を状態毎に比較し認識することが可能となる。このため、後に行う認識段階に HMM での分類結果、さらにはシンボル化する前の生の時系列情報をも含めた、より多くの情報を用いて認識することが可能となる。

一般に HMM の認識に用いられる前向きアルゴリズムでは全ての状態遷移を使い生成確率を求めるため、状態の遷移を一意を決定することはできない。そのため、意味のある複数の状態に時系列を分割することも困難である。しかし、最適な状態系列から生成確率を求める Viterbi アルゴリズムを用いれば、時刻毎に一つの状態に対応付けられるため、状態毎に分割するだけで時系列を複数の区分に分割することが可能となる。また、Viterbi アルゴリズムを用いて状態を分割するためには、時系列を時系列的に

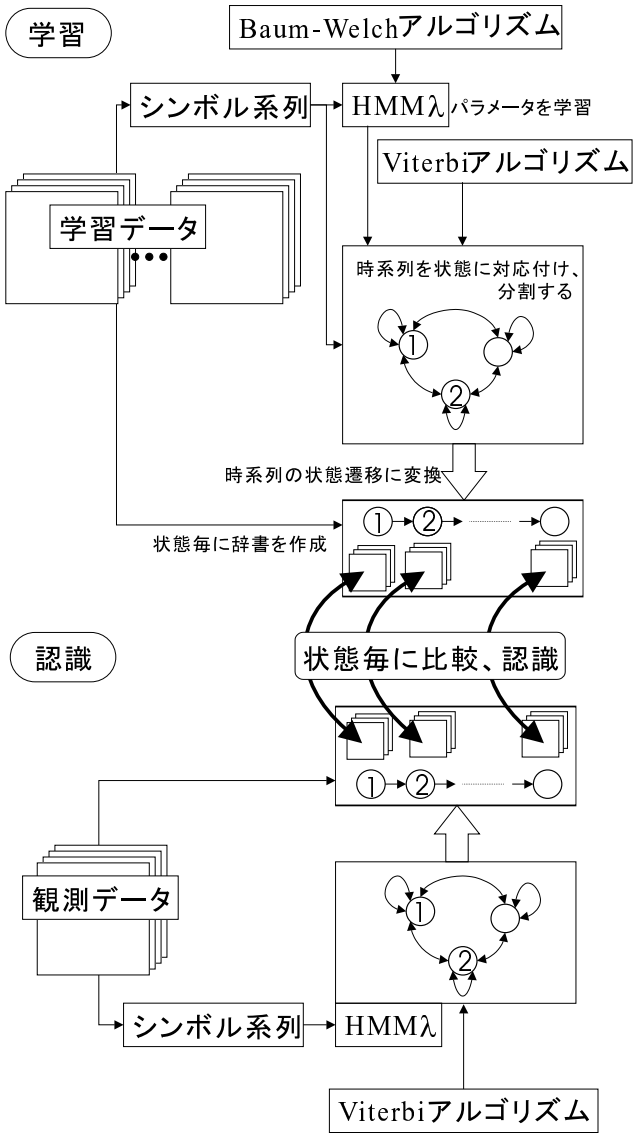

図 2: 時系列フィルタリング

意味のある複数の区分に分割したときの状態 (以下、目標状態群) に対応するように HMM の状態が学習されなければならない。そこで、図 3のような、目標状態群に学習した値を近づける手法について以下に示す。

3.2 学習における初期値

Baum-Welch アルゴリズムは、観測系列の生成確率を最大にするλの局所的最適値を求めるものであるため、初期値の設定は重要である。

初期値の設計方法としては、それぞれの状態に意味を持たせるために、その意味にあった状態を表す出力の場合はシンボル出力確率 $b_j(k)$ の値を高く、表さない出力の場合は確率の値を低く設定する。また、状態遷移確率 a_{ij}は、目標状態群において状態 i から状態 jへの遷移が起こりやすい場合には値は高く、遷移が起こりにくい場合には値を低く設定する。

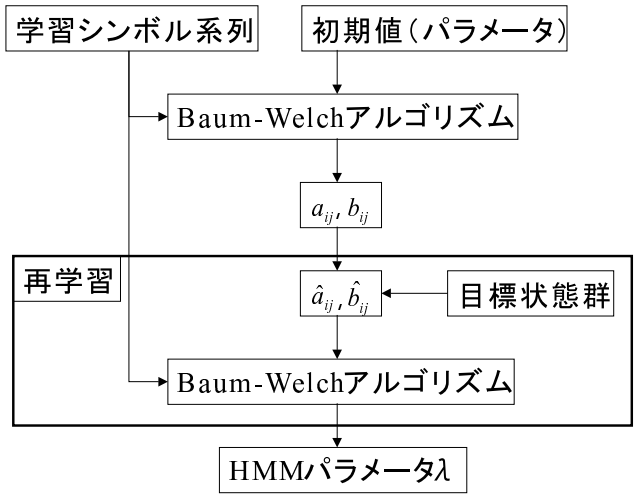

図 3: 再学習

3.3 再学習によるフィルタ設計

設定した初期値を利用して Baum-Welch アルゴリズムを用いて HMM パラメータλを学習させたとしても、それが目標状態群に学習できるとは限らない。そこで、目標状態群に近づけるためパラメータを再学習させる手法について提案する。

まず、状態のもつ 2 つのパラメータである状態遷移確率 a_{ij}とシンボル出力確率 $b_j(k)$ ではどちらが状態に意味を持たせやすいかについて考える。目標状態群のパラメータにするには状態遷移確率 a_{ij}よりもシンボル出力確率 $b_j(k)$ の設定が重要である。これは、状態確率 a_{ij}は、状態の遷移のしやすさを表しているの対して、シンボル確率 $b_j(k)$ は、その状態 S_jのシンボル kの出力のしやすさ、すなわち、その状態 S_jの意味を表しているからである。

設定した初期値を利用して学習した状態遷移確率 a_{ij}とシンボル出力確率 $b_j(k)$ は、学習に用いた時系列に対応した局所的最適値になっている。そのため、値をある程度、目標状態群に近付けたとしても学習に用いた時系列の情報を保持したまま、より目標状態群に近付いた値に再度収束させることができる。

そこで、シンボル出力確率 $b_j(k)$ をより目標状態群のパラメータに近付けるため、再学習時の初期値の状態遷移確率$\hat{a}_{ij}$を一定値に統一する。

$$\hat{a}_{ij} = \frac{1}{N} \quad (1 \leq i, j \leq N)$$

これは、状態遷移確率の値が、シンボル出力確率の値の変化を抑制することを防ぐためである。

シンボル出力確率 $b_j(k)$ では、その状態 S_j に対して意味を表す出力シンボルには $Label : High$ を、意味を表さない出力シンボルに対しては $Label : Low$ を設定する。

学習されたシンボル出力確率 $b_j(k)$ から再学習の初期値 $\hat{b}_j(k)$ の求め方を以下に示す。

1. シンボル出力確率 $b_j(k)$ に対応する $Label$ が $High$ である場合はステップ 2. へ、Low である場合はステップ 3. へ。

2. $b_j(k)$ が閾値 th_{high} 以上あれば

$$\hat{b}_j(k) = b_j(k)$$

とし、閾値 th_{high} 未満であれば

$$\hat{b}_j(k) = th_{high}$$

として、ステップ 4. へ。

3. $b_j(k)$ が閾値 th_{low} 以下あれば

$$\hat{b}_j(k) = b_j(k)$$

とし、閾値 th_{low} より大きい値ならば

$$\hat{b}_j(k) = th_{low}$$

として、ステップ 4. へ、

4. 次のシンボル出力確率に進む。

しかし、このままでは、シンボル出力確率集合 $\mathbf{B}$ の条件である

$$\sum_{k=1}^{M} b_j(k) = 1 \quad (1 \leq j \leq N)$$

を満たさないので

$$\hat{b}_j(k) \leftarrow \frac{\hat{b}_j(k)}{\sum_{k=1}^{M} \hat{b}_j(k)}$$

とする正規化を行う。

再学習の初期値として $\hat{a}_{ij}, \hat{b}_j(k)$ を使い Baum-Welch アルゴリズムにより再学習することで HMM パラメータを目標状態群に近付けることが可能となる。

閾値 th_{low} を小さい値に設定すると、より目標状態群に近づくシンボル出力確率に収束し、大きい値に設定すると、より時系列を考慮したシンボル出力確率に収束する。このため、閾値 th_{low} を適切な値に設定することにより、目標状態群に近づいたシンボル出力確率を得ることができる。

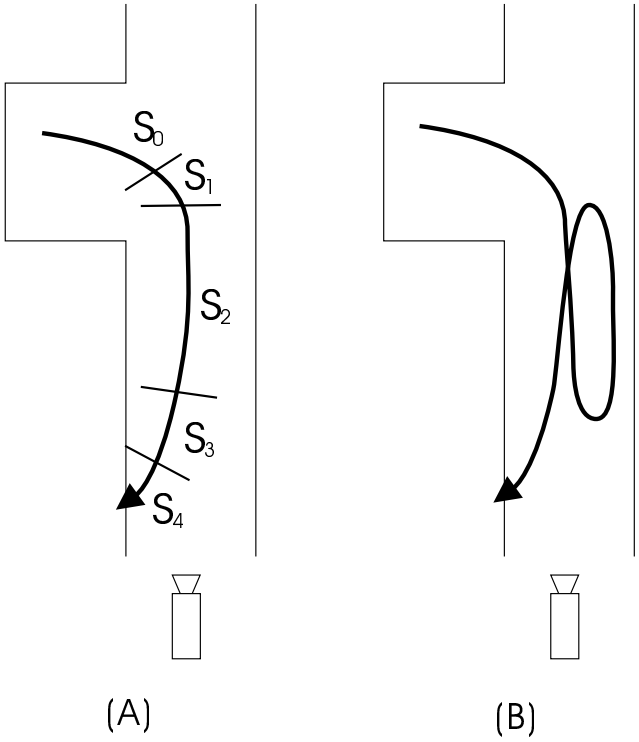

図 4: 撮影状況

4 実験

提案した時系列画像フィルタリングの効果を確認するために、通路の天井に設置した監視カメラから得られた映像を用いて人物識別実験を行った。監視カメラとして SONY 製の 1/2 インチ CCD カラービデオカメラ DXC-107A とキャノン製の PH6X8-1.1-II レンズを用いた。撮影した映像は一度 DV テープに保存し、canopus 製の DVRaptor を用いて取り込み、秒 6 フレームの画素数 640×480 の時系列画像に変換した。また、特別な照明は用いていない。設置してある蛍光灯のみの照明で撮影したため、人物の立っている場所により比較的に明るい場所と暗い場所が存在する。

図 4,5に示すように、5 人に左奥から手前の左の部屋に入室する行動を撮影した。撮影したデータは顔領域を手動により切り出し、32×32pixel に正規化する。顔領域とは、耳、あご、鼻、髪の毛をすべて含む最小の長方形の領域とする。この処理により、図 6に示すような正規化された時系列の顔画像が得られる[1] 。

HMM を使用するには 1 次元のシンボル系列に変

[1]種々の顔領域切り出しアルゴリズムが提案されているが、切り出しは本論文の主題ではないため手動により行った。

図 5: 例 撮影された画像

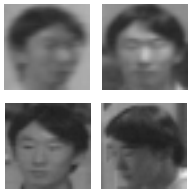

図 6: 例 顔画像

換する必要がある。顔の方向を用いて画像からシンボル系列に変換を行う。まず、事前に監視カメラを使い、照明の比較的明るい場所で一人づつ回転して撮影を行う。撮影された時系列画像から顔領域を切り出し、32 方向の正規化された顔画像を得て、それぞれの顔画像に方向のラベルを付ける。この処理をすべての人物で行うことにより、同じ方向ラベルを持つ複数枚の画像が全方向について得られる。この画像の組を方向辞書画像とする。撮影して得られた正規化された顔画像列と方向辞書画像との相互相関を計算することにより 32 のシンボルを持つ観測シンボル系列に変換する。図 7に方向とシンボルの関係を示す。32 方向のシンボルは、前向き顔を方向 0(図中 下) として、右回りに、左向き顔を方向 8(図中 左)、後ろ向き顔を方向 16(図中 上)、右向き顔を方向 24(図中 右) とした。

図 4(A) に示す移動が行われた場合、監視カメラには、まず、右向きの顔が現れ、次に正面を向き、最後に左を向きながらフレームから出ていく顔が撮影される。このような顔画像を意味のある状態に分ける場合には、顔の向きによって分けることができるため、図 7に示すように状態を右向き顔状態 S_0、

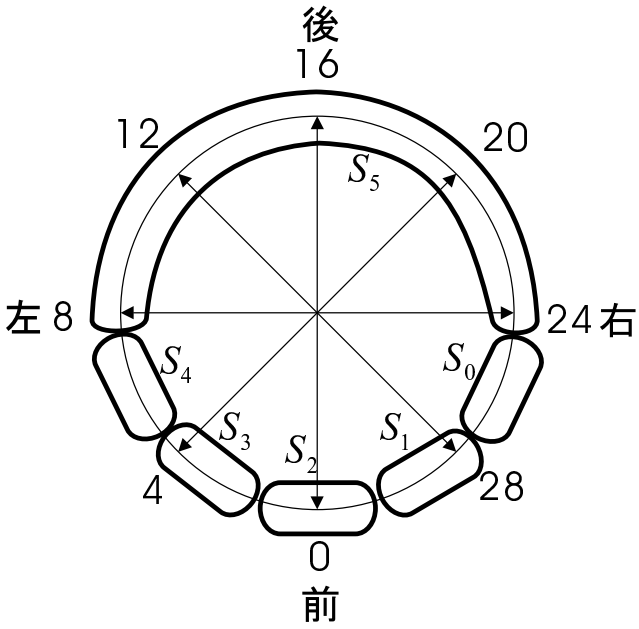

図 7: 方向ラベルと状態

右斜め顔状態 S_1、正面顔状態 S_2、左斜め顔状態 S_3、左向き顔状態 S_4、後ろ向き顔状態 S_5の 6 個に設定して学習させる。この状態に学習することにより、人物認識に不必要な後ろ向き顔を、ひとつの状態に割り当てることができ、容易に取り除くことができる。

学習時のパラメータの初期値は以下のように設定した。初期状態確率π_iは、時系列が右向きの顔から現れるため、$\pi_0 = 1$ とし、他を 0 とした。状態遷移確率 a_{ij}は、現在の状態への遷移を 0.3 とし、隣合う状態への遷移 (例えば、正面顔状態 S_2から右斜め顔状態 S_1と左斜め顔状態 S_3への遷移) を 0.2 とし、他の場合を 0.1 とした。シンボル出力確率 $b_j(k)$ は、正面顔状態 S_2では、正面顔状態を表すシンボルである $31, 0, 1$ のシンボル出力確率 $b_2(31), b_2(0), b_2(1)$ を 0.12 に、その隣のシンボルである $30, 2$ のシンボル出力確率 $b_2(30), b_2(2)$ を 0.05 に、他のシンボル出力確率 $b_j(k)$ を 0.02 とし、他の状態 S_0, S_1, S_3, S_4も同様に設定した。後向き顔状態 S_5は、後向き顔状態を表すシンボルである $8 \sim 24$ のシンボル出力確率 $b_j(k)$ を $0.0411 (= (1 - 0.02 \times 15)/17)$とし、他のシンボル出力確率 $b_j(k)$ を 0.02 と設定した。

再学習時に用いる $Label$は、右向き顔状態 S_0の場合は、方向 25、26、27 を $Label : High$ に、他の方向は $Label : Low$に設定した。他の状態の右斜め顔状態 S_1、正面顔状態 S_2、左斜め顔状態 S_3、左向き顔状態 S_4、後ろ向き顔状態 S_5 については表 1に示す。

表 1: 状態と方向のラベル付け

状態	Label:High	Label:Low
S_0	25,26,27	0∼24,28∼31
S_1	28,29,30	0∼27,31
S_2	31,0,1	2∼30
S_3	2,3,4	0,1,5∼31
S_4	5,6,7	0∼4,8∼31
S_5	8∼24	0∼7,25∼31

5 人のそれぞれ 4 パターンの歩行 (図 4(A)) の顔画像列、計 20 パターンを学習データに用いて学習を行った。それぞれの時系列のフレーム数は約 60 前後である。認識データ (図 4(B)) には、左奥から手前左の部屋に移動する途中に一度 U ターンし歩いてきた方向に移動し、また U ターンして部屋に移動する行動を用いた。これは、同じ 5 人のそれぞれ 1 パターン、計 5 パターンとし、それぞれの時系列のフレーム数は約 80 前後である。

次に、HMM を用いて学習データを状態毎に保存し辞書データを作る。学習データで学習した HMM を使い、認識データから得られたシンボル系列を 6 個の状態に分割する。この処理により、時系列顔画像のそれぞれに対してそれがどの状態であるかを対応付ける。顔画像の状態に対応した状態の辞書データと比較し人物認識を行う。次式の値 G が最大となるカテゴリ i を認識結果とする。

$$G(i) = \sum_{t,i \in D(l)} \max_{l=1,\cdots,L} f(B(t), D(l))$$

ただし、fは相互相関を、$B(t)$ は時刻 t の顔画像を、$D = \{D(l)\}$ は $B(t)$ の状態に対応する辞書データの集合を表す。

フィルタリングなしとは、観測データの正規化された顔画像と学習データの全ての顔画像と比較し、人物認識を行った結果である。人物 A、B、C、D、E の認識実験のフィルタリングなしと再学習を行わないで HMM を用いて状態の分割をしたもの (提案手法 1) と再学習を行って HMM を用いて状態の分割をしたもの (提案手法 2) の実験結果を図 8に示す。

グラフの縦軸は、正解のカテゴリの G と、正解を除いたカテゴリの中での G の最大値との差を表す。すなわち、正しく認識したものは、グラフが中央の太い線より上に、その値は正しく認識されたカテゴリの G と 2 番めに大きな値をとった G との差を表す。誤って認識したものは、グラフが中央の太い線より下に、その値は実際に観測されたカテゴリの G と誤って認識したカテゴリの G との差を表す。人物 B は提案手法を使うことにより正しく認識できるようになった。また、正しく認識されたものは値が大きくなっており、他のカテゴリとの差が大きくなっている。誤って認識されたものでも値が小さくなっており、誤認識されたカテゴリとの差が小さくなっているため、HMM を用いた状態の分割の有効性が確かめられた。また、カテゴリ B、D で、提案手法 1 の値より提案手法 2 の値が小さい。これは、シンボル系列に変換する際の誤差が影響し、誤った状態に対応付けられた認識データの顔画像に対して、提案手法 1 では、誤った状態に対応付けられた辞書と比較し正解のカテゴリが選択されていた。しかし、提案手法 2 では、再学習を行うことにより辞書データの状態の対応付けが改善されたため、提案手法 1 で選択していたカテゴリを選択できなくなったことが原因である。

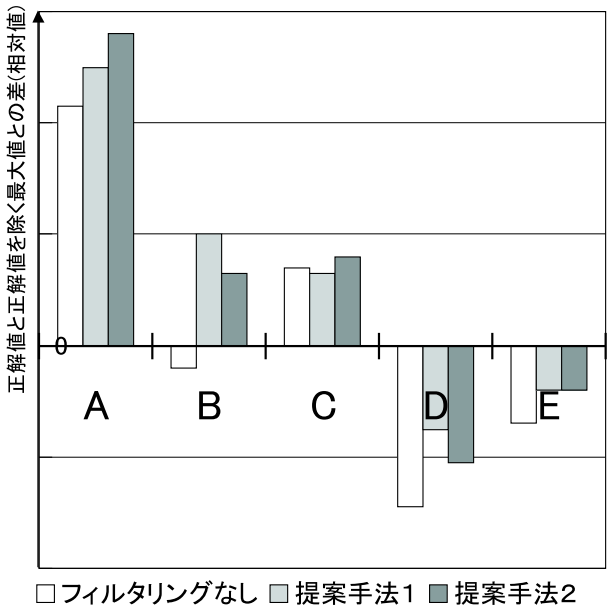

図 8: 実験結果

表 2に人物 A の認識実験を使い、手法による状態の対応づけとカテゴリ選択の違いを示す。フィルタリングなしの相互相関値は、その顔画像が全辞書データに対する相互相関値の最大値であり、カテゴリは、相互相関が最大となった顔画像の属するカテゴリを表す。また、提案手法 1、2 の状態は、HMM を用いた時系列フィルタリングを行い対応付けられた状態を示し、相互相関値は、その顔画像がその状

表 2: 人物 A の認識結果

	顔画像						
フィルタリングなし	相互相関値	0.993	0.951	0.953	0.970	0.972	0.972
	カテゴリ	A	D	C	E	B	A
提案手法 1	状態	S_0	S_5	S_4	S_2	S_5	S_3
	相互相関値	0.993	-	0.951	0.960	-	0.968
	カテゴリ	A	-	A	A	-	A
提案手法 2	状態	S_0	S_5	S_5	S_2	S_3	S_4
	相互相関値	0.993	-	-	0.960	0.963	0.969
	カテゴリ	A	-	-	A	A	A

態に対応付けられた辞書データに対する相互相関値の最大値であり、カテゴリは、相互相関が最大となった顔画像の属するカテゴリを表す。フィルタリングなしと提案手法 1、2 では、フィルタリングを行うことにより適切なカテゴリ選択が行われている。また、提案手法 1 と 2 では、再学習を行うことにより、時系列に対する状態の対応付けがより正確になり、時系列フィルタリングをより適切に行うことができた。

5 まとめ

HMM を用いることにより時系列を複数の状態に分割しフィルタリングを行う手法を提案し、HMM のシンボル出力確率を状態に合わせて再学習する方法についても示した。人物認識実験を行い時系列フィルタリングの有効性を示した。

今後は、状態の遷移を考慮した不必要な領域の状態の削除について、また、提案した時系列フィルタリングに適した認識手法についても検討する予定である。

参考文献

[1] 清水宏明、石井儀雄、谷内田正彦、"HMM を利用したジェスチャー認識"、CVIM144-14、Jan. 1999

[2] 大和淳司、倉掛正治、伴野明、石井健一郎、" カテゴリ別 VQ を用いた HMM による動作認識法 "、信学論 (D-II)、Vol.J77-D-II No.7 pp1311-1318、Jul. 1994

[3] 益満健、小林哲則、" 部分隠れマルコフモデルとそのジェスチャー認識への応用"、PRMU97-203、Jan. 1998

[4] 大塚尚宏、大谷淳、" 連続した表情シーケンス画像からの HMM を用いた個別表情抽出に関する検討"、PRMU97-154、Nov. 1997

[5] 杉山善明、有木康雄、" 部分空間射影による顔領域の追跡と学習"、PRMU97-162、Nov. 1997

[6] 山口修、福井和広、前田賢一、" 動画像を用いた顔認識システム"、PRMU97-50、Jun. 1997

[7] 中川聖一、" 確率モデルによる音声認識"、電子情報通信学会、1988