

## 状態継続長を考慮した分離型2次元格子HMMによる顔画像認識

高橋 良彰<sup>†</sup> 玉森 聡<sup>†</sup> 南角 吉彦<sup>†</sup> 徳田 恵一<sup>†</sup>

<sup>†</sup>名古屋工業大学 情報工学科 〒466-8555 名古屋市昭和区御器所町

**あらまし** 本稿では、状態継続長を考慮した分離型2次元格子HMMを提案する。分離型2次元格子HMMは、顔の位置や大きさの正規化が確率モデルの構造として組み込まれたモデルであり、画像の変動を状態遷移によって表現している。しかし、HMMのモデル構造では、状態継続確率が状態遷移確率のみに依存しており、継続長の増加に伴い指数的に減衰するため、変動を表現する精度が不十分である。そこで、状態継続長を明示的にモデル構造に含むモデルとして隠れセミマルコフモデル (Hidden semi Markov model; HSMM) が提案されている。提案法では、HSMMのモデル構造を分離型2次元格子HMMに組み込むことで、より正確に画像の変動を表現できるモデルの構築を行う。また、変分EMアルゴリズムを用いた提案法の学習アルゴリズムを導出し、顔画像認識実験により提案法の有効性を示す。  
**キーワード** 隠れマルコフモデル, 分離型2次元格子HMM, 状態継続長, 隠れセミマルコフモデル

### Face recognition based on separable lattice 2-D HMM with state duration modeling

Yoshiaki TAKAHASHI<sup>†</sup>, Akira TAMAMORI<sup>†</sup>, Yoshihiko NANKAKU<sup>†</sup>, and Keiichi TOKUDA<sup>†</sup>

<sup>†</sup> Dept. of Computer Science and Engineering, Nagoya Institute of Technology, Gokiso-cho, Syowa-ku, Nagoya, 466-8555, Japan

**Abstract** This paper proposes separable lattice 2-D HMMs with state duration modeling. Separable lattice 2-D HMMs (SL-HMMs) are probabilistic models with invariances to size and location which are represented by state transitions of HMMs. However, the state duration probability of HMMs exponentially decreases with increase of duration because it depends only on transition probability in the HMMs, therefore SL-HMMs cannot represent image variations accurately. To overcome this problem, hidden semi Markov models (HSMMs) have been proposed as a model in which the state duration probability is explicitly modeled by parametric distributions, e.g., Gaussian distributions. In this paper, we propose separable lattice 2-D HMMs with state duration modeling for representing the variations of images accurately. We also derive a training algorithm for the proposed model based on the variational EM algorithm and evaluate the effectiveness of the proposed model in face recognition experiments.

**Key words** Hidden Markov model, Separable lattice 2-D HMM, State duration, Hidden semi Markov model

#### 1. まえがき

画像認識では、さまざまな統計的アプローチに基づく手法が提案されている。しかし、これらの手法では画像が位置や大きさなどの変動を含む場合、認識精度の低下が起これば考えられる。そのため、あらかじめ画像を正規化しておく必要がある。従来法の多くは、画像をヒューリスティックな手法に基づいて正規化していたため、その精度がデータに依存してしまう問題がある。

それに対し、画像の正規化をモデル構造に含む確率モデルとして分離型2次元格子HMM (Separable Lattice 2-D HMM) [1] が提案されている。分離型2次元格子HMMは、縦方向・横方向で独立した状態遷移を持ち、画像の変動に対して矩形を維持しながら、移動・伸縮を表現することが可能となるモデル化手法である。しかし、分離型2次元格子HMMでは、状態の継続

する確率が遷移確率のみに依存しているため、状態継続長を適切に表現することができない。したがって、顔画像が大きく移動したり伸縮した場合、画像の位置や大きさを適切に正規化することができず顔画像のモデル化精度が低下してしまう問題が生じる。

このような問題に対し、HMMの状態継続長を明示的にモデルパラメータに含むモデルとして、隠れセミマルコフモデル (Hidden Semi Markov Model; HSMM) [2] が提案されている。本稿では、HSMMのモデル構造を分離型2次元格子HMMに組み込み、状態継続長を考慮したモデルの構築手法を提案する。HSMMは、HMMの状態が継続して出力する時間を考慮したモデル構造であり、音声合成や音声認識などの分野でその有効性が確認されている [3]。状態継続長を考慮した分離型2次元格子HMMを顔画像のモデル化に用いることで、変動の大きな画像に対する認識精度の改善が期待できる。

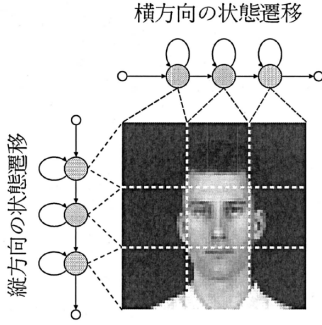


図1 分離型2次元格子HMMのモデル構造

## 2. 分離型2次元格子HMM

2次元に拡張されたHMMの一種として、分離型2次元格子HMMが提案されている。分離型2次元格子HMMは、縦方向・横方向でそれぞれ独立した状態遷移を持つモデルであり、モデルの出力は縦方向と横方向の状態の組み合わせによって決定される。分離型2次元格子HMMは、縦方向・横方向のそれぞれに対し画像の幾何学的な連続性を維持したモデル構造を持つ。これより、学習データが変動した場合にも、縦方向・横方向それぞれにおいて伸縮が吸収可能であるため、位置・大きさを正規化したモデルの構築を行うことができる。以降、縦と横の2つの次元における表現として、次の表記を用いる。

- ( $\cdot$ )<sup>(1)</sup> : 縦方向の状態遷移に関する変数
- ( $\cdot$ )<sup>(2)</sup> : 横方向の状態遷移に関する変数

分離型2次元格子HMMにおける観測ベクトルは次のように表される。

$$\mathbf{O} = \{\mathbf{O}_{t^{(1)}t^{(2)}} | t^{(1)} = 1, 2, \dots, T^{(1)}, t^{(2)} = 1, 2, \dots, T^{(2)}\} \quad (1)$$

ここで、 $\mathbf{O}$ は出力画像を表し、 $t^{(1)}, t^{(2)}$ はそれぞれ画像の縦方向・横方向の座標、 $\mathbf{O}_{t^{(1)}t^{(2)}}$ は2次元座標に格子状に配置された観測ベクトル、 $T^{(1)}, T^{(2)}$ はそれぞれ画像の縦方向・横方向の座標の最大値を表す。

縦方向・横方向の状態数をそれぞれ $K^{(1)}, K^{(2)}$ とし、状態系列における隠れ変数を次のように表す。

$$S_t^{(m)} \in \{1, 2, \dots, K^{(m)}\} \quad m \in \{1, 2\} \quad (2)$$

また、隠れ変数の構造を次のように2つのマルコフ連鎖から構成されるものとして定義する。

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

$$\mathbf{S}^{(m)} = \{S_1^{(m)}, \dots, S_{T^{(m)}}^{(m)}\} \quad (4)$$

図1に分離型2次元格子HMMのモデル構造を、図2にグラフィカルモデル表現を示す。

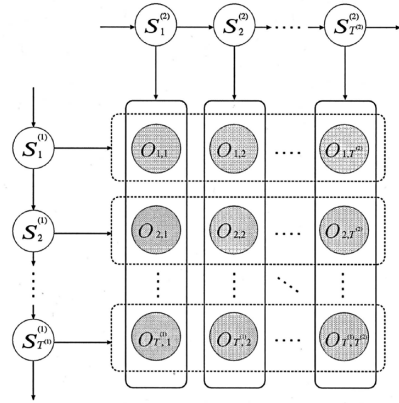


図2 分離型2次元格子HMMのグラフィカルモデル表現

分離型2次元格子HMMの尤度関数は、観測系列 $\mathbf{O}$ と隠れ変数 $\mathbf{S}$ の同時確率として次式で表される。

$$P(\mathbf{O}|\Lambda) = \sum_{\mathbf{S}} P(\mathbf{O}, \mathbf{S}|\Lambda) \quad (5)$$

$$\begin{aligned} P(\mathbf{O}, \mathbf{S}|\Lambda) &= P(\mathbf{S}|\Lambda)P(\mathbf{O}|\mathbf{S}, \Lambda) \\ &= P(\mathbf{S}^{(1)}|\Lambda^{(1)})P(\mathbf{S}^{(2)}|\Lambda^{(2)}) \prod_t P(\mathbf{O}_t|\mathbf{S}_t, \Lambda) \end{aligned} \quad (6)$$

ここで、分離型2次元格子HMMのモデルパラメータを次のように定義する。

$$\Lambda = \{\Lambda^{(1)}, \Lambda^{(2)}, \mathbf{B}\} \quad (7)$$

$$\Lambda^{(m)} = \{\mathbf{\Pi}^{(m)}, \mathbf{A}^{(m)}\} \quad (8)$$

1) 初期状態分布 $\mathbf{\Pi}^{(m)} = \{\pi_i^{(m)} | 1 \leq i \leq K^{(m)}\}$   
 $\pi_i^{(m)} = P(S_1^{(m)} = i|\Lambda)$ は、観測 $\mathbf{O}$ に関する第 $m$ 次元の状態系列において、時刻(座標) $t^{(m)} = 1$ における状態が $i$ である確率を表す。

2) 状態遷移確率 $\mathbf{A}^{(m)} = \{a_{ij}^{(m)} | 1 \leq i \leq K^{(m)}, 1 \leq j \leq K^{(m)}\}$

$a_{ij}^{(m)} = P(S_{t^{(m)}}^{(m)} = j | S_{t^{(m)}-1}^{(m)} = i, \Lambda)$ は、観測 $\mathbf{O}$ に関する第 $m$ 次元の状態系列において、状態 $i$ から状態 $j$ に遷移する確率を表す。

3) 出力確率分布 $\mathbf{B} = \{b_k(\mathbf{O}_t) | k \in K\}$

$b_k(\mathbf{O}_t) = P(\mathbf{O}_t | \mathbf{S}_t = k, \Lambda)$ は、 $K$ を各次元の隠れ変数の直積とすると、2次元状態空間 $K$ における状態 $k$ が観測ベクトル $\mathbf{O}_t$ を出力する確率を表す。また、各状態の出力確率分布は、平均ベクトル $\mu_k$ 、共分散行列 $\Sigma_k$ を持つ単一ガウス分布であると仮定する。

## 3. 状態継続長を考慮した分離型2次元格子HMM

### 3.1 隠れセミマルコフモデル

HMMは、各状態において自己遷移を持つことによりモデル

から出力を継続して行う長さを表現している。これにより、信号列の時間的な特徴変化を考慮してモデル化を行うことができる。HMMにおいて、同じ状態に継続して滞在する確率である状態継続確率は、次式で表すことができる。

$$P_i(d) = a_{ii}^{d-1}(1 - a_{ii}) \quad (9)$$

ここで、 $i$ は状態番号であり、 $a_{ij}$ は状態 $i$ から $j$ に遷移する確率とする。また、 $d$ は状態継続長を表す。図3に、式(9)をグラフ化したものを示す。式(9)及び図3より、HMMでは状態継続確率は遷移確率のみに依存しているため、継続時間の増加とともに指数的に状態継続確率が低下していくことがわかる。したがって、ピークがあるような状態継続長分布は適切に表現することができない。HMMではこのような状態継続に対する確率的な制約により、長い時間継続した信号列に対してはモデル化精度が低下する問題が挙げられる。

上記の問題を解決するために、HMMに明示的な状態継続長分布を与えるモデル構造が提案されている。このモデル構造を隠れセミマルコフモデル (Hidden Semi Markov Model; HSMM) と呼び、図3に示すように、状態継続長分布に応じた状態継続確率が出力できる。隠れセミマルコフモデルは、信号列の時間的な長さを考慮してモデルを構築できるといった利点から、音声合成や音声認識などの分野で有効性が確認されている。

### 3.2 状態継続長を考慮した分離型2次元格子HMM

分離型2次元格子HMMは、画像の変動量が増加した場合、状態継続確率の近似精度が低下し、適切な状態継続長を表現することができない。したがって、顔画像に対するモデル化精度が低下してしまう問題が生じる。

そこで本稿では、より正確に画像の変動を表現できるように、状態継続長を考慮した分離型2次元格子HMMを提案する。HSMMのモデル構造を分離型2次元格子HMMに組み込むことにより、変動の大きな画像に対しても頑健なモデルの構築が期待できる。また提案法では、出力確率分布と状態継続長分布のパラメータを同時に推定するため、generalized forward-backward algorithmにより状態継続長を考慮した期待値計算を行う。また、尤度最大化基準に従い、モデルパラメータを再推定する。

HSMMはモデルパラメータとして状態継続長分布に関するパラメータを持つため、提案法ではHMMにおける状態系列を状態継続長を明示的に表す変数により表現する。提案法では、状態継続長と状態系列との依存関係を考慮して、次式のように変数の対応を定義する。

$$S^{(m)} = \{L^{(m)}, q^{(m)}, d^{(m)}\} \quad (10)$$

$$q^{(m)} = \{q_1^{(m)}, q_2^{(m)}, \dots, q_L^{(m)}\} \quad (11)$$

$$d^{(m)} = \{d_1^{(m)}, d_2^{(m)}, \dots, d_L^{(m)}\} \quad (12)$$

ただし、 $q^{(m)}$ は実際に出力を行った状態番号列(ただし、状態の継続を含まない)を表し、 $L^{(m)}$ は状態番号列の長さを表す。また、 $d^{(m)}$ は状態継続長を表す。例えば、通常のHMMの状

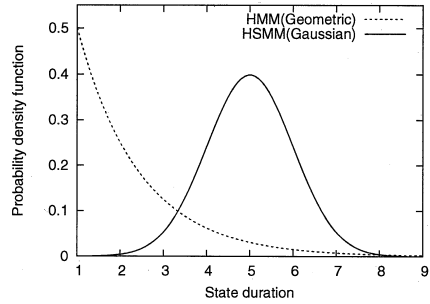


図3 状態継続長分布

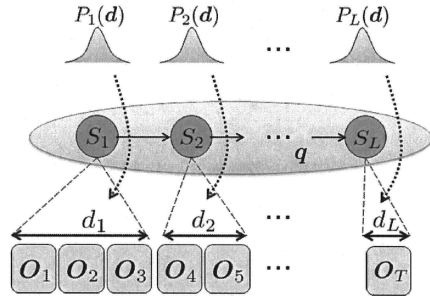


図4 Left-to-right型HSMM

態系列が

$$S^{(m)} = \{1, 1, 1, 2, 2, 2, 2, 3, 3\} \quad (13)$$

であるとき、 $L = 3$ 、 $q^{(m)} = (1, 2, 3)$ 、 $d^{(m)} = (3, 4, 2)$ となる。HSMMでは $d^{(m)}$ を明示的にモデル化することにより、より適切な状態遷移を得ることができる。HSMMのモデル構造と変数の関係を図4に示す。

提案法における尤度関数は次式で定義される。

$$P(O|\Lambda) = \sum_L \sum_q \sum_d P(O, L, q, d|\Lambda) \quad (14)$$

$$\begin{aligned} P(O, L, q, d|\Lambda) &= P(L, q|\Lambda)P(d|L, \Lambda)P(O|L, q, d, \Lambda) \\ &= P(L^{(1)}, q^{(1)}|\Lambda^{(1)})P(d^{(1)}|L^{(1)}, \Lambda^{(1)}) \\ &\quad \cdot P(L^{(2)}, q^{(2)}|\Lambda^{(2)})P(d^{(2)}|L^{(2)}, \Lambda^{(2)}) \\ &\quad \cdot \prod_{t_1, t_2} \prod_{s^{(1)}=t_1}^{t_1-1} \prod_{s^{(2)}=t_2}^{t_2-1} \\ &\quad \cdot P(O_{s^{(1)}s^{(2)}}|L, q, d, \Lambda) \end{aligned} \quad (15)$$

ただし、

$$L = \{L^{(1)}, L^{(2)}\} \quad q = \{q^{(1)}, q^{(2)}\} \quad (16)$$

$$d = \{d^{(1)}, d^{(2)}\} \quad (17)$$

である。また、 $P(L^{(m)}, q^{(m)}|\Lambda^{(m)})$ は状態遷移確率を表し、 $P(d^{(m)}|L^{(m)}, \Lambda^{(m)})$ は状態継続確率を表す。提案法では状態継続長分布をガウス分布で表現し、モデルパラメータに状態継続

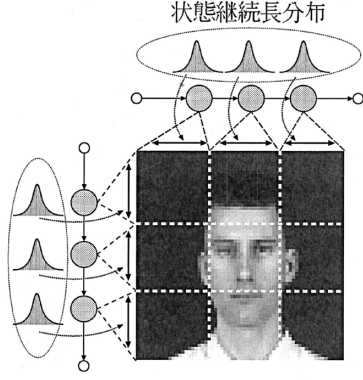


図5 提案法のモデル構造

長分布の平均  $\eta_i^{(m)}$ , 分散  $\sigma_i^{2(m)}$  を追加する。また、本稿ではスキップなしの left-to-right 型の HSM を仮定する。この仮定により、初期状態確率及び遷移確率は 1 となる。図 5 に、提案法によるモデル構造を示す。

### 3.3 学習アルゴリズム

本稿では、EM アルゴリズムにもとづく最尤推定によりモデルパラメータを推定する。EM アルゴリズムでは、次に示す期待値である  $Q$  関数を最大化させることで最尤推定を近似する。

$$Q(\Lambda, \Lambda') = \sum_L \sum_q \sum_d P(L, q, d | O, \Lambda) \ln P(O, L, q, d | \Lambda') \quad (18)$$

EM アルゴリズムでは、状態数やパラメータが多い場合、莫大な計算量を要するため実時間で計算が困難となる。そこで、本稿では、EM アルゴリズムに変分近似を導入した変分 EM アルゴリズム [4] を用いて計算量の削減を図る。

まず、対数尤度の下限  $\mathcal{F}$  を事後確率分布  $P(L, q, d | O, \Lambda)$  の近似確率分布  $Q(L, q, d)$  を用いて定義する。

$$\begin{aligned} \ln P(O | \Lambda) &= \ln \sum_L \sum_q \sum_d P(O, L, q, d | \Lambda) \\ &= \ln \sum_L \sum_q \sum_d Q(L, q, d) \frac{P(O, L, q, d | \Lambda)}{Q(L, q, d)} \\ &\geq \sum_L \sum_q \sum_d Q(L, q, d) \ln \frac{P(O, L, q, d | \Lambda)}{Q(L, q, d)} \\ &= \mathcal{F}(Q, \Lambda) \end{aligned} \quad (19)$$

変分 EM アルゴリズムは  $Q(L, q, d)$  を計算する E ステップと、 $\mathcal{F}$  を  $\Lambda$  に関して最大化する M ステップからなる：

$$(E\text{-step}) : Q_{k+1} = \arg \max_Q \mathcal{F}(Q, \Lambda_k)$$

$$(M\text{-step}) : \Lambda_{k+1} = \arg \max_{\Lambda} \mathcal{F}(Q_{k+1}, \Lambda)$$

これらのステップを繰り返すことで変分 EM アルゴリズムは尤度の上昇が保証される。変分 EM アルゴリズムでは尤度の下限  $\mathcal{F}$  を最大化することにより、尤度を極大に導く。

ここで、事後確率分布  $P(L, q, d | O, \Lambda)$  と近似確率分布  $Q(L, q, d)$  の確率間尺度である KL divergence は次式で与えられる。

$$\text{KL}(Q || P) = \sum_{L, q} \sum_d Q(L, q, d) \ln \frac{Q(L, q, d)}{P(L, q, d | O, \Lambda)} \quad (20)$$

式 (20) を用いて、 $\mathcal{F}$  は次式のように書き換えられる。

$$\begin{aligned} \mathcal{F}(Q, \Lambda) &= \sum_{L, q} \sum_d Q(L, q, d) \ln \frac{P(L, q, d | O, \Lambda) P(O | \Lambda)}{Q(L, q, d)} \\ &= -\text{KL}(Q || P) + \ln P(O | \Lambda) \end{aligned} \quad (21)$$

式 (21) より、 $\ln P(O | \Lambda)$  は  $Q(L, q, d)$  に依存しないため、 $\mathcal{F}$  の最大化は  $P$  と  $Q$  の KL divergence を最小化することと等価である。また、最適な近似分布  $Q$  は変分法の解であるラグランジュ方程式を解くことにより導出できる。

提案法では、変数の増加による期待値の計算量の増加を抑えるため、変数の次元間における独立性を仮定し、近似分布  $Q(L, q, d)$  に対し、次のように制約条件を設ける。

$$Q(L, q, d) = Q(L^{(1)}, q^{(1)}, d^{(1)}) Q(L^{(2)}, q^{(2)}, d^{(2)}) \quad (22)$$

ただし、

$$\sum_{L^{(m)}, q^{(m)}} \sum_{d^{(m)}} Q(L^{(m)}, q^{(m)}, d^{(m)}) = 1 \quad (23)$$

を満たすこととする。 $\mathcal{F}$  の最大化にあたり、最適な近似分布は変分法に基づき次式で与えられる。

$$\begin{aligned} Q(L^{(m)}, q^{(m)}, d^{(m)}) &\propto \frac{1}{Z_Q^{(m)}} P(L^{(m)}, q^{(m)}) P(d^{(m)}) \\ &\quad \cdot \exp \langle \ln P(O | L, q, d, \Lambda) \rangle_{Q^{(m)}} \\ &= \frac{1}{Z_Q^{(m)}} P(L^{(m)}, q^{(m)}) P(d^{(m)}) \\ &\quad \cdot \prod_{t^{(m)}} \prod_{s=t^{(m)}-d^{(m)}+1}^{t^{(m)}} h_s(q_{t^{(m)}}^{(m)}) \end{aligned} \quad (24)$$

ただし、 $Z_Q^{(m)}$  は正規化定数である。

ここで、時刻  $t^{(m)}$  まで  $d^{(m)}$  時間、状態  $k^{(m)}$  に滞在する期待値を  $\gamma_{t^{(m)}}^{d^{(m)}}(k^{(m)})$  とすると、観測確率  $h_{t^{(m)}}^{(m)}(i)$  の対数は次のように表される。

$$\begin{aligned} \ln h_{t^{(m)}}^{(m)}(i) &= \sum_{\bar{i} \in t/t^{(m)}} \sum_{\bar{k} \in k/k^{(m)}} \sum_{\bar{d} \in d/d^{(m)}} \prod_{n=\bar{m}} \gamma_{\bar{i}^{(n)}}^{\bar{d}^{(n)}}(\bar{k}^{(n)}) \\ &\quad \cdot \sum_{s=\bar{i}^{(n)}-\bar{d}^{(n)}+1}^{\bar{i}^{(n)}} \ln b_{q_{\bar{i}^{(n)}}^{(n)}=\bar{k}, q^{(m)}}^{(n)}(O_{s, t^{(m)}}) \end{aligned} \quad (25)$$

式 (25) における期待値は、後に示す generalized forward backward algorithm を用いて計算できる。

提案法において、変分近似を用いることで、計算量は  $O(\prod_m N^{(m)} T^{(m)} D^{(m)})$  となる。 $D^{(m)}$  は最大継続長を示し、 $D^{(m)} = \max_{d^{(m)}} \{d_1^{(m)}, \dots, d_L^{(m)}\}$  である。これより、計算量は多項式オーダーに抑えられるため、実時間で計算が可能となる。

### 3.3.1 Generalized forward backward algorithm

HSMM では、モデルパラメータに状態継続長分布に関するパラメータを導入する。よって、HSMM におけるモデル学習では、状態継続長を考慮した期待値計算を行うことが必要となる。そこで提案法では、generalized forward-backward algorithm [2] を用い、状態継続長を考慮した期待値計算を行う。

観測ベクトル  $\mathbf{O} = (o_1, \dots, o_{T^{(m)}})$  及びモデルパラメータ  $\Lambda$  が与えられた時、観測ベクトル  $\mathbf{o}$  が HSMM より出力される確率は generalized forward-backward algorithm を用いて計算できる。まず、前向き確率  $\alpha_{t^{(m)}}^{(m)}(i)$  を次のように定義する。

$$\begin{aligned} \alpha_0^{(m)}(i) &= \begin{cases} 1 & i = 1 \\ 0 & \text{otherwise} \end{cases} \\ \alpha_{t^{(m)}}^{(m)}(i) &= P(o_1, \dots, o_{t^{(m)}}, q_{t^{(m)}}^{(m)} = i | q_{t^{(m)}+1}^{(m)} \neq i, \Lambda) \\ &= \sum_{j \neq i} \sum_{d^{(m)}=1}^{K^{(m)}} \alpha_{t^{(m)}-d^{(m)}}^{(m)}(j) a_{ji}^{(m)} P_i(d^{(m)}) \\ &\quad \cdot \prod_{s=t^{(m)}-d^{(m)}+1}^{t^{(m)}} h_s^{(m)}(i) \\ &\quad \left( \begin{array}{l} 1 \leq t^{(m)} \leq T^{(m)} \\ 1 \leq i \leq K^{(m)} \end{array} \right) \end{aligned} \quad (26)$$

同様に、後ろ向き確率  $\beta_{t^{(m)}}^{(m)}(i)$  を次のように定義する。

$$\begin{aligned} \beta_{T^{(m)}+1}^{(m)}(i) &= \begin{cases} 1 & i = K^{(m)} \\ 0 & \text{otherwise} \end{cases} \\ \beta_{T^{(m)}}^{(m)}(i) &= P(o_{t^{(m)}+1}, \dots, o_{T^{(m)}}), q_{t^{(m)}}^{(m)} = i | q_{t^{(m)}+1}^{(m)} \neq i, \Lambda) \\ &= a_{iK^{(m)}}^{(m)} \beta_{T^{(m)}}(K^{(m)}) \quad (1 \leq i \leq K^{(m)}) \\ \beta_{t^{(m)}}^{(m)}(i) &= \sum_{j \neq i} \sum_{d^{(m)}=1}^{K^{(m)}-T^{(m)}-t^{(m)}} a_{ij}^{(m)} P_j(d^{(m)}) \\ &\quad \cdot \prod_{s=t^{(m)}+1}^{t^{(m)}+d^{(m)}} h_s^{(m)}(j) \beta_{t^{(m)}+d^{(m)}}^{(m)}(j) \\ &\quad \left( \begin{array}{l} 1 \leq t^{(m)} \leq T^{(m)} \\ 1 \leq i \leq K^{(m)} \end{array} \right) \end{aligned} \quad (27)$$

以上で定義された前向き確率と後ろ向き確率を再帰的に計算することにより、E-step における期待値計算が可能となる。

### 3.3.2 パラメータ再推定式

提案法におけるモデルパラメータは、尤度最大化基準に従って推定することができる。提案法において、出力確率分布の平均  $\boldsymbol{\mu}$ 、分散  $\boldsymbol{\Sigma}$  は次の式で再推定できる。

$$\begin{aligned} \boldsymbol{\mu}_{ij} &= \frac{\sum_{t^{(1)}=1}^{T^{(1)}} \sum_{t^{(2)}=1}^{T^{(2)}} \sum_{d^{(1)}=1}^{t^{(1)}} \sum_{d^{(2)}=1}^{t^{(2)}} \gamma_{t^{(1)}}^{d^{(1)}}(i) \gamma_{t^{(2)}}^{d^{(2)}}(j) \sum_s \mathbf{O}_s}{\sum_{t^{(1)}=1}^{T^{(1)}} \sum_{t^{(2)}=1}^{T^{(2)}} \sum_{d^{(1)}=1}^{t^{(1)}} \sum_{d^{(2)}=1}^{t^{(2)}} d^{(1)} d^{(2)} \gamma_{t^{(1)}}^{d^{(1)}}(i) \gamma_{t^{(2)}}^{d^{(2)}}(j)} \end{aligned} \quad (28)$$

$$\begin{aligned} \boldsymbol{\Sigma}_{ij} &= \frac{\sum_{t^{(1)}=1}^{T^{(1)}} \sum_{t^{(2)}=1}^{T^{(2)}} \sum_{d^{(1)}=1}^{t^{(1)}} \sum_{d^{(2)}=1}^{t^{(2)}} \gamma_{t^{(1)}}^{d^{(1)}}(i) \gamma_{t^{(2)}}^{d^{(2)}}(j) \sum_s \mathbf{O}_s \mathbf{O}_s^T}{\sum_{t^{(1)}=1}^{T^{(1)}} \sum_{t^{(2)}=1}^{T^{(2)}} \sum_{d^{(1)}=1}^{t^{(1)}} \sum_{d^{(2)}=1}^{t^{(2)}} d^{(1)} d^{(2)} \gamma_{t^{(1)}}^{d^{(1)}}(i) \gamma_{t^{(2)}}^{d^{(2)}}(j)} \\ &\quad - \boldsymbol{\mu}_{ij} \boldsymbol{\mu}_{ij}^T \\ \mathbf{s} &= \{s^{(1)}, s^{(2)}\} \\ (s^{(m)} &= t^{(m)} - d^{(m)} + 1, \dots, t^{(m)}) \end{aligned} \quad (29)$$

時刻  $t^{(m)}$  まで  $d^{(m)}$  時間、状態  $i$  にとどまる確率である  $\gamma_{t^{(m)}}^{d^{(m)}}(i)$  は、以下のように計算できる。

$$\begin{aligned} \gamma_{t^{(m)}}^{d^{(m)}}(j) &= \frac{1}{Q^{(m)}(\mathbf{O}|\Lambda)} \sum_{i=1}^{K^{(m)}} \alpha_{t^{(m)}-d^{(m)}}^{(m)}(i) a_{ij}^{(m)} P_j(d^{(m)}) \\ &\quad \cdot \prod_{s=t^{(m)}-d^{(m)}+1}^{t^{(m)}} h_s^{(m)}(j) \beta_{t^{(m)}}^{(m)}(j) \end{aligned} \quad (30)$$

また、正規化項  $Q^{(m)}(\mathbf{O}|\Lambda)$  は任意の  $t^{(m)}$  を用いて次式で計算される。

$$\begin{aligned} Q^{(m)}(\mathbf{O}|\Lambda) &= \sum_{i=1}^{K^{(m)}} \sum_{j=1}^{K^{(m)}} \sum_{d^{(m)}=1}^{t^{(m)}} \alpha_{t^{(m)}-d^{(m)}}^{(m)}(i) a_{ij}^{(m)} P_j(d^{(m)}) \\ &\quad \cdot \prod_{s=t^{(m)}-d^{(m)}+1}^{t^{(m)}} h_s^{(m)}(j) \beta_{t^{(m)}}^{(m)}(j) \end{aligned} \quad (31)$$

提案法では、状態継続長分布はガウス分布でモデル化を行う。ここで、状態  $i$  の状態継続長分布について、平均  $\eta_i^{(m)}$  と分散  $\sigma_i^{2(m)}$  は次のように再推定できる。

$$\begin{aligned} P_i(d^{(m)}) &= \mathcal{N}(d_i^{(m)}; \eta_i^{(m)}, \sigma_i^{2(m)}) \\ &= \frac{1}{\sqrt{2\pi\sigma_i^{2(m)}}} \exp\left\{-\frac{(d_i^{(m)} - \eta_i^{(m)})^2}{2\sigma_i^{2(m)}}\right\} \end{aligned} \quad (32)$$

$$\begin{aligned} \eta_i^{(m)} &= \frac{\sum_{t^{(m)}=1}^{T^{(m)}} \sum_{d^{(m)}=1}^{t^{(m)}} \gamma_{t^{(m)}}^{d^{(m)}}(i) \cdot d^{(m)}}{\sum_{t^{(m)}=1}^{T^{(m)}} \sum_{d^{(m)}=1}^{t^{(m)}} \gamma_{t^{(m)}}^{d^{(m)}}(i)} \end{aligned} \quad (33)$$

$$\begin{aligned} \sigma_i^{2(m)} &= \frac{\sum_{t^{(m)}=1}^{T^{(m)}} \sum_{d^{(m)}=1}^{t^{(m)}} \gamma_{t^{(m)}}^{d^{(m)}}(i) \cdot d^{2(m)}}{\sum_{t^{(m)}=1}^{T^{(m)}} \sum_{d^{(m)}=1}^{t^{(m)}} \gamma_{t^{(m)}}^{d^{(m)}}(i)} - \eta_i^{2(m)} \end{aligned} \quad (34)$$



表 1 実験条件

原画像サイズ	720 × 576
使用画像サイズ	64 × 64, grayscale
学習データ	6枚/1人 × 100名分
テストデータ	2枚/1人 × 100名分
手法	HMM, HSMM
状態数	8×8, 16×16, 24×24, 32×32, 40×40
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)

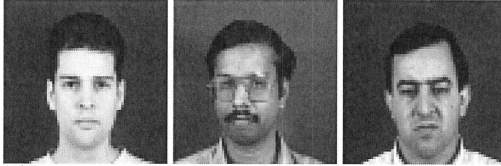


図 6 位置・大きさを正規化した画像



図 7 位置・大きさの変動を加えた画像

#### 4. 認識実験

本稿では、位置や大きさの変動が大きい画像に対する提案法の有効性を確認するため、顔画像認識実験を行った。実験に使用する画像データは、XM2VTS データベース [5] を用いた。使用した画像データの条件及び実験条件を表 1 に示す。実験には、位置・大きさを正規化した画像、及びガウス乱数により位置・大きさの変動を加えた画像を用いた。図 6, 7 にそれぞれ、位置・大きさを正規化した画像と位置・大きさの変動を加えた画像の例を示す。

はじめに、変動を正規化した画像に対して、認識実験を行った。実験結果を図 8 に示す。図 8 より、認識率は状態数が 8×8, 16×16, 24×24, 40×40 において、提案法が従来法より上回る結果が得られた。しかし、状態数が 32×32 においては、提案法が従来法の認識率を若干下回る結果となった。これは、初期値が不適切で局所的な最適解に収束してしまったためであると考えられる。

次に、位置・大きさの変動を加えた画像に対して認識実験を行った。実験結果を図 9 に示す。図 9 をみると、全ての状態において、提案法が従来法の認識率を上回る結果が得られた。状態数が 8×8, 16×16, 24×24 の時、提案法による認識率の改善幅が大きく、特に状態数が 8×8 の時、最大で 7% 認識率が向上した。状態数が多い場合、提案法による認識率の改善幅が小さいが、これは状態数の増加とともにモデルの自由度が高くなり提案法による有効性が表れなかったものであると考えられる。

#### 5. むすび

本稿では、状態継続長を考慮した分離型 2 次元格子 HMM

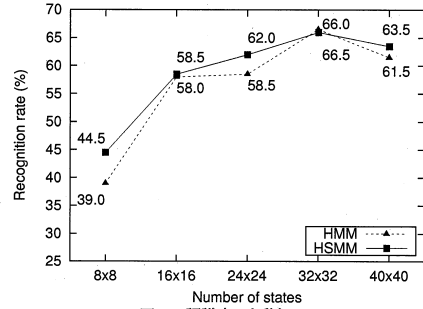


図 8 認識率 (変動無し)

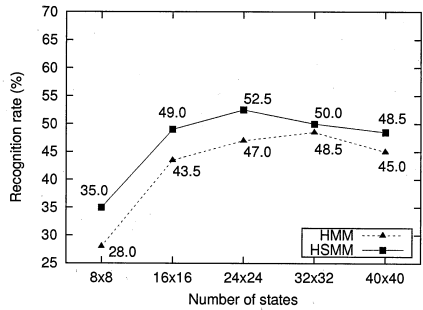


図 9 認識率 (変動有り)

を提案し、画像が大きく変動した場合でも対応可能となるモデル構造について検討した。また、提案法の有効性を確認するため、XM2VTS データベースを用いて顔画像の認識実験を行った。実験の結果、HSMM によるモデルの構築を行う提案法により、位置・大きさを正規化した画像、及び大きく変動を加えた画像に対して、従来法より高い認識率を得ることができた。これより、提案モデルによって変動の大きさによらず頑健な顔画像認識が可能であることが示された。

今後の課題として、確定的アンニーリングによる学習アルゴリズムの改善などが挙げられる。

#### 文 献

- [1] Yoshihiko Nankaku, Keiichi Tokuda, Tadashi Kitamura, and Zoubin Ghahramani, "Separable Lattice Hidden Markov Models," IEICE TRANS, INFORMATION AND SYSTEM, VOL. E00-D, NO.1 Jan. 2003.
- [2] S.E. Levinson, "Continuously variable duration hidden Markov models for automatic speech recognition," Computer Speech and Language, pp.29-45, 1986
- [3] Keiichiro Oura, Haiga Zen, Yoshihiko Nankaku, Akinobu Lee, Keiichi Tokuda, "Hidden Semi-Markov Model Based Speech Recognition System Using Weighted Finite-State Transducer," ICASSP2006, vol.1, pp.33-36, 2006-05.
- [4] 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.
- [5] K. Messer, J. Mates, J. Kittler, J. Luetttin, and G. Maitre, "XM2VTSDB: The Extended M2VTS Database," Audio- and Video-Based Biometric Person.