

HMM を利用したジェスチャー認識の高性能化

清水 宏明 岩井 儀雄 谷内田 正彦

大阪大学大学院 基礎工学研究科 システム人間系専攻

Abstract

隠れマルコフモデル (HMM) は時系列データを有効に扱えるため、ジェスチャー認識によく利用されているが、実際のシステムで用いるためには、計算コストの問題と認識精度の問題が重要である。本論文では、計算コストの問題に対処するため、KL 展開によって特徴量を次元圧縮する手法を提案し、計算時間、認識率について検討する。また、認識精度の問題に対処するため、ジェスチャー間のつながり (コンテキスト) を考慮した HMM によるジェスチャー認識の手法を提案し、その有効性を実験により確かめる。

Performance Improvement of Gesture Recognition using HMM

Hiroaki SHIMIZU, Yoshio IWAI, Masahiko YACHIDA

Department of Systems and Human Science, Graduate School of Engineering Science,
OSAKA University

Abstract

HMMs are often used for gesture recognition because of the robustness. However, the computational cost and accuracy of recognition are important for real applications such as gesture recognition, speech recognition or virtual reality.

In this paper, we propose methods for performance improvement of gesture recognition using HMMs. For the computational cost, we use KL transform to compress the input information and propose a recursive calculation method for the HMMs' probabilities. For the accuracy of recognition, we use an automaton layered up on HMMs to deal with context information of gestures. We also show experimental results to make the efficiency of our methods clear

1 はじめに

近年は、誰もが手軽に計算機と接する機会を得られるようになった。しかし、人間と計算機とのインタラクションにはマウスやキーボードといった間接的で機械的な装置を利用することがほとんどで、操作にはある程度の熟練を要することが多い。そのため、直接的で自然なマンマシンインターフェイスが求められている。その中でもジェスチャーは、人か

らのメッセージを計算機に伝えるのに有効な手法であり、ジェスチャーを認識しようとする研究が広く行われている。ジェスチャーを認識する技術には接触型センサによるものと、非接触型センサとに大別される。接触型センサは大がかりで装着した人に身体的拘束を与えるので、非接触型センサの方が自然なインターフェイスの実現という点で望ましいと言える。この非接触型のセンサの代表的なものにカメラがある。カメラから得られた画像は情報量が多く、

ジェスチャー認識によく用いられる。

これまでも動画像を用いた認識手法は色々提案されているが、このうち時系列データを扱うものには、DP マッチングによるもの [1] と、隠れマルコフモデル (HMM) によるものがある。以前は、主に DP マッチングによる手法がよく用いられていたが、近年では、音声認識の分野で成果を上げている HMM が広く用いられている。HMM は確率的な状態遷移モデルであり、大量の時系列データを統計的に扱えるため学習が可能であり、DP マッチングによる手法よりも、不特定者認識に対して有利であることが示されている [2]。HMM を用いた研究に、カテゴリ別 VQ を用いた動作認識法がある [3] [4]。これは、画像からメッシュ特徴を抽出しベクトル量子化してシンボル系列を生成し、HMM によって認識する手法である。同様に、カテゴリ別 VQ を表情認識に用いた例もある [5]。また、顔と手の 3 次元位置を用いた手法 [6] もある。

画像から得られる特徴は非常に多いため、HMM へ入力するシンボル系列を生成する量子化には時間がかかる。そのため、特徴量を情報圧縮することが考えられる。その手法として KL 展開による情報圧縮がある。KL 展開を用いた認識に、インタラクティブシステム構築のための動画像からの実時間ジェスチャー認識法がある [7]。これは、得られた特徴量を KL 展開することでジェスチャー空間に投影してジェスチャー曲線を作成し、モデル曲線と比較することで認識を行なうものである。また、特徴量として得られたオプティカルフローを KL 展開することでジェスチャー空間を生成するものもある [8]。KL 展開によって圧縮された特徴量を用いればベクトル量子化の処理時間は減るが、情報圧縮に処理を要するため単純に処理時間が減るわけではない。そのため、どの程度 KL 展開によって情報を圧縮するのかを検討することが重要である。

従来型の HMM による認識は、単独の動作や繰り返し動作には非常に有効ではある [8] が、連続する動作の認識に対しては個々の動作の変わり目を判別することが難しく、誤認識や認識率の低下を招く原因となっている。このため、HMM を改良しようという試みがなされている。これに対する研究としては、連続出力確率密度分布を用いた HMM による動画像からの複数人物の表情認識がある [9] [10]。これは、従来の left-to-right 型の HMM にさらに最終状態から初期状態へのパスを設け、連続出力確率密度

分布をもつ HMM を用いることで、動作の分離に成果を上げている。また、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 [11] では、部分隠れマルコフモデルという従来の HMM を改良したものをを用いることによって対処している。他にも HMM を改良した研究としては、Coupled HMM [12] などがある。

本研究では、まず、KL による情報圧縮を行ってからベクトル量子化を行い HMM へ入力するシステムと、画像から得られたオプティカルフローをそのままベクトル量子化し HMM へ入力するジェスチャーの認識システムを認識率、処理速度の点などで比較し、検討を行う。次に、HMM の認識に用いられる Viterbi アルゴリズムを改良した再帰的なアルゴリズムによる処理速度の検討を行なう。最後に、図 1 の認識部に、モデル間のつながりを考慮したオートマトンを追加したシステムを提案し、連続動作の認識に有効であることを確かめる。

2 KL 展開と HMM によるジェスチャー認識

2.1 認識手法の概要

処理の流れを図 1 に示す。まず、動画像から特徴量としてオプティカルフローを抽出し、特徴ベクトルを得る。オプティカルフローはテンプレートマッチング法によって求める。次に、得られた特徴ベクトルを KL 展開により情報圧縮を行い、次元の小さくなった特徴ベクトルを得る。そして、ベクトル量子化してシンボルに変換した後、HMM によって認識する。以降 2.2 ではフローの抽出について、2.3 では KL 展開について、2.4 では HMM について述べる。

2.2 動画像からのフローの抽出

動画像中から得られる画素をそのまま特徴量として用いた場合、背景や照明条件、服装などに影響されるため好ましいとはいえない。そこで、人物の動き情報 (オプティカルフロー) を特徴量とすれば、これらの環境条件に適した認識システムを構築できる。オプティカルフローは物体の移動量と方向とで表され、テンプレートマッチング法によって求められる。テンプレートマッチング法は、計算コストが

KL展開による情報圧縮と特徴抽出部

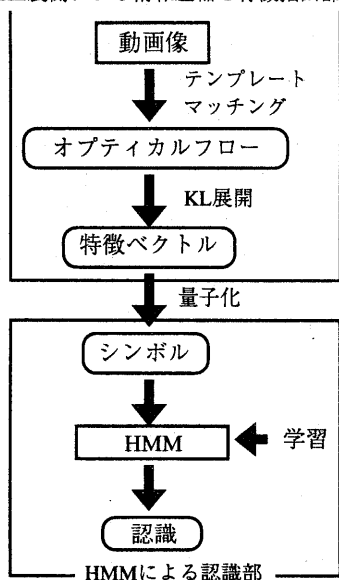


図 1: KL 展開による認識の処理の流れ

大きいことと、オプティカルフローの推定精度が悪いという問題点があるが、前者は専用のハードウェアを用いることで、後者は HMM を用いることで解決することができる。

入力画像の特徴量 X_i は、入力画像を $I \times J$ のブロックに分割し、各ブロック内で求められたオプティカルフローを、ラスタ走査をすることで得る。具体的には、次式 (1) のように並べる。

$$X_i = (x_1, x_2, \dots, x_S, y_1, y_2, \dots, y_S)^T \quad (1)$$

ここで、 X_i の次元数 S は $I \times J$ である。

2.3 KL 展開による情報圧縮と特徴抽出

2.3.1 KL 展開

動画像から得られた特徴量でパターン認識をするためには、動画像から得られた特徴量とモデルの特徴量との相関をとらなければならない。特徴量にオプティカルフローを用いた場合、次元数が大きくなるため、相関にかかる計算コストは大きくなる。この計算コストを小さくするためには、特徴量の次元数を小さくする必要がある。よって、次に示す KL

展開を用いて、得られたオプティカルフローの次元を落とし、低次元のパターン空間を作成する。

2.3.2 情報圧縮と特徴抽出

本論文では、KL 展開に適用する多変量の要素としてオプティカルフローを用いる。前述したように、KL 展開によって、多次元の多数のモデルをパターン空間に記述することが可能となる。具体的には、(1) 式の特徴量を各ジェスチャーの動作からそれぞれ得ることにより、計 N 個を採用し、

$$X = (X_1, X_2, \dots, X_N) \quad (2)$$

とすると、これを KL 展開して、固有ベクトル

$$\phi = (\phi_1, \phi_2, \dots, \phi_K) \quad (3)$$

を得る。ここで、固有値 λ_i に対応する固有ベクトル ϕ_i の要素は次式で表される。

$$\phi_i = (a_{i1}, a_{i2}, \dots, a_{iM})^T \quad (4)$$

ただし、 M は X_i の次元数に等価なので $M = 2 \times I \times J$ である。

また、第 i 固有ベクトル ϕ_i の要素を係数とする合計変量を z_i とすれば、

$$z_i = \phi_i X = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{iM}x_M \quad (5)$$

で表すことができ、これを第 i 主成分と呼び、その分散は λ_i である。

主成分の数に関して、複雑な現象をできるだけ単純化し、かつ、計算コストを小さくするという観点からいえば、少ない数の主成分で元の特徴量が表現できることが望ましい。本論文では、動画像から得られた特徴ベクトルを KL 展開を用いて次元圧縮した場合と、動画像から得られた特徴ベクトルをそのまま用いた場合とを比較し、少ない数の主成分で元の特徴量が表現できているかを実験により確かめる。

2.4 HMM によるジェスチャー認識

2.4.1 隠れマルコフモデル (HMM)

一般に、画像認識に用いられるマルコフモデル [13][14] は、出力シンボルによって一意に遷移状態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る。出力シンボル系列が与えられても状態遷移系列は唯一に決まらない。

観測できるのはシンボル系列だけであることから隠れマルコフモデルと呼ばれている。隠れマルコフモデルは、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りジェスチャー認識には有効な手段である。

画像認識に用いられる HMM M は一般の HMM に最終状態を付加したもので、次の 5 つの組 $M = (S, Y, A, B, \pi)$ で定義される [15]。

- S : 状態の有限集合 ; $S = s_i$
- Y : 出力シンボルの集合
- A : 状態遷移確率の集合 ; $A = \{a_{ij}\}$
 a_{ij} は状態 s_i から s_j への遷移確率
- B : 出力確率の集合 ; $B = \{b_{ij}(k)\}$
 $b_{ij}(k)$ は s_i から s_j への遷移時にシンボル k を出力する確率
- π : 初期状態確率の集合 ; $\pi = \{\pi_i\}$
 π_i は初期状態が s_i である確率

2.4.2 シンボル系列の生成

離散 HMM を用いて認識を行うためには、動画像から抽出された特徴量である多次元ベクトルを、そのまま HMM の入力シンボルとして扱うことができない。そこで、得られた特徴ベクトルをベクトル量子化してシンボルを生成する。

ベクトル量子化では、まず、すべてのモデルに対してクラスタリングを行ってクラスタ代表点を求め、各クラスタにラベル付けをしておく。つぎに、画像から得られた N 次元特徴ベクトル $X = (x_1, x_2, \dots, x_N)$ と、予め求めておいたクラスタ代表点 $C_i = (c_{i1}, c_{i2}, \dots, c_{iN})$ とのユークリッド距離を比較し、最も近いクラスタのラベルをシンボルとする。これを各フレームに対して行うことでシンボル系列を作成する。

2.4.3 HMM によるジェスチャー認識

HMM を用いた認識は、各ジェスチャーに対応する HMM M を予め学習させておき、観測されたシンボル系列 $Y_t = (y_1, y_2, \dots, y_T)$ がジェスチャー i である確率 $p(y|M)$ を各 HMM M に対して算出し、その確率が最大となる HMM を認識結果とする。

このように生成確率を高くするためには、予め HMM パラメータ $\{\pi, A, B\}$ を調整しておく必要がある。このパラメータの学習法として Baum-Welch の再推定式が確立されている [16]。この手法は多量

の学習サンプルを用いて、繰り返し学習することによりパラメータの推定を行う。

HMM の認識アルゴリズム [16][15] は、Forward アルゴリズムが知られている。このアルゴリズムは可能性のある全ての遷移状態について、観測されたシンボル系列を生成する確率の和を求めるものである。しかしながら、全ての状態遷移について計算すると時間がかかるため、計算速度を高速化した手法として Viterbi アルゴリズムがある。Viterbi アルゴリズムは、各々の状態遷移のうち、最も高い確率を与えるパスにおける出力確率を用いるものである。

2.4.4 再帰計算による高速化

Viterbi アルゴリズムは、入力されたシンボル系列の長さの分だけ状態遷移を行うが、毎回更新されるシンボル系列は最初と最後の部分だけであるので、ほとんどは前回と同じパスを通る。そのため、前回と同じ計算を再びする必要がある。そこで、この前回と同じパスを通る部分を毎回計算するのではなく、予め記憶しておくことにより確率計算を再帰的にを行い、高速化する手法を提案する。

時刻 t において、観測されるシンボル系列を $x^t = (x_1, x_2, x_3, \dots, x_T)$ 、Viterbi 経路による遷移状態を $S^t = (\pi, S_1, S_2, \dots, S_T)$ 、状態 π から S_1 へ遷移する時の遷移確率を a_0^t 、その際シンボル x_1 を出力する確率を b_0^t 、また、状態 S_1 から S_2 へ遷移する時の遷移確率を a_1^t 、シンボル x_2 を出力する確率を b_1^t 、HMM の生起確率を P^t とする。次に、時刻 $t+1$ において観測されるシンボル系列は $x^{t+1} = (x_2, x_3, x_4, \dots, x_T, x_{T+1})$ 、Viterbi 経路は $S^{t+1} = (\pi, S_2, S_3, \dots, S_T, S_{T+1})$ となる。この時刻 $t+1$ において、状態 π から S_2 への遷移確率を a_0^{t+1} 、シンボル x_2 を出力する確率を b_0^{t+1} 、状態 S_t から S_{T+1} への遷移確率を a_{T-1}^{t+1} 、シンボル x_{T+1} を出力する確率を b_{T-1}^{t+1} 、HMM の生起確率を P^{t+1} とすると、計算式は次のようになる。

$$P^{t+1} = \frac{a_0^{t+1} b_0^{t+1} a_{T-1}^{t+1} b_{T-1}^{t+1}}{a_0^t b_0^t a_1^t b_1^t} P^t \quad (6)$$

本論文では、実験により、この手法と Forward アルゴリズム、Viterbi アルゴリズムとの比較、検討を行う。

3 HMMの新しい手法

3.1 従来手法の問題点

従来のHMMによるジェスチャー認識の研究では、個々のジェスチャーの認識率を上げることに焦点を当てている。しかし、実際には複数のジェスチャーを続けて行うことが多く、ジェスチャーの分離が難しいため、従来の手法では認識率が低下してしまう。

また、実環境では、あるジェスチャーの後には別のあるジェスチャーを行う可能性が高いといったジェスチャー間のつながり（コンテキスト）が存在することが多いが、HMMではジェスチャー間に相関関係はない。

そこで、本論文では、これらの問題点に対応した新しい手法について提案する。

3.2 提案手法

時刻 t で、HMM M によってシンボル系列 $\mathbf{x}^t = (x_{t-T+1}, \dots, x_{t-1}, x_t)$ が生起する確率 P_M^t は次のように表される。

$$P_M^t = P(\mathbf{x}^t | M) \quad (7)$$

この生起確率は区間 T のシンボル系列によって決定するため、ある連続区間にわたって雑音が混入すると、この確率は小さくなってしまい、誤認識の原因となる。また、連続する異なったジェスチャーを認識しなければならない場合は、HMMの内部状態がモデルに1:1に対応しているわけではなく、モデル毎にHMMは独立しているため、ジェスチャーの分離が難しい。

そこで、ジェスチャーと状態が1:1に対応したオートマトンをHMMとは別に生成する(図2)ことで、雑音の影響を低減しコンテキストを考慮する。このオートマトンは安定状態と不安定状態からなり、まず、入力シンボル x_t を受けて、前回の認識結果(安定状態)から今回の認識候補(不安定状態)へと遷移する。この候補とHMMの出力から今回の認識結果を決定し、オートマトンは次の安定状態(認識結果)へと遷移する。このオートマトンを実現するために、つぎの新しい確率 P_C を考える。時刻 $t-1$ で認識されたモデルを o_{t-1} 、時刻 t で認識されたモデルを o_t とすると、新たな生起確率 $P_N(\mathbf{x}^t)$ を次のように表現する。

$$P_N(\mathbf{x}^t) = \{P_C(x_t | o_t, o_{t-1})\}^d P_M(\mathbf{x}^t | M) \quad (8)$$

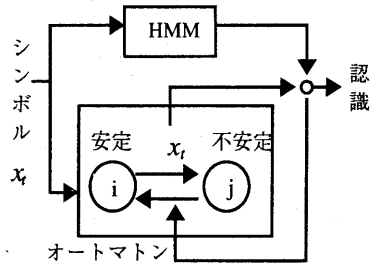


図2: 提案手法

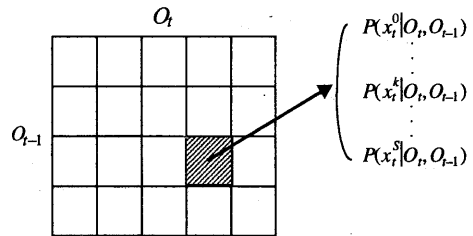


図3: オートマトンの遷移テーブル

ここで、 P_C は、状態を各ジェスチャーとするオートマトンから生起される確率、 d は単独のモデルによる影響とコンテキストによる影響のバランスをとるための減衰係数であり、実験によって決定する。この手法により、同じジェスチャーに対する個人差はHMMによって学習することができ、用途によって異なるジェスチャー間のコンテキストはオートマトンで学習できるため、ジェスチャーの認識では非常に有利である。

3.3 オートマトンの学習方法

オートマトンを学習するため、まず、連続する複数のジェスチャーの動画像からオプティカルフローを抽出しベクトル量子化しシンボルに変換した後、HMMにより学習し認識結果を得る。次に、その認識結果と量子化した際に得られたシンボルから、時刻 $t-1$ において認識モデルが i つまり $o_{t-1} = i$ であり、時刻 t への遷移の際にシンボル x_t が観測され、時刻 t での認識モデルが j である確率 $p(x_t | o_t = j, o_{t-1} = i)$

を次のように計算する。

$$p(x_t | o_t = j, o_{t-1} = i) = \frac{\#\{x_t | o_t = j, o_{t-1} = i\}}{\sum_k \#\{x_t | o_t = k, o_{t-1} = i\}} \quad (9)$$

ここで、 i, j はHMM M を表すモデルの番号であり、 i は前フレームの認識で一意に決定される。 $\#$ は集合の要素数を返す演算子である。これを図3に示した各セルに対して求めることで遷移テーブルを生成する。

HMMの学習は従来通り Baum-Welch の推定アルゴリズムによって行う。

4 実験及び考察

4.1 実験環境

動画からのオプティカルフローの抽出には富士通製カラー-trackingビジョン (TRV-CPW5) を用いた。本実験では 16pixel ごとに 18x10 の計 180 点をのフローを抽出した。このフロー抽出に掛かる処理時間は約 20ms/frame である。これにより生成される特徴ベクトルは、 x, y 軸について求めるため 360 次元となる。

認識に用いたジェスチャーは、楽器 (ギター、バイオリン、カスタネット、ドラム、ピアノ、シンバル、チェロ) を演奏する 7 種類のジェスチャーとする。ジェスチャー空間作成のためのモデル画像数は 1 つのジェスチャーにつき 400 枚、合計 2800 枚用いた。また、HMM の学習に用いたフレーム数は 1 つのジェスチャーにつき 600 フレームであり、学習人数は 7 人である。認識には、この学習に用いたフレームとは異なるものを 14 人分用いた。

4.2 KL 展開による認識実験

KL 展開により動画から得られたオプティカルフローを情報圧縮をした特徴ベクトルを用いて、HMM によりジェスチャー認識を行った場合と、そのまま HMM によるジェスチャー認識を行った場合の認識率を図4に示す。ここでは、主成分の数を 40 とした。また、チェロをするジェスチャーに対して、主成分数を変化させた時の認識率の変化を図5に示し、主成分数を変化させた時の認識時間の変化を図6に示す。

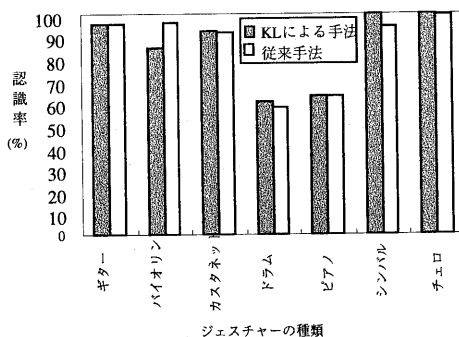


図 4: KL による手法と従来手法の認識率

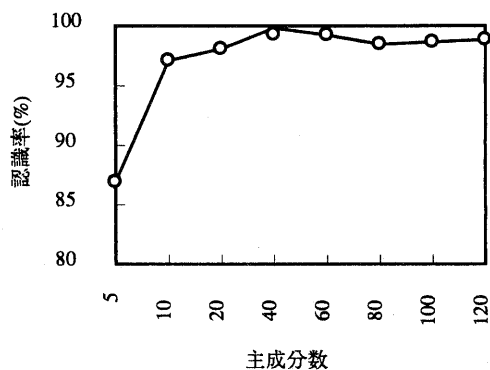


図 5: 主成分数の変化に対する認識率の変化

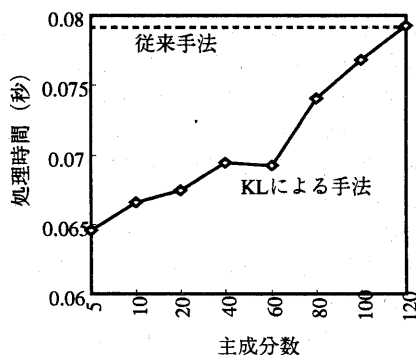


図 6: 主成分数の変化に対する処理時間の変化

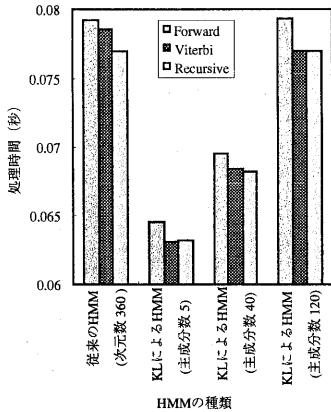


図 7: 各認識アルゴリズムの処理時間

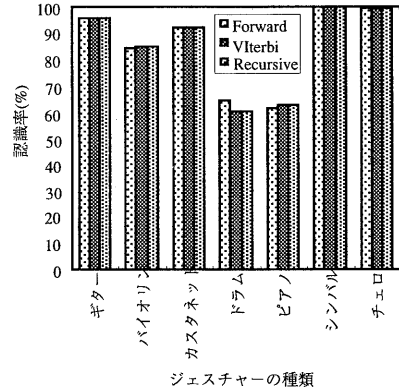


図 8: 各認識アルゴリズムの認識率

4.3 再帰計算を用いた認識実験

認識アルゴリズムに再帰的な計算を用いた場合の処理時間、認識率を、Forward アルゴリズム、Viterbi アルゴリズムを用いた場合の結果と合わせて図 7, 図 8 に示す。

4.4 オートマトンを用いた HMM による認識実験

オートマトンを用いた HMM によるジェスチャー認識では、まず、図 4 の認識実験で使用した認識用のデータを用いて、個々のジェスチャーに対する認識精度を調べた。この結果を図 9 に示す。

次に、コンテキスト情報を考慮した場合の認識精度を調べるため、予め、ギター、バイオリン、カスタネット、ドラム、ピアノ、シンバル、チェロの順に連続して行った動画像 15,000 フレームを用いてオートマトンの学習を行った。そして、学習で用いたものとは別の 3000 フレームに対して、オートマトンを用いた HMM のシステムと HMM のみのシステムとで認識モデルの遷移状態を調べた。この結果を図 10 に示す。

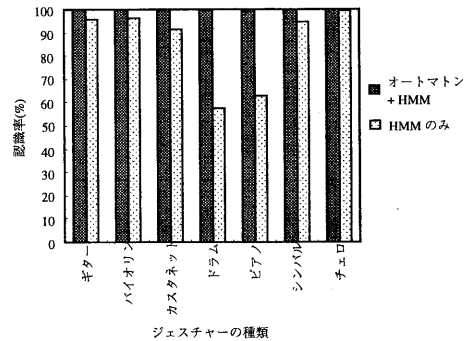


図 9: オートマトンを用いた HMM の個々のジェスチャーに対する認識結果

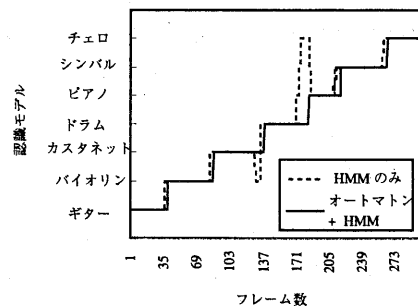


図 10: 連続ジェスチャーに対する認識モデルの遷移

4.5 考察

図 4 から、主成分数を 40 とした場合、つまり 360 次元の特徴量を 40 次元に次元圧縮した場合でも十分に元の特徴を表現できているといえる。ピアノと

ドラマの認識率が特に低いのは、この2つのジェスチャーが似ているため、互いに他方のジェスチャーとして認識されることと、フローの目が荒いため、特徴を得るのに十分にフローが抽出できなかったためであると考えられる。

また、図4、図5、図6から、主成分数が40ぐらいであれば、情報圧縮を行わない手法と同程度の認識精度があり、かつ、処理速度が高速化していることがわかる。今回の実験では、動画像から得られた特徴量の次元数は360次元であったが、より精度の高い認識を行なうたは、オプティカルフローの抽出間隔を密にしなければならない。その場合、特徴量の次元数は今回よりも増加するため、この2つの手法の処理時間の差はますます大きくなることが予想される。したがって、特徴ベクトルの次元数が高いほど、KL展開による手法が処理時間の点で有利といえる。

次に、図7から、今回提案した再帰的な計算による認識アルゴリズムと、Viterbiアルゴリズムとの処理時間の差は、Forwardアルゴリズムとの差に比べて小さく、処理時間短縮にはあまり効果はないことがわかった。また、図8からこの3種類のアルゴリズムによる認識率の差はほとんどないことが確かめられた。

オートマトンを用いたHMMの認識実験では、図9から個々のジェスチャーの認識に対しても、有利であることが確かめられた。ピアノとドラマのジェスチャーに対する認識率が飛躍的に向上したのは、オートマトンが現在の認識モデルを状態として持っているため、互いに他方のジェスチャーに誤って認識されることが大幅に減ったためであると考えられる。

また、図10から、ジェスチャーが移り変わる際に生じる誤認識が抑えられるためジェスチャーの分離を容易に行うことができることが確認された。

5 まとめ

本論文では、KL展開によるHMMを用いたジェスチャー認識の処理時間に対する優位性について述べ、また、コンテキスト情報を用いるため、オートマトンを認識部に取り入れたHMMによるジェスチャー認識の手法について提案し、その有効性を確かめた。今後は、このオートマトンのより優れた学習方法について検討していくつもりである。

参考文献

- [1] 高橋勝彦, 関進, 小島浩, 岡隆一. “ジェスチャー動画像のスポッティング認識”. 信学論 (D-II), Vol. J77-D-II, No. 8, pp. 1552-1561, 1994.
- [2] 武士展照, 春山智, 小林哲則. “HMMを用いた手振り認識”. *PRMU96-8*, May 1996.
- [3] 大和淳司, 大谷淳, 石井健一郎. “隠れマルコフモデルを用いた動画像からの人物の行動認識”. 信学論 (D-II), Vol. J76-D-II, No. 12, pp. 2556-2563, 1993.
- [4] 大和淳司, 倉掛正治, 伴野明, 石井健一郎. “カテゴリー別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, pp. 1131-1318, 1994.
- [5] 坂口竜己, 大谷淳, 岸野文朗. “隠れマルコフモデルによる顔動画像からの表情認識”. テレビジョン学会誌, Vol. 49, No. 8, pp. 1060-1067, 1995.
- [6] L.W.Campbell, D.A.Becker, A.Azarbayejani, A.F.Bobock, and A.Pentland. “Invariant features for 3-D gesture recognition”. In *Proc.International Conference on Automatic Face and Gesture Recognition*, pp. 157-162, 1996.
- [7] Takahiro WATANABE, Chil-Woo LEE, and Masahiko YACHIDA. “Recognition of Complicated Gesture in Real-Time Interactive System”. In *5th IEEE International Workshop on RO-MAN*, pp. 268-273, 1996.
- [8] 梶直志, 岩井儀雄, 谷内田正彦. “動き情報と情報圧縮を用いたロバストなジェスチャー認識手法”. 信学論 (D-II), Vol. J80-D-II, No. 9, pp. 1983-1992, 1998.
- [9] 大塚尚宏, 大谷淳, 中津良平. “連続出力確率密度分布を用いたHMMによる動画像からの複数人物の表情認識”. 信学論 (D-II), Vol. J80-D-II, No. 8, pp. 2129-2137, 1997.
- [10] 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”. In *PRMU97-154*, November 1997.
- [11] 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”. In *PRMU97-203*, January 1998.
- [12] M.Brand, N.Oliver, and A.Pentland. “Coupled hidden Markov models for complex action recognition”. In *CVPR97*, Puerto Rico, 1997.
- [13] 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.
- [14] Lawrence Rabiner. “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”. *Proceeding of the IEEE*, Vol. 77, No. 2, 1989.
- [15] 中川聖一. 「情報理論の基礎と応用」. 近代科学社, 1992.
- [16] 小林哲則. “隠れマルコフモデルに基づく音声認識”. 電気学会論文誌, Vol. 113-C, No. 5, pp. 259-301, 1993.