

連続状態モデル間の相互作用に基づく多視点動作認識

川嶋 宏彰[†] 松山 隆司[†]

Multi-Viewpoint Gesture Recognition by an Integrated Continuous State Machine

Hiroaki KAWASHIMA[†] and Takashi MATSUYAMA[†]

あらまし 複数のセンサから得られる観測データや、マルチモーダルな特徴量からなる複数の時系列データを統合しながら、ジェスチャーなどの動的事象を識別するモデルとして、「Integrated Continuous State Machine (ICSM)」を提案する。これは、各特徴系列を個別に識別する状態遷移モジュールをそれぞれ用意し、それを連続状態空間において動的に統合するものであり、各モジュールの信頼性を逐次計算することにより、クラスの識別にとってより重要な入力へ、動的にフォーカスを切り替えることが可能となる。本論文では、統合に適したモジュールとして、連続状態空間 (実ベクトル空間) 中での状態遷移によってクラス識別を行う Continuous State Machines (CSM) を導入し、続いて、複数の CSM 間における、状態の共起性に基づいた相互作用を導入する。連続状態空間を用いることにより、モジュール内の状態遷移とモジュール間の相互作用を、いずれもエネルギー関数最小化の枠組みで定式化することができる。本手法の有効性を、2 視点カメラを用いた手話認識に適用することによって検証した。

キーワード 情報統合, 連続状態空間, 状態遷移モデル, 状態の共起性, 時系列パターン認識

1. ま え が き

近年、時系列パターン認識の手法として、音声認識には HMM や連続 DP を用いた手法が提案・実用化され、動画像認識の研究においても広く用いられている [1] ~ [3]。しかし、これら多くの提案手法は、識別に必要な観測データが常に得られることを条件としており、実世界の多様に化する環境下では、対象に遮蔽やノイズなどの変動が生じ観測データが不完全となるため、安定した認識が困難である。

我々は、このような実環境下で頑健な認識を行うには、複数のセンサやマルチモーダルな特徴量からなる時系列情報を、認識時に「動的に」統合することが不可欠であると考え。これは、単一のセンサから得られる観測データや単一の特徴量では、識別に有効な情報を、一時的もしくは長期にわたって失っている場合があるためである。このとき、他の観測データから識別に必要な情報を補うとともに、不要な情報を一時的に切り捨てることにより、識別の信頼性を上げるこ

とができると考えられる。そこで、本論文では、各時系列パターンの識別を個々に行うことのできる認識モジュールをそれぞれ用意し、そうしたモジュール間で情報を統合しながら動的な事象の認識を行う、統合型システムアーキテクチャを提案する (図 1)。各認識モジュールへの入力は、異なるセンサ (複数カメラ、マイクなど) から得られるマルチモーダルな観測データや、異なる特徴量 (部分画像、シルエットやフローなど) からなる時系列データである。

一般に、このようなモジュール統合型のアーキテクチャを設計する際には、

(1) 各認識モジュールとして何をを用いるか

(2) 複数のモジュールをどのように統合するかが重要である。本論文では、はじめに、システムを構成する認識モジュールとして、連続空間 (実ベクトル空間) 中での状態遷移を行うモデル Continuous State Machine (CSM) を新たに導入する。

CSM は予め学習されたダイナミクスにしたがって自律的に状態を遷移することが可能であり、外部からの入力の変化が、予め学習された CSM 内部のダイナミクスと同様の変化である場合には、CSM の状態は特定の状態に遷移していく。この状態を観測することにより、入力パターンの識別結果を得ることができる。

[†] 京都大学大学院 情報学研究科
〒 606-8501 京都市左京区吉田本町
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo, Kyoto 606-8501

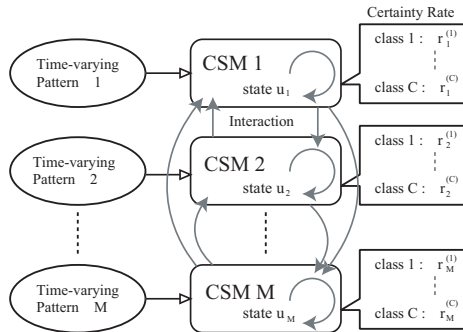


図 1 システムアーキテクチャ
Fig.1 System Architecture.

この点で、従来の HMM [1], [2] や連続 DP マッチング [3] のように入力データとモデルとのマッチングに基づいて識別を行うモデルとは大きく異なる。例えて言えば、この自律的な状態機械は、外部の入力に対して「共鳴・同期」を起こすことで時間的な変化パターンの識別を行うモデルであると考えられる。さらに、状態空間がメトリックを持っているために状態間の類似性が表現可能であり、モジュール内の状態遷移、および次に述べるモジュール同士の統合を、状態間の距離を用いたエネルギー関数の最小化として直感的に定式化できる。

次に、モジュールの統合法としては、複数 CSM 間における動的な相互作用を導入する。マルチモーダルな情報を統合した認識を行う際には、リップリーディングにおける発話内容と口の動きなどのように、同時に起りうる状態の共起性に注目する必要がある。そこで、ここでの相互作用とは、あらかじめ学習された CSM 間の状態の共起性に基づいて、認識時に互いに相手の状態を推定し、状態レベルで情報を統合する機能とする。実世界において、このように複数の時系列情報・モダリティを用いて認識を行う際には、各特徴量の信頼性は時間的に変化しており、一概にどの特徴量の信頼性が相対的に高いということは言えない。そこで、各 CSM が他の CSM に与える影響力を、識別の信頼性に応じて動的に制御することにより、識別に重要な情報への動的なフォーカシングが実現できる。

まず 2. でシステムの概要について述べる。続く 3. では単一 CSM におけるダイナミクスの詳細を、4. では、複数 CSM の統合手法の詳細を述べる。5. では、提案法の有効性を評価し、最後に 6. で、結論と今後の課題について述べる。

2. システムアーキテクチャ

認識システムは M 個の CSM から構成され、各 CSM にはそれぞれ別の観測データが独立に入力される (図 1)。さらに、全ての CSM は動的な相互作用によって互いに結合されている。このような統合されたアーキテクチャを、Integrated CSM (ICSM) と呼ぶことにする。

2.1 Continuous State Machines

連続状態空間 (実ベクトル空間) における状態遷移モデルとしては、既に各種のリカレントニューラルネットワーク [4]~[6] が提案されている。しかし、いずれのモデルも、状態遷移則を決定するにあたり明確な根拠はなく、状態遷移則における外部入力と状態の演算が常に一定であることも合理的でない。さらに、パラメータの調節も経験的になりがちである。一方、カルマンフィルタは、確率過程と最尤推定に基づいて連続状態空間での状態推定式を導くため、外部入力とシステムによる状態予測の信頼性をそれぞれ表現可能であるが、予測やスムージングに用いる場合が大半であり [7], [8], クラス識別の手法としての応用は見られない。本論文で提案する CSM は、カルマンフィルタの逐次最適化手法とリカレントネットワークの時系列パターン認識手法を統合したものであり、外部からの入力と予測に基づいたエネルギー関数を最小化することによって状態遷移則が導かれる (3.1 にて詳述する。以下同様)。

図 2 に示すように、学習時には、各クラスの時系列パターンを、それぞれ軌道アトラクタとして状態空間に埋め込む。さらに、それぞれのアトラクタにしたがって遷移するようなダイナミクスを学習する。基本的な HMM では各クラス毎に状態空間を持ち、それぞれが独立にモデルの遷移確率を計算するが、CSM では全クラスが同じ状態空間中の軌道アトラクタとして埋め込まれるため、軌道の設計法を工夫することによってクラス間の相互抑制などを実現することができる (3.3)。

認識時には、CSM は導出された状態遷移則に従って状態を遷移させる。このとき、入力パターンの変化が、予め学習したダイナミクスと同様の変化であれば、CSM の状態は学習した軌道のいずれかに引き込まれるため、どの軌道に引き込まれたかによって識別結果を得ることができる (3.5)。

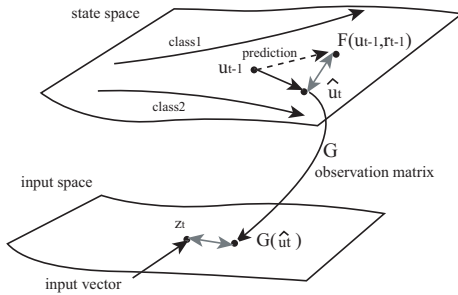


図 2 状態空間中に形成された軌道アトラクタ (2本の軌道がそれぞれのクラスに対応する.)
 Fig. 2 Trajectories of two classes in the state space. (Each trajectory corresponds to each class.)

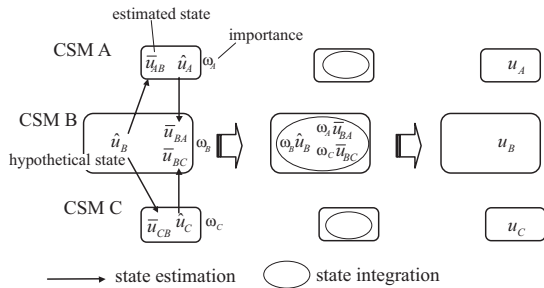


図 3 CSM 間の動的な相互作用 (CSM A, C 間の相互作用は省略している.)
 Fig. 3 The interaction among CSMs. (The interaction between CSM A and C is omitted in this figure.)

2.2 確信度分布

各 CSM においては、各クラスに対する確信が時々刻々変化するものとし、これを測る尺度として「確信度」を定義する。時刻 t において、CSM m は確信度 $r_{m,t}^{(c)}$ をクラス c ($c = 1, \dots, C$) に対して持つとする。すなわち、各 CSM は確信度の分布を、ベクトル $r_{m,t} = (r_{m,t}^{(1)}, \dots, r_{m,t}^{(C)})^T$ (T は転置) として保持し、これを時間と共に変化させていく。ただし、 $\sum_{c=1}^C r_{m,t}^{(c)} = 1$ が成り立つように正規化しておく。

2.3 CSM 間の相互作用

過去に提案されている時系列情報の統合手法としては、Coupled HMM [2] や Wada らの方法 [9] が挙げられ、モジュールとしてはそれぞれ HMM、非決定性有限オートマトンが用いられている。これらの統合方法は、いずれもモジュール同士の状態の共起性に注目し、同時に起こりえない状態の組み合わせに対しては、認識時に評価値が低くなるように設計されたものである。

本論文も、状態の共起性に注目し、状態レベルでの

統合を実現する。しかし、各 CSM によって推定された状態を、全て同じ重みで扱うのは適切でない。これは、システムを構成する CSM のいくつかは、不完全なデータに基づいて認識を行っており、そのような信頼性の低い CSM がシステム全体に悪影響を及ぼす (認識率の低下を招く) 可能性があるからである。そこで、各 CSM の重要度を定義し、これを用いて各 CSM の影響力を制御することによって、より認識にとって重要な観測データにフォーカスする機能を設ける。

以上より、複数の CSM 間での相互作用は次の 2 つの機能によって実現される (図 3)。

他の CSM の状態推定:

いったん各 CSM が状態遷移によって一時的な状態 (仮状態) を計算すると、各 CSM はあらかじめ学習された状態の共起性 (4.1) に基づいて、他の CSM の状態を推定する (4.2)。

推定された状態の統合:

各 CSM は、他の CSM によって推定された状態を自分の仮状態と統合し、状態を更新させる (4.4)。ただし、それぞれの CSM の影響力は、どのクラスに対しても同程度の確信度を持っている CSM に対しては弱く (重要度を低く) する。逆に、ある特定のクラスに対してのみ高い確信度を持っている場合は強く (重要度を高く) する (4.3)。

3. 単一 CSM のダイナミクス

3.1 状態遷移

CSM における状態遷移は、制御システムにおける状態空間モデルと同様に、実ベクトル空間中での遷移を考える。各時刻の入力 z_t の独立性、および状態のマルコフ性を仮定すると、次の時刻の状態は、システムそのものの予測 (トップダウン的情報) と、外部からの入力 (ボトムアップ的情報) によって推定できる。そこで、CSM の状態遷移は、各時刻 t においてエネルギー関数

$$E_{1,t}(u_t) = (1 - \kappa_t) \|u_t - \mathcal{F}(u_{t-1}, r_{t-1})\|^2 + \kappa_t \|Gu_t - z_t\|^2 \quad (1)$$

を最小化することによって決まるものとする (図 2)。ただし、 $\|\cdot\|$ はベクトルのユークリッドノルムをあらわし、 u_t, z_t と r_t はそれぞれ時刻 t における状態、外部入力、確信度分布をあらわす。また、 $\mathcal{F}(\cdot)$ は状態の予測関数であり、時刻 $t-1$ における状態だけでなく、そのときの確信度分布 r_{t-1} から次の時刻 t の状

態を予測する．なお，一般に関数是非線形であってもかまわないとする． G は観測行列であり，状態空間から観測空間への写像を行う． κ_t ($0 \leq \kappa_t \leq 1$) は外部からの入力強度である．

このとき，式 (1) における入力強度はシステムの予測と外部入力の影響力のバランスを決定する．すなわち，式 (1) の右辺第一項は，状態空間における予測との距離を，第二項は観測データからの距離をあらわしており，両者の trade-off から，次の状態が決まる．これは，カルマンフィルタ [7], [8], [10] における最尤推定と本質的に同様の考え方であり，不確実性を表す共分散行列を単にスカラー値として簡略化したものとして考えることができる． κ_t は外部入力と予測の信頼性に応じて動的に制御されるパラメータであるが，今回は簡単化のために， κ_t は時間的に不変であり $\kappa_t = \kappa(\text{const.})$ とする．

ここで， $\frac{\partial E_{1,t}}{\partial \mathbf{u}_t} = 0$ とすることにより，時刻 t において $E_{1,t}$ を最小化する状態 $\hat{\mathbf{u}}_t$ を求めると，次のような状態遷移則が得られる．

$$\hat{\mathbf{u}}_t = K^{-1}\{(1-\kappa_t)\mathcal{F}(\mathbf{u}_{t-1}, \mathbf{r}_{t-1}) + \kappa_t G^T \mathbf{z}_t\} \quad (2)$$

ただし， $K = (1-\kappa_t)I + \kappa_t G^T G$ であり， I は単位行列を表す．今は κ_t を一定としていることから，式 (2) はちょうど Elman 型 [4] と同様のモデルになる．

なお，式 (2) より求まる状態 $\hat{\mathbf{u}}_t$ を，時刻 t における仮状態と呼ぶ．これは，後の 4. で述べるモジュール間の相互作用によってさらに更新されるからである．

3.2 確信度に基づく状態の予測

時刻 t における状態 \mathbf{u}_t を \mathbf{u}_{t-1} から予測するには RBF や MLP [11]，非単調ニューラルネット [5], [6] などの非線形な写像を用いることもできるが，今回は単純に次のような関数を用いる．

$$\mathcal{F}(\mathbf{u}_{t-1}, \mathbf{r}_{t-1}) = \sum_{c=1}^C r_{t-1}^{(c)} \mathcal{F}^{(c)}(\mathbf{u}_{t-1}) \quad (3)$$

ただし， $\mathcal{F}^{(c)}(\mathbf{u}_{t-1})$ は各クラスの条件付き予測関数である．この式は， $\mathcal{F}(\cdot)$ が各クラス c の条件付き予測関数の線形和で近似できると仮定している．ここで，各条件付き予測関数の係数は各時刻 $t-1$ の確信度 $r_{t-1}^{(c)}$ である．さらに，各クラス c の条件付き予測関数 $\mathcal{F}^{(c)}(\mathbf{u}_{t-1})$ は線形であると仮定し，遷移行列 $W^{(c)}$ を用いて

$$\mathcal{F}^{(c)}(\mathbf{u}_{t-1}) = W^{(c)} \mathbf{u}_{t-1} \quad (4)$$

のようにあらわす． $W^{(c)}$ は 3.4 で述べる学習によって決定される．

ここで，確信度を以下のように定式化する．

$$r_t^{(c)} \propto \exp(-d_t^{(c)}/\sigma^2) \quad (5)$$

ただし， $r_t^{(c)}$ ($c = 1, \dots, C$) は $\sum_{c=1}^C r_t^{(c)} = 1$ を満たすように正規化されている．また， $d_t^{(c)}$ は予測と外部入力の二乗誤差

$$d_t^{(c)} = \|\mathbf{z}_t - G\mathcal{F}^{(c)}(\mathbf{u}_{t-1})\|^2 \quad (6)$$

である．式 (5) において，二乗誤差を $[0, 1]$ の範囲の正規化された値に写像するために *softmax* 関数 [11] を用いている．

3.3 状態空間と軌道アトラクタの設計

前節までに述べた状態遷移を用いて時系列パターン認識を行うには，状態空間が次の 2 つの要件を満たすように設計する必要がある．

(1) 現在の状態に応じた次状態の予測

(2) 観測された時系列パターンのクラス識別

ところで，時系列パターンは，それぞれが入力空間における軌道としてとらえることができる．今，各クラスについて 1 つの代表的な時系列パターンを考えると，一般に，入力空間におけるこれらの軌道は，クラス内やクラス間で交差や重なりを持つ．すなわち，入力空間では予測すべき次の入力が一意に決まらず，さらにクラス同士も分離できない．したがって，上の 2 つの要件を満たす状態空間は，各クラスの軌道が，クラス内およびクラス間でともに分離されるような空間であると考えられる．そこで，直接的な方法として，入力空間と判別空間の直積空間を考えることにする．

今， $S_I^{(1)}, \dots, S_I^{(C)}$ が C クラスの学習する時系列パターンであるとする．ここで行列 $S_I^{(c)} = [\mathbf{s}_{I,1}^{(c)}, \dots, \mathbf{s}_{I,L-1}^{(c)}, \mathbf{s}_{I,L}^{(c)}]$ はクラス c の時系列パターンである．ただし，各クラスでパターンの長さ L は一定であると仮定する．このとき， $\mathbf{s}_{I,t}^{(c)}$ はある時刻での入力ベクトルであり，入力空間中の一点となり， $S_I^{(c)}$ は入力空間中の軌道となる．

この $S_I^{(c)}$ に対し，付加的な時系列パターン $S_O^{(c)} = [\mathbf{s}_{O,1}^{(c)}, \dots, \mathbf{s}_{O,L-1}^{(c)}, \mathbf{s}_{O,L}^{(c)}]$ を加える．このとき，パターン全体としては $S^{(c)} = [S_I^{(c)T}, S_O^{(c)T}]^T$ となる．なお， $\mathbf{s}_{I,t}^{(c)}, \mathbf{s}_{O,t}^{(c)}, \mathbf{s}^{(c)} = [\mathbf{s}_{I,t}^{(c)T}, \mathbf{s}_{O,t}^{(c)T}]^T$ の次元数は，それぞれ $N_I, N_O, N (= N_I + N_O)$ とする．

先に述べた 2 つの要件を満たすような $S_O^{(c)}$ のコー

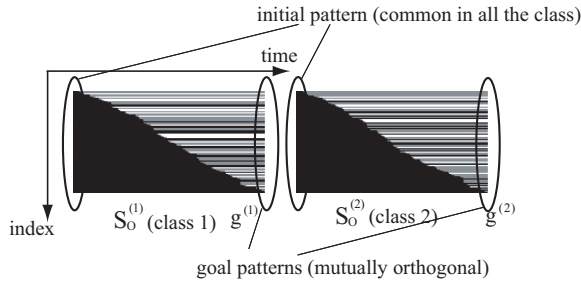


図4 判別空間における時系列パターンの例（2クラスの
場合）．白，黒はそれぞれ $-1, 1$ を表す

Fig. 4 Examples of the spatio-temporal patterns in the discrimination space (2 classes). White and black pixels denote 1 and -1 , respectively.

ディングは，森田 [5] と同様に次の直接的な方法をとる．

(1) 初期パターン o をすべてのクラスで共通とする．また，各クラスのゴールパターン $g^{(1)}, \dots, g^{(C)}$ を互いに直交するように選ぶ（すなわち $g^{(c)T} g^{(c')} = 0 (c \neq c')$ ）．これらは初期パターン o とも直交する．各パターンベクトルのノルムは等しいとする．ただし，これらのパターンは全て 2 値ベクトルとし，実際には乱数を用いて発生させることにより，ほぼ直交するパターンを作成する．

(2) 各クラスについて o と $g^{(c)}$ を単調に補間し， $S_o^{(c)}$ を得る．ただしここでの補間とは，一方のベクトルの要素を端から順に徐々に他方のベクトルと入れ替えていく操作によって行う．これは，線形補間を用いると，線形関数で状態の予測を行うことが困難となるためである（実験で用いた $S_o^{(c)}$ の例を図4に示す）．以上のようなコーディングにより，各クラスの軌道は時系列パターンの開始付近では近く，その後はクラス同士を次第に引き離すことができる．

単純化されたダイナミクス:

以上のように状態空間をデザインした結果，観測行列は $G = [I_{N_I}, O]$ となる．このとき，式 (2) は

$$\begin{cases} \hat{u}_{tI} = (1 - \kappa_t) [\mathcal{F}(u_{t-1}, r_{t-1})]_I + \kappa_t z_t \\ \hat{u}_{tO} = [\mathcal{F}(u_{t-1}, r_{t-1})]_O \end{cases} \quad (7)$$

のようになる．ただし，添字 I と O は N 次元ベクトルの 1 から N_I 番目の要素， $N_I + 1$ から $N (= N_I + N_O)$ 番目の要素を取り出したものをそれぞれ意味し， $\hat{u} = [\hat{u}_I^T, \hat{u}_O^T]^T$ が成り立つ．

3.4 バッチ学習アルゴリズム

学習の目的は，各クラスの条件付き予測関数 $\mathcal{F}^{(c)}$

のパラメータを決定することである．今は，それぞれの予測関数が線形関数であることを仮定しているため (式 (4))，各クラスの遷移行列 $W^{(c)} (c = 1, \dots, C)$ を求めればよい．ここでは，最小二乗誤差基準に従い一般化逆行列を用いて求めることにする．

はじめに，各クラスについて 1 つずつ代表的なサンプルパターン $S^{(c)} = [S_I^{(c)T}, S_O^{(c)T}]^T$ が与えられた場合を考える．このとき， $W^{(c)}$ は $s_t^{(c)}$ をかけることによって $s_{t+1}^{(c)}$ を出力するように決定したい．すなわち， $W^{(c)}$ は

$$S'^{(c)} = W^{(c)} S^{(c)} \quad (8)$$

を満たす．ここで， $S'^{(c)} = [s_2^{(c)}, \dots, s_L^{(c)}, s_L^{(c)}]$ である．なお， L 番目のパターン $s_L^{(c)}$ に対してはそれ自身が出力されるようにしている．

式 (8) を解くことにより，クラス c に対する遷移行列は

$$W^{(c)} = S'^{(c)} S^{(c)+} \quad (9)$$

となる．ただし， $S^{(c)+}$ は $S^{(c)}$ の一般化逆行列である．

次に，複数のサンプルを用いて学習する場合を考える．簡便な方法としては，一旦全てのサンプルパターンの長さを線形に正規化し，各時刻毎に平均ベクトルを求めることによって，各クラスにおける代表的なパターン系列を作成することが考えられる．これに上で述べた単一サンプルによる学習法を適用することにより，多サンプル学習を実現できる．しかし，一般にサンプルパターン間は時間方向に非線形な関係があり，このように線形に正規化することは適切でない場合がある．この解決方法については 6. でふれる．

3.5 クラスの決定過程

一般に時系列パターンの認識において，いつ，何を評価基準としてクラスを決定するかは，パターン区間の時間的なセグメンテーションと密接に関係するため，それ自身が研究対象となりうる．例えば，常に評価値を計算し，極値をとった段階でパターン区間とクラスを決定する方法 [3] や，パターン区間がセグメンテーションされていることを仮定して，パターン入力完了後の評価値（もしくはその累積）からクラスを決定する方法 [1] などがある．本論文では，認識前にパターン区間が $[t_{\text{start}}, t_{\text{end}}]$ に切り出されているものと仮定し，次の 2 つの評価値を比較することにする．

(A) 類似度:

CSM の状態 u_t は，外部入力に応じて学習した軌道の 1 つに沿って遷移する．よって，状態空間のうち

3.3 で付加した空間，すなわち判別空間における状態 $u_{O,t}$ と，各クラスのゴールパターンとの類似度を，クラスを決定する際の評価値として用いることができる。

ここで類似度としては正規化相関

$$e_t^{(c)} = \frac{g^{(c)}}{\|g^{(c)}\|} \cdot \frac{u_{O,t}}{\|u_{O,t}\|} \quad (10)$$

を用いる。類似度 $e_t^{(c)}$ は -1 から 1 の値をとりうるが，初期状態はゴールパターンと直交するように選ぶため，正解クラスに対する類似度は 0 から次第に 1 に近づいていく。また，各ゴールパターンは互いに直交するように選んでいるため，2 つ以上のクラスに対する類似度が同時に 1 になることはない。このとき，クラスは $\arg \max_{(c,t)} e_t^{(c)}$ ($t = t_{\text{start}}, \dots, t_{\text{end}}$)，によって決定される。

(B) 累積予測誤差:

予測誤差は式 (6) によって計算される。これを単純に累積をとることにより，累積予測誤差

$$D_t^{(c)} = \sum_{t'=t_{\text{start}}}^t d_{t'}^{(c)} \quad (11)$$

を評価値として用いる。このとき，クラスは $\arg \min_c D_{t_{\text{end}}}^{(c)}$ によって決定される。

3.6 予備実験: 単一 CSM の性能評価

ここで，人工的に作成した時系列パターンを用いてシミュレーションを行い，単一 CSM の性能評価を行った。この予備実験は，CSM の詳細な評価を行うためではなく，CSM の持つ能力およびパラメータに対する特性を定性的に示すことを目的とする。

はじめに，評価に用いるパターンを作成するために， ± 1 を値とする 150 次元の 2 値ベクトルを互いに直交するようにランダムに選び， A, B, C, D, E, F とし，このうち 3 パターンを単調に補間することにより，系列長 90 の時系列パターン ABC, ABD, DAC, DBC を作成した。すなわち，入力空間の次元 $N_I = 150$ である。なお，補間の仕方としては判別空間における時系列パターンと同様の方法を用いた (3.3)。次に，判別空間の次元 N_O は 100 とし，付加的な時系列パターンを作成した。判別空間の次元数がこのように高次元である理由は， 150 次元の入力パターンに対して全順序をつけ，かつ異なったクラスのパターン間では距離が離れるように，3.3 で述べた 2 値パターンの符合理化を行う必要があるからである。

学習時には，作成したパターンを用いて 3.4 に基

づいて遷移行列の計算を行った。なお，式 (9) における一般化逆行列の計算には特異値分解を用いた。

次に，認識時のパラメータの決定方法であるが，予測および外部入力の影響のバランスは，式 (1) から分かるように，入力強度 κ と N_I, N_O によって決まる。仮に，全てのベクトルの要素値が正規化されており，さらにいずれの要素にも同程度の誤差が生じるとするならば，以下の関係を満たすように入力強度を決定することによって，予測と外部入力の影響を同程度にすることができる。

$$(1 - \kappa)(N_I + N_O) = \kappa N_I \quad (12)$$

本実験の条件では $\kappa = (N_I + N_O)/(2N_I + N_O)$ はおよそ 0.6 となるが，予測の影響を外部入力に対してやや強め， $\kappa = 0.5$ とした。ところで，いずれのクラスに対してもほぼ同じ確信度を持つ場合は，CSM の状態は複数のアトラクタの中間に引き込まれ，いずれの類似度も低下して誤認識が生じうる。式 (5) におけるパラメータ σ を小さくするほど，予測誤差の高いクラスと 0 に近いクラスで確信度の差が大きくなる。そこで， 1 位と 2 位以下の確信度の差を大きくするために， σ に関しては予測誤差のノルムに対して比較的小さな値を経験的に選んでいる。実験では $\sigma = 1.0$ を用いた。

なお，認識開始時における状態ベクトルの初期値は，いずれの実験も $u_0 = [0^T, \sigma^T]^T$ とした。

入力の時間的変動に対する頑健性:

認識時には，学習系列の一つである系列 DBC を時間軸方向に非線型に伸縮し，これを入力しながら CSM を動作させた。この時の類似度の変化を図 5 に示す。グラフの横軸は，入力開始時からの CSM の状態更新回数を，縦軸は類似度を表す。入力したパターンのクラスに対応する類似度が最終的に最も高くなり，正しく認識されたことが確認できる。このように，入力系列に時間的伸縮が加わった場合でも，正しいクラスを識別することが可能である理由は，CSM が，式 (1) のエネルギー関数の第二項により，外部入力に追従してボトムアップ的に状態を遷移させようとするためである。

入力パターン系列の伸長に対しては，系列長を 10 倍以上とした場合でも識別可能であった。一方，入力パターン系列の短縮に対しては，系列長を $1/3$ よりも短くした場合に，CSM の状態遷移が追従できず誤認識が起こった。このように，入力に追従できなくなる変動の最大値は，入力強度 κ や入力・判別空間の次元

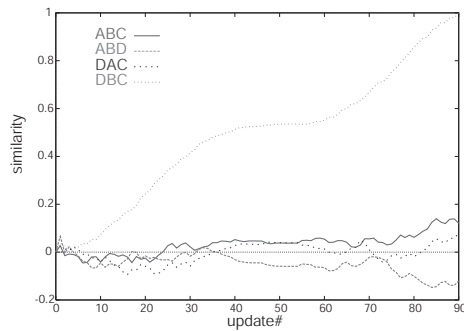


図 5 非線形な時間伸縮を加えた入力に対する類似度の時間変化 (横軸は CSM の遷移回数を表す)

Fig. 5 Robustness against non-linear temporal fluctuation.

に依存する．すなわち， κ が $(N_I + N_O)/(2N_I + N_O)$ に対して小さくなるほど，CSM の状態は学習されたダイナミクスにしたがって自律的に遷移するようになり，対応可能な時間伸縮の範囲も小さくなると考えられる． $\kappa = 0.6$ とした場合には，系列長を 1/5 に短縮したパターンの識別も可能であった．

ノイズに対する頑健性:

認識時には，学習系列の一つである系列 *DBC* に対して，平均で 20% の要素の符合を反転させたパターンを作成し，これを入力しながら CSM を動作させた．この時の類似度の変化を図 6 に示す．

グラフより，入力したパターンのクラスに対応する類似度が最終的に最も高くなり，正しく認識されたことが確認できる．このように，学習パターンと認識時の入力パターンは，完全に一致していなくても，識別可能である．これは，式 (1) のエネルギー関数の第一項により，予測を用いてトップダウン的に状態を遷移させようとするためであり，入力が (学習パターンに対して) 不完全である場合には，補間しながら状態遷移を続けることが可能である．

学習外パターンの棄却能力:

学習モデルに無いパターン系列 *DEF* が入力されたときの，CSM の類似度の変化を図 7 に示す．*DEF* は，前半部は *DAC*, *DBC* との相関が高く，系列の後半になるほどいずれの学習パターンとの相関も小さくなるような変化をする時系列パターンであり，パターン *D, E, F* を学習パターン同様に補間して作成した．

グラフから分かるように，学習モデル外のパターンが入力されたとき，CSM の類似度はいずれのクラスに対しても低いままである．入力パターンが学習され

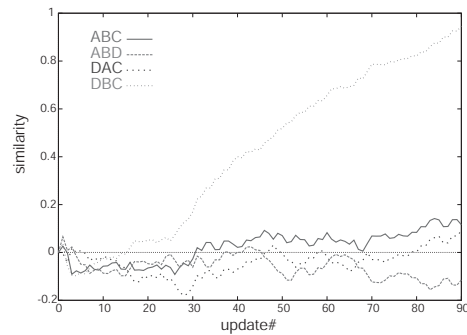


図 6 ノイズを加えた入力に対する類似度の時間変化

Fig. 6 Robustness against noise.

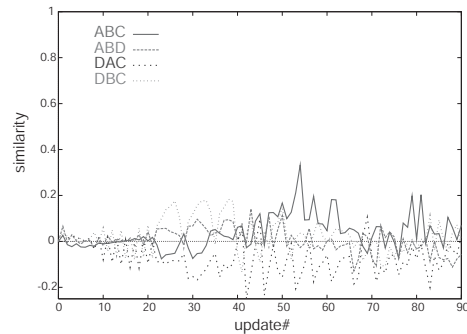


図 7 学習モデルに無い入力パターンに対する類似度の時間変化

Fig. 7 Rejection of the input pattern.

たパターンのいずれとも異なる変化をする場合には，式 (6) の予測誤差がいずれも同程度の値をとり，いずれかのクラスの確信度 (式 (5)) が続けて高くなることは無い．このため，CSM の状態は各クラスのアトラクタのちょうど中間当りに引き込まれ，いずれの最終状態からも遠ざかることになる．したがって，学習外パターンを棄却する必要がある場合には，類似度に閾値を設け，最大類似度がこの閾値を越えないには，入力パターンをいずれのクラスにも属しないと判定することができる．

4. 複数 CSM 間の相互作用

4.1 状態の共起性の学習

離散状態を扱う際には，状態の共起性は行列としてあらわすことができる．しかし，今は状態ベクトルを考えているために，単に行列であらわすことは不可能である．そこで，2 つの CSM における状態ベクトルの共起性を，多次元ヒストグラムによって求めることにする．たとえば，クラス c の時系列パターンを入

力した際の、状態 $u_{m,t}$ と $u_{l,t}$ の出力空間における共起性は、次の手順によって求められるヒストグラム $\mathcal{H}(u_{O,m,t}, u_{O,l,t}|c)$ によって表される。

(1) あらかじめ個別に学習を済ませた各 CSM に対して、同時に観測されたクラス c の時系列パターンを並列に入力し、状態を遷移させる。ただし、このときの CSM の確信度は、学習するクラスを 1 とし、残りを全て 0 とする。

(2) 時刻 t に同時に起こった状態の組 $u_{O,m,t}, u_{O,l,t}$ をヒストグラム $\mathcal{H}(u_{O,m,t}, u_{O,l,t}|c)$ に投票する。

(3) 全てのクラスおよびサンプルパターンに対して、1, 2 を繰り返す。

実際には *kd-tree* [12] を用いて多次元ヒストグラムを実現した。

4.2 他モジュールの状態の推定

いったん各 CSM が各自の入力と予測によって仮状態 $\hat{u}_{m,t}$ を計算すると (式 (2)), 各 CSM はこれに基づいて他の CSM の状態を推定し、どの状態へ更新させるかを定める。この推定はあらかじめ学習された状態の共起性によって行われる。

CSM m の仮状態から CSM l の状態を推定する際に、 u_m の状態が条件付き確率密度関数 $p(u_{O,m,t}|u_{O,l,t} = \hat{u}_{O,m,t}, c) = N(\bar{u}_{O,ml,t}^{(c)}, \Sigma_{O,ml,t}^{(c)})$ によって与えられると仮定する。ただし、 $\bar{u}_{O,ml,t}^{(c)}$ は平均ベクトルであり、 $\Sigma_{ml,t}^{(c)}$ は共分散行列である。この分布は共起性ヒストグラム $\mathcal{H}(u_{O,m,t}, u_{O,l,t}|c)$ より、 $\hat{u}_{O,l,t}$ の近傍にある q 個の点を探索することによって求まる (図 8)。

これは各クラスについて求まるため、確信度によって重み付けされた線形和を考えることによって、CSM l による CSM m の推定状態を

$$\bar{u}_{O,ml,t} = \sum_{c=1}^C r_{l,t}^{(c)} \bar{u}_{O,ml,t}^{(c)} \quad (13)$$

とする。

4.3 モジュールの影響力のコントロール

より信頼性の高い CSM の影響力を高めるために、各 CSM がどれだけ特定のクラスに対して確信を持っているかを表す尺度として「重要度」を導入する。1 位のクラスに対する確信度のみが高いときに 1 に近づき、上位 2 つのクラスが同程度の確信度を持つ場合は 0 に近づくものとし、時刻 t におけるモジュール m の重要度 $\omega_{m,t}$ を次式のように定義する。

$$\omega_{m,t} = (d_{m,t}^{2nd} - d_{m,t}^{min}) / d_{m,t}^{2nd} \quad (14)$$

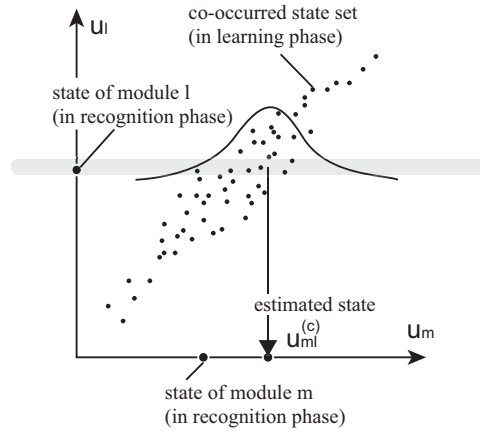


図 8 条件付き状態密度関数と推定された状態 (横軸と縦軸はそれぞれ CSM m, l の N_O 次元出力空間を表す)

Fig. 8 Conditional probability density function and estimated state. (vertical and horizontal axes denote N_O -dimensional state space of CSM m and l , respectively.)

ここで、 $d_{m,t}^{min}$ は時刻 t における各クラスに対する予測誤差 $d_{m,t}^{(c)}$ (式 (6)) のうち最小のもの、 $d_{m,t}^{2nd}$ は 2 番目に最小のものである。

4.4 推定された状態の統合

各 CSM は、4.2 で述べたように、学習された状態の共起性を用いて相手の状態を推定する。ある CSM m が、時刻 t において自分の状態 (仮状態) および他の CSM から推定された状態を統合をする際には、次の評価関数を最小化するように状態を更新する。

$$E_{2,m,t}(u_{m,t}) = (1 - \gamma)\omega_{m,t} \|u_{m,t} - \hat{u}_{m,t}\|^2 + \gamma \sum_{l=1, l \neq m}^M \omega_{l,t} \|u_{O,m,t} - \bar{u}_{O,ml,t}\|^2 \quad (15)$$

ここで、 γ ($0 \leq \gamma \leq 1$) は他の CSM に対する依存度である。右辺の第一項は自身の仮状態からの二乗誤差をあらわす。第二項は、他の CSM の推定値からの二乗誤差を、それぞれの CSM のもつ重要度によって重み付けして足し合わせたものである。信頼性の低い CSM に関しては、式 (14) によって重要度が低下するため、CSM m が状態を更新する際の影響も低くなる。なお、 $E_{1,t}$ (式 (1)) では、CSM 内における外部入力と予測との trade-off を表していたのに対し、 $E_{2,m,t}$ は、CSM 間における、状態の引っぱり合いを表した

ものになっている。

ここで, $\frac{\partial E_{2,m,t}}{\partial \mathbf{u}_{m,t}} = 0$ とすることにより, 時刻 t における CSM m の統合された状態 $\mathbf{u}_{m,t}$ は, 次の式によって得ることができる。

$$\begin{cases} \mathbf{u}_{I,m,t} = \hat{\mathbf{u}}_{I,m,t} \\ \mathbf{u}_{O,m,t} = \alpha_{m,t}^{-1} \left\{ (1-\gamma)\omega_{m,t}\hat{\mathbf{u}}_{O,m,t} + \gamma \sum_{l=1, l \neq m}^M \omega_{l,t}\hat{\mathbf{u}}_{O,m,t} \right\} \end{cases} \quad (16)$$

ただし, $\alpha_{m,t} = (1-\gamma)\omega_{m,t} + \gamma \sum_{l=1, l \neq m}^M \omega_{l,t}$ である。

5. 実験結果

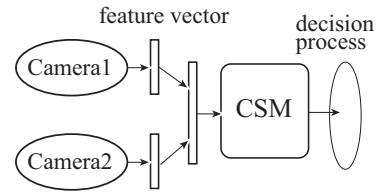
5.1 認識率の比較 (完全な入力)

評価実験として, 2 視点カメラによる手話 7 単語の識別を行った。はじめに, 被験者の正面と側面に設置されたカメラを用いて, 各フレームで同期させて撮影を行い, 曜日をあらかず手話単語を, それぞれ 21 サンプル (3 被験者 \times 7 サンプル) 得た。次に, 各フレーム (192 \times 192 画素) において, 背景差分によって人物のシルエット領域を取り出して 2 値画像とし, さらに 24 \times 24 画素にダウンサンプリングしてラスタースキャンすることにより, 576 次元 ($= N_I$) の実特徴ベクトルを得た。すなわち, 各要素の大きさは元画像の 8 \times 8 画素のブロックにおける人物領域の割合を意味し, 各要素の値は -1 から 1 に正規化を行った。なお, 判別空間の次元および入力強度はそれぞれ $N_O = 200, \kappa = 0.5$ とした。

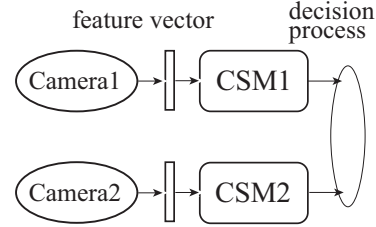
学習には, 21 サンプルのうち 12 サンプルを選び, 残りの 9 サンプルを用いてテストを行った。学習に用いるサンプルの選び方は 10 通りで変化させたため, 認識率は, 630 試行のうちで正しく認識した試行数の割合である。

比較のため, 次の 3 つの統合型アーキテクチャについて性能評価を行った (図 9)。

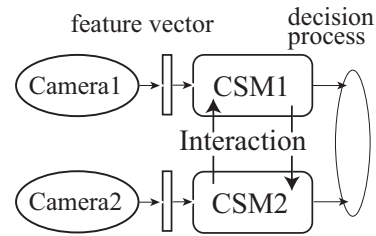
(a) **Feature-level Integration (FI)**: 単一の CSM を用いる (図 9(a)). CSM の入力は, 2 つのカメラから得られた特徴ベクトルを, 各フレームで連結したものである。このとき, 特徴ベクトルレベルで信頼性の低い入力を抑制するために, 次のような方法を用いた。まず, 一時刻前の状態から予測された入力ベクトルと, 実際に入力された特徴ベクトルの誤差ベクトルを求める。次に, 各カメラに対応する要素ごとでまとめて平均二乗誤差を求め, この誤差の低さを入力の信頼性の基準とする。平均二乗誤差が一定の閾値を越えた場合



(a) Feature-level Integration (FI)



(b) Decision-level Integration (DI)



(c) State-level Integration (SI)

図 9 比較に用いる 3 つのアーキテクチャ
Fig. 9 Three architectures for comparison.

(実験では 0.5 と設定) は, そのカメラに対応する特徴ベクトルの信頼性が低いものとし, 値を全て予測値に置き換える。以上の操作を全ての時刻において行う。

(b) **Decision-level Integration (DI)**: 2 つの独立な CSM を用いる (図 9(b))。各 CSM の入力, 各カメラから得られた特徴ベクトルである。それぞれの結果は, クラスを決定する段階で統合する。すなわち, 類似度の平均 $e_t^{(c)} = \frac{1}{M} \sum_{m=1}^M e_{m,t}^{(c)}$, もしくは累積予測誤差の平均 $D_t^{(c)} = \frac{1}{M} \sum_{m=1}^M D_{m,t}^{(c)}$ を用いて, 3.5 にしたがってクラスを決定する。

(c) **State-level Integration (SI)**: 前章で述べた ICSM を用いる (図 9(c))。これは上の DI に対して, CSM の状態空間における動的な相互作用を加えたものであり, 他の CSM に対する依存度 γ は $0 < \gamma < 1$ である。 $\gamma = 0$ のとき, 2 つの CSM は独立に動作し, DI に一致する。

これらの 3 つのアーキテクチャを用いた場合の認識

率を表 1 に示す。ただし、表の CSM1,2 は、それぞれ正面カメラ、側面カメラに対応する単体の CSM における認識率である。

表 1 より分かるように、CSM2 (側面カメラ) の認識率は、CSM1 (正面カメラ) よりもかなり低かった。これは、手話は元々正面を向き合って対話するためのものであり、側面から見ただけでは、どの手話単語も非常に類似した映像になるためであると考えられる。上で述べた 3 つの統合方法のうち、FI と DI の認識率は CSM1 と CSM2 の間の値になっているが、特にクラス決定過程として、3.5 で述べた (A) 類似度を用いた場合には、SI が最も良い識別結果を示した。これは、いずれのデータも同じ重みで処理する FI と DI では、側面からの入力データに引っ張られるが、一方 ICSM は、信頼性の低い CSM の重要度を低下させることによって、認識により重要なデータにフォーカスできるためであると言える。表 1 の SI における括弧内の数字は、各モジュールの重要度を動的に変化させず、全て一定値 $\frac{1}{M}$ に保持した場合の認識率であるが、重要度を動的に制御した場合に比べ大きく低下しており、このことから、動的な統合が有効であることが分かる。ただし、単体の CSM1 に対する SI の認識率の上昇は、2 ポイントにとどまっている。これは、多くの入力サンプルパターンに対して、全時刻を通じて CSM2 の信頼性が CSM1 に対して低く、CSM2 が CSM1 に与える影響が少なかったためであると考えられる。

なお、この実験では他の CSM に対する依存度 γ (式 (15)) を経験的に求めている。図 10 に γ を 0.0 から 0.3 まで 0.05 刻みで変えたときの認識率を示す。ただし、識別結果は (A) 最大類似度に基づいて決定した。このグラフから分かるように、依存度 γ には適切な値が存在し、表 1 の評価実験では、もっとも高い認識率を示した $\gamma = 0.1$ を用いた。

ところで、クラス決定過程として (B) 累積予測誤差を用いた場合は、(A) に比べて DI 以外で認識率が低くなっている。これは、(B) が入力空間での誤差を用いていることから、ノイズの影響を受けやすいことに起因すると考えられる。

5.2 遮蔽による不完全な入力

ICSM における重要度の変化を見るために、認識対象系列の一つに対し、前半 (フレーム 11 から中間まで) には正面の映像に遮蔽を人工的に加え、後半 (中間から終了の 10 フレーム前まで) には側面の映像に同様の遮蔽を加えた。ここでの遮蔽とは、特徴ベクト

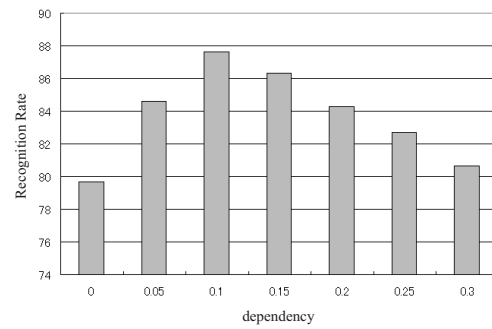


図 10 依存度 γ に対する認識率の変化の様子

Fig. 10 Recognition rate against dependency γ .

表 1 完全な入力を加えた場合の認識率の比較 (%) (括弧内の数字は、重要度を一定にした場合の ICSM の認識率)。 (A), (B) は 3.5 で述べた 2 つの決定過程を意味する。

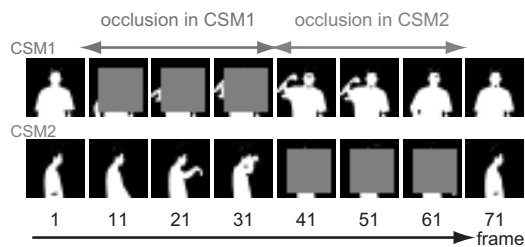
Table 1 Recognition rates for complete inputs(%). (The numbers in the parenthesis denote the recognition rates of the ICSM without dynamic control of importance.) (A) and (B) denote two decision processes described in sec.3.5.

Decision Process	CSM1	CSM2	FI	DI	SI
(A) Max.Sim.	86	53	79	80	88 (76)
(B) Min.Err.	85	51	77	83	83 (81)

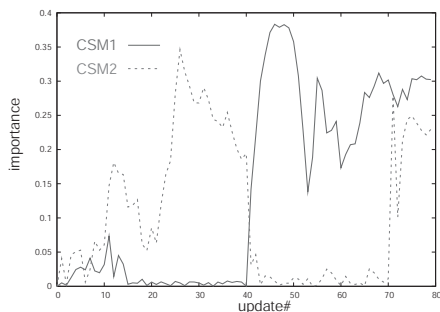
ルのうち 18×18 の矩形領域に相当する要素の値を 0 とすることを意味する。各 CSM の重要度が認識時に変化する様子を、図 11 のグラフ (横軸が時刻、縦軸が重要度) に示す。グラフから、フレーム 11 から 39 においては、遮蔽付きの入力を加えた CSM1 の重要度が遮蔽のない CSM2 に比べて低くなっていることが分かる。フレーム 41 から 70 では両者の大小関係が入れ替わっており、これは遮蔽の加え方の変化に対応している。

入力に遮蔽が起こると、CSM は特定のクラスとして予測することが困難となり、1 位と 2 位の予測誤差および確信度の差は小さくなる (式 (6))。そのため、遮蔽の起こった CSM の重要度は式 (14) にしたがって低下し、遮蔽のない CSM の状態更新に与える影響は小さくなる。このように、ICSM では信頼性の低い CSM を動的に抑制し、識別に重要な CSM の影響力を高めることができる。

表 2 は、5.1 で用いた全ての入力サンプルデータに対して同様の遮蔽を加えた際の認識率である。認識時のパラメータおよび認識率の計算に用いた試行数は前節と同様とした。表 2 より、決定過程として最大類似



(a) Input Sequences



(b) Importance of consisting CSMs

図 11 入力に遮蔽を加えたときの重要度の変化

Fig. 11 Response of importance against temporally occluded input data.

表 2 不完全な入力を加えた場合の認識率の比較 (%)

Table 2 Recognition rates for incomplete inputs(%).

Decision Process	CSM1	CSM2	FI	DI	SI
(A) Max.Sim.	73	26	57	61	77 (67)
(B) Min.Err.	73	36	63	69	73 (68)

度 (A) を用いた際の SI の認識率が最も高く、さらに、完全なデータを加えた場合 (表 1) と比べても、認識率の低下が最も小さいことが分かる。以上より、観測データが不完全になった際も、ICSM は他の統合法に比べて安定な認識を行うことが可能であるといえる。

6. む す び

異なる特徴量の時系列情報を動的に統合し、認識に必要な入力へ動的にフォーカシングを切り替える機能をもつ認識システムとして ICSM を提案した。これは、複数の状態遷移モジュール (CSM) 間の相互作用によって実現され、連続状態空間を用いているために、CSM 内の状態遷移および CSM 間の相互作用は、それぞれエネルギー関数の最小化によって定義される。CSM 内の状態遷移は、予測と外部入力の trade-off であり、ノイズなどによる不完全な入力を補完すると同

時に、入力の時間的な変動に追従することを可能にする。CSM 間の相互作用は、は自モジュールの状態と他モジュールによって推定された状態との trade-off であり、モジュール間で情報を補完し合うと同時に、識別により重要な入力へのフォーカシングを実現する。

さらに、今後の課題としては次のようなものが考えられる。

(a) 状態空間の構成法：現在は入力空間と判別空間の直積をとって状態空間としているため、状態ベクトルの次元数が非常に大きくなり、計算量や誤識率が上昇しうる (Hughes phenomena)。このため、圧縮された状態空間を用いることを検討する必要がある。すなわち、識別と予測を行うという基準にしたがって、サンプル時系列パターンから、状態空間を自動的に構成することが望ましい。予測に適した空間の構成法に関しては、Rao [7] によって提案されているが、これは動作の状態を短期的にしか保持しておらず、反復を含んだ動作の表現やクラスの識別には適用できない。現在、識別と予測の両方に適した状態空間を構成するための学習アルゴリズムを検討中であり、これに関しては稿を改めて報告する。

(b) 動画認識における空間的な変動の吸収：今回提案したモデルでは空間的な変動に対処できない。この問題に対する解決方法としては、空間的な変動に普遍な特徴量を用いることも可能であるが、もともと CSM が状態空間から高次元な入力空間へフィードバックを持っていることから、次のような機構が考えられる。各 CSM が、対象の局所的な部分特徴 (肩や頭、手などの部分、もしくは単に格子状に分割された部分画像など) をそれぞれ入力・学習する。次に認識時には、入力空間にフィードバックされた予測値をテンプレートとして各 CSM で局所部分の追跡を行い、これと同時に識別を行う。このとき、対象の空間的な位置やスケール変化を吸収すると共に、識別に重要であるという基準にしたがって、動的な注視を実現できる。

(c) オンライン学習の導入：3.4 で述べたバッチ学習では、各クラスのサンプルパターンを時間方向に線形に正規化して学習しているが、同一クラスのサンプルパターン間において、時間方向に非線形な変動がある場合、これは適切でない。そこで、CSM の時間変動を吸収する能力を活かし、オンライン学習を導入する必要がある。これは、一旦代表的なパターンを学習した後に他のサンプルパターンを入力し、この時の状態ベクトルの履歴から式 (9) の遷移行列を逐次計算す

ることにより実現できると考えられる。

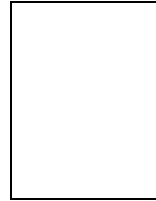
謝辞 本研究は科学研究費補助金 特定領域研究 133224051 の補助を受けて行った。

文 献

- [1] J. Yamato, J. Ohya, and K. Ishii. Recognizing human action in time-sequential images using hidden markov model. Proc. CVPR, pp. 379–385, 1992.
- [2] M. Brand and N. Oliver. Coupled hidden markov models for complex action recognition. MIT Media Lab Vision and Modeling TR407, 1997.
- [3] 高橋勝彦, 関進, 小島浩, 岡隆一. ジェスチャー動画のスポットニング認識. 信学論, Vol. J77-D-II, No. 8, pp. 1552–1561, 8 1994.
- [4] J.L. Elman. Finding structure in time. Cognitive Science, Vol. 14, pp. 179–211, 1990.
- [5] 森田昌彦, 村上聡. 非単調神経回路網による時系列パターンの認識. 信学論, Vol. J81-D-II, No. 7, pp. 1679–1688, 7 1998.
- [6] Masahiko Morita. Memory and learning of sequential patterns by nonmonotone neural networks. Neural Networks, Vol. 9, No. 8, pp. 1477–1489, 8 1996.
- [7] R. Rao. Dynamic appearance-based recognition. In Proc. CVPR, pp. 540–546, 1997.
- [8] H. Segawa, H. Shioya, N. Hiraki, and T. Totsuka. Constraint-conscious smoothing framework for the recovery of 3d articulated motion from image sequences. Proc. FG, pp. 476–481, 9 2000.
- [9] T. Wada and T. Matsuyama. Multiobject behavior recognition by event driven selective attention method. IEEE Trans. on PAMI, Vol. 22, No. 8, pp. 873–887, 2000.
- [10] B.D.O. Anderson and J.B. Moor. Optimal Filtering. Prentice-Hall, 1979.
- [11] C.M. Bishop. Neural Networks for Pattern Recognition. Oxford Univ, 1995.
- [12] J.L. Bentley. Multidimensional binary search trees used for associative searching. Commun. ACM, Vol. 18, No. 9, pp. 509–517, 9 1975.

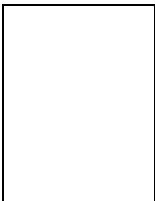
(平成 x 年 xx 月 xx 日受付)

松山 隆司 (正員)



1976 京大大学院修士課程了。京大助手、東北大助教授、岡山大教授を経て 1995 より京大大学院電子通信工学専攻教授。現在同大学院情報学研究科知能情報学専攻教授、学術情報メディアセンター長、評議員。工博。画像理解、人工知能、分散協調視覚の研究に従事。1980 情報処理学会創立 20 周年記念論文賞、1990 人工知能学会論文賞、1993 情報処理学会論文賞、1994 本会論文賞、1995 第 5 回国際コンピュータビジョン会議 Marr Prize、1996 国際パターン認識連合 Fellow、1999 本会論文賞、2000 画像センシングシンポジウム優秀論文賞各受賞。

川嶋 宏彰 (正員)



1999 京大・工・電気電子卒。2001 同大大学院修士課程了。2002 同大学院博士後期課程中退し、同大学院情報学研究科知能情報学専攻助手。時系列パターン認識、リカレントニューラルネットに関する研究に従事。