

## 科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 6 月 3 日現在

機関番号：17104  
 研究種目：若手研究（B）  
 研究期間：2011～2012  
 課題番号：23700271  
 研究課題名（和文） 時空間パターンの非線形マニフォールドによる表現とジェスチャー認識への適用  
 研究課題名（英文） Representation of Spatio-temporal Patterns with Nonlinear Manifold and its Application to Gesture Recognition  
 研究代表者  
 堀尾 恵一（HORIO KEIICHI）  
 九州工業大学・大学院生命体工学研究科・准教授  
 研究者番号：70363413

研究成果の概要（和文）：時空間パターンは特徴空間内のベクトル集合（マニフォールド）で表現でき、グラフ距離に基づく自己組織化マップを用いてこのマニフォールドを近似し、さらに高階の自己組織化マップにより学習したマニフォールド間の類似関係を解析する。ジェスチャーの中には、特異な形状のマニフォールドを構成するものがあり、高階自己組織化マップのアルゴリズムの変更を行うことで、学習安定性の向上を実現した。

研究成果の概要（英文）：Spatio-temporal patterns is considered as a set of feature vectors in a feature space called as a manifold. In this study, the manifolds were adequately approximated using graph distance based self-organizing map (GSOM), furthermore, similarity between manifolds could be analyzed by using higher rank SOM. In the gesture recognition, some manifolds represent peculiar forms, and the learning of higher rank SOM often becomes unstable. The learning algorithm was modified to realize a stable learning.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	2,400,000	720,000	3,120,000

研究分野：ソフトコンピューティング

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：時空間パターン認識，マニフォールド，近似，高階自己組織化マップ

### 1. 研究開始当初の背景

音声や動画などの時空間パターンの認識に関する研究は、古くは動的計画法（DP）や隠れマルコフモデル（HMM）に始まり、ニューラルネットワークを利用したものなど数多く報告されている。これらのほとんどの手法は、各時刻のパターンを状態と捉え、その状態の移り変わりを予測・観測することで認識するもの、また、時間情報を空間情報へ変換し静的パターンとして取り扱うものであった。これを実問題（例えば背景や照明条件に関し制限のないジェスチャー認識など）へ適用するためには、パターンの空間的な変形（平行移動や回転，拡大縮小）と時間的な変形（時間方向への伸縮）の双方を個別

に考え、それぞれに対策を講じる必要があった。一方で、時空間パターンは、パターン（特徴）空間内のパターンの軌跡の集合、つまりパターン空間内の非線形マニフォールドと捉えることができる。これにより、時空間パターンの空間的および時間的な変形は、マニフォールドの平行移動や密度変化として表現することが可能であると考えられ、双方の変形に関しマニフォールドの変形のみを考慮すればよいことになる。

申請者は、これまでにグラフ距離に基づき非線形マニフォールドを学習する自己組織化マップ（SOM）を開発した。これは基本的に1つのマニフォールドの位相関係を低次元空間で可視化するものであった。一方で、

古川らは機能モジュールの類似関係を低次元空間で可視化するモジュラーネットワーク SOM (mnSOM) を提案している。モジュールとして SOM を採用した高階型 SOM である SOM2 は、下位の各 SOM がそれぞれ非線形マニフォールドを学習し、上位の SOM は類似度に応じて下位 SOM の類似性を可視化するものである。これらの研究を踏まえ、各時空間パターンをグラフ距離に基づく SOM で学習し、それをさらに高階の SOM で学習する時空間パターン認識システムが期待できる。

## 2. 研究の目的

本研究で取り扱う時空間パターンとして手首より先を使ったジェスチャーを考える。時空間パターンを取り扱う場合に重要なことは、パターンの空間的な変形と時間的な変形の双方を考えないとならないことである。以下、それぞれに対して問題点を紹介し、解決への方策を述べる。

まず考慮しないといけないことは、(1) 如何にパターン空間を定義すればよいかという問題である。これは、画像からどのような特徴を抽出すべきかという問題に帰着できる。ここで重要なことは、画像内での平行移動や回転、拡大縮小が、パターン空間内での平行移動として表現できるような特徴とすることである。例えば、 $M \times M$  画素の画像を  $M \times M$  次元の輝度ベクトルとして表現した場合、画像内での平行移動は、パターン空間内で非線形な移動として観測されるので、位置ずれのある同じジェスチャーは、非線形伸縮を含む全く形状の異なる非線形マニフォールドとなる。よって、本研究では、画像内での平行移動や伸縮がある場合でも、パターン空間内で線形変換として捉えられるような特徴を明らかにする。

次に考慮すべきことは、(2) 如何にマニフォールドを表現するかという問題である。ジェスチャーの途中で速度が変化した(時空間パターンの時間伸縮がある)場合、マニフォールド内でデータ密度が変化することと対応する。従来から研究されている状態遷移に基づく手法では、速度が大きく変化した場合、次時刻の状態の予測が困難になり認識が困難になることが多いが、ここでは、それをデータの粗密という比較的取り扱いやすい変化となるので、従来手法と比較して時間伸縮に頑健であると考えられる。本研究では、形状が同じであるが、領域毎に密度がことなるマニフォールドを同一のものと表現する方法を明らかにする。

次に、(3) 非線形マニフォールド間の距離系の設計が必要である。ここでは、マニフォールドの平行移動や回転、拡大縮小などの変形を吸収し、形状の違いのみを考慮した距

離が必要となる。申請者らが過去に取り扱ったマニフォールドをグラフで表現する方法がここで活用可能であると考えているが、線形変換を吸収するための工夫が必要である。平行移動や拡大縮小、回転などを制約として取り入れてグラフを構築し、それらの間の特徴量を距離とする新しい距離系を明らかにする。

上記3つの課題を解決し、頑健で実応用可能なジェスチャー認識器の開発を目的とする。

## 3. 研究の方法

### (1) 画像の前処理手法の検討

背景や照明条件にできる限り制約を用いないジェスチャー認識器のためには、画像の前処理が非常に重要となる。ここでの前処理として、背景差分による背景除去、画像内輝度を一定にするための照度変換、対象物体(ここでは手)を抜き出すための肌色抽出などが考えられる。将来的には、移動ロボットへの搭載を考えているので、背景差分は適切ではないと考えられる。また、特徴抽出との関連性も含め、システムの他手法との相性も十分考慮しなければならない。さらに、処理速度の向上も必要である。本研究では、パーティクルフィルタと呼ばれる手法を用いて画像内の手の領域を抽出、追跡することを行った。

### (2) 特徴抽出手法の検討

画像内での対象物体(ジェスチャーの場合は手首より先)の平行移動や回転、拡大縮小に影響を受けないような特徴として、対象物体の形状のみを表現する方法が必要である。その候補としては、輪郭やエッジ情報を利用する方法や、手のモデルを画像内の手と整合させる方法、対象物体の存在領域を点集合で表現する方法などが考えられる。それぞれの方法に対し、画像内での変形と特徴空間内で変形との関連性を調査する。ここでは、画像内での変形が特徴空間内での線形変換で表現できることが必要である。また、動画像を処理するために、高速処理が要求される。上記各手法の高速化を行い、1枚の画像を10msec以内での処理を実現する。必要に応じ、FPGAの利用なども考える。本研究では、FPGAでの実装は実現できなかったが、対象物体の特徴として、単純に画素値を用いることで十分であることを確認した。

### (3) マニフォールド表現方法の検討

時空間パターンの認識における問題点の1つとして、ジェスチャー中に速度の変化が生じることである。通常、均一のサンプリング周波数で測定されるパターンは、速度の遅い時間帯ではパターンがほとんど変化せず、速度の速い時間帯ではパターンの変化が著しい。状態遷移に基づく認識手法では大きな問題となる。パターンの軌跡をマニフォールド

で表現する本申請研究では、状態遷移に基づく手法と比較するとさほど大きな問題とはならないが、速度が変化することはデータ密度に粗密が生じるため、密度変化に影響を受けないマンフォールドの表現方法が要求される。マンフォールドを表現する方法として、離散的なグラフを利用する方法、各データにエネルギー密度を定義してその和を存在確率とする方法、カーネル法などを利用して関数近似する方法などが考えられる。密度に影響を受けない方法とすると、グラフを用いる方法が適当であると考えられるが、24年度以降に取り掛かる予定のマンフォールド間の距離も考慮に入れ、(1)と同様に処理速度(できれば1枚の画像を10msec以内での処理を実現する。必要に応じ、FPGAの採用なども考える)なども含め、適切な表現方法を総合的に検討した。

#### (4) マンフォールド間の距離系の検討

特徴空間内の非線形マンフォールドで表現された時空間パターンを識別・認識するためにマンフォールド間の距離を定義する。(1)で述べたように、特徴空間内の線形変換は無視されるべきであるので、マンフォールドの平行移動、回転、拡大縮小は無視され、形状のみで距離が決定されることが要求される。マンフォールドの表現方法として(2)でグラフを採用したと仮定すると、グラフのノードの対応付けやノードの次数、リンクの角度などに基づいてマンフォールドの距離を定義することが考えられる。また、(2)で存在確率を採用したとすると、Kullback-Leibler 情報量を用いることが有用と思われる。ここでは、必要に応じ、(3)マンフォールド表現方法を見直すことも考えないといけない。また、計算速度も高速化が必要である。10msec以内に処理を行うことができれば、1枚の画像に対し(1)～(4)の処理で30～40msec程度となり実時間処理が可能になる。ここでは、識別のため多くのテンプレートとの比較が必要となるので、並列処理を念頭においたFPGAでの実装が有効になると考えている。上記の事項を考慮し、適切な距離系の検討を行った。結果として、以後の処理も考慮してグラフで表現することが妥当であると判断した。

#### (5) 各要素技術間の関連性を見直し

上記(1)～(4)までを実現した後、要素技術間の関連性を考えて各要素技術の改良を行う。例えば、(4)マンフォールド間の距離で、マンフォールドの線形変換のみではなく非線形変換にも対応できるものが実現できれば、(1)特徴抽出手法において制約が緩くなるので、計算時間をより重視した手法に修正可能である。また、例えば、(2)においてデータの粗密に影響を受けないようなマンフォールド表現方法が実現できな

ければ、(1)においてサンプリング周波数を随時適切に変化される手法を考える必要がある。上記の例のように様々な関連性を考えながら、システム全体として、精度および計算速度の両面において最適化を検討した。

#### (6) ジェスチャー認識システムの構築および検証

上記(1)～(5)を踏まえ、ジェスチャー認識システムの構築を行った。背景や照明条件の変化の存在するシミュレーションを様々な状況下で行い、性能評価を行った。

## 4. 研究成果

本研究の目的は、ジェスチャーの認識であり、具体的には、動画像から手首より先の部分を抽出、追跡し、その形状の時間変化から動作の意味を認識することである。2年間の研究期間の1年目である平成23年度は、(1)対象の追跡、(2)対象を表現する特徴量の検討、(3)ジェスチャーをマンフォールドで表現する際の表現方法、について検討を行った。

### (1) 対象の追跡

パーティクルフィルタと呼ばれる手法を用いて対象の追跡を試みた。ジェスチャーの場合は、対象が高速かつ複雑に移動する場合も多く、通常のパーティクルフィルタでは、高精度で追跡することができない場合が多々あった。本研究では、対象の動きをモデル化してパーティクルに埋め込む手法を採用しているが、様々な動きを想定したモデルをパーティクルに埋め込み、異なるモデルが混合したパーティクルフィルタを提案した。シミュレーションにより、高速移動や急激な方向変化にも従来手法と比較して頑健に追跡できることを確認した。

### (2) 対象を表現する特徴量

対象である手首より先の画像に対して、認識に適した特徴量について検討した。従来から認識問題でよく用いられているSIFTやHOGと呼ばれる特徴量について検討したが、この問題の場合、SIFTでは対象の形状が表現されていないこと、HOGでは、対象の回転や並行移動に対応できないこと、があり、良好な認識結果を得ることができなかった。SIFTとHOGのそれぞれの特長を活用した新しい特徴量について検討を行った。後述のマンフォールドの表現方法と合わせて考慮すると、画像の輝度値を用いることが適当であると結論付けた。

### (3) マンフォールドの表現方法

マンフォールドの表現方法として、まず、データ点の集合体として表現する方法を検討したが、データ点に疎密があり、適切な表現はできなかった。混合ガウス分布を用いることで各マンフォールドを適切に表現することができた。しかしながら、マンフォールド

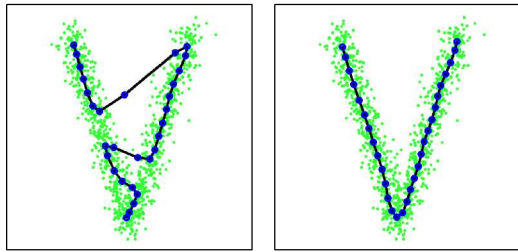


図 1：自己組織化マップによるマンフォールドの学習例. (a) 通常 (b) グラフ距離に基づく

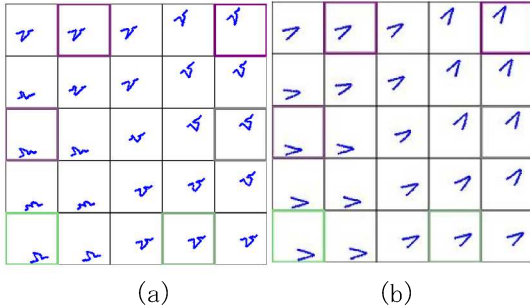


図 2：複数のマンフォールドの分類結果. (a) 高階 SOM, (b) グラフ距離に基づく高階 SOM

間の類似性を考慮しなければならないので、データの分布をグラフで表現し、グラフ上の距離に基づく自己組織化マップでマンフォールドを表現することとした。図 1 にデータ分布を通常の自己組織化マップとグラフ距離に基づく自己組織化マップで学習した際の例を示す。

2 年目の 24 年度は、非線形マンフォールドを学習するグラフ距離に基づく自己組織化マップの学習安定性について、その問題点の明確化および改善策の提案を行い、ジェスチャー認識への応用の可能性について検討した。詳細を以下に示す。

(4) マンフォールド間の距離系の検討  
マンフォールドの表現方法をも考慮して、高階型自己組織化マップを採用した。つまり、下位の自己組織化マップでは、データ分布を近似したグラフ上の距離に基づいてマンフォールドを近似し、上位の自己組織化マップでは、下位の学習結果を分類することができる。これにより、ジェスチャーが構成するマンフォールド間の類似性に基づいた分類が実現できた (図 2 参照)。

(5) 各要素技術間の関連性の見直し  
上記まで、ジェスチャーが特徴空間内にデータ分布としてマンフォールドを構成すること、また、高階の自己組織化マップを用いることで、マンフォールド間の類似性に基づいた分類が可能であることを示した。しかしながら、下位にグラフ距離に基づく自己組織化マップを利用していることに起因する学習

の不安定性が生じた。この要因を特定するために、下位および上位の自己組織化マップの学習パラメータを様々に変更し、学習の様子を観測した。その結果、下位のグラフ距離に基づく自己組織化マップの学習と上位の自己組織化マップの学習を交互に行う際、上位の結果を下位の学習に利用するコピーバックと呼ばれる操作が主要因であることが分かった。これは、下位の自己組織化マップにおいてグラフ距離を用いているため、十分な学習を行わないと学習結果に歪みが生じ、非線形マンフォールドを学習できないことである。よって、下位の自己組織化マップを適切に学習するために、下位と上位の学習量を適切に配分するパラメータ決定法を実現した。これにより任意の形状の非線形マンフォールドの学習が適切に行うことができ、ジェスチャー認識という実問題への適用可能性が十分であることを示唆した。

#### (6) ジェスチャー認識システムの構築および検証

ジェスチャー認識システムとして、対象領域である手の自動追跡および高階型自己組織化マップによる識別の一連の動作を行うソフトウェアを開発した。しかしながら、処理速度が十分ではなく、実時間での処理は実現できなかったが、録画した動画像に関しては、高い識別率を実現した。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 1 件)

- ① Hideaki Misawa, Keiichi Horio, Nobuo Morotomi, Kazumasa Fukuda, Hatsumi Taniguchi, “Extrapolation of Group Proximity from Member Relations Using Embedding and Distribution Mapping,” IEICE Trans. on Information and Systems, 査読有, Vol. E95-D, No. 3, pp. 804-811, 2012.

〔学会発表〕 (計 5 件)

- ① 藤田典宏, 堀尾恵一, “グラフ距離に基づく高階自己組織化マップの学習安定性の改善,” 第 14 回 自己組織化マップ研究会, 香川, 3 月 21 日, 2013.
- ② 吉田慎平, 萩原悟一, 今村律子, 石川秀大, 磯貝浩久, 堀尾恵一, “多段 SOM によるデータ集合間の関係性解析～チームスポーツ解析への応用～,” 第 14 回 自己組織化マップ研究会, 香川, 3 月 21 日, 2013.
- ③ Ying Li and Keiichi Horio, “Stressed State Classification Based on PPG Dynamics Using Neo Fuzzy Neuron Module

mnSOM,” Proc. of Int. Workshop on Smart Info-Media Systems in Asia (SISA2012), pp.185-188, Sep. 6-8, Bangkok, Thailand, 2012.

- ④ Ying Li and Keiichi Horio, “Stress State Estimation Based on Photoplethysmogram Dynamics,” Proc. of 2012 Int. Workshop on Nonlinear Circuits, Communication and Signal Processing (NCSP’ 12), pp.756-759, Mar. 4-6, Honolulu, U.S.A., 2012.
- ⑤ Hideaki Misawa, Keiichi Horio, Nobuo Morotomi, Kazumasa Fukuda, Hatsumi Taniguchi, “Relational Higher-Rank SOM for Bacterial Flora Analysis,” Proc. of 2011 Int. Workshop on Smart Info-Media Systems in Asia, pp.167-170, Oct. 31-Nov.2, Nagasaki, Japan, 2011.

## 6. 研究組織

### (1) 研究代表者

堀尾 恵一 (HORIO KEIICHI )

九州工業大学・大学院生命体工学研究科・准教授

研究者番号：70363413