

## [招待講演] 動画像を用いた手指の三次元形状の推定

島田 伸敬<sup>†</sup> 今井 章博<sup>††</sup> 浜田 康志<sup>†</sup> 白井 良明<sup>†</sup>

<sup>†</sup> 立命館大学 情報理工学部 〒 525-8577 滋賀県草津市野路東 1-1-1

<sup>††</sup> 大阪大学 大学院工学研究科 〒 565-0871 大阪府吹田市山田丘 2-1

E-mail: †shimada@ci.ritsumeai.ac.jp

あらまし 人間の手は多関節構造物として非常に複雑な物体であり、その形状や姿勢を画像列から非接触に計測することは、人体全体の形状姿勢計測とならんで長年コンピュータビジョン（画像解析）の分野におけるチャレンジングな課題となっている。本稿では、著者らのグループでこれまでに研究してきた画像列から三次元手指形状を推定する手法について述べる。3次元形状モデルを画像に照合する手法と、画像の見えを記録しておいて複雑背景下で領域抽出と照合を行う手法について、それぞれについて手法の概要と今後の課題について述べる。

キーワード 手指形状姿勢推定、画像列、非接触計測、関節物体

## 3-D Hand Posture Estimation Using Image Sequence

Nobutaka SHIMADA<sup>†</sup>, Akihiro IMAI<sup>††</sup>, Yasushi HAMADA<sup>†</sup>, and Yoshiaki SHIRAI<sup>†</sup>

<sup>†</sup> College of Information Science and Engineering, Ritsumeikan University

Nojihigashi 1-1-1, Kusatsu, Shiga, 525-8577 Japan

<sup>††</sup> School of Engineering, Osaka University Yamadaoka 2-1, Suita, 565-0871 Japan

E-mail: †shimada@ci.ritsumeai.ac.jp

**Abstract** Since the human hand has the highly articulated structure, it has been a challenging problem in the computer vision field to estimate its posture from an image sequence in a touchless way. This paper introduces the 3-D hand posture estimation methods using image sequence which the authors has been developing. There are two major categories: one is the appearance-based method and the other is the 3-D-shape-model-based method. The overview and problems to be solved are described for each method.

**Key words** 3-D hand posture estimation, image sequence, touchless sensing, articulated objects

### 1. ま え が き

ヒューマンインターフェイスの一つとして、ジェスチャ認識が注目されている。バーチャルリアリティにおいては、特別な手袋を用いずに手指の状態を入力できるように、TVカメラからの画像から手指の三次元形状と姿勢を非接触に認識することが望まれている。また、コンピュータの操作や手話による対話などにも直観的なジェスチャインターフェースが有効な場合がある。

コンピュータの進歩により、画像に基づくジェスチャ認識手法は10年ほど前から研究が盛んになり、簡単なジェスチャに限れば実時間で認識できるようになったが、道具や物体を手で操作する様子を識別したり、手話のような複雑な手の形と動きを推定することは、背景と手指領域の峻別や、手指とカメラの位置関係による見え方の変化、ジェスチャ自体に揺らぎがあることなどから、単純な見えの照合では非常に困難である。この問題

に対処するために、これまで大きく分けると3-D-model-basedと2-D-appearance-basedの二種類のアプローチが研究されてきた。

前者の手法は、画像から局所的な特徴を抽出して、その特徴に対して三次元形状モデルを当てている[2][13]。この方法は、最小二乗基準に基づいた高精度な姿勢の推定を試みているが、セルフオクルージョンなどによって見えが多様に変化する手の姿勢に対してはロバスト性に欠ける。また、画像の局所的な特徴ではなく多視点画像から作成されたボクセルに対する三次元モデルの当てはめにより姿勢を推定する方法が提案されている[3]。しかしセルフオクルージョンが起るとボクセルが正しく作成されず、姿勢が推定できない。

後者の方法では、記録されている対象物体の様々な姿勢の二次元の見えから入力の手指の見えと最も照合するものを選びだされる[4]。これは特徴の抽出を行わず、直接入力の見えと記録されているモデルの見えを比較するのでセルフオクルージョン

に対してロバストである。見えの画像を主成分分析 (PCA) で圧縮すれば、計算時間の短縮が可能である [5]。しかし、これらの方法は入力を少数のパターンに分類しており、三次元情報の抽出は行っていない。Black ら [6] はこのアプローチを二次元の位置と方向の推定まで拡張したが、三次元までは至っていない。見えと姿勢パラメータとの間の関係を学習しておけば高速に姿勢を推定することができるが [9] [12]、3 次元的な見えの変化について考慮されていないため、学習用画像サンプルが数多くいるという問題を抱えている。

著者らの研究グループでは、これまで 3 次元形状モデルと入力画像のシルエットマッチングによる関節物体の姿勢を計測する手法について研究してきた [14] [15] [10] [11] [17] [18]。3-D-model-based と 2-D-appearance-based の方法との橋渡しをする考え方として、起こりうる見えを三次元形状モデルから生成し、その見えと入力の見えとの照合を行う “Estimation by Synthesis” [7] [8] の考え方にに基づき、三次元モデルから生成した見えを効率よくデータベースに登録し、三次元的な見えの変化を考慮した照合を行うことによって、従来の単純なジェスチャだけではなく、多様に変化する手指の形状を三次元的に推定することができる。

また上記の手法は関節物体のシルエットがきれいに抽出できることを前提にしているため、実際のジェスチャ推定に適用するには手指領域抽出問題を解決することが必要となるが、複雑な背景下では手指の形の推定結果を手がかりにしないと雑音や解釈のあいまい性のために領域抽出の誤りが頻発する。著者らのグループでは、手話単語の認識を対象に、見えの可能な変化を遷移ネットワークに登録し、領域抽出と手形状識別を並行的に行いながら手形状変化の系列として手話単語を認識する研究を行っている [1] [19] [23]。

ここではこれら著者らの研究内容についてまとめて報告する。

## 2. 手領域輪郭に基づく実時間手指形状計測システム

3 次元構造モデルを利用して、あらかじめありうる手指の見え方とその時間変化を学習させておき、入力画像と照合することで手指形状を識別しつつ画像上の領域を追跡することが可能となる。

あらかじめモデルから生成した手形状シルエット輪郭をデータベースに登録する。姿勢推定時には、入力画像のシルエットから輪郭を抽出し、輪郭形状がマッチする候補をデータベースから検索し、見つかった候補を生成する際に用いた形状モデルの関節パラメータが入力画像の推定姿勢パラメータであるとする (図 1)。次の時刻における画像フレームの推定では、姿勢パラメータが大きく変動しないという仮定に基づいて近傍を探すことにより効率的な識別を行う。また、一時的な識別誤りに対応するため、上位いくつかの候補を残しながら並列探索を行なうビーム探索を用いる。

ここで、手指の関節の自由度は二十数自由度と高く、さまざまな仮定によって指の動きを限定したとしても 8 自由度程度はあるため、すべての見え方モデルをデータベースに登録することは現実的でない。したがって、各自由度ごとにある程度粗

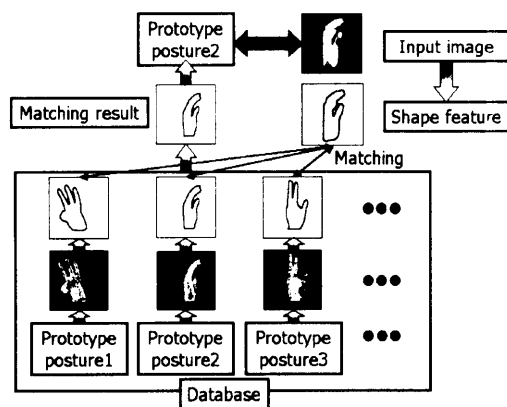


図 1 姿勢推定方法

Fig. 1 Overview of posture estimation

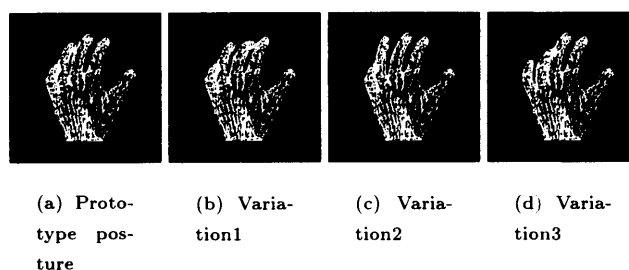


図 2 典型姿勢とバリエーション

Fig. 2 Prototype posture and its variations

くサンプリングしたスパースな見え方データベースとなる。すると、入力画像の照合時には必ずしも見え方モデルとよく一致する見え方が得られるとは限らず、登録したモデルの中間的な姿勢が入力されることが多くなる。これを従来法のまま照合すると照合の誤りが非常に多くなる問題点がある。

そこで本手法では、データベースに登録された見え方モデルの近傍で 3 次元姿勢パラメータ (関節角度) を変動させ (図 2)、見え方のシルエットがどのように変動するかをあらかじめ学習し (図 3)、許容される変動とそうでない変動にわけてから照合を行う方法をとる。これにより 3 次元構造モデルから許される見え方変動であれば多少大きくても許容して照合し、逆に許されない見え方変動は小さくても棄却することが可能となり、照合の精度が向上した (図 4)。

隠蔽が発生する候補では、微小な関節角度変化が大きな見え方変化をもたらす、見え方変動が単純な正規分布モデルでは記述できないことがわかった。そこで、輪郭変形の学習時にクラスタリングをおこなって複数の正規分布の混合モデルで記述する方法を採用し、また輪郭上に等間隔にとったサンプル点が手形状の個人差によってずれることを考慮して学習する方法と組み合わせてより頑健な照合が可能となった。

以上のアルゴリズムを、ハイエンド PC に大容量メインメモリを搭載し、相互に Gbit 高速 LAN で接続した PC クラスタシステム上に実装した。これにより約 16 万通りの手指姿勢を 10fps 程度で推定するシステムを構築することができた [17] [18]。

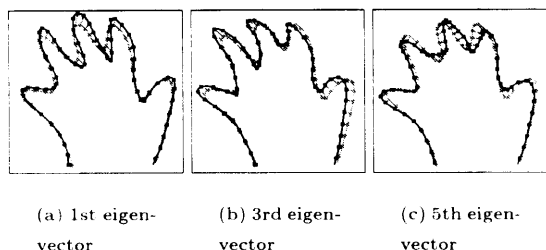


図3 学習された許容変動の例 (固有ベクトル)  
Fig.3 Obtained possible variations (eigenvectors)

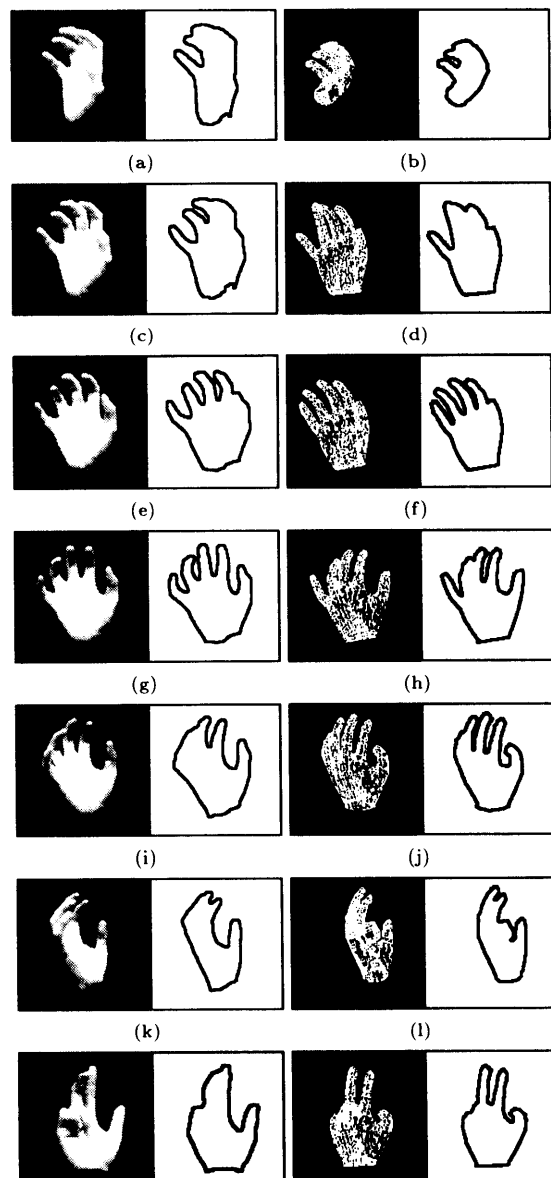


図4 許容変動を考慮した照合結果  
Fig.4 Matching results by considering possible variations

### 3. 形状モデルの不確かさを補償するオンライン式モデル詳細化手法

本手法で用いる手指の3次元形状モデルは実際の人体をレー

ザレンジファインダで取り込んだものをもとにしているが、各パーツの形状や関節の位置が固定のジェネリックなものであり、実際のユーザの手とは完全に一致しない。そのため、データベース中の手形状輪郭も実際の入力画像とはそもそも完全に一致しないので、姿勢（関節角度）の推定精度は一定以上には向上しない。

しかし、推定時には次々と画像が入力されていくので、推定中に各部の寸法が変化しないという仮定を用いれば、姿勢を推定しながらオンラインでモデル形状を詳細化してユーザにあわせていくことが可能となる。3次元形状を測定するには一般に2つ以上の方向から見た画像が必要とされるが、本研究ではこれを1つのカメラで行うことを検討した。物体の形状と姿勢がまったく変化しなければ1つのカメラでも物体を回転して見る方向を変えながら画像を入力すれば形状測定が可能であるが、関節物体は関節が動いたため、毎時刻変化した姿勢の画像が入力される。本質的にこの種の推定問題は不良設定であり、奥行きのないあいまい性のため解を求めることができない。しかし本研究の結果、この場合でも、

- (1) 関節の動きの早さに上限がある（動きの連続性）
- (2) 各部の寸法（長さ、太さなど）が変化しない（部分剛体性）
- (3) 関節の可動範囲、ならび各関節どうしの相関が既知
- (4) 各部の寸法の上限と下限、ならびに各部の寸法どうしの相関が既知

などの条件が仮定できる場合には、単眼視動画像からでも関節物体の形状を詳細化できることがわかった。以上4つの条件を形状と姿勢のパラメータ空間における不等式制約として記述し、制約と画像から得られる観測値（誤差を含む）の両方を満足するパラメータを探す。この計算は多次元空間では極めて大きな計算量を必要とするため、カルマンフィルタにおける平均値と共分散楕円の組 [15]、あるいは多次元楕円体と平行超平面の組 [16] によって近似的に記述し、時系列観測と上述の不等式制約条件を分布切取法 (図5) によって情報統合することによって、パラメータのとりうる範囲を狭く限定する。これによりオフラインではあるが現実的な計算量で、単眼動画像から手指の形状（長さや太さ）と姿勢（関節角度）を推定することができた。図6は合成画像による正解 (a)、通常のカルマンフィルタを用いた推定結果 (b)、本手法による不等式制約を用いた推定結果 (c) である。(b) では関節の角度の曲がる方向が奥行きのないあいまい性のために間違っているが、(c) では正しく推定できている。

手指実動画像に対する実験例を図7に示す。この例では、各関節の動きが30度/フレーム程度の速い動きの入力画像を用いた。図7(a)-(c)は入力シルエットとそれから抽出された指先の突起特徴である。はじめに大まかな形状モデルをあたえ、突起特徴を利用したシルエットマッチング [14] によりオクルージョンを解決し、大まかな姿勢推定を得る。それを元に各指節の対応する中心軸を突起特徴から切り出し、大まかな推定結果を初期値として、形状と姿勢を同時に推定する。推定結果が (d)-(f) である。図8に、最初に与えた大まかな形状モデルが本手法によって修正された結果を示す。

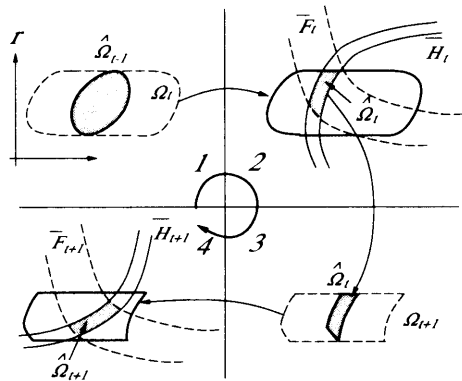
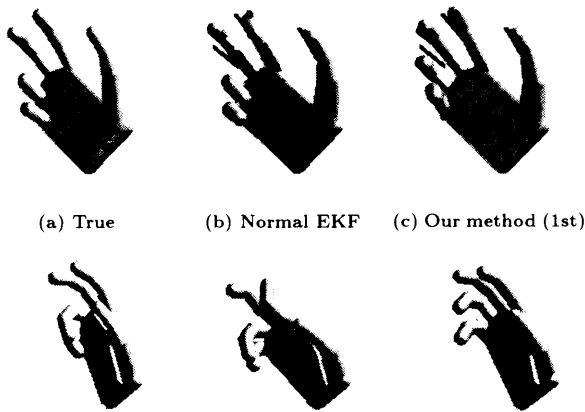


図5 パラメータ領域の逐次的な更新

Fig. 5 Incremental update of the parameter region



(d) Other view of (a) (e) Other view of (b) (f) Other view of (c)

図6 三次元の関節物体の推定

Fig. 6 Estimaion of 3-D articulated object

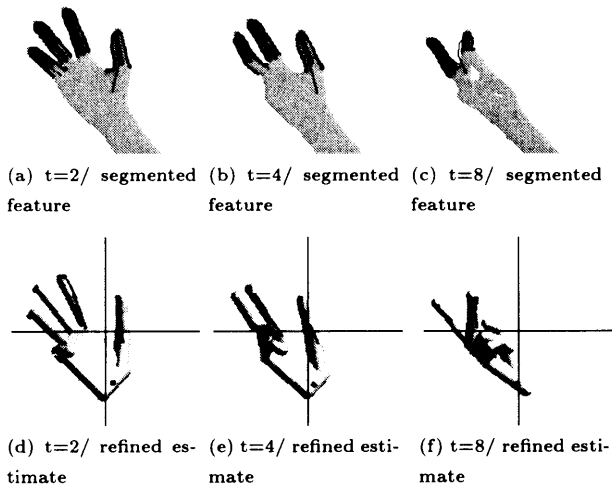
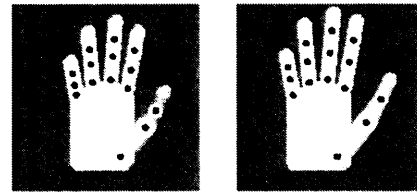


図7 実画像からの推定例

Fig. 7 Estimation results for real images

#### 4. 遷移ネットワークを用いた複雑背景下での手指領域追跡と手話単語の認識

手話単語の認識では、発生するジェスチャの数がかぎられているので、必要な単語をHMMモデルによって学習し、推定時の探索範囲を限定することでより高速に形状推定、ならびに単



(a) 初期形状 (b) 形状推定結果

図8 実画像からの形状推定結果

Fig. 8 The shape estimation result for real images

語の認識を行った[22].

しかし背景を単純にしたり、肌色をキーに領域分割をしてから追跡を行う方法では、机の上や本などの極めて複雑な背景下では切り出しに失敗して認識できない. そのような場合に対応するために、あらかじめ登録されたジェスチャモデルに基づいて手の形状を背景から切り出しつつ形状姿勢推定を行う手法を提案している.

手の形状はダイナミクスにしたがって連続的に変化するが、ダイナミクスパラメータはジェスチャごとに離散的に変化する、との仮定から、同時に複数のダイナミクスによる予測を考慮しつつフィルタリングによる予測と推定を行う Switching Linear Modelを導入した[20][21]. このモデルでは確率的に複数のジェスチャの可能性を考慮しつつ、連続的に変化する動的輪郭を追跡でき、複雑な背景下での手指ジェスチャを追跡、識別するのにこのモデルが有効であることを確認した. しかし数個のジェスチャを登録するのに多量の学習計算を必要とし、認識時にも実時間化が困難な計算量を必要とした. そこで複数の手話単語学習シーケンスから見えの遷移ネットワークを構成し、遷移ネットワークを可能な遷移をたどることによって効率的に照合を行う手法を提案した[23].

実際に演じられる手話の画像は以下のような特徴がある.

[高速性] 高速な手指の形状変化と移動を伴うため、画像が動きでぼける.

[複雑背景] 画像上で手指と似た色の背景と重なる.

つまり、従来のモデルフィティングに基づく追跡手法や、色、動きベクトルによる領域抽出法の適用が困難な状況といえる.

本手法では、あらかじめ手話単語のサンプルシーケンスを用意しておき、ありうる二次元的見え方形状とそのありうる変化を遷移ネットワークとして学習させ、これを用いて照合と追跡を行う.

手指の高速な移動と変形を伴うジェスチャ動画像に対して、二種類のモデルから構成される遷移ネットワーク(図9)をあらかじめ生成しておき、そこから見えのマッチング候補を選択する. 手指形状が重要な意味を持つ瞬間では手指の移動速度は遅い. 一方、手指の移動速度が速い瞬間では手指輪郭は不明瞭となるが、手指形状は重要でない. そこで、手指の移動速度が遅いときは形状と位置と速度の三つを特徴とするモデルを登録する. また、手指の移動速度が速いときは明瞭な輪郭形状が得られないので、位置と速度のみを特徴とするモデルを登録する. 速度に応じて適宜選択された特徴を持つモデルから構成される

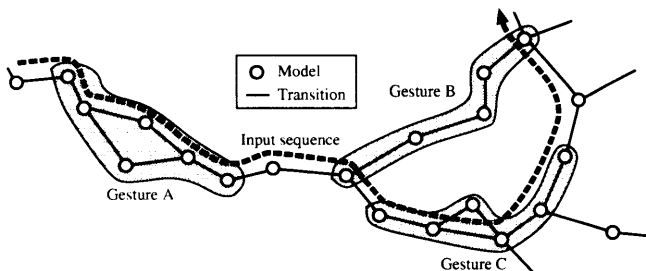


図9 遷移ネットワークに基づくジェスチャ認識



図10 学習する形状の例

遷移ネットワークを生成することで、高速な手指移動と変形を伴うジェスチャの手指形状追跡を行う。

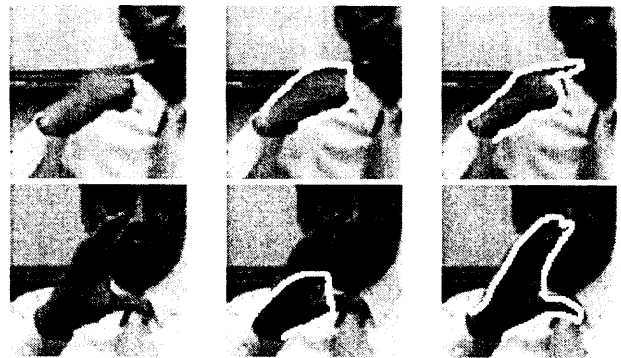
遷移ネットワークを拡張することで、ジェスチャ区間の事前抽出を必要としないジェスチャ認識を行うことができる。遷移ネットワークは遷移区間も含むすべての学習用サンプル画像列から一つの大規模なネットワークが構築される。入力画像の形状推定結果の系列はこのネットワーク上で経路を構成し、また各ジェスチャは遷移ネットワーク上の部分経路に対応する。そこで、形状推定結果の経路がジェスチャの部分経路を通過したことを検出することで、ジェスチャを認識する(図9)。遷移ネットワークに基づく形状推定によりジェスチャ間の遷移区間も含めて常に形状変化の追跡が為され、その推定結果からジェスチャの開始、追跡、終了を検出するので、あらかじめジェスチャ区間を切り出す必要がない。

遷移ネットワーク中のモデルと画像の照合では、手指輪郭および背景からエッジ点が観測される確率に基づく評価基準によって複雑背景下で正しい照合を実現した。とくに、従来の評価では、あるモデルの評価にはそのモデルの照合度のみを考慮し、他のマッチング候補が間違いであるかどうかは評価していないが、本手法では真の手指輪郭上、および背景中のエッジ点の存在確率に基づき、すべてのマッチング候補のモデル輪郭点を評価する評価基準を尤度分布として定義し、ベイズ推定の枠組みにより各モデルを評価する。

手話単語20単語(各単語3シーケンス)の計2390フレームの時系列画像から、形状を持つモデル数が187、形状を持たないモデル数が101の合計288モデルからなる遷移ネットワークを自動生成し、速い動作や顔と手の重なりを含む手指形状の追跡を正しく行うことができた。図10に学習画像に含まれる手指形状の例を示す。

図11に入力画像に対して推定されたモデルを示す。

図12に時系列画像の推定結果例を示す。結果画像はビーム探索で探索された経路うち、最終フレームで推定されたモデルに到達可能な経路のモデルを描いている。



(a) Input image (b) Selected model using average of contour contrast (c) Selected model using our criterion

図11 提案手法による照合結果

テスト画像列60個に対して単語の認識実験を行った結果、60個のテスト画像列のうち、44個について正解ジェスチャのみが検出された。また、複数のジェスチャが検出された場合には、時系列の連続性を考慮したジェスチャ照合の評価基準に基づいて最適ジェスチャを選択することで、13個について正解ジェスチャが検出された。

## 5. まとめ

本研究では、複雑な関節物体としての手指の形状と姿勢を画像列から非接触に推定する手法について、

- (1) 3次元手指構造モデルを用いた形状変動の学習に基づく手指形状識別
  - (2) 形状モデルの不確実さを補償するオンライン式モデル詳細化手法
  - (3) 手話知識を利用した複雑背景下における手指領域の追跡と単語認識
- の3つのトピックを取り上げ概要を述べた。

3次元手指構造モデルを用いた形状変動の学習に基づく手指形状識別では現在約16万通りの姿勢パターンが登録されているが、手指関節の自由度を8自由度に制限しているため、自由度を増やして実際の手指のあらゆるな形を識別しようとするパターン数が指数的に増大するという問題点がある。これを解決するには、手指輪郭全体をパターンマッチするのではなく、モデル中の独立に稼働する部分ごとに画像輪郭との対応づけをおこない、部分ごとのマッチングを行う必要がある。

形状モデルの不確実さを補償するオンライン式モデル詳細化手法については、従来不可能とされてきた単眼視時系列画像から関節物体の腕の長さの推定について、関節の可動範囲や各部の長さの相関を利用することにより推定できる場合があることを原理的に示すことができた。しかし、多次元空間におけるパラメータ領域の更新には極めて多大な計算量を費やし、楕円をもちいた近似計算をもちいると大まかには計算できるが、個人の手指形状の違いを正確に区別し計測するほどの精度を得るには至っていない。

遷移ネットワークを利用した複雑背景下における手指領域の抽出と追跡については、手指の三次元姿勢変化が微妙であっても、カメラと手話演者との位置関係によって輪郭形状がしば

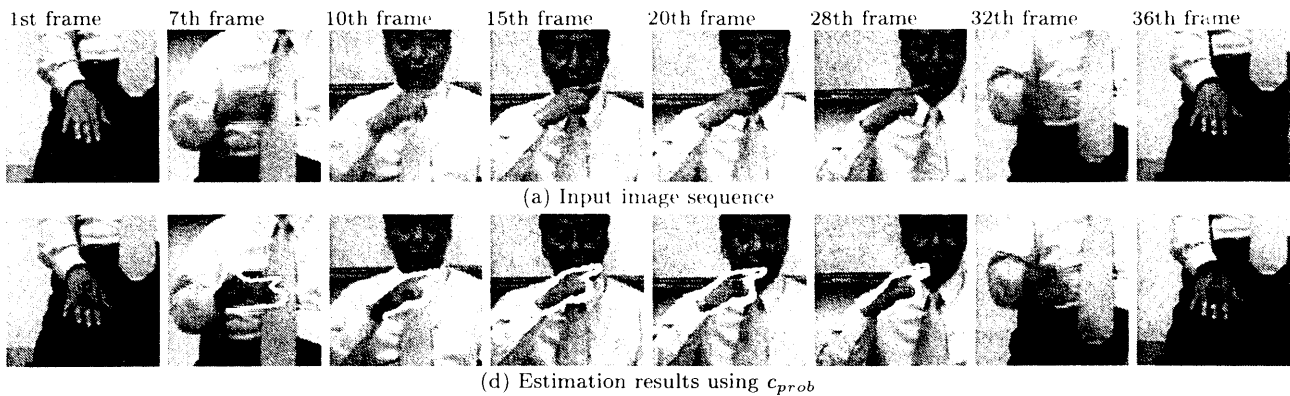


図 12 動画の形状推定結果

しば大きく変化してしまうことが原因で照合・追跡に失敗するケースが見つまっている。現在の遷移ネットワークのモデルノードには手指形状が二次元輪郭モデルとして格納されているため、あらかじめ学習に利用した手話サンプルのなかに存在しなかった 3 次元的な輪郭の変形は考慮することができない。3 次元手指モデルを用いた形状変動の学習方法を適用することが課題である。個別の手話単語ではなく手話文を認識することも課題である。

本格的な手指形状推定システムを構築するには、これらの要素技術を現実的な計算資源の中で統合することが必要であるが、現時点では個々の要素技術が個別にできあがった段階であり、これらを実用的なシステムとするには、単語の認識と特徴抽出や追跡の処理が相互に補完しあう構造に組み直す方が今後必要である。

#### 文 献

- [1] 浜田康志, 島田伸敬, 白井良明, "手話認識のための複雑背景における手指形状推定", 信学技報, PRMU2003-152, pp. 7-12, Nov. 2003.
- [2] 岩井儀雄, 八木康史, 谷内田正彦, "単眼動画画像からの手の 3 次元運動と位置の推定". 信学論 D-II Vol.J80-D-II, No.1, pp.44-55, 1997.
- [3] 藪上勝宏, 上田悦子, 松本吉央, 小笠原司, "PC クラスタを用いたリアルタイム手形状推定". インタラクション 2004 論文集, pp.199-200, 2004
- [4] X.Liu, K.Fujimura. "Hand Gesture Recognition using Depth Data". In *Proc. 6th Int. Conf. on Automatic Face and Gesture Recognition*, pp. 529-534, 2004.
- [5] K. Imagawa, R. Taniguchi, D. Arita, H. Matsuo, S. Lu, S. Igi. "Appearance-based Recognition of Hand Shapes for Sign Language in Low Resolution Image". In *ACCV2000*, pp. 943-948, 2000.
- [6] M. J Black and A. D. Jepson. "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation". *Int.J.of Computer Vision* 26(1), pp. 63-84, 1998.
- [7] 亀田 能成, 美濃 導彦, 池田 克夫, "シルエット画像からの関節物体の姿勢推定法". 信学論 D-II Vol.J79-D-II, No.1, pp.26-35, 1996.
- [8] J.Cui, Z.Sun. "Visual Hand Motion Capture for Guiding A Dexterous Hand". In *Proc. 6th Int. Conf. on Automatic Face and Gesture Recognition*, pp. 729-734, 2004.
- [9] R. Rosales and S. Sclaroff. "Algorithms for Inference in Specialized Maps for Recovering 3D Hand Pose". *Proc. 5th Int. Conf. on Automatic Face and Gesture Recognition*, pp. 143-148, 2002.
- [10] N. Shimada, K. Kimura, and Y. Shirai and Y. Kuno. "Hand Posture Estimation by Combining 2-D Appearance-based and 3-D Model-based Approaches". In *ICPR'00*, pp.709-712, 2000.
- [11] N. Shimada, K. Kimura, and Y. Shirai. "Real-time 3-D Hand Posture Estimation based on 2-D Appearance Retrieval Using Monocular Camera". In *Proc. Int. Workshop on RATFG-RTS*, pp. 23-30, 2001.
- [12] V. Athitsos and S. Sclaroff. "An Appearance-based Framework for 3D Hand Shape Classification and Camera View-point Estimation". In *Proc. 5th Int. Conf. on Automatic Face and Gesture Recognition*, pp. 40-45, 2002.
- [13] J. M. Rehg and T. Kanade. "Visual Tracking of High DOF Articulated Structures: an Application to Human Hand Tracking". *ECCV'94*, pp. 35-46, 1994.
- [14] 島田伸敬, 白井良明, 久野義徳, "確率に基づく探索と照合を用いた画像からの手指の三次元姿勢推定". 信学論 (DII) vol.J79-D-II No.7, pp. 1210-1217, 1996.
- [15] 島田伸敬, 白井良明, 久野義徳, 三浦純 "緩やかな制約知識を利用した単眼視動画画像からの関節物体の形状と姿勢の同時推定". 信学論 (DII) vol.J81-D-II No.1, pp. 45-53, 1998.
- [16] N. Shimada, Y. Shirai and Y. Kuno, "Model Adaptation and Posture Estimation of Moving Articulated Objects Using Monocular Camera", *Proc. of Int'l Workshop on Articulated Motion and Deformable Objects (LNCS 1899)*, pp.159-172, 2000.
- [17] 島田伸敬, 今井章博, 白井良明, "単眼画像入力による非接触ビデオレート手指形状推定システム". 第 8 回画像センシングシンポジウム, pp.313-318, 2002.
- [18] 今井章博, 島田伸敬, 白井良明, "輪郭の変形の学習による 3-D 手指姿勢の認識". 信学論 (DII) vol.J88-D-II No.8, pp. 1643-1651, 2004.
- [19] Y. Hamada, N. Shimada and Y. Shirai, "Hand Shape Estimation under Complex Backgrounds for Sign Language Recognition", *Proc. of 6th Int. Conf. on Automatic Face and Gesture Recognition*, pp.589-594, 2004.
- [20] M. H. Jeong, Y. Kuno, N. Shimada and Y. Shirai, "Recognition of Shape-Changing Hand Gestures", *IEICE Transactions Division D, Vol.E85-D, No.10*, pp.1678-1687, 2002.
- [21] M. H. Jeong, Y. Kuno, N. Shimada and Y. Shirai, "Recognition of Two-Hand Gestures Using Coupled Switching Linear Model", *IEICE Transactions Division D, Vol.E86-D, No.8*, pp.1416-1425, 2003.
- [22] 谷端 伸彦, 島田 伸敬, 白井良明, "複雑背景下における手指特徴抽出と手話認識", *MIRU2002, Vol.II*, pp.105-110, 2002.
- [23] 浜田康志, 島田伸敬, 白井良明, "遷移ネットワークに基づく複雑背景下での手指ジェスチャの認識", *情報処理学会CVIM研究会研究報告 2005-150-2*, pp.9-16, 2005.