

人体ジェスチャの画像計測手法とその応用

島田伸敬 白井良明

立命館大学情報理工学部知能情報学科

〒 525-8577 滋賀県草津市野路東 1-1-1 E-mail: shimada@ci.ritsumei.ac.jp

人間の体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、長年コンピュータビジョン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手指形状を計測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照合する方法、さらに複雑背景で領域抽出と照合を行う手法を紹介する。またこれらの手法を手話認識やスポーツスキル診断への応用した例について報告する。

Image-based Measurement of Human Gesture and Its Applications

Nobutaka Shimada and Yoshiaki Shirai

Dept. of Human and Computer Intelligence, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu, 525-8577, Japan

Since the human body 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 have been developing. The methods are divided into two major categories: one is the appearance-based method and the other is the 3-D-shape-model-based method, which can be suitably selected according to applications. The paper also introduces their applications to sign language recognition and sport skill diagnosis.

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久しい。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人体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。

画像に基づくジェスチャ認識手法は十数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が[6]、セルフオクルージョンが起こるとボクセルが正しく作成されない問

題もある。そこで、近年では3次元モデルから見えを生成して入力画像との照合尤度を計算し、particle filterなどの探索手法を応用して、照合のロバスト性を確保しつつ広いパラメータ空間を効率的に探索する手法が提案されてきている[3][4][5]。

一方後者の方法では、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が選び出される[7]。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[8]。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[9]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[12][19]、三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もある。

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][17][18][20][21]。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”[10][11]の考え方に基づき、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

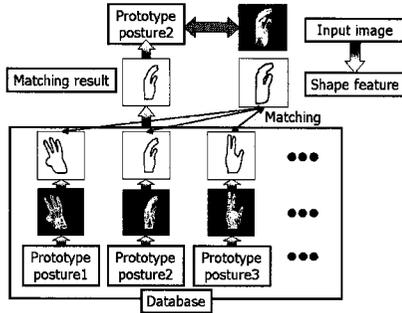


図 1: 姿勢推定方法



(a) Prototype posture (b) Variation1 (c) Variation2 (d) Variation3

図 2: 典型姿勢とバリエーション

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人体領域抽出問題を解決することが必要となるが、複雑な背景下では人体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][22][26]。

これらの研究に加えて、画像に基づく人体形状の計測技術を用いた実用システムとして、ゴルフスイングの自動診断システムを最近メーカーと共同開発した [27]。

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



(a) 1st eigenvector (b) 3rd eigenvector (c) 5th eigenvector

図 3: 学習された許容変動の例 (固有ベクトル)

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る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、ビーム探索による並列探索を今回の実装では用いたが、近年よく用いられるサンプリング法である particle filter を利用することもできる。

ここで、手指の関節の総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 5~8 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。

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ることで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 [20][21]。

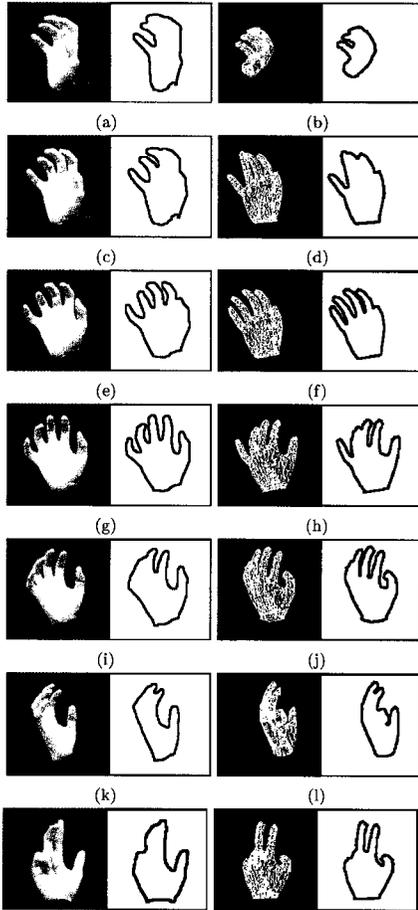


図 4: 許容変動を考慮した照合結果

3 形状モデルのオンラインモデル詳細化

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

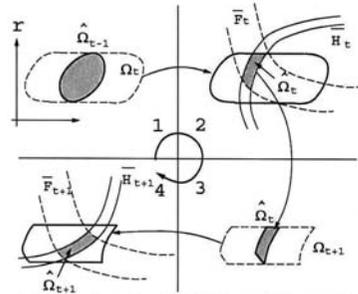


図 5: パラメータ領域の逐次的な更新

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くため、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、興行きの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、ならびに各関節どうしの許容角度差が既知
4. 各部の寸法の上限と下限、ならびに各部の寸法の許容差が既知

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つの条件はいずれも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され、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] によって可能なパラメータ群を近似的に記述する。前時刻でのパラメータ群と時系列観測、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すだいに限定する。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方向が興行きの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

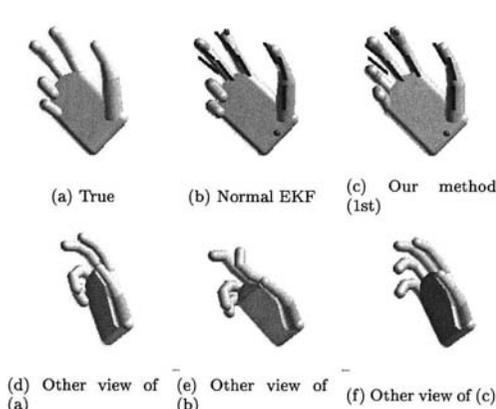


図 6: 三次元の関節物体の推定

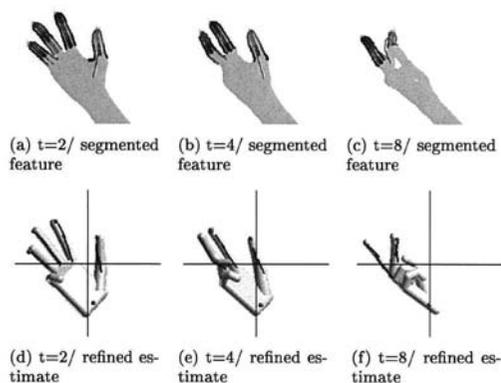


図 7: 実画像からの推定例

ら抽出された指状の突起特徴である。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。

4 複雑背景下での手指追跡と手話単語認識

手話単語の認識では、発生するジェスチャの数がざらられているので、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、ならびに単語の認識を行った [25]。

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、机の上や本だなの前など

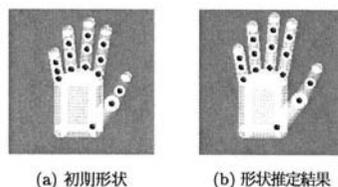


図 8: 実画像からの形状推定結果

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 [23] [24]。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が必要であった。

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 [26]。

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。

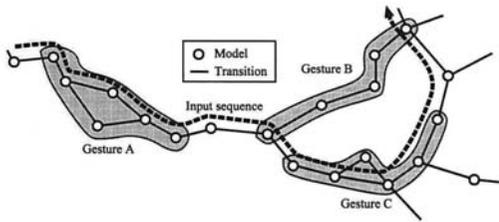


図 9: 遷移ネットワークに基づくジェスチャ認識

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く照合を実現した。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。

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候補



図 10: 学習する形状の例

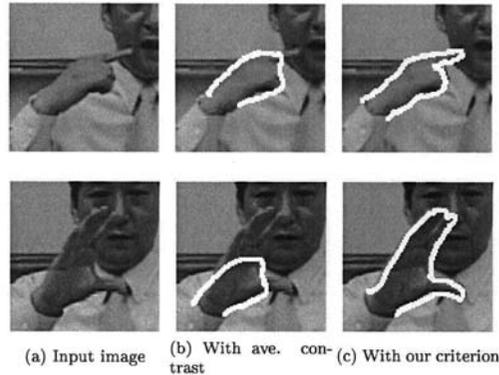


図 11: 提案手法による照合結果

として検出された場合では、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。残りの 3 個は、3 次元的な見えの変化のために、正しく照合するモデルが存在せず追跡に失敗したものである。

5 画像解析に基づくゴルフスイングの自動診断システムの開発

画像に基づくジェスチャ推定手法の応用として、ゴルフスイングを例にとり、ゴルフファーが自分でスイングの善し悪しを簡便にチェックできるよう、コンピュータが自動でスイングを診断するシステムを住友ゴム(株)およびSRI 研究開発(株)と共同開発した[27]。

コンピュータに自動でスイングの良否を判断させるには、スイングを何らかの方法で数値化する必要がある。ゴルフに限らずスポーツの分野では、フォームを分析するために古くから映像が用いられてきた。最近では、モーションキャプチャーシステムを用いて映像解析を自動で行えるようになった[28]が、複数台の特殊なカメラを必要とするために、簡易に店舗設置しづらい。本システムはスイングにおける関節の動き全てを逐次追跡するのではなく、ゴルフスイングに特化することで、ゴルフスイングを診断するために必要な画像だけを抽出し、それらの画像について、画像処理を組み合わせることにより、マークを付加していない場合や、マークが隠れている場合でも関節

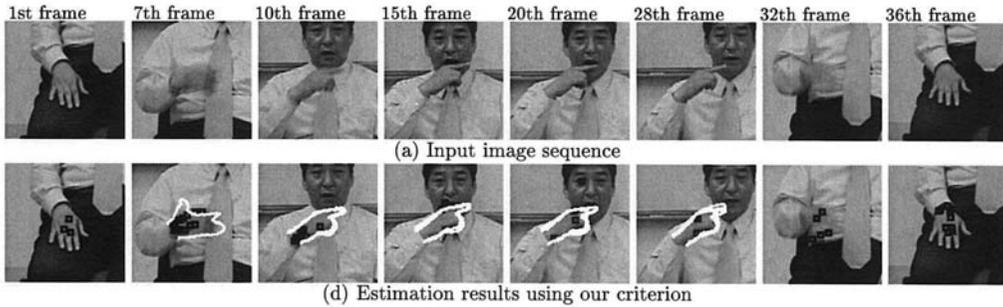


図 12: 動画画像の形状推定結果



図 13: ゴルフスイング診断システム

などの位置をスイングを診断するのに必要な精度で推定する。これにより、ゴルファーにとっては診断用の上着を着用する程度でスイングを計測でき、小規模なシステムでゴルフスイングの特徴を自動で数値化できるようになる。さらに、取得した位置情報によりスイングを分析し、プロのトレーナーによるトレーニングプログラムに基づいて、修正ポイントや練習方法を自動で提示してくれるシステムとした。開発したシステムのハード構成図を図 13 に示す。

5.1 画像特徴の抽出

本システムでは、ダンロップトレーニングプログラムに基づき、撮影した 180 枚のスイング画像から診断に用いる 12 枚のスイング画像（以下、チェック画像）を抽出する（図 14）。シャフトに取付けたマークを追跡し、ゴルフシャフトの向きを時計の針にみたてた 8 時や 9 時などの画像を抽出する。また、テンプレートマッチングを利用して、左腕が水平（地面と平行）になる画像を抽出する。

チェック画像毎に診断に必要な関節などの位置情報を取得するために、以下のような画像処理を行う。まず、シルエットを背景差分法で抽出する。スイング画像を撮影する前に、予め背景画像を 30 枚撮影し、その平均を背景として生成する。この画像を用いて背景差分を行う。

輪郭・曲率・直線・エッジを抽出した結果を基に、以下の画像処理により関節などの位置情報を取得する（図 15,16）。

1. 輪郭情報を用いて、左右の肩が存在しそうな範囲を

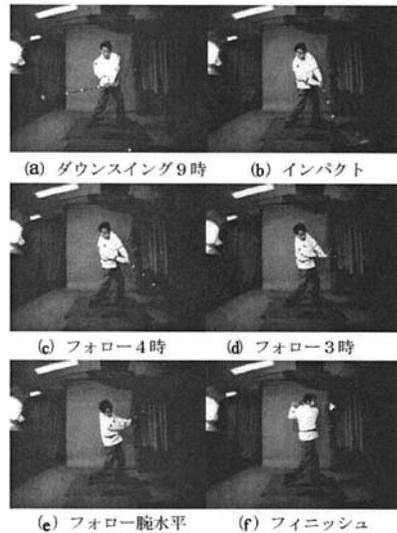


図 14: チェック用キーフレーム画像（正面）

推定し、その範囲内で HSV 成分を利用して色抽出により、位置情報を取得する。色が抽出できない場合は、曲率情報を用いて左右の肩を推定する。

2. 輪郭情報を用いて、アドレスでのスタンスの幅を抽出する。ただし、左足の端は、ライトの影響で人の影が発生し輪郭情報だけでは特定できないことがある。そこで、エッジ情報を用いて左足の端を推定する。
3. 輪郭情報（楕円枠内）を用いて、直線近似により背骨のラインを取得する。背骨ラインの始点は、顔を肌色抽出することにより推定する。
4. 曲率情報の極値を用いて、側面画像の右足先・右足首・左足先を推定する。
5. 正面画像から腕で隠れているために色抽出できないベルト（臍）の位置を、側面画像から推定する。
6. 右膝付近の直線情報を用いて、直線の交点から右膝頭を推定する。

以上の処理方法で、関節などの位置を推定した結果を図 17 に示す。アドレス画像について、正面は 13 箇所、側面

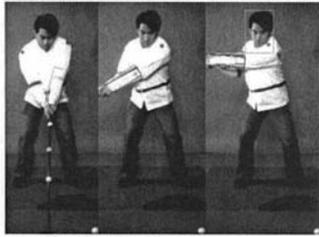


図 15: 腕と顔の抽出と追跡

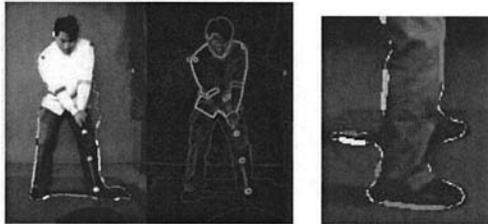


図 16: 各部の輪郭抽出結果 (左上: 正面画像 右上: 側面画像 (脚部) 下: 側面画像 (背中))

は 15 箇所の関節やスタンスの幅、グリップと体までの距離などのスイングの特徴を表わす情報を取得している。

5.2 スイングスキル診断

得られたチェック画像毎のスイングの特徴を表わす情報を利用して、トレーニングプログラムに基づきスイングを分析する。分析するチェックポイントは正面が 34 箇所、側面が 35 箇所である。モデルスイングと比較し、チェックポイント毎に定められた許容範囲から外れておれば、スイングの改善が必要である。改善が必要なチェック画像と改善すべき点の一例を図 18 (a) に示す。

スイングを分析した結果、改善点が多い場合には、どこから修正すればよいのかが受診者には分かりにくいので、ゴルフスクールでの指導要領に基づき、優先順の高い点からアドバイスを表示する。また、受診者が個人で練習できるように、アドバイス毎に練習方法 (ドリル) も音声付動画画像で提供する。さらに、ゲーム感覚で楽しみ、スキルアップの励みにもつながるよう、スイングの分析結果および弾道の結果について、点数付けして評価している。アドバイスの表示やドリル、点数結果などの一例を図 18 (b) に示す。

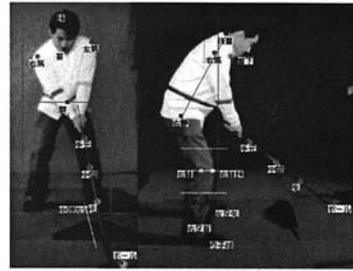


図 17: 抽出された特徴点群 (正面)



(a) 要改善点表示 (b) 総合診断結果

図 18: 診断結果表示

6 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. 形状モデルの不確か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
4. 画像解析に基づくゴルフスイングの自動診断システムの開発

の 4 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とパターンの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。

形状モデルのオンライン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、不等式制約がそもそも不正確にしか与えられない場合には得られる推定パラメータ範囲も不正確となるため、テクスチャの対応などより多くの制約を考慮する必要があると考えられる。

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、学習に利用した手話サンプルが2-Dの輪郭であるため、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ある。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。

画像解析に基づくゴルフスイングの自動診断システムの開発では、特徴追跡用マーカーと背景差分による輪郭解析の併用によって診断ポイントとなる人体各部の姿勢を計測しているが、屋外練習場やゴルフコース内での診断を可能とするには激しく変化する日照条件や風などによる背景変動のもとで人体領域を抽出する必要がある。またマーカーもできるだけ使用しない形態にする必要がある。

本格的な人体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。

参考文献

- [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] 岩井徹雄, 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] Stenger, B., Thayananthan, A., Torr, P. and Cipolla, R.: Filtering using a tree-based estimator, International Conference on Computer Vision, Vol.2, pp.1063-1070, 2003.
- [4] Wu, Y., Lin, J. and Huang, T. S.: Analyzing and Capturing Articulated Hand Motion in Image Sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, No.12, pp.1910-1922, 2005.
- [5] R. Okada, B. Stenger, T. Ike, and N. Kondoh, "Virtual Fashion Show Using Real-Time Markerless Motion Capture," Proc. of Asian Conference on Computer Vision, Vol. 2, pp. 801-810, 2006.
- [6] 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". インタラクシオン 2004 論文集, pp.199-200, 2004
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 亀田 能成, 美濃 導彦, 池田 克夫, "シルエット画像からの間接物体の姿勢推定法". 信学論 D-II Vol.J79-D-II, No.1, pp.26-35, 1996.
- [11] 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.
- [12] 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.
- [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] 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 *CPR'00*, pp.709-712, 2000.
- [18] 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.
- [19] V. Athitsos and S. Sclaroff. "An Appearance-based Framework for 3D Hand Shape Classification and Camera Viewpoint Estimation". In Proc. 5th Int. Conf. on Automatic Face and Gesture Recognition, pp. 40-45, 2002.
- [20] 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.
- [21] 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 手形状の認識". 信学論 (DII) vol.J88-D-II No.8, pp. 1643-1651, 2004.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 植田勝彦, 白井良明, 島田伸敬, 大貫正秀, "TV カメラからの映像を用いたゴルフスイング自動診断システム", DIA2006 動的画像処理実用化ワークショップ予稿集, CDROM, 2006
- [28] MOORHOUSE, 中川, "自動動作分析システムのスポーツへの応用", バイオメカニクス研究機関誌, Vol.8 No3, pp.193-200, 2004.