

複数カメラを用いた実時間三次元手形状推定

陳, 維英

九州大学システム情報科学研究院知能システム学部門

藤木, 隆司

九州大学システム情報科学研究院知能システム学部門

有田, 大作

九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎

九州大学システム情報科学研究院知能システム学部門

<https://hdl.handle.net/2324/5878>

出版情報：画像の認識・理解シンポジウム. 2006, pp. 328-333, 2006-07

バージョン：

権利関係：

複数カメラを用いた実時間三次元手形状推定

陳 維英[†] 藤木 隆司[†] 有田 大作[†] 谷口 倫一郎[†]

[†]九州大学大学院システム情報科学府 〒816-8580 福岡県春日市春日公園 6-1

E-mail: †{chen,fujiki,arita,rin}@limu.is.kyushu-u.ac.jp

あらまし 本稿では、ヒューマンインタフェースやインタラクティブシステムへの応用を目的としたビジョンベースの実時間三次元手形状推定手法を提案する。人の手は多くの関節を持つため形状のわずかな変化でも見えが大きく変化するだけでなく、手の回転によりセルフオクルージョンを引き起こすことが推定を困難にしている。本研究では、複数カメラで得られた画像から、オクルージョンが少なく、手形状の情報量が最も多いカメラ画像を2枚選んで、特徴点の三次元位置及び手の回転角度を求め、それに基づいて手の3次元形状を推定する。手指の形状推定では、画像特徴からモデルベースによる制約知識を利用した逆運動学手法 (IK) による推定と、IKの結果を受け仮説を立てた画像解析という様に、相互の情報を利用することで手指形状推定を行っていく。複雑背景における実画像を用いた実験結果は本手法の有効性を示している。

キーワード 三次元手形状推定, 逆運動学, 手の制約知識, 複数カメラ

Real-Time 3D Hand Shape Estimation using Multiple Cameras

Weiyang CHEN[†], Ryuji FUJIKI[†], Daisaku ARITA[†], and Rin-ichiro TANIGUCHI[†]

[†] Department of Intelligent System, Kyushu University 6-1 Kasuga-koen, Kasuga, Fukuoka 816-8580 Japan

E-mail: †{chen,fujiki,arita,rin}@limu.is.kyushu-u.ac.jp

Abstract This paper proposes a novel method for real-time 3D hand shape estimation that can be used for vision-based human interface and interactive applications. In this method, to reduce the influence of self-occlusion, hand shape features are acquired from the multiple camera images and two proper cameras are automatically selected for stereo vision. Then, all joint angles are estimated by inverse kinematics. Considering that the motion constraints of the articulated hand can reduce the estimation complexity arising from the high degrees of freedom of the hand, physical hand constraints heuristically introduced are combined with Inverse Kinematics to make the estimation more efficient. The effectiveness of our approach is demonstrated with experiments on a number of different hand motions with finger articulation and global hand rotation under complex background.

Key words 3D hand shape estimation, Inverse Kinematics, Hand Constraints, Multiple Cameras

1. はじめに

近年、ヒューマンインタフェースやインタラクティブシステムへの応用を目的として、ユーザにマーカやセンサ等の装着を強要することなしに任意の姿勢における三次元手形状を実時間で計測し、システムへの入力へ利用しようという試みが行われている。これは、日常生活における様々な役割に基づき、直感的かつ自由度が高いインタラクションが可能となると考えられているからである。具体的には、CG キャラクタの手の動き生成やロボットハンドの遠隔操作など手の動きをそのまま利用したシステムや、推定結果から手話やジェスチャの認識するシステムに利用することができる。

この技術を実現する手段としてビジョンベースの手法が研究されており、アプローチの仕方から大きく二種類に分けること

ができる。

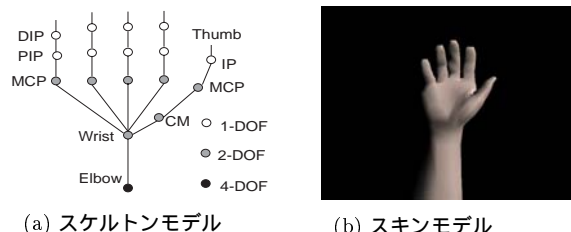
一つは二次元における手の様々な姿勢での見えを登録しておく、入力画像と最も類似するものを選択するアピアランススペースのアプローチである。このアプローチには様々な手の姿勢を事前にデータベースに登録しておくためセルフオクルージョンの問題に対処することができる。またマッチングにかかる時間を短くするようなアルゴリズムも提案されている。例えば、類似した見えを持つテンプレートは同じ親を持つような階層構造で表し、テンプレートと入力画像との類似度を事後確率を用いて計算することで前時刻の結果を考慮し、なおかつ可能性の少ない場合はその子孫のテンプレートに対して探索を行わないことによって計算を削減する手法 [1] や、形状特徴を手の形状パラメタや視点の類似したものが近傍にくるように考慮した特徴空間上で表し、前時刻で推定された点の近傍に対してビームサー

手を行うことによって計算量を削減する手法 [2] である。しかし、アルゴリズムの高速化を図ってとしても手の自由度およびカメラの視点依存性という問題が存在するため、任意の三次元手形状に対応するテンプレートを用意した場合、テンプレート数が非常に多くなってしまい、実時間での推定は困難である。また、指の関節角まで推定する場合、見えだけの情報で十分な精度が得られるかという問題もあると考えている。

もう一つは、画像から局所的な特徴を抽出し、その特徴に合うような手の三次元モデルの形状パラメタを推定するモデルベースのアプローチである。岩井ら [3] は最小自乗基準に基づいた高精度な姿勢の推定を行うことができるが、観測方向によってセルフオクルージョンが起こり、任意の手の姿勢に対してはロバスト性に欠ける。そこで、内海ら [4] はセルフオクルージョンの問題に対処するために、多視点画像から得られる、手の動き・変形による見え方に対してロバストな画像特徴を使って手の位置・姿勢を安定に推定し、掌を正面から観測する最適な視点の画像を用いて手形状を認識する手法を提案したが、精緻な三次元姿勢情報までは至っていない。亀田ら [5] は手指の 3 次元形状モデルを用いて考えられる姿勢候補を積極的に生成し、その中で入力画像のシルエットに対する適合度が最も高い候補を探索する手法を提案した。この手法は一つのシルエット画像からでもオクルージョンを含む姿勢を推定できるが、人間の手指の自由度は大きいために探索空間は広く、一般に全探索は現実的ではない。そこで島田ら [6] は広大な探索範囲を削減するために、動きによる姿勢予測に基づく事前確率の順に探索し、シルエットから抽出される形状特徴を用いて候補を限定する手法を提案した。

このように、ビジョンベースの三次元手形状推定には多くの姿勢に対応できることと計算時間のトレードオフが存在する。また、画像から求まる形状特徴としても手形状を知るために有効な情報とロバストに抽出できることといった点でのジレンマが存在する。

本研究では、計算コストの少ない手法を目指し、複数カメラを用いて画像特徴を抽出し、制約知識を利用した逆運動学手法を用いて三次元手形状推定を行うモデルベースのアプローチを提案する。低レイテンシーやエラー回復の容易性を考え、各フレーム単位でなるべく独立に処理を行う。基本的には、二視点の画像で特徴点の三次元位置を求めることができるが、手の回転やカメラ視点による手のセルフオクルージョンにロバストではない。したがって、より広範囲の回転に対応するため 3 台以上のカメラを用いて異なる視点から見た画像を取得し、オクルージョンが少なく手形状の情報量が最も多い画像を 2 枚自動的に選ぶ。そうすることにより、手形状の特徴点の三次元位置がロバストに抽出でき、手の回転角度の計算が簡単になる。手は多くの関節を持つため手形状推定が複雑になる一方、手には関節角の可動範囲や依存関係など様々な制約が存在する [10] ~ [12]。これらの研究は主により自然なアニメーションや精巧なモデルの構築を目的としてなされているが、手形状推定に利用することも有効であると考えられる。よって、それだけから手形状を知るには不十分であるが比較的ロバストに求まる画像特



(a) スケルトンモデル (b) スキンモデル

図 1 手の三次元形状モデル

徴から、制約知識を利用した逆運動学解法 (以降、IK と呼ぶ) を用いて実時間で三次元手形状を推定する。さらに、推定結果からより詳細な仮説を立てることで、ロバスト性の劣る手領域内エッジ情報も利用可能になり、これを用いてより正確な手形状推定を行なう。

2. 手の三次元形状モデル

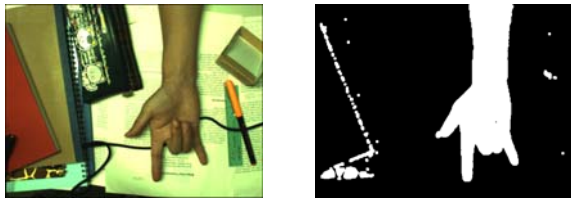
手は本来非剛体の物体であるが、本研究では剛体の関節物体として近似し、手形状推定に用いる。モデルには骨格のみで表現するスケルトンモデル (図 1(a)) と骨格のデータに付随して三次元の手の見えを表現するスキンモデル (図 1(b)) を用意するが、スキンモデルは手形状推定には用いずに結果の表示のためだけに利用した。

各指 (親指を除く) の DIP 関節および PIP 関節に 1 自由度、MP 関節に 2 自由度、計 4 自由度を考える。親指の場合は、IP 関節に 1 自由度、MCP 関節と CM 関節に 2 自由度、計 5 自由度を考える。手首の関節にはヨー、ピッチ方向回転の 2 自由度、肘の関節には平行移動成分と回転移動成分の 6 自由度を考える。ただし、ヨー、ピッチ、ロールは、それぞれ関節において左右の首振り、上下の首振り、左右の傾きの回転を表す。一般に手形状推定する場合は、手首と肘のパラメタは手首のパラメタとして一つにまとめて表されることが多いが、実際は手のロール方向の回転は肘によって制御されている。記述することができる手形状に本質的に違いがないが、今回は現実に近いパラメタの設定を採用した。また、これらのモデルは肘を最上位として、それ以降接続する関節を子とする階層構造を持ち、親の関節のパラメタの変更はこの関節に影響を与える。

手の形状パラメタは、上で述べた関節の屈曲角度と肘の三次元位置との動的なパラメタと掌の大きさや指の長さなどの静的なパラメタから成り、これらパラメタを用いて三次元上での手形状を記述することができる。本研究では、静的なパラメタはあらかじめ測定し手動で設定されるものとする。よって三次元手形状推定とは関節の屈曲角度と肘の三次元位置の総数 29 のパラメタを逐次推定することとなる。

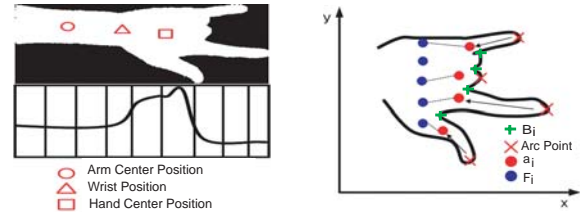
3. 色領域分割による手領域抽出

多視点による解析の有効性を確認するため、今回はとりあえず 3 台のカメラを使用する。複雑背景下で 3 枚の画像を取り込み、各々の入力画像中において手領域を抽出する。入力画像としては RGB カラー画像を取り込み、HSV (色相・彩度・明度) カラー画像に変換したものを用いる。最初に、手動で抽出



(a) 入力画像 (b) 二値画像

図 2 色領域分割による肌色領域抽出



(a) 手首・掌中心・腕中心位置

(b) 弧と指の対応

図 3 特徴点抽出

した手領域の色相と彩度の 2 次元ヒストグラム $h_{skin}(h, s)$ を計算し、事前情報としてシステムに与える。それに基づいて、入力画像における画素が肌色であるか判定し、肌色領域を表す二値画像を求める。ベイズ定理によって入力画像における画素 (h, s) が肌色である確率 $p(skin|h, s)$ を式 1 のように近似的に求めることができる [9]。

$$p(skin|h, s) = \frac{h_{skin}(h, s)}{h_{input}(h, s)} \quad (1)$$

ここで、 $h_{input}(h, s)$ は入力画像の色相と彩度の 2 次元ヒストグラムである。画素の肌色確率 $p(skin|h, s)$ が閾値より大きいと、肌色画素として識別する。ただし、得られた肌色領域を示す二値画像には手領域境界のノイズが多い。それを解決するため、モフォロジーフィルタとガウシアンフィルタを用いて画像を平滑化する。図 2 に平滑化後の二値画像を示す。また、肌色背景により複数の領域が抽出されるため、最大領域を手領域として抽出する。さらに、後の工程で行う画像処理を簡単なものとするために、手領域の主軸が画像の横軸と平行になるように回転の正規化を行う。

4. 手の特徴抽出

任意の手形状を推定するためには、どのような形状に対しても安定して特徴が抽出される必要がある。そこで、複数台カメラから画像を取り込み、オクルージョンが少なく手形状の情報が最も多い画像を 2 枚選ぶ。そして、その 2 枚の画像を用いてステレオ視で特徴点の三次元位置を求める。手のセルフオクルージョンが小さくなるのは、掌の正面に近い位置のカメラで得られた画像であり、そのような画像では、抽出される手領域が大きくなる。したがって、ここでは手の領域が最も大きい 2 枚のカメラ画像を特徴量抽出に用いる。以下に、特徴抽出の詳細を述べる。

4.1 手首・腕中心・掌中心の三次元位置検出

図 3(a) の上部の画像から、画像上での腕中心、手首、掌中心位置を求める。ただし、ここでの腕中心とは画像に含まれる前腕部分の中心位置のことをいう。

始めに手首位置 (x_{wrist}, y_{wrist}) を求める。まず、画像を横幅 n の区間に分割する。図 3(a) の下部において、左端から各区間での肌色画素数の和を求める。一般に、人の腕の幅はほぼ一定で、手首から掌にかけて幅が広がっていく。このことを利用して、左端から隣接する区間での肌色画素数の和を比較していき、ある区間での肌色画素数の和が前区間のものを定数倍したものより大きくなる位置を見つける。そのときの区間の始まりの x

座標を x_{wrist} とする。 y_{wrist} は、 x_{wrist} 上の点を上から下に走査し、手首の両端となる点を検出した上で、それらの中心位置を計算することで求める。このとき、手首の幅も計算しておく。

求めた x_{wrist} を境に、左側の領域を腕領域、右側の領域を掌領域とする。腕中心位置は腕領域の肌色画素の重心とする。一方、掌中心は、掌領域の肌色画素の中で輪郭からの距離が最も大きくなる点とする。掌中心に掌領域の重心を用いないのは、重心の位置は指を広げているか曲げているかに大きく影響され安定していないからである。

4.2 手のロール・ヨー・ピッチ方向の回転角度

4.2.1 ロール方向の回転角度

本研究では、3 台カメラを図 4 のように配置している。図 4 のように手がカメラ 1 に向けて回転する場合は、カメラ 3 の画像にはオクルージョンが生じやすいため、カメラ 1 とカメラ 2 の画像を用いて、ロール方向の回転角度 θ_{roll} を計算する。ここで、各カメラ投影が弱透視投影であると仮定する。よって、手領域の幅 x_i を手の幅 x_0 を用いて求めることができる。

$$x_1 = k_1 x_0 \cos(\theta_{roll} - \alpha_1) \quad (2)$$

$$x_2 = k_2 x_0 \cos(\theta_{roll} - \alpha_2) \quad (3)$$

ただし、 k_i はカメラ位置によるスケーリング係数、 α_i はカメラが垂直軸からの回転角度である。式 2 と式 3 から、ロール方向の回転角度を次のように計算することができる。

$$\theta_{roll} = \tan^{-1} \frac{x_1 k_2 \cos \alpha_2 - x_2 k_1 \cos \alpha_1}{x_2 k_1 \cos \alpha_1 - x_1 k_2 \cos \alpha_2} \quad (4)$$

ここでは、実験環境下でのカメラから手への距離に基づき、 $k_1 = 0.95 k_2$ とした。また、カメラ 2 とカメラ 3 が選ばれる場合も、 θ_{roll} は同じように計算できる。

4.2.2 θ_{roll} による特徴点位置の修正

4.1 節では、手首や腕や掌が平面物体と仮定する上で特徴点を抽出した。しかし、手首や腕や掌が平面物体ではないため、ロール方向の回転によって得られた特徴点には誤差が生じる。よって、画像から抽出された特徴点の 2 次元位置を修正する必要がある。ここで、得られるロール方向の回転角度を用いて手首や腕中心や掌中心の画像座標を修正する。手首は式 5 のように修正した。また、他の特徴点も同じように修正した。

$$\mathit{wrist}'_{2D} = \mathit{wrist}_{2D}(1 - \gamma \sin \theta_{roll}) \quad (5)$$

ただし、 γ は手首の厚さによるパラメータである。

4.2.3 ヨー・ピッチ方向の回転角度

手首・腕中心・掌中心の三次元位置を用いて手首のヨー・ピッ

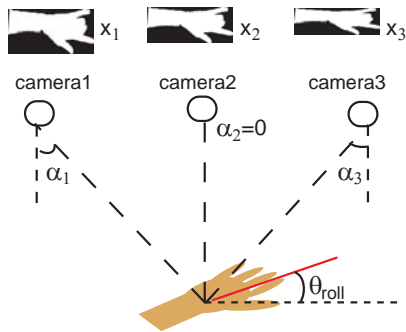


図 4 カメラの配置

手方向の回転角度を簡単に計算することができる。肘から手首の向きを持つ単位ベクトル及び手首から掌中心の向きを持つ単位ベクトルをそれぞれの平面に投影し、投影した 2 組のベクトルのなす角がヨー・ピッチ方向それぞれの回転角度である。

4.3 先頭関節検出

4.3.1 輪郭上の弧検出

掌領域の輪郭上から曲率を用いて弧を検出する (図 3(b))。このとき、検出された弧は指の指先または関節に相当するはずであり、本稿ではこの関節のことを先頭関節と呼ぶ。IK を用いて三次元手形状推定することを考えた場合、関節の先端に相当する指先を求めることが望ましい。しかし、掌方向に指を曲げた場合、指先は同じ肌色に隠れてしまうため検出が困難となる。一方、先頭の関節位置は手領域を表す輪郭の上に存在するためマスクに抽出することができる。そこで、本稿では先頭の関節位置を検出し、不足した情報を補うために手の制約知識を取り入れた IK により先頭の関節角がどの関節であるか決定し、指の関節角の推定に用いる。

4.3.2 弧と指の対応

もし、検出された弧の数が 5 個の場合、全ての指に対応する弧が見つかったと判定し、検出された弧の輪郭上の順番を基に、親指、示指、... と対応づけていく。一方検出された弧の数が 5 個ではなかった場合、検出された順番だけではどの指に対応するか判断できない。

検出された弧が $t (t < 5)$ 個の場合、各指の予測される位置を計算し、検出された弧の位置と近いものを見つけることで弧と指の対応を求める。なお、弧の x 座標の値は指の屈伸によって大きく変わるので、 y 座標の値のみを利用して以下のように行う (図 3(b))。

まず、画像上での各指の候補位置 $F_i (i = 1, \dots, 5)$ を、これまでに得られた画像特徴とあらかじめモデルのパラメタとして求まる指の間隔を用いて計算する。ただし、添え字の i は、順番に親指、示指、...、小指に対応している。

次に 1 から 5 までの数字から t 個選んだ組み合わせ $C(num_1, \dots, num_t)$ の中から次式が最小となる組み合わせ C_{min} を見つける。

$$\sum_{j=1}^t \|F_{num_j} - a_j\| \quad (6)$$

ここで a_j とは検出された弧の y 座標である。しかし一般に、指

先の y 座標の値は内転・外転によって変わりやすく、指の付け根の位置は影響を受けにくい。よって、輪郭点を走査していくことで、伸ばした状態の指については付け根の位置を求め、その y 座標の値を a_j とすることで指の内転・外転の影響を少なくする。具体的には、検出された弧点から隣接する輪郭点を両方向に一画ずつ走査していき、どちらか一方で x 座標の極小点が見つかった場合、各点の y 座標の midpoint を求め a_j の値とする。

求めた C_{min} により、弧点 a_j が $num_j \in C_{min}$ 番目の指と対応していることを求めることができる。この方法は、弧点の検出された並びを考慮することで計算する弧点と指の候補点の組み合わせを減らし、なおかつ一対一の対応を保証している。また、検出された弧の数が 5 個よりも多かった場合も同様の考え方により対応を求める。さらに、このとき検出できなかった指に対しては、凹点を用いて推定する。

4.3.3 凹点を用いた弧の推定

検出された弧が $t (t < 5)$ 個の場合、検出できなかった弧の推定が必要である。本稿では、輪郭上から親指と小指の間の凹点を曲率を用いて抽出し、弧の推定に用いる。ここで、隣接指の間の凹むところの中心を凹点 B_i とする (図 3(b))。検出された凹点の数が 4 個の場合、全ての指に対応する凹点が見つかったと判定し、検出された凹点の輪郭上の順番を基に、 $B_{thumb-index}, B_{index-mid}, \dots$ と対応づけていく。そして、輪郭に弧と凹点の相対位置関係を用いて検出できなかった弧の位置を推定することができる。例えば、中指に対応する弧が凹点 $B_{index-mid}$ と凹点 $B_{mid-ring}$ の間に存在するはずである。この位置制約を用いて検出できなかった弧の位置を推定することができる。一方、検出された凹点の数が 4 個ではなかった場合、4.3.2 節のように検出された凹点がどの指とどの指の間のものであるか判定する。そして、それらの凹点を用いて弧の推定を行う。ただし、凹点の情報を用いて弧の位置を推定できない場合、選択されなかった F_i と同じ y 座標の値の点を輪郭上から見つけ、そこから局所的に曲率が大きくなる点を探すことで近似的に推定する。

5. 制約知識を用いた IK による手形状推定

前節で述べた特徴量を基に、逆運動学解法 (IK) を用いて手指の関節角度を推定する。IK とはゴールにターゲットが一致するために必要な関節の角度を推定する手法である。なお、IK の解法には Cyclic Coordinate Descent (CCD) 法 [13] を用いた。しかし、IK の処理は逆問題を解いており、一般に複数の解を持つことになる。よって、本手法では一意で最適な解を得るために、以下のような手の制約知識を利用した。

5.1 手の制約知識

5.1.1 関節角の可動範囲

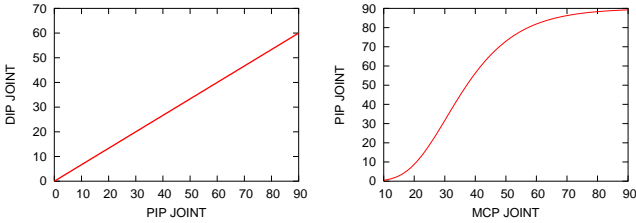
各関節には可動範囲がある [14]。指同士でも、示指・小指と比較して中指・薬指は伸展しづらいなど特徴がある。また、指は屈曲させると、同時に内転が起こるため、MP 関節の屈曲角度に伴い、線形的に内転・外転の可動範囲を制限する。これら知識に基づいて関節角に可動範囲を設定した。それぞれの関節に対する可動範囲を表 1 に示す。ここで、 $JOINT - F$ は指の

表 1 関節角の可動範囲

| | DIP-F | PIP-F | MCP-F | MCP-AA |
|----|------------|------------|------------|------------|
| 小指 | 0° ~ 70° | 0° ~ 90° | -30° ~ 90° | -20° ~ 20° |
| 環指 | 0° ~ 70° | 0° ~ 90° | -20° ~ 90° | -20° ~ 20° |
| 中指 | 0° ~ 70° | 0° ~ 90° | -20° ~ 90° | -20° ~ 20° |
| 示指 | 0° ~ 70° | 0° ~ 90° | -40° ~ 90° | -20° ~ 20° |
| | IP-F | MCP-F | CM-F | CM-AA |
| 母指 | -20° ~ 80° | -20° ~ 40° | -20° ~ 70° | -20° ~ 35° |

表 2 処理時間

| 各アルゴリズム | 時間 (msec) |
|-----------------------------|-----------|
| RGB 画像から HS 画像への変換 (3 台カメラ) | 11*3 |
| 色分割による手領域抽出 (3 台カメラ) | 43*3 |
| 手の特徴量の計算 | 6 |
| IK による手形状推定 | 19 |
| 合計 | 187 |



(a) PIP-DIP 関節間の依存関係 (b) MP-PIP 関節間の依存関係
図 5 指内の関節角の依存関係

屈曲の可動範囲, $JOINT - AA$ は MCP 関節や CM 関節の内転・外転の可動範囲である。ただし、親指の MCP 関節に対して、内転・外転の運動を考慮していない。よって、親指は他の指と同じように 4 自由度となる。

5.1.2 指内の関節角の依存関係

同じ指における関節に対して、図 5 に示すような関節角の依存関係がある。

- DIP と PIP の関節角度には比例関係 ($\theta_{DIP} = \frac{2}{3}\theta_{PIP}$) が見られる [12]。
- MP と PIP の関節角度には S 字型の相関関係がある [15]。

5.2 CCD による指の関節角の推定

指の関節角度を推定するには、まず 4.3 節で求めた先頭の関節位置がどの関節 (指先を含む) のものであるか特定する必要がある。CCD 法はターゲットとゴールの距離を最小とするような最適化問題を反復計算により解を求める発見的な手法である。一回の計算コストが小さく、前時刻の姿勢から次の姿勢を推定するので、時系列のデータを実時間で扱うのに有効である。また、CCD 法には動的に制約を付け加えることが簡単であるという特徴が挙げられる。

5.2.1 先頭関節の特定

先頭の関節の特定法としては、ある指において、抽出された先頭の関節位置 (ゴール) に指定した関節位置 (ターゲット) が一致するようにターゲットより根元側の関節 (MP 関節まで) を変化させていく。このとき、ターゲットとして全ての関節 (指先を含む) の場合を考え、ゴールとターゲットの距離が最小となる関節を先頭の関節として選択する。しかし、指定された関節より根元側の関節が手首より遠い位置にある場合は、候補として外される。これは、先頭の関節は輪郭上の弧として得られたので、手首と最も遠い位置に位置する関節である必要があるからである。

5.2.2 指の関節角の推定

上で述べたように、先頭関節を特定する同時に指の関節角を

推定するが、先頭関節より根元側の関節の関節角度しか推定することができない。よって、先頭の関節が指先であると判定された場合を除いては、さらに末端側の関節の角度を推定する必要がある。この問題を解決するために、5.1.2 節に述べた指内の関節角の依存関係を用いる。この関係を利用して、先頭の関節より末端側の関節角度を推定する。ただし、先頭の関節が MP 関節と判定された場合、どの関節の関節角度も求めることができない。しかしながら、MP 関節が輪郭上の弧として検出されるときは指を最大限に曲げている場合が多い。よって、この場合の各関節角度は可動範囲内で最大となる角度とした。

5.3 手領域内部のエッジによる関節角推定の修正

これまで三次元手形状推定に用いてきた、手の制約条件は一般的なもので、個人差や意図的に制約以外の動作した時のことは考慮していない。なかでも、MP-PIP 関節の関係式は把握動作における一般的な関係で他の制約条件と比べると強い制約ではない。よって、手の制約条件を用いた推定には誤差が生じる。しかし一方では、推定結果から大まかな手形状を知ることができるため、より詳細な仮説を立てることが可能となる。このため、画像から得られる形状特徴として輪郭情報だけでなく手領域内部のエッジ情報の利用することが可能となる。具体的には、曲げ状態の指において、指先の推定位置周辺で指先に対応する形状特徴を探索し、もし指先が求まった場合は再び IK を用いてより精密な関節角度を推定する。

6. 実験

本手法を用いて複雑背景における三次元手形状推定の実験を行った。実験には Pentium IV, 3.2GHz の CPU を搭載した計算機を 1 台、IEEE1394 規格のカメラ (Point Grey Research Flea) を 3 台使用した。また、画像サイズは 640×480 pixel である。実験では、3 台のカメラを図 4 のように配置している。これらは、固定カメラで、計測空間中に設置したワールド座標系とカメラ座標系の間であらかじめキャリブレーションしておくものとする [16]。

図 6 に本システムを用いた三次元手形状推定結果を示す。図 6(a) は掌方向に指を曲げた姿勢の入力画像及びそれに対する姿勢推定結果である。図 6(b) は手の回転を行った姿勢の入力画像及びそれに対する姿勢推定結果である。図 6(c) は手を回転しながら指を曲げた姿勢の入力画像及びそれに対する姿勢推定結果である。この推定結果の精度を評価するために、推定結果のモデルのスケルトンを入力画像に投影した。投影されたスケルトン (白線) を入力画像の手形状と比較することによって、多くの姿勢に対応する三次元手形状推定ができることを確認した。

ただし、手領域内部のエッジ時からより有効な特徴点を検出できない場合には、近似的な制約関係を用いたため指の関節角度の推定には誤差が発生したことがある。

表 2 に 1 フレームあたりの計算時間を示す。色領域分割による手領域抽出には時間がかかるため、全体の処理時間は 187msec/frame である。

7. おわりに

本稿では、複数カメラを用いて抽出される有効な画像特徴に、手の制約知識を組み合わせることで三次元手形状推定を行うモデルベースのアプローチと共に、推定結果から詳細な仮説を立てることで、より正確な推定をする手法を提案した。これにより、手の回転と指の関節運動の姿勢に対応できるシステムを構築した。複雑背景における手の実画像を用いた実験結果は提案手法の有効性を示している。

将来の課題としては以下の点が挙げられる。

- 本システムでは、複雑背景における画像処理は非常に時間がかかる。実際のヒューマンインタフェースに応用するために、より計算時間の少ないシステムを構築する。
- 今回利用した手の制約知識の他に、ある指を曲げると隣接する指も曲がってしまうなどの指間の関節角の依存関係も存在する。推定精度を向上するために、指内制約や指間制約などを組み合わせてよりロバスタな制約モデルを作成する。
- 指の長さや掌の大きさなど自動的にそれらのパラメータを推定することで、より手軽に利用できるシステムを構築する。

謝 辞

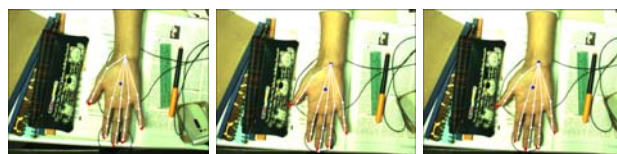
本研究の一部は、文部科学省特定領域研究「直接操作を基にした個人適応能力を有する誘発型ユーザインタフェースに関する研究」の補助を受けた。

文 献

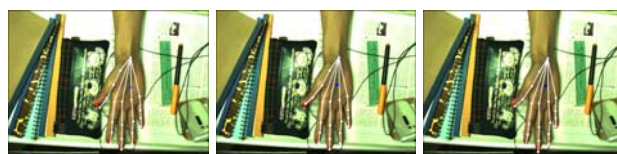
- [1] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Filtering using a tree-based estimator", in ICCV, pp.1063-1070, 2003.
- [2] N. Shimada and Y. Shirai, "Hand Posture Estimation based on 2D Appearance Retrieval Using Monocular Camera", In Proc. Int. Workshop on RATFG-RTS, pp. 23-30, 2001.
- [3] 岩井儀雄, 八木康史, 谷内田正彦, "単眼動画からの手の3次元運動と位置の推定", 電子情報通信学会論文誌, Vol. J80-DII, No. 1, pp. 44-55, Jan. 1997. (DID. 136).
- [4] 内海章, 大谷淳, 中津良平, "多数カメラを用いた両手手振りの検出", 情報処理学会論文誌, Vol.40, No.8, pp. 3143-3154, 8 1999.
- [5] 亀田, 美濃, 池田, "シルエット画像からの関節物体の姿勢推定", 電子情報通信学会論文誌, DII Vol. J79 -DII No.01 pp.26-35, 1996.
- [6] 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3次元姿勢推定", 電子情報通信学会論文誌 D-II Vol.J79-D-II No.7, pp.1210-1217, 1996.
- [7] Shan Lu, Dimitris Metacass, Dimitris Samaras, John Oliensis, "Using multiple cues for hand tracking and model refinement", IEEE Conf. pp443-450, 2003.
- [8] Etsuko Ueda, Yoshio Matsumoto, Masakazu Imai and Tsukasa Ogasawara. "Hand Pose Estimation for Vision-based Human Interface", IEEE Transactions on Industrial Electronics. Vol.50, No.4, pp.676-684, August, 2003.
- [9] K. Schwerdt and JL Crowley, "Robust face tracking using color", in Proc. of 4th International Conference on Automatic Face and Gesture Recognition, Grenoble, France, pp90-95, 2000.
- [10] C. Häger-Ross, and M. H. Schieber. "Quantifying the Independence of Hand Finger Movements: Comparisons of Digit, Hands, and Movement Frequencies", The Journal of Neuroscience, pp8543-8550, 2000.
- [11] Kamper, DG, EG Cruz, and MP Siegel. "Stereotypical fingertip trajectories during grasp", J Neurophysiol, pp3702-3710, 2003.
- [12] George ElKoura, Karan Singh, "Handrix: Animating the Human Hand", SIGGRAPH Symposium on Computer Animation, 2003.
- [13] Wang, L. T. and Chen, C. C, "A combined optimization method for solving the inverse kinematics problem of mechanical manipulators", IEEE J. Robotics and Automations. pp489-499, 1991.
- [14] 鎌倉矩子, "手のかたち 手のうごき", 医歯薬出版株式会社, 1989.
- [15] 市川 亮, "物体の把握・操作時における指関節の動きモデル化", 和歌山大学学士學位論文, 2002.
- [16] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", Transactions of Robotics and Automation, Vol.3, No.4, pp.323-344, 1987.



(a) 指関節運動に対応する手形状推定



(b) 手の回転に対応する手形状推定



(c) 指関節運動と手の回転に対応する手形状推定

図 6 三次元手形状推定結果