

## Real-time Hand Pose Estimation for Human Interface

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

吉本, 廣雅  
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

<http://hdl.handle.net/2324/5909>

---

出版情報 : 火の国情報シンポジウム, 2004-03  
バージョン :  
権利関係 :



# ヒューマンインターフェースのための実時間手形状計測

藤木隆司, 吉本廣雅, 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 春日市春日公園 6-1

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

あらまし 本論文では、自然なインタフェースを実現するための、ビジョンを用いた実時間手形状計測について述べる。手は多数の関節をもつために、そのモデルは高い自由度をもつことになるこのことが、計算量の増大やオクルージョンなどの発生を引き起こし、形状の推定を難しくしている。しかし一方で、手の大きさは不変で関節角には可動範囲があるなど様々な制約も存在する。本研究では画像から手首、指先の三次元位置を2台のカメラで計測し、手の可動範囲を利用することによって、逆運動学解法を用いて手の姿勢を推定する手法について述べる。本論文では研究の第一段階として画像から手首、指先位置の特徴点をロバストに計測できるように簡単な色マーカーを装着する。

キーワード 実時間手形状計測, 逆運動学, コンピュータービジョン

## Real-time Hand Pose Estimation for Human Interface

Ryuji Fujiki, Hiromasa Yoshimoto, Daisaku Arita and Rin-ichiro Taniguchi

Graduate School of Information Science and Electrical Engineering

Kyushu University

6-1 Kasuga-koen, Kasuga, 816-8580 Japan

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

In this paper, we propose a method of realtime hand pose estimation for natural user interface. Since a human's hand has many joints, it has geometrically high degree of freedom. This makes hand pose estimation difficult. For example we have to deal with high computation and self occlusion problems. On the other hand, a human's hand has many constraints, i.e each joint has a movable range. We solve the problem of hand pose estimation by Inverse Kinematics using these constraints. As the first step, we attach color markers fingers and a wrist in order to estimate feature points robustly

**Key words** Real-time Hand Pose Estimation, Inverse Kinematics, Computer Vision

## 1. はじめに

近年、誰もが使いやすい自然なインターフェースの研究が盛んに行われている。なかでも、人間の身体の動きを利用するものが有効であると考えられている。特に、人間の手の動作は物を掴む、放す、操作するなどの働きや、コミュニケーションの道具として重要な役割を果たしている。このことから、手の形状や動きをコンピューターの入力手段とする方法は、より自然なインタフェースを実現することが期待でき、手話認識や仮想現実感などへの応用が考えられている。

手形状を推定する手法として、装着型の測定装置であるデータグローブを用いる方法がある<sup>[2]</sup>。この方法の場合、ユーザーの指の関節角の曲げ角度、手の向き、位置を直接取得することが出来るため特徴量が比較的容易に得られる点が優れているものの、データグローブの接続により操作範囲が制限され自然な操作が出来にくいという問題がある。操作性を考えれば、手形状を非接触で推定できる手法が望ましい。

非接触型の手形状推定の手法として、カメラを用いたビジョンベースの手法が考えられる。このビジョンベースの手法のアプローチとしては、大きく二つの種類に分けることができる。

### • 手形状を認識する手法

この手法は手話やジェスチャ認識のために手形状の種類を識別するというものである。

Athisos, Sclaroff らは手の三次元モデルからエッジや突起など画像特徴を持ったモデル画像を生成し、入力画像とマッチングを行う手法を提案している<sup>[3]</sup>。この手法では、いくつかの手形状のシンボルに対して様々な角度から見た画像を用意するためにデータベースが非常に大きくなってしまふ。また、手形状を認識するという手法は、あらかじめ用意した形状にしか対応できないという問題が考えられる。

### • 手形状を計測する手法

これは手の三次元モデルを扱うために画像特徴やトラッキングを用いて関節角のパラメータを計測する手法である。

島田らは考えられる手の形状と姿勢をパラメータ空間で表現し、関節角度やリンクの長さ、太さなどに関する緩やかな制約知識を用いることで単

眼視動画像から手の形状と姿勢の両方の推定をおこなう手法提案している<sup>[4]</sup>。しかし、単眼のカメラからは距離情報を得ることが出来ないため手の姿勢を正確に推定するのは難しい。

上田らは、複数のカメラから得られるシルエット画像から視体積交差法を用いてのボクセルモデルを生成し、手のモデルとフィッティングを行うことで手の姿勢を推定している<sup>[5]</sup>。この手法はオクルージョンの問題に対応できるという利点があるが、姿勢推定の精度や計算コストが高いなどの課題がある。

本論文では実時間で三次元的な手形状を計測するシステムの実現を目指す。研究の第一段階として、手首・指先の6点に装着した色マーカーから得られる特徴点を基に手首、指先の三次元位置を求め、その点をモデルの手首、指先位置に投影し、モデルに対して逆運動学解法を用いることで、手の姿勢を計算するという手法を開発した。逆運動学解法は、一般に逆問題であるため複数の解を持つが、手の関節角の可動範囲や長さが不変という制約条件と時系列での姿勢推定により一意の解を得ることができる。

## 2. 手の三次元モデル

### 2.1 手のモデル

手は本来非剛体の物体であるが、本論文では手のひらと指が関節によって接続された剛体の関節物体としたモデルで近似する。モデルには、骨格のみで表現するスケルトンモデルと(図1)、骨格のデータに付随して三次元での手の見えを表現するスキンモデル(図2)の二つを用いる。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つ。本論文では、スケルトンモデルを用いて、逆運動学解法により姿勢を推定する。一方、スキンモデルは推定で得られた結果から手の見えを表示するために用いた。

### 2.2 関節角の可動範囲

各指はDIP, PIP関節に1自由度, MP関節に2自由度の計4自由度, 母指にはIP関節は1自由度, CM関節には3自由度で考える。手首の自由度は、手の動きを考えた場合、平行・回転移動の6自由度が考えられる。しかし実際の人の手を考えてみると、手首の平行移動と左右の傾き方向の回転移動は肘によって制御されていることがわ

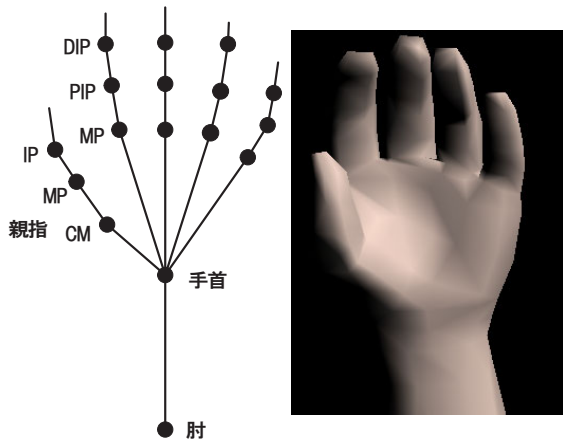


図 1: スケルトンモデル



図 2: スキンモデル

表 1: 関節の可動範囲

		roll	pitch	yow
指	DIP	×	-80 ~ 0	×
	PIP	×	-100 ~ 0	×
	MP	×	-90 ~ 45	-25 ~ 25
母指	IP	×	-80 ~ 0	×
	MP	×	-60 ~ 5	-60 ~ 20
	CM	-60 ~ 60	-3 ~ 55	-32 ~ 5
手首		×	-90 ~ 50	-30 ~ 30
肘		-60 ~ 10	-50 ~ 5	-100 ~ 60

単位: 度

かる．よって，本論文では人間の関節角の可動範囲を用いて逆運動学解法によって解くという観点から，より現実に近い手首に 2 自由度，肘に 3 自由度とし，肘の位置を固定として手の姿勢を考える．以上より，手のモデルは総数 27 の関節角の屈曲角度によって姿勢を記述することができる．

また，それぞれの関節角には可動範囲が設定されている．これにより指が反対の方向に回ってしまうなど不自然な姿勢を制御することができる．

関節角の可動範囲の例を表図 1 に示す．ただし，roll, pitch, yaw の回転は，それぞれの関節において左右の傾き，上下動，左右の首振りを表す．

### 3. 姿勢推定処理

手の姿勢推定の流れを図 3 に示す．

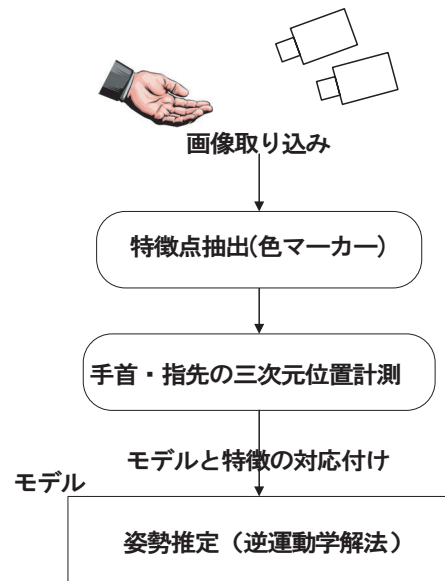


図 3: 姿勢推定の流れ

#### 3.1 特徴点抽出

本論文では手首と指先位置の特徴点をロバストに得るためにマーカーを用いる．マーカーは手首，指先の 6 点に装着し，それぞれ色の異なったものを用いる．(図 4) このことにより，求めたマーカー領域がどの点に対応するか容易に判断することができる．入力画像としては，2 台のカメラから取り込んだ RGB カラー画像 (640 × 480) を用いる，256 × 256 × 256 の RGB 空間で，それぞれのマーカーの色領域に対応する色空間にラベルをつけ LUT を作成する．そして，注目画素の値を，LUT を参照することで，対象領域を決定する．

#### 3.2 三次元位置推定

3.1 で求めたそれぞれの領域の画素から，重心点を計算する．

$$(X, Y) = \sum_{i=1}^{num} \frac{(x, y)}{num} \quad (1)$$

1 式を用いて求めた重心点から三角測量の原理を用いて手首，指先の三次元位置を推定する．また，オクルージョンにより対象領域に対応する画素がなかった場合は，単位時間に化する移動量は小さいとみなし前時刻に推定された重心点を用いる．尚，三次元位置を得るためにはカメラキャリブレーションを行う必要があるが，本システム



図 4: マーカー (手首:白, 親指:青, 示指:赤, 中指:緑, 薬指:水色, 小指:黄)

では Intel の OpenCV programming library を用いて行った。<sup>[6]</sup>

### 3.3 三次元位置とモデルの対応づけ

手首, 指先位置を求めた座標系とモデルの座標系は異なるので座標変換行列 (平行, 回転移動, スケール) を掛けることによってモデルの手首, 指先の各点に投影する。しかし, マーカーで求めた点は手指の表面上の点であり, 関節位置と異なるのでその分補正してから座標変換を行う。

### 3.4 逆運動学解法による姿勢推定

姿勢推定には逆運動学解法を用いる。関節物体やロボットの動きを解析するために運動学という学問が用いられる。運動学には大きく分けて二つに分けられる。

順運動学 (FK) 関節の角度から手先の位置を求める計算

逆運動学 (IK) 手先の位置や角度を指定して, 必要な関節角を求める計算

画像から手形状を推定する場合, 手先位置を求めることができれば, 逆運動学を用いることでそれぞれの関節角を求め, 手の姿勢を推定することができる。本論文では, CCD (Cyclic-Coordinate Descent) 法を用いる。CCD 法は逆運動学解法の一つで, 手先位置と目的地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。この手法は 1 回の計算コストが小さく, 前時刻の姿勢から次の時刻の姿勢推定を行うので, 時系列のデータを実時間で扱うのに有効と考えられる。

具体的な計算手順を以下に示す。図 5 で A, B, C は関節で, この順番で親子関係をもつとする。

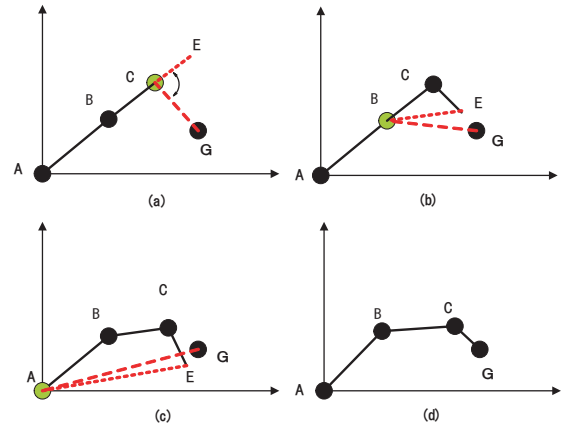


図 5: CCD 法を用いた IK

いま図中の先端の E (ターゲット) を目的地 G (ゴール) の位置まで持っていきたい場合以下の手順で計算を行う。

step1 C の関節位置から, E と G へのベクトル  $\vec{E}$ ,  $\vec{G}$  を求める。

step2 step1 で求めたベクトルを正規化し  $\vec{E}$  から  $\vec{G}$  へ回転させる回転角を算出する。

step3 関節 C に step2 で求めた回転角を加える。このとき関節の角度制限を行うため, 関節の可動範囲を新しく求めた関節角と比較する。関節角の可動範囲を超えていれば, 可動範囲の上限 (または下限) を新しい関節角として設定する。図 5(a)

step4 新しく回転角を用いた場合の, ゴールとターゲットの距離が移動前より大きくなった場合, 元の関節角に戻す。

step5 C の親の関節 B で step2 ~ 4 を繰り返す。以上の作業を指定された親まで繰り返す。図 5(b), (c)

step6 ゴールとターゲットの距離が閾値より小さくなるか, 設定された反復回数だけ step2 ~ 5 を繰り返す。図 5(d)

また CCD 法には, 複数のターゲットがある場合, どのターゲットを 1 番近づくように回転するか決

定するために重みを用いたり，一度に関節が大きく移動しないように damping(粘性減衰) を付け加えることができる．

このアルゴリズムを以下のように関節に適応することによって手の姿勢を推定する．

- (1) 肘の関節に対して，推定で求めた 6 個の点をそれに対応する手首，指先のゴールとし，手首，指先の位置が近づくように関節角を推定する．
- (2) 1 と同様に手首の関節に対して指先の 5 点の誤差が小さくなる関節角を推定する．
- (3) 各指に対して，親の上限を MP 関節までとして関節角を推定する．

(1)～(3) を指定した閾値よりゴールとターゲットの距離が小さくなるか，設定された反復回数だけ繰り返す．

#### 4. 実験と考察

実画像を用いて手の姿勢推定と計算処理時間の評価実験を行った．実験方法としては，画像から提案手法を用いて手形状計測を行って求めた関節位置と，実際の画像から関節の点を手動で位置を求め，関節位置を計測したものと距離を求めて誤差を計算した．結果を図 7 に示す．

指の番号はそれぞれの関節の階層の順番であり，番号が小さいほうが親であり，手の平からの付け根にあたり，4 は指の先端，つまりターゲットを表す．それぞれの関節位置の誤差を平均すると 15.0mm であり，先端の平均は 12.9mm の誤差が見られた．示指・中指・薬指の MP 関節で大きな誤差が見られるのが，それらの関節は画像から関節の位置を見極めるのが困難であったことが影響していると考えられる．

被験者の手は手首から中指の先端までの長さが 200mm であり，カメラキャリブレーションや手動で求められる関節位置の精度から生じる誤差を考慮すれば良好な結果が得られたと考えられる．

1 フレーム辺りの計算時間を求めた結果を示したものが表 2 である．本システムの全体での処理時間が約 14fps であるので，十分に実時間で処理できることがいえる．実験には Pentium4, 2GHz の CPU, 512MB のメモリを搭載した計算機を使用した．

表 2: 処理時間

カメラ画像取り込み	20
特徴点抽出	32
CCD 法による手の姿勢推定	5.37
モデル表示	16
total	72

単位: msec

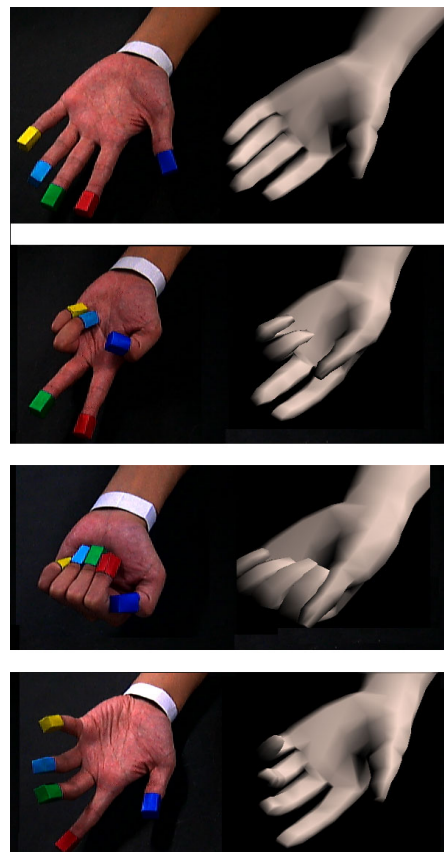


図 6: 姿勢推定結果の例

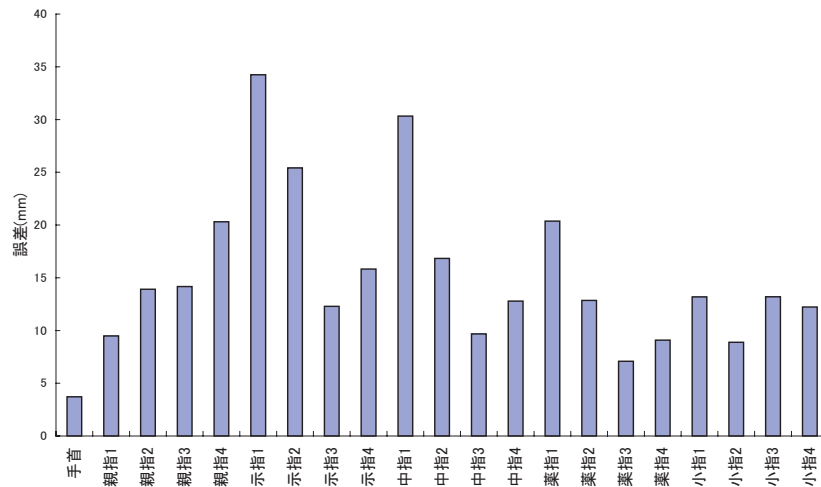


図 7: 関節位置の誤差

図 6 は姿勢推定の結果の例である．得られた結果が画像と対応していることがわかる．

## 5. おわりに

本論文では時系列画像から三次元モデルを用いて手形状を計測する手法を紹介した，姿勢推定に CCD 法を用いることで良好な結果が得られ，実時間に十分対応できるシステムを構築することができた．この手法は，手が平行移動する場合や回転移動する場合でも対応できる．今回は研究の第一段階として色マーカーを手首，指先位置に装着して画像からの特徴点を求めたが，今後はマーカーなしでの手形状計測を行について研究を進める予定である．

## 参考文献

- [1] C.lee, Y. Xu, "Online, Interactive Learning for Gesture for Humam/Robot Interfaces", IEEE International Conferences on Robotics and Automation, vol.4, pp.2982-2987, 1996.
- [2] Masumi Ishikawa, Recognition of hand-gestures based on self-organization using a DataGlove, Australian Journal of Intelligent Information Processing Systems, Vol.6, No.2, pp.65-71 2000
- [3] V.Ahitsos and S.Scaroff, "An appearance-based framework for 3D hand pose shape

classification and camera viewpoint estimation", In Automatic Face and Gesture Recognition, 2002.

- [4] 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からの関節物体の形状と姿勢の同時推定", 電子情報通信学会論文誌 vol.J81-D-II, No.1, pp.45-53, 1998
- [5] 上田悦子, 松本吉央, 今井正和, 小笠原司"多視点シルエット画像を用いた手の形状推定", 情報処理学会研究報告, 2001-CVIM-128, pp.25-31, 2001.
- [6] Intel Corporation. Open Source Computer Vision Library Reference Manual, 2001.
- [7] 鎌倉矩子, "手のかたち 手のうごき" 医歯薬出版株式会社, 1989.
- [8] 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