幾何構造と見えの同時制約に基づく複数多関節物体の運動推定

小川原光一[†] 亀崎 康之^{††} 倉爪 亮^{†††} 長谷川 勉^{†††}

† 九州大学工学研究院 〒 819-0395 福岡市西区元岡 744 †† 九州大学システム情報科学府 〒 819-0395 福岡市西区元岡 744

+++ 九州大学システム情報科学研究院 〒 819-0395 福岡市西区元岡 744

E-mail: †{ogawara,kurazume}@ait.kyushu-u.ac.jp, ††{kamechan,hasegawa}@irvs.is.kyushu-u.ac.jp

あらまし 本論文では,関節構造や接触情報など物体の幾何構造に基づく制約と画像センサから得られる距離情報や 輪郭情報など見えに基づく制約を同時に用いて運動パラメータを最適化し,既知の3次元形状モデルを用いて多関節 物体の運動追跡を行う手法を提案する.従来手法では,幾何構造を陰に含んだ運動パラメータを見えの制約のもとで 最適化する手法が一般的であったが,局所解に陥りやすいという問題があった.提案手法では,多関節物体の各部位 を独立した剛体とみなし,部位間の幾何構造の制約と観測データから得られる見えの制約を両方含んだ制約式のもと で個々に平行して収束計算を行うことによって,全体として双方の制約を同時に満たす運動を推定する.また,物体 とそれを操作する多指多関節ロボットハンドの同時運動推定実験において提案手法の有効性を確認した. キーワード 運動推定,多関節物体,幾何構造に基づく拘束,見えに基づく拘束,ICP

Motion Estimation of Multi-Joint Rigid Objects under Geometric and Appearance Constraints Koichi OGAWARA[†], Yasuyuki KAMEZAKI^{††}, Ryo KURAZUME^{†††}, and Tsutomu HASEGAWA^{†††}

† Faculty of Engineering, Kyushu University Motooka 744, Nishi-ku, Fukuoka-shi, 819-0395 Japan
†† Graduate School of Information Science and Electrical Engineering, Kyushu University
††† Faculty of Information Science and Electrical Engineering, Kyushu University
E-mail: †{ogawara,kurazume}@ait.kyushu-u.ac.jp, ††{kamechan,hasegawa}@irvs.is.kyushu-u.ac.jp

Abstract In this paper, we propose a method that estimates the motion of several multi-joint rigid objects under geometric and appearance constraints. Most of the previous works try to recover motion parameters which implicitly involves geometric constraints under appearance constraints, however these methods are easily stuck at a local minimum. In the proposed method, each part in multi-joint objects is treated as an independent object and motion parameters of all the parts are simultaneously recovered under both geometric and appearance constraints. Constraints on relative position and orientation between consecutive joints are used as geometric constraints, and depth from stereo and occluding boundary edges are used as appearance constraints. The motion parameters for all the joints are iteratively updated by robust and parallel ICP. The proposed method was evaluated by estimating motion parameters of a manipulation task operated by a multi-fingered robot hand.

Key words Motion Estimation, Multi-Joint Object, Geometric Constraint, Appearance Constraint, ICP

1. はじめに

本研究では,多関節物体を含む複数の物体が接触を 伴って運動するときに,これを画像センサで観測し,各 物体の3次元位置姿勢を時間軸にそって逐次推定する運 動追跡問題を扱う.

画像を入力とした多関節物体の運動追跡手法に関する 研究は,これまで主に人体の全身運動計測や手指の運動 計測を目的として盛んに行われてきた[1].

これらの研究は,大量の事例に基づき入力画像から対象の位置姿勢への写像を直接学習する方法[2],[3]と,関節構造を持った対象の形状モデルを既知としてこれと入力との誤差を勾配法によって最小化する方法[4]~[8]の2つに大きく分けることができる.

後者については,モデルを2次元に投影し入力画像と の誤差を最小化する方法[4],モデルとステレオ視から得 られる距離データおよび画像エッジとの誤差を最小化す る方法[5],多視点から得られる拘束とモデルとの誤差 を最小化する方法[6]などがよく用いられるが,同時に 推定する運動パラメータの数が多く高次元空間での探索 問題に帰着するため,局所解に陥りやすいという問題が あった.この問題を回避する方法として,モデルの階層 性を利用した逐次位置合わせ手法[8]や,パーティクル フィルタを用いた確率的な勾配法[7]などが提案されて いるが,次元の高さに起因する問題を回避するには至っ ていない.

一方,多関節物体の隣接する部位間の幾何的な制約 を弱め,各部位を独立した運動パラメータを持つ剛体 として扱い姿勢推定を行う手法が提案されている[9]~ [12].これらの手法では,部位をノードとするグラフ構 造によって人体もしくは手指を表現し,入力画像に対す る各部位の尤度と隣接する部位間の幾何構造に関する事 前確率を同時に最大化する各部位の運動パラメータを求 めることにより姿勢推定を行う.このとき,各部位の位 置が全く未知の場合には別途学習した部位識別器によっ て画像から部位の候補を検出し[9],[11],動的計画法[9], Non-parametric Belief Propagation (NBP)法[11],[12], Iterative Closest Point (ICP)法[10] などの最適化手法 を用いて姿勢推定を行う.

各部位の運動は隣接する部位から幾何的な制約を受け るものの独立した6次元空間の探索問題に帰着できるた め,例えば上記の階層的な逐次位置合わせ手法と比べて, 末端の部位の探索範囲が上位の部位の推定結果に拘束さ れることなく,かつ低次元空間の組合わせにより効率の よい探索が可能になる.

本論文では,上記の幾何的な制約を弱める手法に対し て以下の拡張を提案する.

まず,上記の手法はいずれも対象物体が1つの場合を 想定しているが,本論文では,多関節物体を含む複数の 物体が接触を伴って運動する場合の姿勢推定問題を扱う. そのため,画像センサから得られる入力データと各部位 との対応付けを計算し,その上で部位間の関節構造に関 する幾何的な誤差,部位同士の接触に関する幾何的な誤 差,および入力画像とモデルとの見えの誤差の3つの項 からなる誤差関数を定義し,並列 ICP 法により各部位 の運動パラメータの最適化を行う.

また,本研究では特に,多指多関節ロボットハンドに よる物体操り作業の運動計測を提案手法の対象タスクと して想定している.そこで,誤差関数の各項の重みを変 化させることによって,幾何的な制約をほぼ厳密に満た した運動パラメータの推定を行い,各物体の3次元位置 姿勢および各関節の角度情報を求める手法を提案する. 最後にロボットハンドと操作物体の同時運動推定実験を 行い,提案手法の有効性を確認した. 2. 幾何構造に基づく制約

2.1 部位間の関節構造に関する幾何的な制約



(a) 部位間の幾何制約を (b) 部位間の幾何制約を弱めた場合 厳密に満たす場合

図1 多関節物体における部位の関係

本論文で扱う多関節物体は、剛体である複数の部位から構成され、部位同士は1自由度の回転軸によって連結されているとする、図1に示すように、部位座標系の原点を回転軸上に取り、部位iの座標系から世界座標系への変換を以下の同次変換行列 T_i で定義する、このとき、 $R_{i,z}$ は回転軸に平行となるように取り、 $p_i = (R_i, t_i)$ を部位iの運動パラメータとする、

$$\boldsymbol{T}_{i} = \begin{bmatrix} \boldsymbol{R}_{i} & \boldsymbol{t}_{i} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}_{i,x} & \boldsymbol{R}_{i,y} & \boldsymbol{R}_{i,z} & \boldsymbol{t}_{i} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(1)

また,図1(a) に示すように,部位間の関節構造に関する幾何的な制約が厳密に満たされた状態で,部位i-1の座標系において,部位i-1の原点から部位iの原点へ向かうベクトルを $l_{i-1,i}$ とする.

ここで,位置に関する誤差と回転軸の向きに関する誤差の和として誤差関数 E_j を以下のように定義する. E_j は図 1(a)の状態では0になるが,制約を弱めた図 1(b)の状態では0より大きな値を取る.

$$E_{j} = \sum_{i} \| \mathbf{R}_{i-1} \mathbf{l}_{i-1,i} + \mathbf{t}_{i-1} - \mathbf{t}_{i} \|^{2} + \alpha \| \mathbf{R}_{i-1,z} - \mathbf{R}_{i,z} \|^{2} (2)$$

2.2 部位同士の接触に関する幾何的な制約



本研究では,対象物体を構成する部位は変形を起こさ ない剛体であることを想定しており,運動推定の結果図 2 に示すように部位同士が重なる場合には,これを解消 する必要がある.また,指先と操作物体など,特定の部 位同士が接触していることが分かっている場合には,部 位間の距離を0にする必要がある.これを後述の誤差関 数最小化の過程に組み込むため,重なりおよび接触の有 無に応じて,誤差を解消する勾配ベクトル $(\overline{m_k r_k})$ を求 める.

しかし, Oriented Bounding Box (OBB) のように単 純な凸包で近似する方法では精度不足であり, また vclip [13] のようにポリゴン同士の最近傍点を求める方法 では対象が非凸の場合にうまく求まらないことがある. そこで本研究では,計算時間がポリゴンの頂点数に対し て線形な以下に示す近似的な手法を用いる.



図 3 部位同士の重なり度合いの計算

図 3 に示すように,部位 i と部位 j の距離を求めたい とする.このとき,事前に部位 j の重心を中心としてこ れを取り囲む均一なポリゴン面を設定する.本研究では, 正 20 面体の各三角パッチをさらに4 つずつ 3 回分割し た計 1280 面のポリゴン球を使用した.次に,部位 j の重 心 o_j を通りポリゴン球の各面 f の中点に向かう直線を 考え,面 f の方向を正として,重心から自分自身と交差 する点までの符合付き距離 d(f) のうち最も大きいものを そのポリゴン面までのベクトルを球座標で表し,2つの 偏角 (longitude φ , latitude θ) をキーとして対応するポ リゴン面を求めるルックアップテーブル $f = LUT(\varphi, \theta)$ を計算しておく.

部位 i と部位 j の距離を求めるには,部位 i の各頂点 m_k ごとに,部位 j の座標系に直したベクトル $R_j^{-1}\overline{o_jm_k}$ から,偏角 φ_k, θ_k を求め,ルックアップテーブルを用い て部位 j の重心からの距離を求める.これに基づき,部 位同士の接触に関する誤差関数 E_c を以下のように定義 する.

$$E_c = \frac{1}{\#S_c} \sum_{k \in S_c} \| \overline{\boldsymbol{o}_j \boldsymbol{m}_k} - d(LUT(\varphi_k, \theta_k)) \frac{\overline{\boldsymbol{o}_j \boldsymbol{m}_k}}{|\overline{\boldsymbol{o}_j \boldsymbol{m}_k}|} \|^2(3)$$

ただし S_c は,重なり状態にある $\overline{o_j m_k} < d(LUT(\varphi_k, \theta_k))$ となる頂点集合と,接触を伴う部位のうち $\overline{o_j m_k} \neq d(LUT(\varphi_k, \theta_k))$ となる頂点集合の和である.また,部位 i座標系での頂点の位置を m'_k (固定値)とすると, m_k は運動パラメータの関数として $m_k = R_i m'_k + t_i$ と求められる. 3. 見えに基づく制約

本章では,見えに基づく制約として,ステレオ視など により得られる3次元情報を利用した制約と,入力画像 上の輪郭エッジなど2次元情報を利用した制約の2つの 場合を考える.

3.1 3次元情報に基づく見えの制約



図 4 3次元点群情報に基づく対応点探索

運動追跡問題の場合は,部位の大まかな位置姿勢は前 時刻の推定結果より既知であるため,図4に示すように, 前時刻の位置姿勢においた3次元形状モデルの可視の頂 点 m_k と,それに対応するデータ点 r_k との差が推定誤 差であるとみなせる.

対応点を探索する方法には,データ点からモデル点を 探索する方法と,モデル点からデータ点を探索する方法 の2通りが考えられる.本研究では,後述するように前 景すなわち物体に対応する全てのデータ点が分かってい るため,データ点からモデル点を探索する方法を採用す る.この場合,誤差関数 *E*_{3d} は以下のように定義される.

$$E_{3d} = \sum_{k \in N_{3d}} \| \boldsymbol{m}_k - \boldsymbol{r}_k \|^2$$
 (4)

ただし, N_{3d} は物体に対応するデータ点の番号である.

3.2 2次元情報に基づく見えの制約



図 5 輪郭情報に基づく対応点探索

2次元情報として輪郭エッジを扱い,画像に投影した モデルの輪郭エッジ mi_k と画像エッジ ri_kの対応を考 える.ここでも,画像エッジからモデルエッジを探索す る方法と,モデルエッジから画像エッジを探索する方法 の2通りが考えられる.本研究では,5.章で述べるよう に,対応する画像エッジが存在しないモデルエッジが多 く出てくるるため,画像エッジからモデルエッジを探索 する方法を採用する.

対応点を見つけたあとは,カメラからの距離が対応するモデルエッジ m_k と同じになるように画像エッジを3次元空間に逆投影し,3次元座標 r_k を得る.誤差関数 E_{2d} は以下のように定義される.

$$E_{2d} = \sum_{k \in N_{2d}} \| \boldsymbol{m}_k - \boldsymbol{r}_k \|^2$$
 (5)

ただし, N_{2d} は検出された画像エッジの番号である.

4. 運動推定

4.1 誤差関数

2. 章および 3. 章で定義した誤差関数の重み付き和として, 誤差関数 *E* を定義する.

$$E = w_j E_j + w_c E_c + w_{3d} E_{3d} + w_{2d} E_{2d} \tag{6}$$

ここで,これらの誤差がガウス分布に従う場合には, 最小二乗法によって式(6)を最小化することにより妥当 な運動パラメータを推定することができる.しかし,幾 何構造の制約に基づく誤差関数についてはこれが当ては まるが,見えの制約に基づく誤差関数については誤対応 の影響により外れ値の頻度が高くなるため,式の最小化 が不安定になる.

この外れ値の問題を解決するために,本研究では Wheeler の手法 [14] にならいロバスト推定であるM推定 法を導入する.

4.2 ロバスト推定

 E_{3d} と E_{2d} を以下のように書き直す.

$$E_{3d} = \sum_{k \in N_{3d}} \rho(\| \boldsymbol{m}_k - \boldsymbol{r}_k \|^2)$$
(7)

$$E_{2d} = \sum_{k \in N_{2d}} \rho(\| \boldsymbol{m}_k - \boldsymbol{r}_k \|^2)$$
(8)

ただし, $\rho(z)$ は誤差 zの任意の関数である.

式 (7) を最小化する部位 i の運動パラメータ $p_i = (\mathbf{R}_i, \mathbf{t}_i)$ は,式 (9) を解くことによって求められる.

$$\frac{\delta E_{3d}}{\delta \boldsymbol{p}_i} = \sum_{k \in N_{3d}} \frac{\delta \rho(z_k)}{\delta z_k} \frac{\delta z_k}{\delta \boldsymbol{p}_i} = 0 \tag{9}$$

ただし, $z_k = \parallel m_k - r_k \parallel^2 = \parallel R_i m'_k + t_i - r_k \parallel^2$ である.

ここで,式 (10) に示す誤差項を表す重み関数 e(z) を 導入する.

$$e(z) = \frac{1}{z} \frac{\delta \rho}{\delta z} \tag{10}$$

すると,式 (9) は式 (11) のように書き換えることがで きる.ここで,e(z)がzの関数であることを無視すると, これは重み付き最小二乗法の形となる.

$$\frac{\delta E_{3d}}{\delta \boldsymbol{p}_i} = \sum_{k \in N_{3d}} e(z_k) z_k \frac{\delta z_k}{\delta \boldsymbol{p}_i} = 0 \tag{11}$$

本研究では、外れ値を考慮した誤差分布としてローレンツ分布を用い、重み関数 e(z) を式 (12) のように定義した.

$$e(z) = \left(1 + \frac{1}{2} \left(\frac{z}{\sigma}\right)^2\right)^{-1} \tag{12}$$

式(8)についても同様に変形する.

4.3 並列 ICP 法に基づく運動パラメータの推定 運動推定の流れを図6に示す.



図 6 運動推定の流れ

入力データを受け取ると,データ点からモデル点への 対応点探索を効率よく行うために,部位の形状モデルか ら KdTree を用いて探索木を生成する.対応点探索は *E_c*, *E_{3d}*, *E_{2d}* の各計算で必要になるが,*E_c* はすべてのモ デル頂点を,*E_{3d}* は可視のモデル頂点を,*E_{2d}* は可視の 投影輪郭エッジを構成要素として,それぞれ独立に探索 木を生成する.ここで求めた探索木は,その回の繰り返 し計算では更新せずに用いる.

次に,部位ごとに式(6)を偏微分して勾配ベクトルを 計算し,続いて部位ごとに独立して一次元探索を行い各 部位の運動パラメータを更新する.

式 (6) の重みは,繰り返し計算の初期には関節構造の 幾何拘束に関する重み w_j を0として見えと接触の制約 のみを用いて運動パラメータを推定する.その後, w_j を 徐々に増やし,関節構造に関する誤差が十分小さくなる まで収束計算を繰り返す.

5. 実 験

5.1 実験環境



図7 視覚と3本指多関節ハンド

提案手法の有効性を検証するために,図7に示す10 自由度の3本指多関節ロボットハンドを用いて物体の操 り作業を行い,これを4台のカメラから構成される視覚 を用いて観察し,ハンドと物体の同時運動推定実験を 行った.

5.2 入力データの処理



図 8 入力データ処理

本実験では,参照カメラを用いた背景差分処理[15]に よって得られる前景領域(図8(b))をハンドと物体であ るとみなし,この輪郭エッジ(図8(d))を2次元情報と して用いた.物体と背景の間の輪郭エッジを用いること によって,物体や背景の模様に起因するエッジの影響を 受けずにモデルエッジとの対応を取ることが可能になる.

同時に,相関演算に基づくマルチベースラインステレ オアルゴリズム[16] と Belief Propagation (BP) による 平滑化処理[17] を組み合わせたアルゴリズムによって距 離データを計算し,このうち前景領域に含まれるデータ (図8(c))を3次元情報として用いた.

5.3 実験結果

実験は,直方体を3本指で把持しあらかじめ決められ た動きに沿って回転させる動作を行い,同時に視覚デー タを記録しておき,オフラインで運動推定を行った.今 回の実験では,サーボの関節角度の真値の情報も同時 に取得可能であったため,推定結果と真値との比較を 行った.



図 9 推定された関節角度の平均および最大誤差

推定された関節角度と真値との平均および最大誤差を プロットした結果を図9に示す.提案手法以外に,物体 の運動パラメータとして重心の並進・回転および全関節 の角度を取りこれらを同時に最適化する手法を実装して 比較を行い,提案手法の方がおおむね良い推定結果が得 られることを確認した.

計算は Xeon 3.0 GHz 搭載の計算機で行った.重みは, w_{3d}, w_c は1, w_{2d} は10とし, w_j は前半は0,後半は 1から繰り返しの度に値を倍に増やした.繰り返し回数 の平均は,前半・後半ともに10回程度であった.1フ レームの運動推定にかかる時間は,3次元復元処理も含 め,提案手法では11秒程度,比較手法では10秒程度 であった.

入力画像と3次元空間に投影した推定結果の比較を図 10に示す.

6. ま と め

本論文では,部位間の関節構造および接触情報など幾 何構造に基づく制約と,画像センサから得られる距離情 報および輪郭情報など見えに基づく制約を同時に用い, 幾何的な制約を弱め多関節物体の各部位の運動パラメー タを独立に推定することによって,複数の物体の接触を 伴う運動を逐次推定する手法を提案した.

また,物体とそれを操作する多指多関節ロボットハンドの同時運動推定実験において提案手法の有効性を確認した.

本実験ではハンドと視覚をそれぞれ三脚に固定した単 純な状況を扱ったが,提案手法が想定しているタスクは,



(a) 入力画像



(b) 提案手法



(c) 幾何拘束を含むパラメータを使用した手法

図 10 入力画像と3次元空間に投影した推定結果の比較

ロボットアームのたわみなどに起因するハンドの位置ず れの補正や,他のロボットハンドとの協調操りにおける 全体運動の同時推定などより複雑な運動であり,この検 証が今後の課題である.

文 献

- T. B. Moeslund, A. Hilton and V. Kruger: "A survey of advances in vision-based human motion capture and analysis", Computer Vision and Image Understanding: CVIU, **104**, 2, pp. 90–126 (2006).
- [2] R. Rosales and S. Sclaroff: "Combining generative and discriminative models in a framework for articulated pose estimation", International Journal of Computer Vision: IJCV, 67, 3, pp. 251–276 (2006).
- [3] A. Fathi and G. Mori: "Human pose estimation using motion exemplars", Proc. of IEEE International Conference on Computer Vision: ICCV (2007).
- [4] M. de La Gorce and N. Paragios: "Model-based hand tracking with texture, shading and self-occlusions", Proc. of IEEE Conference on Computer Vision and Pattern Recognition: CVPR (2008).
- [5] R. Plankers and P. Fua: "Articulated soft objects for multi-view shape and motion capture", IEEE Trans. on Pattern Analysis and Machine Intelligence, 25, 9, pp. 1182–1187 (2003).
- [6] Q. Delamarre and O. D. Faugeras: "3d articulated models and multi-view tracking with silhouettes", IEEE International Conference on Computer Vision: ICCV, pp. 716–721 (1999).
- [7] J. Deutscher, A. Blake and I. Reid: "Articulated body motion capture by annealed particle filtering", Proc. of IEEE Conference on Computer Vision and Pattern Recognition: CVPR, pp. 126–133 (2000).
- [8] R. Kehl, M. Bray and L. V. Gool: "Full body track-

ing from multiple views using stochastic sampling", IEEE Conference on Computer Vision and Pattern Recognition: CVPR, pp. 129–136 (2005).

- [9] B. Daubney, D. Gibson and N. Campbell: "Real-time pose estimation of articulated objects using low-level motion", Proc. of IEEE Conference on Computer Vision and Pattern Recognition: CVPR (2008).
- [10] L. Mundermann, S. Corazza and T. Andriacchi: "Accurately measuring human movement using articulated icp with soft-joint constraints and a repository of articulated models", Proc. of IEEE Conference on Computer Vision and Pattern Recognition: CVPR (2007).
- [11] L. Sigal, S. Bhatia, S. Roth, M. Black and M. Isard: "Tracking loose-limbed people", Proc. of IEEE Conference on Computer Vision and Pattern Recognition: CVPR, pp. 421–428 (2004).
- [12] E. Sudderth, M. Mandel, W. Freeman and A. Willsky: "Visual hand tracking using nonparametric belief propagation", Proc. of the 2004 IEEE CVPR Workshop on Generative Model based Vision (2004).
- [13] B. Mirtich: "V-clip: Fast and robust polyhedral collision detection", Technical report, Mitsubishi Electric Research Laboratory (1997). TR-97-05.
- [14] M. D. Wheeler: "Automatic Modeling and Localization for Object Recognition", PhD thesis, CMU (1996).
- [15] J. Sun, W. Zhang, X. Tang and H.-Y. Shum: "Background cut", Proc. of ECCV (2006).
- [16] R. Yang and M. Pollefeys: "A versatile stereo implementation on commodity graphics hardware", Real-Time Imaging, 11, 1, pp. 7–18 (2005).
- [17] P. F. Felzenszwalb and D. P. Huttenlocher: "Efficient belief propagation for early vision", Proc. of CVPR, pp. 261–268 (2004).