形状の個体差に適応する精密なジェスチャ推定

島田 伸敬　白井良明　久野義徳　三浦 純

大阪大学工学部 電子制御機械工学専攻 〒565 大阪府吹田市山田丘 2-1

Tel. 06-879-7333 FAX. 06-879-7247 E-mail: shimada@mech.eng.osaka-u.ac.jp

本研究では個体間でそれぞれに異なる人体形状に適応する物体モデルを単眼視動画像に当てはめることによって，人間のジェスチャを推定する手法について提案する．人体のような関節物体を単眼視動画像で観測した場合には，奥行きがわからないためモデルの形状と姿勢の間には曖昧さが生じる．そこで，関節角度や各部の長さ，太さに関する制約知識を利用して，この曖昧性を解決する．制約知識と現在までの観測結果を統合するために，可能な形状と姿勢のバリエーションをパラメータ空間における閉領域で表現する．最初に大まかな形状と姿勢の範囲を与えた上でさまざまな時系列観測を入力すると，この閉領域がしだいに小さくなっていき，単眼視画像から物体の形状と姿勢が次第に正確に推定されていくことを，実験例を使って示す．

Precise Gesture Estimation with Automatic Shape Adjustment to Individual User

Nobutaka Shimada, Yoshiaki Shirai, Yoshinori Kuno and Jun Miura

Dept.of Computer-Controlled Mechanical Systems, Osaka Univ.

This paper proposes a method to estimate precise human gesture adjusting a 3-D shape model to individual user's body from a monocular image sequence. Due to lack of depth information in a monocular image, the shape and pose of the articulated model cannot be uniquely determined. In order to resolve this problem, our method utilizes constraint knowledge about joint angles, lengths and widths of each part of human body. Possible variations of the shape and pose are represented as a closed set in the parameter space to integrate observations from images and the constraint knowledge. Given initial rough shape model and various observations, the possible parameter set becomes small gradually and the gesture and the shape of model are precisely estimated.

1　はじめに

人間がコミュニケーションをとる手段として最も基本的なものであるジェスチャを，画像を用いて機械に認識させる手法 [1][2][3] が注目されている．その多くは，予め形状モデルを与えておき，モデルの各関節を動かして入力された画像特徴にフィッティングするものであるが，モデルの形状が不正確な場合，ジェスチャは粗い精度でしか推定できない．長さや太さ，さらには表面形状といった人体の形状には個体差があるため，単一の形状モデルでは高い推定精度が期待できない．幾つかのコマンドを伝えるだけならば粗い精度でも十分であるが，仮想空間のなかで物を摑んだり動かしたりできるシステムを作るには，3 次元的なジェスチャが高い精度で推定できる必要がある．今後このような技術に対するニーズが，特に TV ゲーム・アミューズメント方面で高まっていくと予想される．

モデルの形状とジェスチャの両方をともに推定するには，複数カメラを使って得られる特徴点の距離情報 [4] を用いたり，時系列にわたる多くの特徴点の対応を使う方法がある [5]．しかし人体の場合，自己隠蔽によってあるカメラからは観測できない部分がよく生じる．また，画像からモデルに対応づけられるような安定な特徴点を多数抽出することは難しい．これらの手法では，もしデータが欠落したときは形状もジェスチャも推定できない．

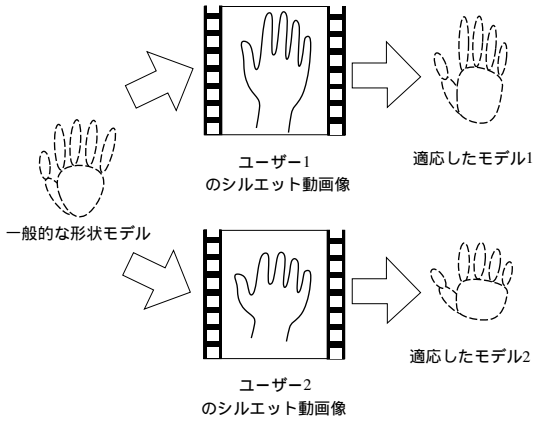


図 1: 個体差に適応する形状モデル

そこで本研究では，単眼シルエット画像列を入力とし，はじめに物体の関節構造と大雑把な形状モデルと姿勢を与えて，その姿勢を推定しつつ，形状モデルも徐々に修正していくことで，正確な形状を得る手法を提案する (図 1)．こうすれば，はじめは一

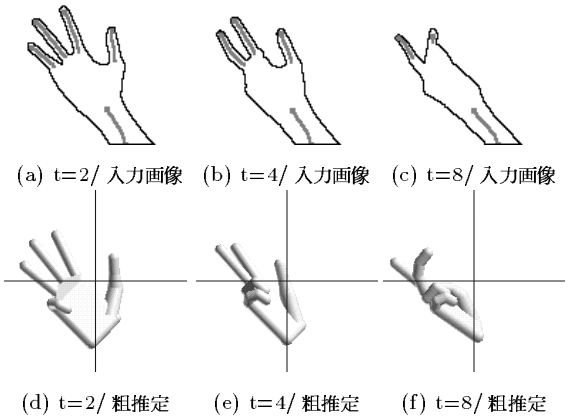


図 2: 大まかな姿勢推定結果

般的な粗いモデルのため精度はよくないが，大まかなジェスチャが推定される．また信頼できるデータが得られれば形状の推定を行ない，しだいにモデルをユーザーに適応させていくので，徐々に精度の高いジェスチャ推定が得られるようになる．

単眼画像列からは一般に奥行きがわからないため，物体の形状と姿勢の間に曖昧性を生じる．しかし，この時，物体各部の長さや太さ，各関節角度に関する制約知識がわかっていれば，推定精度をより向上できる可能性がある．そのような制約知識は，不等式で表される緩やかなものであることが多い．そこで本研究では，関節物体の可能な形状と姿勢をパラメータ空間中の領域で表現し，予測パラメータ領域のうち，観測値と制約をともに満たす領域を求めることで推定を行う枠組を提案する．

2 おおまかな姿勢の推定

本研究では，まず大まかな形状モデルを与えて，考えられる物体姿勢画像を次々と合成し，入力シルエットと比較することによって，掌の向きや指関節の角度を大まかに推定する [3]．本手法では

1. 突起特徴を利用した対応づけ候補の限定
2. 動き予測に基づく事前確率順の探索
3. 確率に基づく候補の評価
4. 入力画像の曖昧性を扱うための複数候補保持

を用い，自己隠蔽や速い動きに対応しつつ，計算量の増加を抑え適切な推定を行なう．この大まかな姿勢推定の結果，形状モデルの各部と入力画像特徴との対応づけが解決できる．図 2に大まかな姿勢推定の結果を示す．

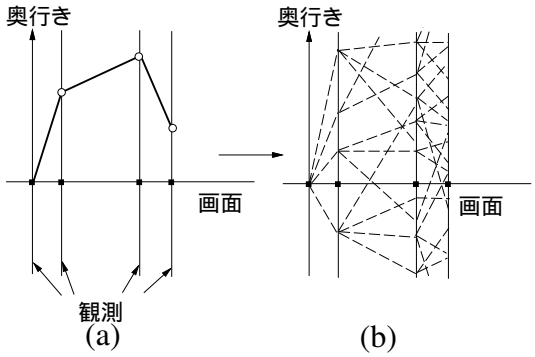


図 3: 単眼視における奥行きの曖昧性

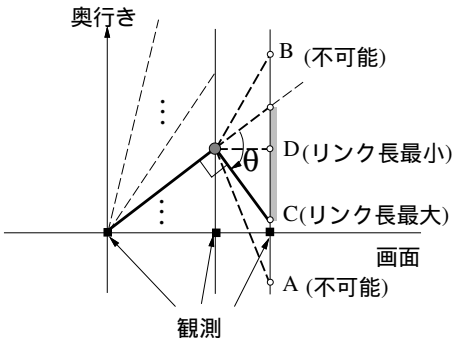


図 4: 制約知識による考えられる形状と姿勢の限定

3 制約知識による奥行きの曖昧性の限定

3.1 単眼視動画像からの形状推定における不等式制約の効果

関節構造が既知である関節物体を単眼視動画像で観測して，そのリンク長と関節角度を推定することを考えよう．画像からは，図 3(a) に示すように線分とその端点が正射影によって観測できるとする．このとき，単眼視画像では奥行きが観測できないため，線分の三次元的傾きや長さが一意に決定できない (図 3(b))．

しかし，人体のような物体を考えると，幾つかの関節角度は伴って変化し，各部の長さや太さなどには相関関係がある．これらの関係を考慮すれば，ある部分の観測が他の部分のパラメータを制約することによって，曖昧性を限定できる可能性がある．ここでは関節物体についての制約知識として以下の 3 つを考える．

1. 各パラメータの最大値と最小値
2. 単位時間に許される各パラメータの変化量（各部の長さは不変）
3. 各時刻における各パラメータ間の関係

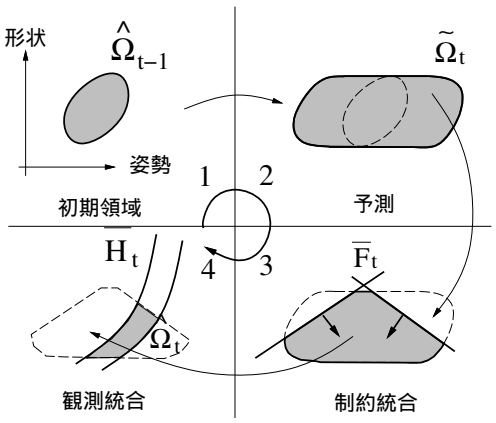


図 5: パラメータ領域の逐次的な更新

これらの制約から，以下のような場合に形状と姿勢の範囲が限定される．図 4の関節物体について，関節角度 θ が $0 \leq \theta \leq \pi/2$ に制約されている場合を考える．図中の A や B は制約を満たさないので不可能であり，リンク長は最大でも C の場合の長さにしかならないことがわかる．またリンク長の最小値は D の場合に相当することがわかる．この例から，奥行きの曖昧さを不等式制約によって減少できることがわかる．また他の関節角との制約が与えられると，さらに曖昧さを減少できる可能性がある．

3.2 形状，姿勢，制約知識の領域表現

本研究では，ある時点での形状や姿勢の可能性を，パラメータ空間における領域として考え，時系列観測と物体の制約知識を用いて可能なパラメータ領域を限定する．まず時刻 $t-1$ での推定パラメータ領域 $\widehat{\Omega}_{t-1}$ から関節のダイナミクスを考慮して，t における予測領域 $\widetilde{\Omega}_t$ を考える（図 5の 2）．次に，形状と姿勢に関する制約式を満たす領域 $\bar{F}_t$ を考える (図 5の 3 に示す帯状の領域)．同様に時刻 t での観測を満たすパラメータ領域 $\bar{H}_t$ を考える (図 5の 4 に示す帯状の領域)．これらから，時刻 t での可能性のあるパラメータ領域は

$$\widehat{\Omega}_t = \widetilde{\Omega}_t \cap \bar{H}_t \cap \bar{F}_t \qquad (1)$$

として逐次的に得られる．位置や関節角などの姿勢パラメータは時間とともに変化するが，長さや太さなどの形状パラメータは変化しないとすると，時系列観測を得るたびに形状パラメータの範囲は狭まるので，これが記憶の役割を果たし，全パラメータの可能な範囲が次第に限定されていくことになる．

O'Rourke ら [6] は各パラメータがとりうる範囲を最大値と最小値による区間で表したが，各パラメータ間の相関を考慮しないため，実際にはありえない各関節位置の組合せも含まれてしまう (図 6の矩形領域)．そこで，本研究はパラメータ領域を図 6に示す

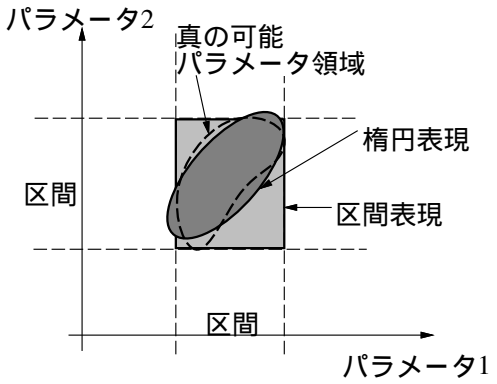


図 6: 可能パラメータ領域の区間表現と楕円表現

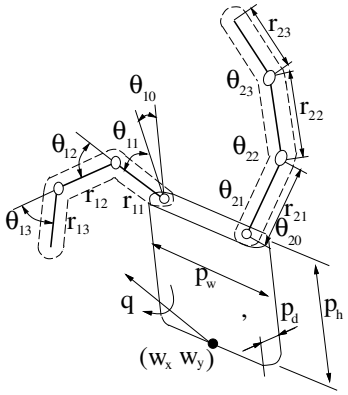


図 7: 関節物体モデル

楕円領域で表すことを考える．この楕円領域を各時刻における観測と不等式制約によって更新していくことにより，人体の姿勢と形状が次第に正確になっていく．今回は楕円領域の表現として，パラメータの確率分布を考え，その平均と共分散による記述を用いた．次節では，拡張カルマンフィルタ (EKF)[7] を用いた各時刻の楕円領域の更新方法について述べる．

4 EKF と分布切り取り法による観測と制約の統合

4.1 関節物体のモデリング

例として，図 7に示す手状の関節物体をモデリングする．簡単のため指の数を 2 本としているが，5 本の指にそのまま拡張できる．各関節は独立に回転し，全体が 3 次元平行移動と剛体回転するとする．投影は weak perspective を考える．ここで，関節物体の状態パラメータ $\boldsymbol{x}$ を，各リンクの長さ r_i，関節角度 θ_i，関節角速度 $\dot{\theta_i}$ (i は関節の指標)，掌の幅 r_w，高さ r_h，厚み r_d，並行移動成分 $\boldsymbol{t}$=$(w_x, w_y)^T$，掌の

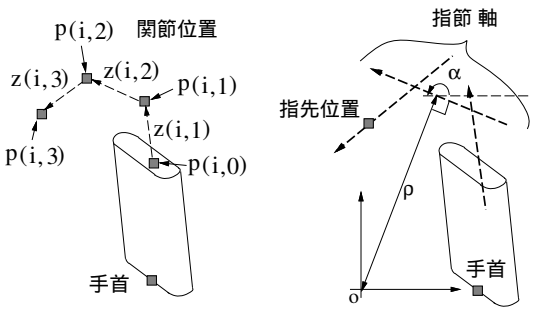


図 8: 観測できる特徴

向きと weak perspective におけるスケール率を表す四元数 $\boldsymbol{q}$[8] を用いて

$$\boldsymbol{x} = (\boldsymbol{t}^T, \boldsymbol{q}^T, \theta_{10}, \cdots, \theta_{23}, \dot{\theta_{10}}, \cdots, \dot{\boldsymbol{q}}^T, \dot{\theta_{23}}, r_{11}, \cdots, r_{23}, p_w, p_h, p_d)^T \quad (2)$$

と定義する．ただし形状は時間不変とする．

まず，状態の遷移方程式と観測方程式をそれぞれ，

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{u}_t \quad (3)$$

$$\boldsymbol{y}_t = \boldsymbol{h}(\boldsymbol{x}_t) + \boldsymbol{w}_t \quad (4)$$

と表現する． $\boldsymbol{u}_t$ は予測誤差， $\boldsymbol{w}_t$ は観測誤差であり，それぞれ平均 0，分散 $\boldsymbol{U}$， $\boldsymbol{W}$ の確率分布に従う白色雑音と仮定する． $\boldsymbol{U}$ の θ 成分は θ の加速度範囲から決定するが， r 成分は r が時間不変なので 0 とする．等速度予測を用いれば，

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{I}_n \end{bmatrix} \quad (5)$$

とできる．ここで $\boldsymbol{I}_m$ は $m \times m$ の単位行列である．また m は時間変化するパラメータ $(\boldsymbol{t}, \boldsymbol{q}, \theta_i)$ の個数，n は時間不変なパラメータ (r_i) の個数である．

次に，観測方程式 (4) を以下のようにモデリングする． ここでは図 8のように，手首位置 $\boldsymbol{t}$ と指の関節位置，指の中心軸と掌の輪郭を直線近似したものが観測できるとする．指 i の第 j 関節 (j=3 のときは指先位置) の投影位置 $\boldsymbol{p}(i,j)$ は

$$\boldsymbol{p}(i,j) = (p_x(i,j), p_y(i,j)) = \boldsymbol{L}\boldsymbol{R}(\boldsymbol{q})\boldsymbol{R}(\theta_{i0}) \cdot [\boldsymbol{d}_{ij} + \boldsymbol{p}_{i0}(r_w)] + \boldsymbol{t} \quad (6)$$

$$\boldsymbol{d}_{ij} = \sum_{k=1}^{j} r_{ik} \left(\cos(\sum_{l=1}^{k} \theta_{il}), \sin(\sum_{l=1}^{k} \theta_{il}), 0 \right)^T \quad (7)$$

とできる． $\boldsymbol{L}$ と $\boldsymbol{R}$ は平行投影と三次元の回転行列であり， $\boldsymbol{p}_{i0}$ は掌に対する指の付け根の相対位置である．

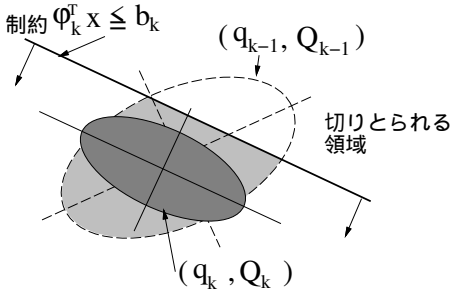


図 9: 分布切り取り法

ここで， i 指の j 節に対応する中心軸直線が，画像の X 軸からの角度 α_{ij} と原点までの距離 ρ_{ij} として観測されるとすると，式 (6) を用いて

$$\alpha_{ij} = \arctan(z_{ij}^{(y)}/z_{ij}^{(x)}) \quad (8)$$

$$\rho_{ij} = (z_{ij}^{(y)} p_{ij}^{(x)} - z_{ij}^{(x)} p_{ij}^{(y)})/|\boldsymbol{z}_{ij}| \quad (9)$$

$$\boldsymbol{z}_{ij} = (z_{ij}^{(x)}, z_{ij}^{(y)}) = \begin{cases} \boldsymbol{p}_{ij} - \boldsymbol{p}_{ij-1} & \cdots j \neq 0 \\ \boldsymbol{p}_{ij} & \cdots j = 0 \end{cases} \quad (10)$$

と表現できる．掌輪郭の観測も同様である．したがって，観測方程式 $\boldsymbol{h}$ は

$$\boldsymbol{h}(\boldsymbol{x}_t) = (\boldsymbol{p}(1,1), \cdots, \alpha_{11}, \cdots, \rho_{11}, \cdots)^T \quad (11)$$

となる．オクルージョンにより観測できない成分があるときはそれを $\boldsymbol{h}$ から除く．

最後に物体の制約知識を次のように線形不等式で表現する．

$$\theta_{min,i} \leq \theta_i \leq \theta_{max,i} \quad (12)$$

$$|\theta_i - \theta_j| \leq \Delta\theta_{ij} \quad (13)$$

$$r_{min,i} \leq r_i \leq r_{max,i} \quad (14)$$

$$|r_i - r_j| \leq \Delta r_{ij}. \quad (15)$$

式 (12)-(15) は一次式だから

$$\varphi_k^T \boldsymbol{x} \leq b_k \qquad (k = 1 \cdots K). \quad (16)$$

と一般化できる．

4.2 分布切り取り法による修正 EKF

前節で導いた遷移，観測方程式と制約不等式から可能なパラメータの範囲を推定する．もし制約がなかった場合には，観測方程式 $\boldsymbol{h}$ は非線形なので EKF を用いることによって，観測 $\boldsymbol{y}_t$ が得られたときの状態推定 $\widehat{\boldsymbol{x}}_t$ と推定分散 $\boldsymbol{P}_t$ が，

$$\widehat{\boldsymbol{x}}_t = \widetilde{\boldsymbol{x}}_t + \boldsymbol{K}_t\{\boldsymbol{y}_t - \boldsymbol{h}(\widetilde{\boldsymbol{x}}_t)\} \quad (17)$$

$$\boldsymbol{P}_t = (\boldsymbol{I} - \boldsymbol{K}_t \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_t}\right|_{\tilde{\boldsymbol{x}}_t})(\boldsymbol{A}\boldsymbol{P}_{t-1}\boldsymbol{A}^T + \boldsymbol{U}) \quad (18)$$

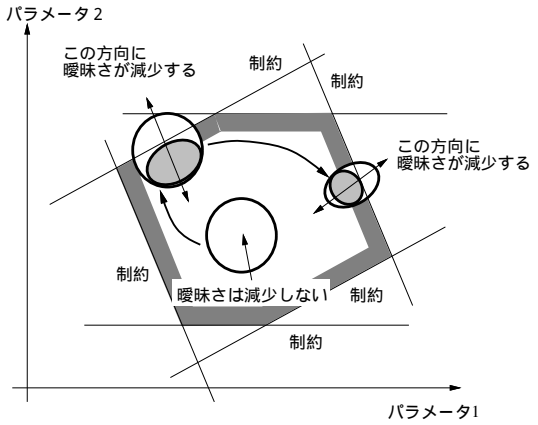


図 10: 分布切り取りによる曖昧性の減少

と得られる．上式で，$\boldsymbol{K}_t$ はカルマンゲインであり，$\tilde{\boldsymbol{x}}_t = \boldsymbol{A}\hat{\boldsymbol{x}}_{t-1}$ とおいた．偏微分の値は式 (6), (8), (9) から計算できる．

本研究では奥行きの曖昧性を減らすために，通常の EKF によって予測の平均と分散を求め，これに制約不等式 (16) を統合した分布を得る．次に得られた分布に対して，通常の EKF による更新式 (17)(18) を用いて観測結果を統合する．

以下では，EKF による予測分布に不等式制約を統合する手法を述べる．まず，制約を使わずに通常の EKF により予測の平均 $\boldsymbol{q}_0$ と分散 $\boldsymbol{Q}_0$ を得，パラメータの分布を正規分布 $N(\boldsymbol{q}_0, \boldsymbol{Q}_0)$ で近似する．次にこの分布のもとで，K 個の式 (16) を同時に満たす部分の平均と分散を求めるが，この計算は数値積分を必要とする．そこで，式 (16) で表される超平面で一つずつ順番に分布を切り取る近似計算を行うと，以下のように解析的に計算ができる．

図 9のように，平均 $\boldsymbol{q}_{k-1}$，分散 $\boldsymbol{Q}_{k-1}$ の正規分布を，k 番目の制約 $\varphi_k^T \boldsymbol{x} \leq b_k$ で切り取ることを考えよう．線形変換

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{T}^T(\boldsymbol{x} - \boldsymbol{q}_{k-1}) \tag{19}$$

を施すと，座標軸に平行な平面と平均 0，分散が単位行列，制約不等式 $x_1' \leq c_k$ の場合に帰着できる．ただし $\boldsymbol{R}, \boldsymbol{T}$ は直交行列, $\boldsymbol{W}$ は対角行列であり，

$$\boldsymbol{T}\boldsymbol{W}\boldsymbol{T}^T = \boldsymbol{Q}_{k-1} \tag{20}$$

$$\boldsymbol{R}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{T}^T\varphi_k = (1, 0, \cdots, 0)^T \tag{21}$$

$$c_k = \frac{b_k - \varphi_k^T \boldsymbol{q}_{k-1}}{\left(\varphi_k^T \boldsymbol{Q}_{k-1}\varphi_k\right)^{\frac{1}{2}}}. \tag{22}$$

を満たす．c_k は，$\boldsymbol{q}_{k-1}$ から式 (16) で表される超平面までの距離を共分散 $\boldsymbol{Q}_{k-1}$ によって正規化したものである．このとき，切り取ったあとの $\boldsymbol{x}'$ の分布の平均 $\boldsymbol{\mu}_k$ と分散 $\boldsymbol{S}_k$ は次のように計算される．

$$\boldsymbol{\mu}_k = (\nu_k, 0, \cdots, 0)^T \tag{23}$$

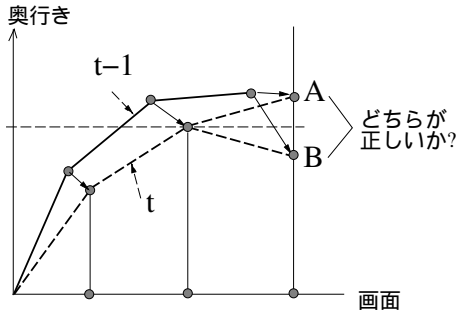


図 11: 奥行きの曖昧さによる推定の多義性

$$\boldsymbol{S}_k = \mathrm{diag}\{1 + c_k\nu_k - \nu_k^2, 1, \cdots, 1\} \tag{24}$$

$$\nu_k = -\sqrt{\frac{2}{\pi}}\exp(-\frac{c_k^2}{2})/(1 + \mathrm{erf}(\frac{c_k}{\sqrt{2}})). \tag{25}$$

ここで，$\mathrm{erf}(\cdot)$ は誤差関数であり，$\mathrm{diag}\{a, b, \cdots\}$ は対角成分が $a, b, \cdots$ である対角行列である．これを式 (19) を用いて，もとの $\boldsymbol{x}$ の座標に逆変換すれば

$$\boldsymbol{q}_k = \boldsymbol{T}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{R}^T\boldsymbol{\mu}_k + \boldsymbol{q}_{k-1} \tag{26}$$

$$\boldsymbol{Q}_k = \boldsymbol{T}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{R}^T\boldsymbol{S}_k\boldsymbol{R}\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{T}^T. \tag{27}$$

と k 回目の切り取り後の平均と分散が求まる．これをすべての k について再帰的に実行すれば，分布切り取り後の平均 $\hat{\boldsymbol{x}}_t = \boldsymbol{q}_K$ と分散 $\boldsymbol{P}_t = \boldsymbol{Q}_K$ が得られる．ただし，正規分布仮定のもとでは分布の裾野が無限に広がっているため，分布切り取りを繰り返すと共分散楕円が小さくなり続ける．そこで，c_k が一定値より大きい制約は可能なパラメータ領域から離れており分布を切りとる効果が少ないので，分布切りとりを行わないようにする．このようにして得られた制約統合後の予測分布と観測を，最後に EKF によって統合し，推定平均と推定分散を得る．

以上の手法により，対象がある制約の限界に近い姿勢をとった場合には，図 10に示すように可能なパラメータ範囲がある方向に限定される．対象がさまざまな制約の限界に近い姿勢をとるたびにパラメータ範囲が限定され，結果として形状と姿勢の両方が正確に推定できる．

5 複数候補の生成による多義性の扱い

制約知識を利用してもなお，形状と姿勢の推定に多義性が生じることがある．図 11のような場合，A, B の両方に解の可能性がある．しかし，通常の EKF ではどちらかの推定しか出力できない．そこで本研究では，EKF による推定の際に複数の候補の存在を考慮する．ある剛体部分 i の観測 $\boldsymbol{h}_i$ と，i のすぐ根元の関節パラメータ θ_i に関して

$$\left.\frac{\partial \boldsymbol{h}_i}{\partial \theta_i}\right|_{\tilde{\boldsymbol{x}}_t} \simeq \mathbf{o} \tag{28}$$

表 1: 実験で用いた制約知識

姿勢に関する制約	形状に関する制約
$\lvert\theta_{i2}-\theta_{i3}\rvert \le 10,$	$0 \le r_{i1}-r_{i2} \le 30,$
$-20 \le \theta_{10} \le 0,$	$\lvert r_{i2}-r_{i3}\rvert \le 10,$
$0 \le \theta_{20} \le 20,$	$55 \le r_{i1} \le 70,$
$0 \le \theta_{ij} \le 90,$	$30 \le r_{i2} \le 50,$
$\lvert\theta_{11}-\theta_{21}\rvert \le 35$	$25 \le r_{i3} \le 50,$
	$\lvert r_{0j}-r_{1j}\rvert \le 5$

$i=\{1,2\}, j=\{1,2,3\}$
θ の単位 deg, r の単位 pixel

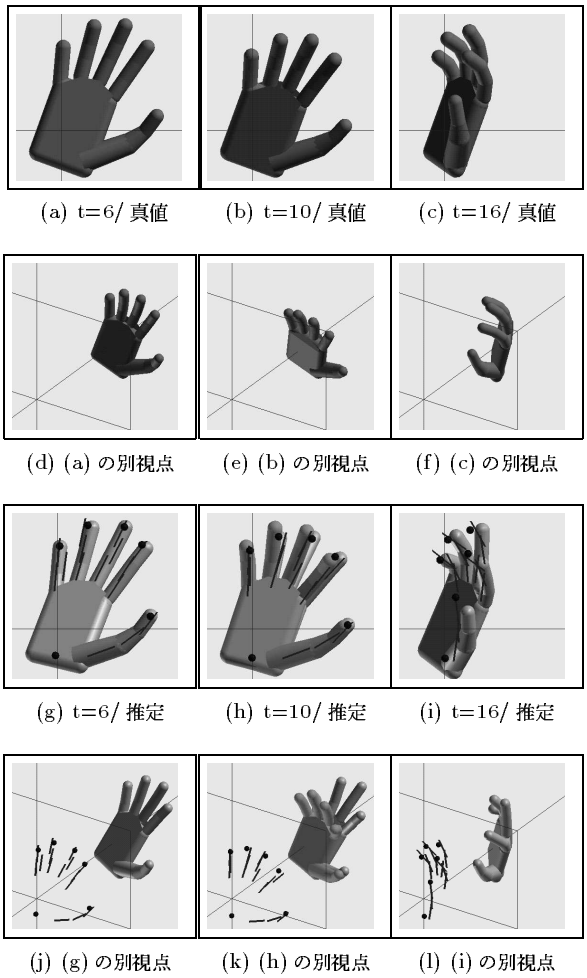
が成り立つときは，θ_i が変化しても $\boldsymbol{h}_i$ はほとんど変化しないので，図 11の A， B の 2 点で観測の線形化を行う．式 (28) が成立する場合には，

1. EKF の予測 $\widetilde{\boldsymbol{x}}_t$ を得る．
2. 投影が画面に関して $\widetilde{\boldsymbol{x}}_t$ と対称になる姿勢 $\widetilde{\boldsymbol{x}}_t^{sym}$ を，θ_i だけを変化させることで生成する．
3. $\widetilde{\boldsymbol{x}}_t$,$\widetilde{\boldsymbol{x}}_t^{sym}$ の位置で観測をそれぞれ線形化する．
4. 3 で線形化されたそれぞれの観測式を用いて，予測分布と観測を EKF により統合し，二つの解を得る．
5. 4 で得られた解について c_k を計算し，c_k がしきい値より小さくなる制約式がある (制約からのはみ出しが大きい) 場合は，間違いとみなして棄却する．
6. 5 で得られた推定に対し，制約による分布切り取りを行う．

図 12に複数の推定候補を用いて推定の多義性が解決される例を示す．第 10 フレーム (図 12(h),(k)) において，奥行きの多義性のために複数の推定が生成されている（そのうち 2 つを図示）．その後，第 16 フレーム (図 12(i),(l)) まで進むと，一方の推定は，指関節の曲がる方向についての制約によって棄却された．

6 実験結果

本手法を合成画像列と実画像列に対して適用した結果を示す．図 13(a)-(f) は，二本の指をもつ手状物体の合成画像列である．軸，指先位置と掌の輪郭線が weak perspective で観測できるとし，各観測値にガウスノイズを加えたものを観測値として計算して，形状と姿勢を推定した結果が図 13(g)-(l) ある．制約知識と初期推定を表 1，?? に示す．また図 15-17は推定パラメータの時間変化であり，真値を実線で，推定値を白丸で，推定の標準偏差の 2 倍を縦線でそれぞれ表した．これらの図から，物体の姿勢が正しく推定できたことがわかる．次に形状の個体差への



(a) t=6/ 真値　(b) t=10/ 真値　(c) t=16/ 真値
(d) (a) の別視点　(e) (b) の別視点　(f) (c) の別視点
(g) t=6/ 推定　(h) t=10/ 推定　(i) t=16/ 推定
(j) (g) の別視点　(k) (h) の別視点　(l) (i) の別視点

図 12: 複数の推定候補による多義性の解決

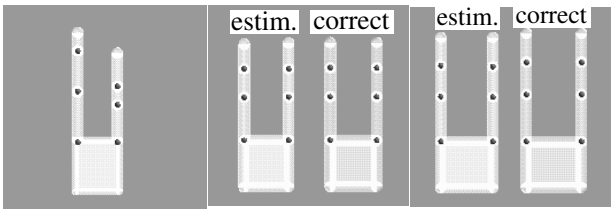
適応結果を示す．図 14(a) は初期推定形状，(b)(c) は 55 フレームにおける推定形状をそれぞれ示した図である．(c) は指の長さと掌の幅が (b) と違うものを観測した場合の結果であり，(b)(c) における指長さ推定の時間変化を図 17,18に示す．(b)(c) とも同一の初期推定と制約知識を用いながら，さまざまな観測値を得ることで，形状の個体差に適応できたことを示している．

最後に実画像列に対する実験例を図 19に示す．この例では各部の関節角と指の長さについてのみ推定を行い，投影は正射影を仮定し，スケール率は固定とした．(a)-(c) は入力シルエットとそれか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 [3] により大まかな姿勢推定 (d)-(f) を得る．大まかな推定から得られたモデルと画像特徴との対応を元に，各指節の対応する中心軸 (g)-(i) を突起特徴から切り出し，本研究で述べた手法により得た形状と姿勢の推定結果が (j)-(l) である．図 20に，最初に与えた大まかな形状モデルと本手法によって修正された形状モデルの比較を示す．

7 まとめ

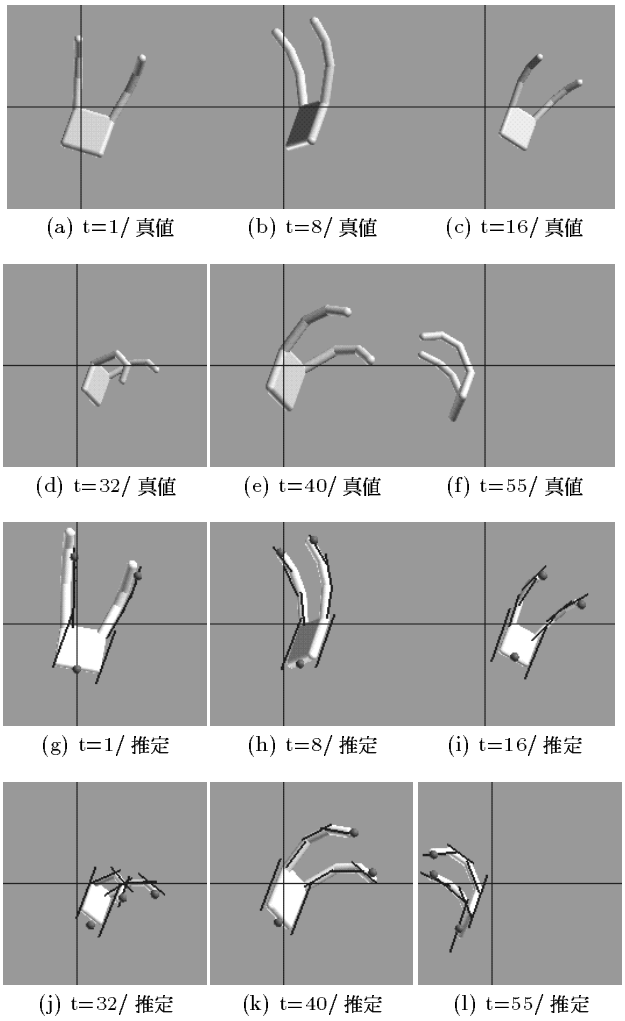
本稿では形状の個体差に適応して関節物体のジェスチャを精密に推定する手法について述べた．可能な形状と姿勢をパラメータ空間における確率分布として表現し，動画像からの観測と不等式で表現した緩やかな制約知識を統合することにより，可能性を限定し，形状の個体差を推定できることを示した．この手法を応用して，四つ足の動物の一般モデルから，猫や馬の動画像を入力することでそれらの詳細な形状モデルを生成する技術の開発が期待できる．

本手法を実際のシステムに応用する場合には，現状では大まかな姿勢を推定する部分に多くの時間がかかっているので，実時間性の確保が大きな課題となる．また，推定精度の観点からは，類似する観測データが非常に長い時間入力された場合に不当に推定分散が小さくなることが問題として残っている．新たな情報が得られていないにも関わらず，観測式の線形化誤差とノイズによる観測値の変化によって推定分散が小さくなる．この影響は，実際の観測値の分散が，EKF の構成時に見積もった分散よりも大きいときに顕著になる．合成画像を用いた実験では，実際に与えたノイズ分散を，EKF における観測誤差の分散の１／２に設定した．実験では，入力画像を十分変化させることにより，適正に形状が推定できた．より長い画像列についても適正に推定ができるように，この線形化の誤差を補正することを今後の検討課題とする．また，各部の長さや太さだけではなく，表面の形状を推定することで，より正確な形状モデルとジェスチャの推定を行なうことを考えている．



(a) 初期形状　(b) 推定結果 1 ($r_1$3=35)　(c) 推定結果 2 ($r_1$3=40)

図 14: 異なる形状に対する推定結果の比較



(a) t=1/ 真値　(b) t=8/ 真値　(c) t=16/ 真値

(d) t=32/ 真値　(e) t=40/ 真値　(f) t=55/ 真値

(g) t=1/ 推定　(h) t=8/ 推定　(i) t=16/ 推定

(j) t=32/ 推定　(k) t=40/ 推定　(l) t=55/ 推定

図 13: 合成画像からの形状と姿勢の推定結果

参考文献

[1] J. J. Kuch and T. S. Huang. “Virtual Gun: A Vision Based Human Computer Interface Using the Human Hand”. In *MVA'94*, pp. 196–199, 1994.

[2] 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.

[3] 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”. 電子情報通信学会論文誌 vol.J79-D-II No.7, pp. 1210–1217, 1996.

[4] Y. Hel-Or and M. Werman. “Recognition and Localization of Articulated Objects”. In *Proc. of Workshop on Motion of Non-Rigid and Aritic-ulated Objects '94*, pp. 116–123. IEEE, 1994.

[5] J. Costeira and T. Kanade. “A Multi-body Factorization Method for Motion Analysis”. In *Carnegy mellon Univ. Tech. Report*, pp. CMU–CS–TR–94–220, 1994.

[6] J. O'Rourke and N. I. Badler. “Model-Based Image Analysis of Human Motion Using Constraint Propagation”. *IEEE Trans. of Pattern Anal. and Machine Intell.,PAMI-2, No.6*, pp. 522–536, 1980.

[7] 片山徹. 応用カルマンフィルタ. 朝倉書店, 1983.

[8] J.Craig 著, 三浦宏文, 下山勲訳. ロボティクス - 機構・力学・制御 -. 共立出版, 1991.

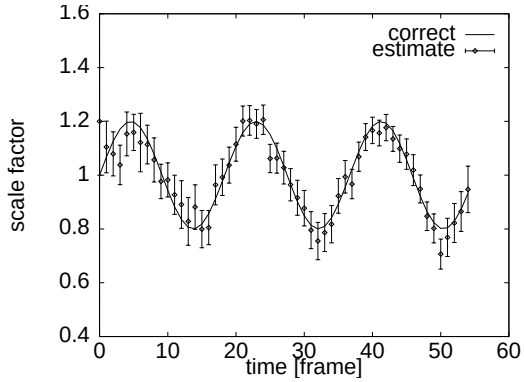


図 15: スケールの推定と分散

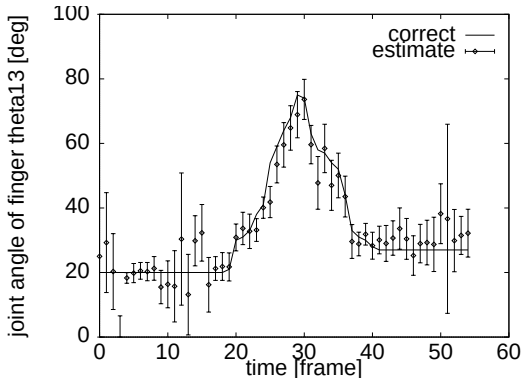


図 16: 関節角 θ_{13} の推定と分散

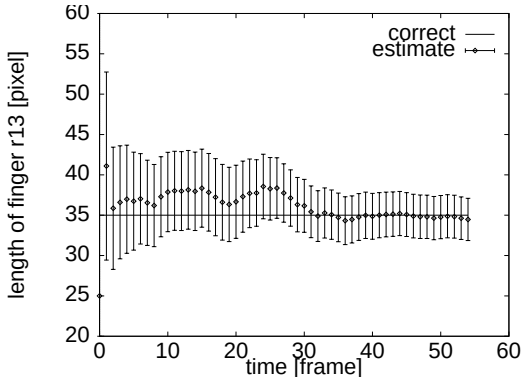


図 17: 指長さ r_{13} の推定と分散 (r_{13}=35 の場合)

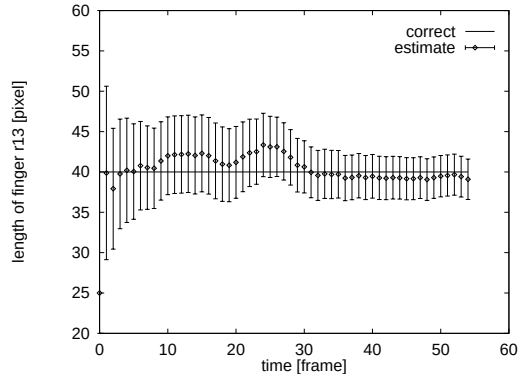


図 18: 指長さ r_{13} の推定と分散 (r_{13}=40 の場合)

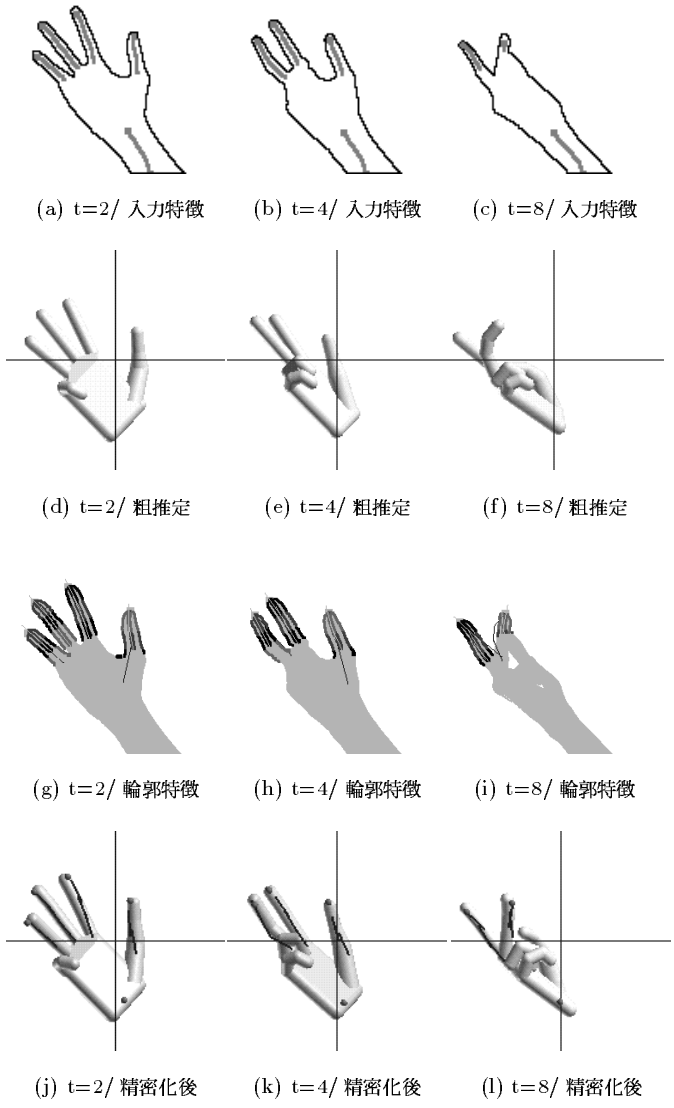


図 19: 実画像からの推定例

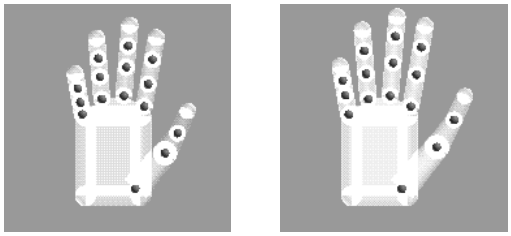


図 20: 実画像からの形状推定結果