

変形可能モデルを用いた3次元多関節物体の形状・姿勢パラメータの推定

米元, 聡
九州大学システム情報科学研究院知能システム学部門

鶴田, 直之
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/2324/5822>

出版情報：電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解. 97 (205), pp.141-148, 1997-07-25. 電子情報通信学会

バージョン：

権利関係：

変形可能モデルを用いた3次元多関節物体の 形状・姿勢パラメータの推定

米元 聡 鶴田 直之 谷口 倫一郎

九州大学大学院システム情報科学研究科

〒816 福岡県春日市春日公園 6-1

092-583-7618

{yonemoto,tsuruta,rin}@limu.is.kyushu-u.ac.jp

あらまし 我々は実空間からのデータ獲得に基づく高リアリティ現実の実現を目指している。本報告では、複雑な3次元多関節構造を有する物体に内在する自己隠蔽や、その動きによる様相の変化に対処可能とするため、複数の変形可能モデルを部位とした対象物モデルを構成し、そのモデル投影像を基にした画像合成による解析 (AbIS) の枠組を提案する。その枠組において、多視点動画像から多関節物体の形状・姿勢パラメータ系列を推定するための方法を述べる。合成データによる実験シミュレーションから本手法の有効性を示す。

キーワード パラメータ推定 モデルフィッティング 多関節物体 自己隠蔽 多視点画像

Shape and Pose Parameter Estimation of 3D Multi-Part Objects

Satoshi YONEMOTO, Naoyuki TSURUTA and Rin-ichiro TANIGICHI

Graduate School of Information Science and Electrical Engineering, Kyushu University

6-1, Kasuga-koen, Kasuga, Fukuoka 816 JAPAN

+81-92-583-7618

{yonemoto,tsuruta,rin}@limu.is.kyushu-u.ac.jp

Abstract

This paper presents an analysis-by-synthesis framework of shape and pose estimation of 3D multi-part objects, whose purpose is to map objects in the real world into virtual environments. In general, complex 3D multi-part objects cause serious self-occlusion and non-rigid motion. To deal with the occlusion among them, we employ both multiple calibrated cameras and time-varying sequences, since there is enough information to estimate the parameters in the sensory data. In our framework, to minimize the error between the selected measurements and the estimated model parameters, we proceed model fitting process based on proper numerical analysis.

key words parameter estimation, model fitting, multi-part objects, self-occlusion, multiple views

1 はじめに

我々は、実空間の観測・計測により得られる情報を実時間でシームレスに仮想空間に融合することにより、リアリティの高い仮想空間を手軽に構築できるシステムの開発を目指している。特に、仮想空間にマッピングすべき実空間(映像)からのデータ獲得を中心に研究を進めているが、具体的な例として、従来の研究ではその複雑さのためにあまり取り上げられなかった動物(人間も含む)を本研究では対象とし、獲得した様々な動物の動き、形状を仮想空間内にマッピングすることを目指している。

「仮想空間内に提示するために必要な情報は事前情報と獲得情報によって決まる」という枠の中で、事前情報と獲得情報をどこで分けるかが重要である。このような視点から、必要以上に情報獲得が難しくならず、しかも手軽に事前情報を与えることができるシステムを作成する。そのためには、仮想空間内に提示するために必要な情報をどのようにモデル化することも重要である。そこで我々は、観測対象を形状・動きなどのシーンパラメータに抽象化してモデル化することで、事前情報と獲得情報の切り分けを行い、すべての獲得情報は事前情報(モデル)によって取捨選択されるものと仮定する。

対象物に関する3次元形状モデルを用いて2次元情報から3次元情報を獲得する方法として以下の2つが挙げられる。

- *Shape From X*に代表される手法によって、2次元情報から3次元情報を陽に計測した上で3次元モデルと3次元上のマッチングを行なう方法。
- 3次元モデルを2次元上に投影し、2次元上でのマッチング結果から3次元情報を推定する方法。

前者の方法では、その3次元情報計測の手法の開発に重点がおかれ、安定して良好な結果を得ることができる手法は存在しない。後者の方法は計算に比較的時間を要するという欠点があるが、2次元情報から3次元情報を復元する問題が逆問題であるのに対し、3次元情報から2次元情報を求める問題はいったんシーンに関するパラメータがわかれば射影幾何(*projection geometry*)を用いて順問題として解くことができるという利点がある。

映像からの3次元形状・動きの解析における最近の主な課題としては以下の2つが挙げられる [6] [7]。

- 人間や動物のような複雑な3次元多関節構造を有する非剛体に対する解析
- 多関節構造体の部位間に存在する隠蔽の問題への対処

特に前者においては、事前情報としてあらかじめシステム作成者が用意した形状モデルを利用する方法 [9]、動きのモデル [12] を仮定して解析する方法などが提案され

ている。従来の方法では、剛体仮説を用いた方法が一般的であるが、自然界においては対象物のほとんどが非剛体であり、そのようなフレキシブルな対象物に対しても形状・動きを解析することが重要となってくる [2][3][5][4]。

一方、後者においては、映像が2次元の投影で、可視の部分が制限されるため、多視点を考慮することによってデータの冗長性を利用する方法が近年盛んに研究されている [10] [6][7][9]。しかし、多視点からの情報を冗長にすることで隠蔽の問題を回避しようとするものがほとんどであり、[8]を除いて厳密な隠蔽の問題を考慮した方法はあまり存在しない。

以上述べた2つの課題に対処する方法として、我々は *Analysis by Image Synthesis*(AbIS)[11]による形状・動きパラメータ推定の枠組(後述)を提案する。第2~5章では、その枠組について述べ、第6章でその実験シミュレーションを示す。

2 AbISの枠組

2.1 AbISの基本原則

*Analysis by Synthesis*とは、ある状態(パラメータ)にあるモデルを観測データと同じ次元に射影し、その射影されたモデルサンプルと観測データの間の誤差を評価関数として最小化することで、モデルの状態を観測データに適應させる、モデル当てはめの原理である。

観測したデータからそのモデルパラメータを推定する問題は、逆問題と呼ばれ、出力(観測データ)をもとに未知の入力(パラメータ)を求める問題として、一般には図1の(b)のように表現される(順問題は(a)に示してある)。

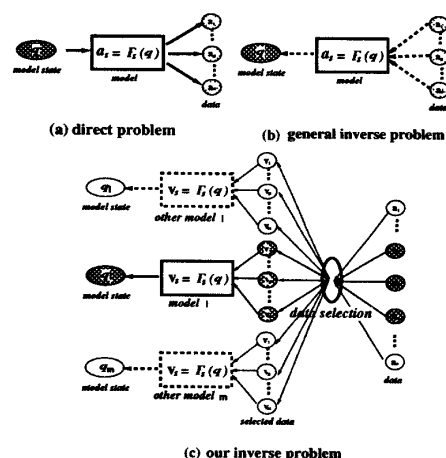


図1: 複数モデルの当てはめ問題

我々の *Analysis-by-Synthesis* (モデル当てはめ)においては、具体的にモデルは3次元パラメータ表現モデル、観測データは2次元画像(特徴点)である。この原理は一般

に、3次元幾何形状モデルの見え方ベースのマッチングと呼ばれる。しかし、本手法では形状モデルとして多関節構造体を扱っているため、複数の部位モデルが当てはめ時に同時に存在し、他の部位モデルが当てはめに影響を及ぼす。そのため他のモデルによって影響を受けることによって当てはめ対象（モデルサンプルとデータの対応ペア）から外れるものを除去する必要があるが、部位モデル間の関係はその構造によりわかっているため、モデル側で適切な当てはめ対象の選択を行なうことができる。そこで我々は、図1の(c)に示すように、事前に適切な当てはめ対象のみを選択する処理を導入したモデル当てはめの原理 (*Analysis-by-Image-Synthesis*) を提案する。

観測データの特徴点集合とした場合、そのモデルの投影像は、対応する可視モデルサンプル点集合で表現される。ここで、可視モデルサンプル点とはモデル投影画像の合成によって生成される「可視」である（すなわち最も視点に近い）モデルサンプル点のことを表す。

2.2 3次元多関節物体追跡システムの概要

AbISの原理を用いて多視点時系列画像から多関節物体の形状・位置・姿勢に関するパラメータ系列を獲得する概要を述べる（図2）。

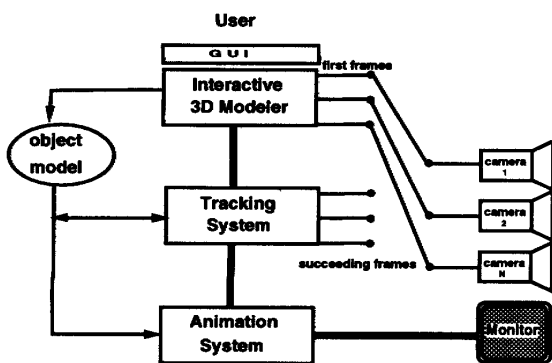


図2: 3次元多関節物体追跡システムの構成

1. 対象物モデルの作成と初期情報の獲得

3次元形状モデラ（後述）を用いてユーザ支援による半自動的な、実画像を利用した概略的なモデル（全体を構成する部位の形状とその構造）の作成と初期モデルパラメータの推定を行なう。

2. 各多視点入力フレームにおける観測データの追跡

本手法では、時系列に渡って処理することを想定しているため、観測データは連続するフレーム間で追跡されている必要がある。本報告では連続2フレーム間での観測データの追跡が適当な特徴点追跡の方法[13]により、達成されうることを仮定する。

3. 多関節構造体のモデルパラメータの推定

4. 3.の結果をもとに各部位のモデルパラメータを更新

5. 次の時刻のフレームに対して2.~4.を繰り返す

上記のように時系列に渡ってパラメータ推定を行なうことで、対象物の形状・位置・姿勢に関するパラメータ系列を獲得する。その獲得したパラメータ系列を、3次元アニメーション部に入力することで、実空間から獲得した対象物データを計算機内の仮想空間において再現する。

3 多関節構造体の形状モデリング

3.1 3次元形状モデラ

様々な対象物に対してシステムを適応させるためには、事前にその各々に対するモデルを用意しておく必要があるが、従来のシステムにおいては、システム作成者のみがモデリングのプロセスに関与でき、システムとしての柔軟性が不足していた。そこで、ユーザがモデリングに関与することができれば、的確なモデルを作成することができ、さらにモデルの初期パラメータをも決定できるという利点がある。また、我々の目標でもある「手軽な事前情報の付与」という観点からも、このモデリングにはユーザの介在が最小であることが望まれるので、負担を軽減するために、GUIベースのインタラクティブな3次元形状モデラの開発を行っている[15]。

3.2 対象物モデルの記述

本手法においては、パラメータ表現された3次元形状モデルであればどのような記述方法でも導入可能である。今回、本報告では数少ないパラメータで表現可能で、変形能力を有する *Deformable SuperQuadratics*（以下、*DSQ*）を採用した。また、*DSQ*にも様々な種類が提案されているが、*tapering* や *bending* といった変形能力を容易に扱える、*Metaxas*らの用いた方法を利用した[4]。

3.2.1 DSQの記述

媒介変数を η, ω とすると、*SQ*上の点 $e(\eta, \omega)$ は以下のように表される。

$$e(\eta, \omega) = \begin{pmatrix} e_1(\eta, \omega) \\ e_2(\eta, \omega) \\ e_3(\eta, \omega) \end{pmatrix} = a \begin{pmatrix} a_1 \cdot C_\eta^{\epsilon_1} \cdot C_\omega^{\epsilon_2} \\ a_2 \cdot C_\eta^{\epsilon_1} \cdot S_\omega^{\epsilon_2} \\ a_3 \cdot S_\eta^{\epsilon_1} \end{pmatrix} \quad (1)$$

ここで、 $-\frac{\pi}{2} \leq \eta \leq \frac{\pi}{2}$, $-\pi \leq \omega < \pi$ で、 $a \geq 0$ はスケール、 $0 \leq a_1, a_2, a_3 \leq 1$ はアスペクト比、 $\epsilon_1, \epsilon_2 \geq 0$ は四辺形度のパラメータであり、 $C_\omega^\epsilon = \text{sign}(\cos \omega) |\cos \omega|^\epsilon$, $S_\omega^\epsilon = \text{sign}(\sin \omega) |\sin \omega|^\epsilon$ である。

*SQ*上の点 e を用いて、*DSQ*上の点 s は、以下のように表される（以下、 (η, ω) は省略）。

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} (\frac{t_1 \epsilon_3}{a a_3} + 1) e_1 + b_1 \cos(\frac{\epsilon_3 + b_2}{a a_3} \pi b_3) \\ (\frac{t_2 \epsilon_3}{a a_3} + 1) e_2 \\ e_3 \end{pmatrix} \quad (2)$$

ここで、 $-1 \leq t_1, t_2 \leq 1$ は x, y 軸方向の *tapering*, b_1 (positive or negative) は *bending* の大きさ, $-1 \leq b_2 \leq 1$ は *bending* が適用される z 軸上の位置, $0 < b_3 \leq 1$ は *bending* の影響範囲を表すパラメータである。

3.2.2 多関節構造体の記述

3次元形状モデル (*DSQ*) を部位とした多関節の木構造で対象物モデルを構成する (図3)。ワールド座標系を

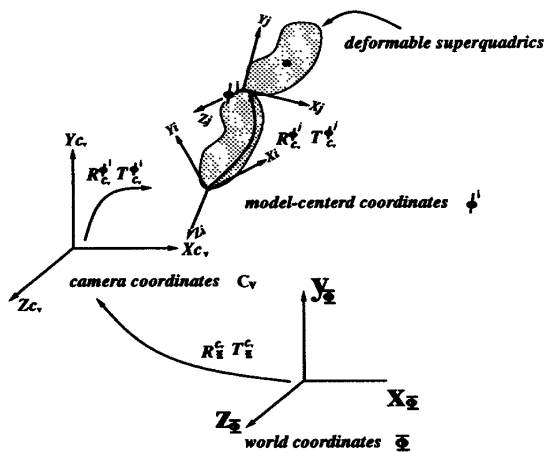


図3: 多関節構造モデルの構造

Φ , 各部位 $part_i$ の部位座標系を各々 ϕ_i とする。

部位座標のワールド座標への変換:

ワールド座標系 Φ における部位 $part_i$ の点の位置 \mathbf{p}^Φ は,

$$\mathbf{p}^\Phi = \mathbf{T}_{\phi_i}^\Phi + \mathbf{R}_{\phi_i}^\Phi \mathbf{p}^{\phi_i} \quad (3)$$

ここで、 $\mathbf{T}_{\phi_i}^\Phi$ は Φ に関する ϕ_i の原点の位置, $\mathbf{R}_{\phi_i}^\Phi$ は Φ に関する \mathbf{p}^{ϕ_i} の方向, \mathbf{p}^{ϕ_i} は ϕ_i に関する部位 $part_i$ 上の点の位置を表す。

上記の定義により、各部位 $part_i$ の形状・位置・姿勢パラメータは以下のように表現される。

$$\mathbf{q}^i = (a, a_1, a_2, a_3, \epsilon_1, \epsilon_2, t_1, t_2, b_1, b_2, b_3, c_1, c_2, c_3, r_1, r_2, r_3)^T \quad (4)$$

ここで、 c_1, c_2, c_3 は並進, r_1, r_2, r_3 は回転を表す位置・姿勢パラメータである。

3.2.3 透視投影幾何

次の規則に従って、各点をワールド座標系からカメラ座標系に変換し、透視投影仮定の画像面に点を投影する。

1) ワールド座標のカメラ座標への変換:

Φ に関する点の位置 \mathbf{p}^Φ を、カメラ座標系 C_v (v は各視点) に関する点の位置 \mathbf{p}^v に変換する。

$$\mathbf{p}^v = (\mathbf{R}^v)^{-1} (\mathbf{p}^\Phi - \mathbf{T}^v) \quad (5)$$

ここで、 \mathbf{R}^v は Φ に関する C_v の方向, \mathbf{T}^v は Φ に関する C_v の原点の位置を表す。

2) カメラ座標の画像座標への投影変換:

透視投影下では、 C_v に関する点 $\mathbf{p}^v = (x^v, y^v, z^v)^T$ は、以下の関係式により画像面上の点 $P^I = (U, V)^T$ に射影される。

$$\begin{aligned} U &= f \frac{x^v}{z^v} \\ V &= f \frac{y^v}{z^v} \end{aligned} \quad (6)$$

ここで、 f は焦点距離である。

4 多関節構造体のパラメータ推定

4.1 初期モデルの獲得

モデラにより作成した対象物モデルを用いて初期フレームにおける位置合わせ (初期パラメータ推定) を行なう方法を述べる (図4)。

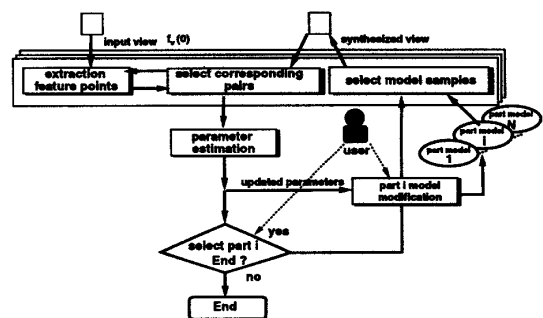


図4: 初期パラメータ推定

各視点 v からの初期フレーム群 $\{f_v(0) | v = 1, \dots, V\}$ (V は視点の総数) について以下の処理を行なう。

Step 1. 半自動の初期フレームでの位置合わせ

モデラで作成した (あるいは作成中の) 各部位モデルを用いて、ユーザの介入により半自動的に、初期フレームでの位置合わせを行ない、初期の位置・姿勢、形状パラメータを推定する。

Step 2. 観測データの選択

各視点の初期フレームにおいてそれぞれ、観測データとしての特徴点を選択する。

$$\{({}^v \tilde{U}_k(0), {}^v \tilde{V}_k(0))^T | k = 1, \dots, {}^v O(0)\}$$

ここで、 ${}^vO(0)$ はフレーム $f_v(0)$ の観測データの総数。

Step 3. モデルサンプルと観測データの初期対応決定

位置合わせの結果をもとに、モデルのサンプル点と観測データ（特徴点）との初期対応を決定する。対応仮説は、各特徴点に対して初期対応ペア間の距離が最短である、部位のモデルサンプル点を対応づけることにより立てる。モデルのサンプル点は媒介変数 (η, ω) を適当な間隔でサンプリングすることにより得られたものである。

各部位 $part_i (i = 1, \dots, N; N$ は部位の総数) の、各視点 v からの初期フレーム $f_v(0) (v = 1, \dots, V)$ について、以下の対応仮説生成の処理を行なう。

各部位、各視点の対応仮説生成

1. 可視モデルサンプル点の選択

各視点 v ごとに、モデル投影像を合成することによって、その視点において可視になるモデルのサンプル点の集合

$$\{({}^vU_{m_j^i}(0), {}^vV_{m_j^i}(0))^T | j = 1, \dots, {}^vM^i(0)\}$$

を求める（詳細は 4.3.2 節）。ここで、 ${}^vM^i(0)$ はサンプル点の総数である。

2. 対応仮説の検証による対応更新

各特徴点 $({}^v\tilde{U}_k(0), {}^v\tilde{V}_k(0))^T$ に対し、部位 $part_i$ の可視モデルサンプル点 $({}^vU_{m_j^i}(0), {}^vV_{m_j^i}(0))^T$ との対応仮説を立て、その仮説が以前の対応部位での仮説よりも高い評価値ならば（あるいは未対応の場合）、その特徴点の対応仮説を更新する。対応のついた特徴点 $({}^v\tilde{U}_k(0), {}^v\tilde{V}_k(0))^T$ を部位 $part_i$ 上のモデルサンプル j の対応特徴点として $({}^vU_j^i(0), {}^vV_j^i(0))^T$ と表す。

全部位に渡って、上記の対応仮説生成を行なった後、各特徴点に対する対応仮説（どの部位上の、どのモデルサンプル点に対応するか）を受理する。

この結果得られる初期フレーム $f_v(0) (v = 1, \dots, V)$ におけるモデルサンプル点と各部位の観測データ点との対応ペアは、

$$\left(({}^vU_{m_j^i}(0), {}^vV_{m_j^i}(0))^T, ({}^vU_j^i(0), {}^vV_j^i(0))^T \right)$$

ここで、 $j \in {}^vP^i(0)$ 。

である。 ${}^vP^i(0)$ は対応のついたモデルサンプルの添字の集合である。

以上求めた初期パラメータおよび観測データとモデルの初期対応は、時系列にわたって推定するための良い初期値を与える。

4.2 構造制約と推定戦略

一般に、多関節構造を有する対象物は、その接続関係（構造）による強い制約を有している。また、部位全体において、各部位ごとのパラメータを一度に推定しようとする、膨大なパラメータ探索空間となり、現実的な解として求めることができない。したがって本手法においては、その構造から来る制約を利用した、トップダウンな準最適解を求めるパラメータ推定戦略をとる。

4.2.1 構造制約

本手法では以下の構造制約を仮定している。

- 自然界における動物の構造は階層的な部位をノードとした木構造で表現できる
- 動物の部位は関節（1点）を中心に互いに接続されている

4.2.2 推定戦略

推定戦略の基本方針

子の姿勢は部位座標系で表現されており、その親の姿勢に依存するため、子の姿勢を求めるには親の姿勢が先に決定している必要がある（ $R_{\phi_i}^*$, $T_{\phi_i}^*$ の計算に必要）。従って、親から子の順にトップダウンに決定する必要があり、上記の方針を満足させる推定方法として、以下のアルゴリズムを考える。

推定戦略

1. $i = \text{root}$ とする。
2. 部位モデル $part_i$ のパラメータを推定する。
3. もし部位 $part_i$ の子供 $c_j^i (j = 1, \dots, \text{childnum})$ が存在するならば、その子についても同様に再帰的に推定する（2. へ）。
4. そうでなければ再帰的推定を終了。

4.3 時系列におけるパラメータ推定

初期フレーム群 $\{f_v(0) | v = 1, \dots, V\}$ での初期対応結果を利用して、フレーム群 $\{f_v(t) | v = 1, \dots, V\} (t = 1, \dots)$ で各部位のパラメータを推定する方法を述べる（図5）。

- 1) 観測データの追跡 特徴点追跡処理により全特徴点を追跡する。
- 2) 推定戦略の順に各部位のパラメータを推定 4.3.2 節。

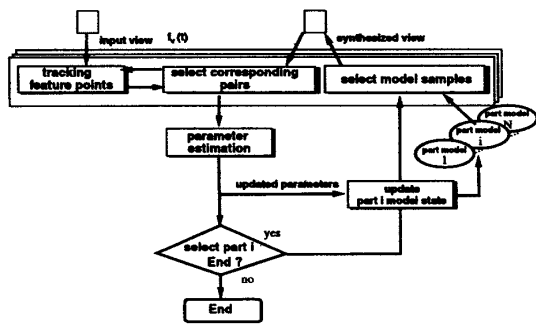


図 5: 時系列におけるパラメータ推定

4.3.1 モデル当てはめ問題

我々の *Analysis-by-Image-Synthesis* において、パラメータ推定はモデルにより適切に選択された観測データへのモデル当てはめ問題に帰着されることはすでに述べた。一般に、モデルとデータの誤差が d_j である時、モデル当てはめ問題の評価関数 E は以下の式で表される。

$$E = \text{metric} \sum_{j \in P} w_j \rho(d_j) \quad (7)$$

本報告では簡単に、 $\rho(x) = x^2$, $w_j = 1 (j \in P)$ とし、*metric* として *min* を用いた (d_j はモデルとデータの値の差)。なお P は事前選択された当てはめ対象の集合を表す¹。

具体的には、モデルと観測データとの間の誤差を以下の評価関数によって定義し、それを最小化することで推定を行なう (s は反復推定の時間を表し、以下、 t は省略する)。

評価関数

$$E_i(s) = \sum_{v=1}^V \sum_{j \in {}^v P^i} \left(({}^v U_{m_j^i}(s) - {}^v U_j^i)^2 + ({}^v V_{m_j^i}(s) - {}^v V_j^i)^2 \right) \quad (8)$$

ここで、 ${}^v P^i$ は、部位 $part_i$ の視点 v における対応ペアの存在するモデルサンプルの添字の集合。

4.3.2 各部位のパラメータ推定

Step 1. モデルサンプルと観測データの対応決定

各部位 $part_i$ の、各視点 v からのフレーム $f_v(0)$ ($v = 1, \dots, V$) について、以下の処理を行なう。

(1-1) 部位 $part_i$ の追跡データの決定 前フレームにおいて部位 $part_i$ 上のモデルサンプル点に対応していた特徴点に対し、特徴点追跡処理で追跡可能であったもの (あるいは新たに出現した特徴点) を求める。また追跡されなかった特徴点に対する対応は破棄する。

¹式 (7) は、例外値を除去するため当てはめ対象を事前選択し、さらに除去できなかった例外値に対してロバスト推定することを意味する。

(1-2) 可視モデルサンプル点の選択 対象物が多関節構造を有する場合、モデルの投影像は以下の2種類の自己隠蔽を生じる可能性がある。

- その部位自身 (の姿勢変化など) によって引き起こされる隠蔽。
- 他の部位によって引き起こされる隠蔽。

従ってこの隠蔽に対処するため、可視となるモデルサンプル点だけを選択する。

具体的なモデルの画像合成手順は以下の通りである。

- *Z-buffer* アルゴリズムを用いて、部位全体のモデルの投影像を作成する。
- その投影像から部位 $part_i$ 上の可視モデルサンプル点を求める。

(1-3) モデルサンプル点と観測データとの対応決定

以下の対応と対応仮説、

1) 前フレーム $f_v(t-1)$ での推定によるモデルと観測データ間の対応

$$\left(({}^v U_{m_j^i}(t-1), {}^v V_{m_j^i}(t-1))^T, ({}^v U_j^i(t-1), {}^v V_j^i(t-1))^T \right)$$

ここで、 $j \in {}^v P^i(t-1)$

2) 観測データの追跡による、前フレームのデータと現在のフレームのデータ間の追跡対応仮説

$$\left(({}^v U_j^i(t-1), {}^v V_j^i(t-1))^T, ({}^v U_j^i(t), {}^v V_j^i(t))^T \right)$$

を時系列に渡る連続推定により得ることができる。

観測データの追跡処理により、フレーム $f_v(t)$ でのデータ点 $({}^v U_j^i(t), {}^v V_j^i(t))^T$ が前フレームにおいて追跡されていることから、以下のようなフレーム $f_v(t)$ における可視モデルサンプル点との対応仮説を立てることができる。

$$\left(({}^v U_{m_j^i}(t), {}^v V_{m_j^i}(t))^T, ({}^v U_j^i(t), {}^v V_j^i(t))^T \right)$$

ここで、 $j \in {}^v P^i(t)$ 。

Step 2. パラメータの反復推定 各視点 v からのフレーム $f_v(t)$ ごとに求めた Step 1. の当てはめ対象すべてに対し、式 (8) を適用する。

以下の手順により、パラメータを反復改良することで推定を行なう。

1. 最急降下の方向に各パラメータ $\hat{\alpha}_k(s)$ を微小調整する ($k \in \{a, \dots, b_3, c_1, \dots, r_3\}$)。

$$\hat{\alpha}_k(s) = \alpha_k(s) - \beta(s) \frac{\partial E_i(s)}{\partial \alpha_k(s)} \quad (9)$$

ここで、 $\beta(s)$ は刻み幅。

2. 位置が更新されたモデルサンプル点と観測データとの対応

$$\left(({}^vU_{mj}^i(s+1), {}^vV_{mj}^i(s+1))^T, ({}^vU_j^i, {}^vV_j^i)^T \right)$$

ここで、 $j \in {}^vP^i; v = 1, \dots, V$.

に対して評価関数 $E_i(s+1)$ を計算し、1.へ(収束していれば終了)。

Step 3. 出現した可視モデルサンプル点への対処

パラメータ推定後、新たな隠蔽によって消滅するモデルサンプル点に対する対応ペアは破棄する。誤差が大きくなったものも追跡誤り(偽対応)として破棄する。また、モデルサンプル点にはモデルベースによる予測により対応する観測データが存在しないモデルサンプル点が登場する。従って、次フレーム $f_v(t+1)$ での処理のために、この新しい可視モデルサンプル点に対応する観測データを初期フレームでの対応の時と同様に(存在すれば)求める。

5 実験結果

提案手法は次の手順で達成される: まず、多視点画像系列を獲得するため実カメラのキャリブレーション及びモデルの合成画像を生成するための仮想カメラの位置合わせを行なう。次に、ユーザが3次元形状モデルを用いて初期モデルの作成及び初期パラメータの推定を行なう。作成した初期モデルを用いて、連続フレームに対して推定していくことで、パラメータ系列を獲得する。

今回、提案手法の基本的な性能を評価するため、合成データ(真値モデル)に対して実験を行なった。3直交配置にした仮想カメラを用いて、モデルに適当なテクスチャ画像を張り付け、レンダリングすることで合成データを作成した。深刻な隠れを有する対象物の例として、2つの関節で連結した3部位から成る腕を想定した。図6は、仮想カメラにより作成した真値モデルの合成画像(入力)を示している。図7は本手法により推定したモデルを入力と同様の視点から合成した画像を示している。図8は真中の部位(2)の推定モデルパラメータと真値モデルパラメータの一部(b_1, r_2)を示している。部位2が深刻な隠蔽の存在下でも、推定・追跡されていることがわかる。図9は、部位2において、新たに出現した対応ペアの数、隠蔽により破棄された対応ペアの数、追跡された対応ペアの数の時系列に渡る推移を示す。このように、モデルベースの予測によれば、ボトムアップな観測データの追跡のみでは不可能な、当てはめ対象の適切な取捨選択を行なうことができる。

6 まとめ

実空間からのデータ獲得を中心に、仮想空間に提示するために必要な情報を1)事前情報として変形可能モデル、2)獲得情報として形状・姿勢などの推定モデルパラメータとすることで、切り分け・モデル化を行なった。複雑な3次元多関節構造を有する非剛体の自己隠蔽の問題に対処することができるのが本手法の最大の特長である。また、真値モデルによる合成データを用いた実験により提案した枠組の手法が有効であることを確認した。

今後の課題としては、多視点実画像系列を実時間で得るための分散マルチカメラシステムの開発、実時間処理(処理時間の低減)のための予測機構の導入、分散協調アルゴリズムとしての推定戦略の実装などが挙げられる。

参考文献

- [1] A.H. Barr, *Global and Local Deformations of Solid Primitives*, Computer Graphics, Vol.18, No.3, pp21-29, 1984.
- [2] A. Pentland and S. Sclaroff, *Closed-Form Solutions for Physically Based Shape Modeling and recognition*, PAMI-13, No.7, 1991.
- [3] A. Pentland and B. Horowitz, *Recovery of Nonrigid Motion and Structure*, PAMI-13, No.7, 1991.
- [4] D. Terzopoulos and D. Metaxas, *Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics*, PAMI-13, No.7, 1991.
- [5] D. Metaxas and D. Terzopoulos, *Shape and Non-rigid Motion Estimation through Physics-Based Synthesis*, PAMI-15, No.6, 1993.
- [6] I.A. Kakadiaris and D. Metaxas, *3D Human Body Model Acquisition from Multiple Views*, ICCV, 1995.
- [7] I.A. Kakadiaris and D. Metaxas, *Model-Based Estimation of 3D Human Motion with Occlusion Based on Active Multi-Viewpoint Selection*, CVPR, 1996.
- [8] J.M. Rehg, T. Kanade, *Model-Based Tracking of Self-Occluding Articulated Objects*, ICCV, pp612-617, 1995.
- [9] D.M.Gavrila, L.S.Davis, *3D model-based tracking of Humans in action: a multiview approach*, CVPR, pp73-80, 1996.
- [10] J.Ohya, F.Kishino, *Human posture estimation from multiple images using genetic algorithm*, 12th ICPR, pp750-753, 1994.
- [11] R. Koch, *Dynamic 3-D Scene Analysis through Synthesis Feedback Control*, PAMI-15, No.6, 1993.
- [12] 石井, 望月, 岸野, 人物合成のためのステレオ画像からの動作認識法, 電子情報通信学会論文誌 D-II, No.8, pp1805-1812, 1993
- [13] J.Shi, C.Tomasi, *Good Features to Track*, CVPR, pp593-600, 1994.
- [14] 金子, 堀内, 本多, 多重分散カルマンフィルタによる3次元推定, 電子情報通信学会論文誌 D-II, No.5, 1996.
- [15] 一木, 他, 多関節非剛体物体のモデル獲得 -画像からのパラメータ推定機能を持った形状モデラ-, 第55回情報処全大(発表予定), 1997.

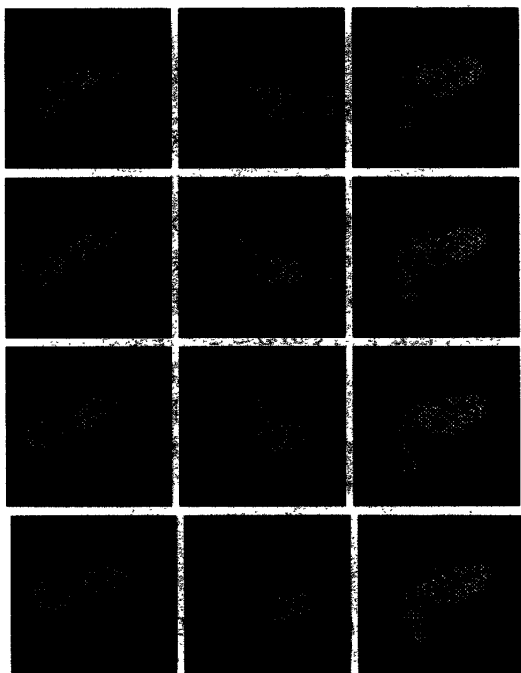


図 6: 入力画像系列 (0,2,4,6): 各行は視点 (前, 左, 上) 方向, 各列は時間方向を表す.

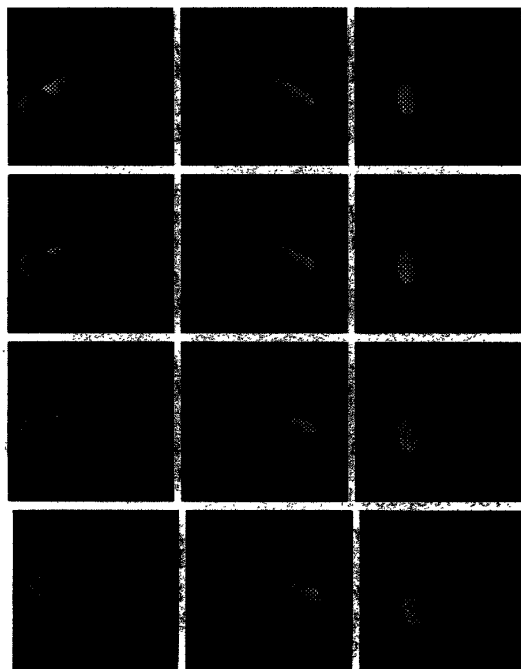


図 7: 推定モデル画像系列 (0,2,4,6): 各行は視点 (前, 左, 上) 方向, 各列は時間方向を表す.

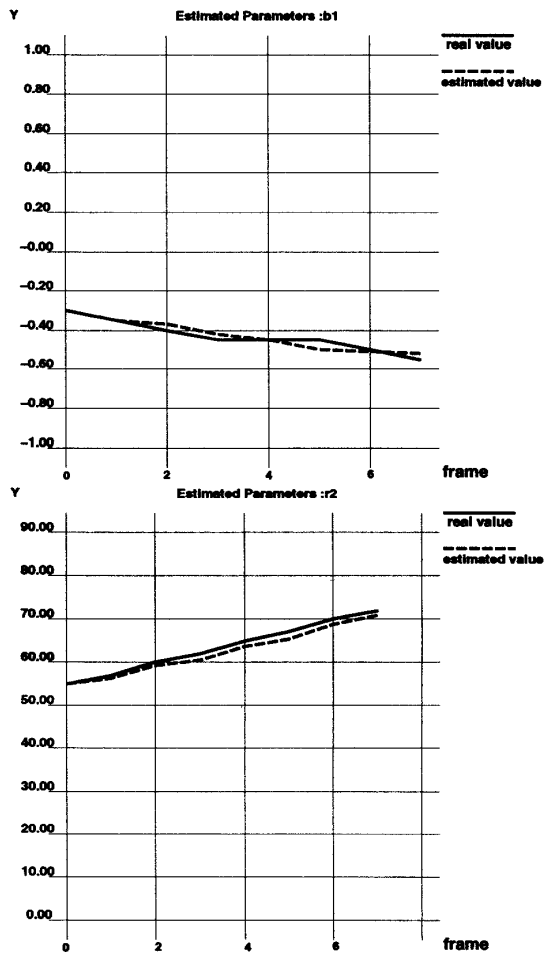


図 8: パラメータ推定結果の真値との比較 (部位 2)

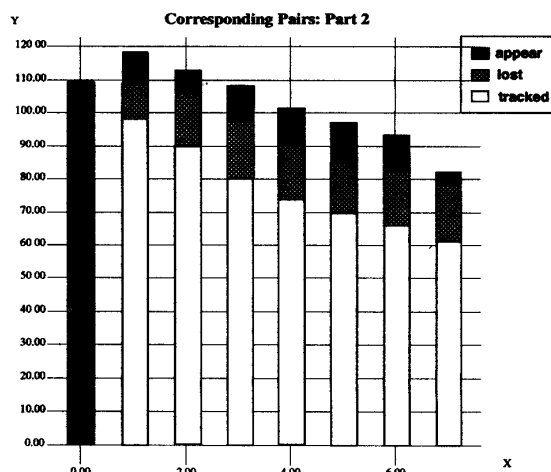


図 9: 対応ペア数 (隠れ・出現・追跡) の推移 (部位 2)