

3次元形状モデルを用いたカメラワークプランニング

西山 学 松山 隆司

京都大学大学院情報学研究科知能情報学専攻

概要 多視点ビデオ映像から対象の3次元形状を復元する手法に、視体積交差法がある。この方法で復元される形状の精度は、対象のシルエットの空間解像度に依存する。本研究では、人物の高精度3次元形状を求めるための方法として、(1) まず人体を複数のパーツからなる多関節物体としてモデル化し、(2) 初期の視体積交差法で得られた全身のボクセルデータに対してモデル当てはめを行い、対象の姿勢を推定した後、(3) 各パーツに対してその形状を精度良く観測できるカメラ群および撮影パラメータを求め、(4) 各カメラを制御してパーツ毎の詳細な映像を撮影し視体積交差法を適用する、という能動的多視点映像撮影方式を考案した。本報告では3の過程をカメラワークプランニングと呼び、そのアルゴリズムとマネキンを対象とした実験による性能評価を行う。

Camerawork-Planning Using 3D Shape Model

Manabu NISHIYAMA and Takashi MATSUYAMA

Department of Intelligence Science and Technology,
Graduate School of Infomatics, Kyoto University

abstract A 3D shape reconstruction from a multi-viewpoint video sequence is computed with volume intersection. Resolution of the shape depends on the resolution of the silhouette taken with cameras. In this report, we present a method for reconstructing the detailed shape of a human body. It is processed as follows: (1) describe the human body as an articulated object, (2) fit the model to the reconstructed volume, estimate the pose of the performer, (3) divide cameras into groups corresponding to a part of articulated model, determine camera parameters in order to reconstruct detailed shape, (4) control each cameras and compute detailed volume of the parts with volume intersection. In this report, we call the third process “camerawork-planning”. We introduce its algorithm and evaluate the performance.

1 はじめに

実世界のシーンを記録するためのデバイスとして、カメラがよく用いられている。特に最近ではデジタルカメラの急速な普及により、簡単に実世界の映像をデジタルデータの形で保存できるようになっている。しかし、カメラによる映像の記録は、3次元の情報を2次元の画像に縮退させてしまう。そこで、より忠実に実世界を記述するため、カメラで撮影された2次元画像から、如何にしてシーンあるいは物体の3次元情報を獲得するかという研究が行われてきた。特に、対象の3次元形状を復元し、実写映像から得られる対象表面の色、テクスチャ情報を、復元した3次元形状に張り合わせることによって、任意の視点

映像を作成する3次元ビデオは、動きのある実世界を記述する新たなメディアとして期待されている。

複数のカメラを用いて撮影された多視点映像から、対象の全周囲3次元形状を獲得する手法として、視体積交差法 [2][3][4][5] がある。これは、まず、個々のカメラに映った対象のシルエット像を3次元空間中に逆投影する。これにより、3次元空間中での対象の存在可能領域を特定することができる。次に、それらの共通領域を求め、これを対象の3次元形状とする。このため、カメラを増やすほど、正確な対象の形状を復元が可能である。しかし、形状復元可能な領域は、全てのカメラの共通視野に限られるので、カメラが常に対象全体のシルエットをとらえていないと、復元された3次元形状の一部が欠けると

いった不都合が生じる。したがって、撮影されるシルエット像の空間解像度には、カメラの性能上の限界に加えて、新たな制限が加わってしまう。これでは対象の形状を細部まで取得するのは難しい。

そこで、対象をいくつかのパーツに分割し、カメラ群をパーツごとのグループに分ける。各グループは担当するパーツだけが共通視野に収まるようにカメラをコントロールすることを考案した。これにより、全体像をとらえていたときよりも空間解像度の高いシルエット画像を取得できるようになり、より詳細な形状復元が可能になる。このとき、“3次元形状取得のためには、どのカメラがどのパーツに注目すればよいか”を判断し、パーツに注目するためのカメラコントロールを行うことを、本論文では“カメラワークプランニング”と呼ぶ。カメラワークプランニングを実現するには以下の問題点を解決する必要がある。

1. カメラ画像からのパーツの認識
2. カメラのグループ分けの方法

本研究では、静的環境下での人体をカメラワークプランニングの対象とする。そして、上記1の問題については、対象の3次元形状モデルを利用して各パーツの3次元位置を求める。一方、2の問題については、カメラごとに、各パーツがどのくらいよく撮影されているか、カメラ群の組合せは形状復元にとって適切か、を決める評価基準を定義し、これをもとにカメラワークプランニングを行う。

2 3次元形状モデルの導入

カメラワークプランニングによって、“どのカメラがどのパーツに注目するか”を決定するためには、初めにパーツの3次元位置を知る必要がある。しかし、対象の全身を撮影した画像を見ただけでは、どの部分がどのパーツに対応するかが分からない。そこで、個々のパーツの形に関する情報、および、各パーツの接続関係の情報を合わせて対象の3次元形状モデルと呼び、復元対象の形状モデルをあらかじめ持っておく。そして、

- Step 1 複数のカメラで対象の全身像を撮影する。
- Step 2 背景差分で対象のシルエットを得る。
- Step 3 視体積交差法を用いて、複数の視点のシルエットから対象の3次元形状データ(ボクセル

データ)を取得する。

- Step 4 モデルをStep 3で得た形状データとマッチングする[1]。これによって形状データの姿勢情報、すなわち、全身像を撮影したときの対象のポーズが分かる。
- Step 5 マッチングの結果とモデルを使って、パーツの3次元位置を得る。

という手順でパーツの位置情報の取得を実現する。以下、これについて詳しく述べる。

2.1 3次元形状モデルの定義

本研究で使用する3次元形状モデルとは、形状復元の対象を分割したパーツについての情報を記述したものである。記述されているのは、大きく分けて、

- パーツの形状、大きさといった個々の属性
- パーツの接続関係、動きの制約といったパーツ間の関係

という2種類の情報である。これらをどのように記述するかについて、本研究で使用する人体の形状モデル(図1)を例に挙げて説明する。



図1: 人体の3次元形状モデル

パーツ個々の属性の定義

前提として、パーツ自体の形状が変形することはない、つまり、剛体であるとする。人体のように動きのある物体については、胸・腰・頭・上腕・下腕・・・のように関節で分けられる部位をパーツとし、パーツ間の関節の回転によって動きを表現する。

パーツの形状の記述には、三角形メッシュデータを用い、全ての面が三角形で構成される多面体によって3次元形状を表現する。

パーツ間の関係の定義

パーツ間の接続関係は木構造で表現する。したがって、枝で結ばれた2つのパーツ間には親子関係が存在するが、これがパーツの運動の従属関係を表しているとする。すなわち、親にあたるパーツが動くと、その子孫のパーツ達も一緒に動く。図2に、今回使用した人体の形状モデルで使われるパーツの接続関係を示す。

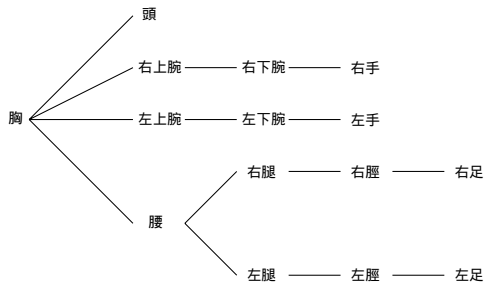


図2: 人体の形状モデルのパーツの接続関係

接続関係を表す木構造において、枝で結ばれた2つのパーツは“関節”で接続されている。この関節を回転させることで子のパーツを動かすことが可能である。このとき、回転させる軸の数を限定したり、回転させる角度の範囲を決めておくことで、モデルの関節の動きを実際の人間のそれに近付けることができる。また、1つの関節の状態を決めるには、その関節の自由度 (= 関節を回転させることのできる軸の数) だけ角度を指定すればよいことになる。したがって、全ての関節の自由度の合計分だけ角度を指定することで、モデルのポーズが決まることになる。

図1の形状モデルは全部で15個のパーツで構成されており、関節の自由度の合計は21である。さらに、基準となる胸部の3次元位置と向きを決定するのに6個のパラメータが必要であるから、合計27個のパラメータを指定することにより、モデルの位置およびポーズが決まる。

2.2 モデルと形状データのマッチング

視体積交差法で得られた形状データと形状モデルを比較し、形状データの姿勢に近づくようにモデルの関節の回転角度を調節する。姿勢をうまく合わせることができれば、モデルからパーツ間の位置関係を把握することができる。

このマッチングの前提条件として、形状モデルの

姿勢の初期値には形状データの姿勢に近いものが与えられていると仮定する。この仮定により、求めるべき姿勢の最適解は初期値から近いところにあると考えることができる。

形状モデルの姿勢が形状データの姿勢にどれだけ近いかを評価する基準としては、“各パーツのメッシュデータ内にある、形状データの3次元点の総数”を用いる。これを N_v とおく。最適解に近い初期値からスタートする場合、モデルの位置および姿勢が形状データに近づくほど、 N_v は大きくなると考えられるから、初期値から出発したときの N_v の極大値を求めて、これをマッチングの解とする。具体的には、以下のようにする。

Step 1 まず接続関係の木構造において、根にあたるパーツの位置を決める。このパーツは位置と向きを変えることができるので、合計で6個のパラメータを決定する。

Step 2 位置を確定させたいパーツを P とする。初期状態において、 P の内部に含まれる形状データの3次元点の数 $N_v(P)$ を計算する。

Step 3 P の位置を決めるパラメータを変動させたときの $N_v(P)$ を計算する。現時点での $N_v(P)$ よりも増加している場合、そのときの状態を新たなパラメータとする。

Step 4 $N_v(P)$ が極大値に達するまでStep 3を繰り返す。

Step 5 パーツ P に繋がっているパーツを新たに P としてStep 2に戻る。全てのパーツの位置が確定すれば終了する。

これによって、形状データに近い姿勢を持ったモデルが得られ、パーツについての情報を持たない形状データから、パーツを認識することが可能になる。形状データとモデルのマッチングを実際に行った実験とその結果については、第4節で紹介する。

3 カメラワークプランニング

前節では、各パーツの位置を取得する手法を提案した。次に、“どのカメラがどのパーツをとらえるのが良いか”という問題を解決する必要がある。この問題は、

$$C = \{C_i \mid 1 \leq i \leq m\} : \text{カメラの集合}$$

$\mathbf{P} = \{P_j | 1 \leq j \leq n\}$: パーツの集合
 $\mathbf{E} = \{(i, j) | C_i \in \mathbf{C}, P_j \in \mathbf{P}\}$
 : カメラ C_i とパーツ P_j を結ぶ枝の集合
 $\mathbf{W} = \{w_{ij} | (i, j) \in \mathbf{E}\}$
 : C_i で P_j を注目するときの評価値の集合

で定義される重み付き二部グラフ $(C, P, E; W)$ (図3) に対して,

$$x_{ij} = \begin{cases} 1 & : C_i \text{ で } P_j \text{ を注目するとき} \\ 0 & : C_i \text{ で } P_j \text{ を注目しないとき} \end{cases}$$

という変数を導入する. このとき, 各パーツに注目するカメラを割り当てる問題は, $I_j (j = 1, 2, \dots, n)$ が与えられているとして, 条件

$$\begin{aligned} \sum_j x_{ij} &\leq 1 \\ \sum_i x_{ij} &\leq I_j \quad \text{の下で} \\ \sum_{i,j} w_{ij} x_{ij} &\longrightarrow \max \end{aligned}$$

とする x_{ij} を決定する最適化問題だとみなせる. x_{ij} に関する2つの条件は, 1つのカメラが注目するパーツは1つだけであること, パーツ P_j を注目するカメラの数は I_j 台までであることを表す. この I_j をパーツ P_j の重要度と呼ぶことにする.

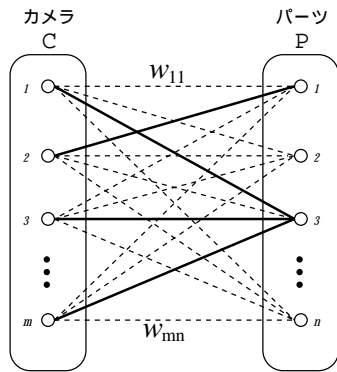


図3: 重み付き二部グラフ $(C, P, E; W)$

この最適化問題を解くには, まず \mathbf{W} を決めなければならない. x_{ij} の値の組合せ, すなわち, カメラ群のパーツごとのグループ分けがパーツの形状を復元するのに都合が良い場合に, 目的関数の値が大きくなるように \mathbf{W} を設定する必要がある. この評価の基準としては次の2つが考えられる.

基準1 それぞれのカメラから目的のパーツがうまく観察できること.

基準2 1つのパーツに対して, 同じような視点を持った複数のカメラが注目していないこと.

基準1はカメラとパーツの組合せに対して評価を行うことを意味する. 一方, 基準2は1つのパーツに注目するカメラ群の組合せを評価することを意味している. したがって, 基準1と基準2の両方でカメラのグループ分けを評価しようとする, w_{ij} はグラフの各枝に対して一定ではなく, x_{ij} の値の組合せによって変化してしまう. このため, 問題は複雑になり, 簡単には解けなくなる.

ここではまず, 基準1の評価として, 前節の手法で得られたモデルの姿勢に関する情報から, パーツの見え方を定量的に判断する方法について述べる. 次に, 基準2のカメラの組合せに対する定量的な評価について述べ, 最後にこれら2つの評価基準を使用した, カメラワークプランニングのアルゴリズムを説明する.

3.1 カメラからのパーツの見え方の評価基準

“パーツをうまくとらえることができる”と判断する根拠として, カメラから見て目的のパーツが他のパーツに隠されていない, という基準を用いるのは, 形状復元に使えるシルエット画像を得るという観点から考えると妥当だといえる. 第2節で述べたように, 形状モデルと3次元形状データのマッチングを行うことにより, モデルの姿勢を3次元形状データに近付けることが可能である. したがって, カメラで撮影したときの人間の姿勢を, モデルを使って仮想空間中に再現することが可能である. そこで, マッチングで得られた姿勢情報を持ったモデルを仮想空間中に置いて, 実際のカメラ群と同じ位置と向きから撮影したときの画像を生成し, 解析することで評価を行う. 具体的には, カメラ C_i からパーツ P_j を見たとき, 他のパーツに隠されずにどれだけ見えているかを示す評価量 $V_{i,j}$ を

$$V_{i,j} = S_{2i,j} / S_{1i,j} \quad (1)$$

で求める. 式(1)において, $S_{1i,j}$ は, 仮想空間中にパーツ P_j だけを置いて, 現実世界の C_i と同一の位置と向きを持った仮想カメラでこれを撮影したとき, 得られる画像に対して P_j が占める面積を表す.

$S_{2i,j}$ は、全てのパーツを仮想空間中に置いた上で撮影したときに得られる画像中で、 P_j が占める面積を表す。したがって、 $V_{i,j}$ はカメラ C_i の画像上で P_j が本来占めるはずの面積のうち、どのくらいが実際に映っているかの割合を表している。

カメラワークプランニングを行うのは、パーツに注目することで空間解像度の高いシルエット像を得たいからである。このため、あるパーツに対して、より高解像度のシルエット像が得られるカメラほど、そのパーツをより良く観察できると評価できる。そこで、パーツがうまく観察できるかの評価基準にカメラとパーツの距離も入れることにする。

モデルによってパーツの位置は分かるので、カメラとパーツの間の距離の計算にもモデルを利用する。カメラ C_i の中心からパーツ P_j の形状を表すメッシュデータの頂点群の重心（以下これをパーツの重心と呼ぶ）までの距離を $D_{i,j}$ とする。この $D_{i,j}$ をパーツがうまく観察できるかの評価に組み込む。

これまで述べた、カメラ C_i でパーツ P_j に注目するときの 2 つの基準 $V_{i,j}$ および $D_{i,j}$ を使って、カメラからのパーツの見え方に対する最終的な評価値 $A(C_i, P_j)$ を次のように定義する。

$$A(C_i, P_j) = \alpha f(V_{i,j}) + \beta g(D_{i,j}) \quad (2)$$

式 (2) において、 $f(V)$ 、 $g(D)$ はともに 0 から 1 の値をとる関数で、 α 、 β は、それぞれの値を重み付けするための定数である。今回行った実験では、 $\alpha = 2.0$ 、 $\beta = 1.5$ とした。関数 f は $V_{i,j}$ に対して単調非減少にしたいので、

$$f(V_{i,j}) = \begin{cases} 0 & V_{i,j} < v_1 \\ \frac{1}{v_2-v_1}V_{i,j} - \frac{v_1}{v_2-v_1} & v_1 \leq V_{i,j} < v_2 \\ 1 & V_{i,j} \geq v_2 \end{cases} \quad (3)$$

とする。 v_1 と v_2 は定数である。実験では、 $v_1 = 0.4$ 、 $v_2 = 0.8$ とした。また、関数 g については、パーツがカメラから近いほど高解像度のシルエットが得られると期待できるので、

$$g(D_{i,j}) = D_{i\min}^2 / D_{i,j}^2 \quad (4)$$

とする。式 (4) で、 $D_{i\min}$ は、カメラ C_i の中心から最も近い重心を持つパーツまでの距離を表す。

3.2 カメラ群の組み合わせの評価基準

視体積交差法では、複数の視点からのシルエット像を手がかりとして対象の存在可能領域を限定し、

対象の 3 次元形状を復元する。このため、同じ台数のカメラを用いても、カメラと対象の位置関係によっては存在可能領域の限定の度合いが異なり、得られる形状データの質が変わる。特に、ほとんど同じ視点からのシルエット像群を使用しても対象の存在可能領域はあまり限定されず、使ったカメラの台数の割には非常に効率の悪い結果になってしまう。

そこで、同じパーツに注目する k 台のカメラ群 $C_g \subset C$ に対して、視体積交差法でうまく形状が復元できるかどうかの定量的評価を行う。評価基準には、カメラからパーツの重心へのベクトルのばらつきを用いる。具体的には、このベクトルと同じ向きを持った単位ベクトルを視線ベクトルと呼び、その x 、 y 、 z 成分のそれぞれの分散の和を使うことにする。したがって、パーツ P_j に注目するカメラ群 C_g の組合せに対する評価値 $A_g(C_g, P_j)$ を次のように定義する。

$$A_g(C_g, P_j) = V_x + V_y + V_z \quad (5)$$

式 (5) において、 V_x 、 V_y 、 V_z はそれぞれ、カメラ群 C_g に属するカメラの視線ベクトルの x 、 y 、 z 成分の分散を表す。

3.3 カメラワークプランニングのアルゴリズム

これまで述べた 2 つの評価基準を用いて、各パーツに注目するカメラ群を決定する。前述したように、この問題の最適解を求めるのは難しい。そこで、最適解を求めるのではなく、2 つの評価関数 $A(C_i, P_j)$ と $A_g(C_g, P_j)$ の両方ある程度大きくすることを目標に、パーツに注目するカメラ群を決めていくことにする。

以下で述べるアルゴリズムでは、カメラの台数が m 台、パーツの数が n 個、パーツ P_j の重要度を I_j とし、重要度が k 番目に高いパーツを P_{Ik} とする。また、

$$\sum_{j=1}^n I_j = m$$

の関係が成り立つものとする。つまり、全てのカメラはそれぞれ、どれか 1 つのパーツに注目している、一方、それぞれのパーツに注目するカメラの数は、ちょうどそのパーツの重要度に一致する。

Step 1 カメラ C_i ($1 \leq i \leq m$) について、パーツ P_j ($1 \leq j \leq n$) に対する見え方の評価値とそ

のパーツの重要度の積 $A(C_i, P_j) \cdot I_j$ を求める．
これにより $m \times n$ のサイズの評価値の表 T が
できる．

Step 2 表 T に対し，カメラごとに $A(C_i, P_j) \cdot I_j$
が降順に並ぶように，パーツ P_1, P_2, \dots, P_n を
ソートする．各カメラはソートの結果先頭に来
たパーツに注目しようとする．しかし，各パー
ツに割り当てられるカメラの数は決まっている
ため，これだけでカメラが注目するパーツが決
定するわけではない．

Step 3 $q := 1$ とする．

Step 4 パーツ P_{I_q} に注目しようとするカメラの数
を s とする．

Step 5 s と P_{I_q} の重要度 I_{I_q} (= 注目するカメラの
数) の大小関係に応じて場合分けする．

i) $s < I_{I_q}$ のとき
 s 台のカメラ全てを P_{I_q} に注目させる．こ
れでも $(I_{I_q} - s)$ 台足りないので 残りはStep
6 で決める．

ii) $s = I_{I_q}$ のとき
 s 台のカメラ全てを P_{I_q} に注目させる．こ
のパーツに注目するカメラは全て決まる．

iii) $s > I_{I_q}$ のとき
 s 台のカメラから I_{I_q} 台を選ぶ全ての組合
せについて，カメラ群 C_g の組合せに対す
る評価値 $A_g(C_g, P_{I_q})$ を計算する．この値
が最大となるカメラ群を P_{I_q} に注目するさ
せる．残った $(s - I_{I_q})$ 台のカメラは，Step
2 のソートの結果が次に大きなパーツに注
目しようとする．

Step 6 $P_{I_1}, P_{I_2}, \dots, P_{I_{(q-1)}}$ のうち，注目するカメ
ラがまだ重要度に満たないパーツについて，注
目しようとするカメラが新たに現れていないか
を確認する．もしあれば，そのカメラを割り当
てる．

Step 7 $q := q + 1$ とし， $q > n$ となるまでStep 4
からStep 6 の処理を繰り返す．

これにより，全てのカメラがどのパーツに注目す
るかが決まる．次に，決定したパーツに注目させ
るため，カメラのパン，チルトの角度で視線方向を決
め，ズームを掛ける必要があるが，これは形状モデ

ルを使って以下のように行う．

Step 1 モデルを使って注目するパーツの重心を求
め，カメラの中心からこの重心へのベクトルを
新たな視線方向にする．(この視線方向から，パ
ン，チルト角が求められる．)

Step 2 新たな視線方向でパーツを見たとき，どのく
らいの大きさになるのかを，パーツを仮想空間
中に置いて確認する．この結果を用いて，パー
ツ全体がカメラ画像内に収まるようなズームの
値を決定する．

4 実験

4.1 実験環境

第 2 節および第 3 節で提案した手法を用いて，マ
ネキンに対してカメラワークプランニングを行った．
実験環境は以下の通りである．

カメラ 視点固定型パン・チルト・ズームカメラ
SONY EVI-G20 \times 9 台
計算機 CPU Pentium III 600MHz \times 2, Main Mem-
ory 512MB \times 9 台

カメラと計算機は 1 対 1 で接続されていて，計算
機からパン，チルト角による視線制御およびズーム
の調整が可能である．計算機は互いにネットワー
クを介して接続されている．9 台のカメラの配置は図
4 のようになっている．

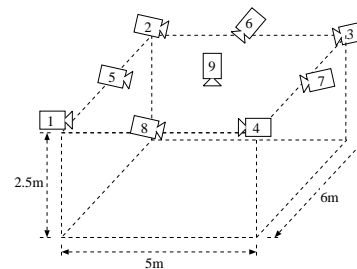


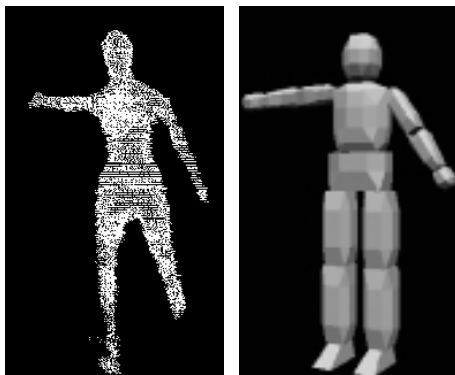
図 4: カメラ群の配置

4.2 モデルと形状データのマッチング実験

マネキンをカメラ 9 の下に置き，9 台のカメラか
らの全身のシルエットを使って並列視体積交差法 [6]
による 3 次元形状復元を行った．このとき得られた
ボクセルデータは図 5(a) のようになった．

このデータと、図 1 の形状モデルをマッチングする。モデルの初期値には図 5(b) の姿勢を与えた。

このマッチングの結果は、図 5(c) のようになった。手や足の先が形状データから若干ずれているものの、各パーツの位置を求めるには十分な結果が得られた。なお、マッチングには 9 台の計算機のうち 1 台を使用した。計算に要した時間は約 15 秒であった。



(a) ボクセルデータ

(b) 初期値



(c) 結果

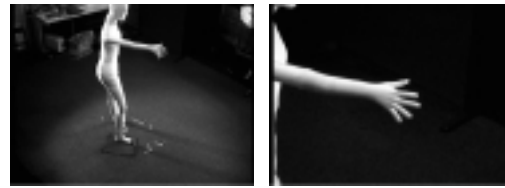
図 5: 形状モデルと形状データのマッチング

4.3 カメラワークプランニング実験

前節で取得したパーツの位置情報を元に、カメラワークプランニングを行った。パーツの重要度は右手が 6、頭が 2、左手が 1 とし、挙げた右手に注目するカメラが多くなるようにした。

この結果、図 4 のカメラ 1, 2, 3, 4, 6, 8 が右手に、カメラ 5, 7 が頭に、カメラ 9 が左手に注目することになり、それぞれが担当するパーツを注目する

ためのパン、チルト角、およびズームの値が得られた。図 6(a) と図 6(b) はそれぞれ、カメラワークプランニング前と後にカメラ 3 で撮影した画像である。カメラワークプランニング後の画像は全て、カメラの最大値までズームを掛けたものになっていた。



(a) プランニング前

(b) プランニング後

図 6: カメラワークプランニング前後のカメラ 3 の画像

4.4 決定されたカメラワークによる 3 次元形状復元

カメラワークプランニングで得られた画像から、従来よりも詳細な形状の復元が可能であることを示すために、右手を撮影した 6 台のカメラ画像に視体積交差法を適用して、3 次元形状の復元を行う。

まず、図 5(a) の全身形状データの右手部分を図 7(a) に示す。全身シルエット像からの視体積交差法では空間解像度の関係上、1cm ボクセルの精度までしか形状を復元することができなかった。これに対し、カメラワークプランニングの結果得られた画像を用いて 2.5mm 間隔で形状復元を行った結果を図 7(b) に示す。図 7(a) と図 7(b) は、形状データを同じ視点から見たものであるが、カメラワークプランニングを使って得られた形状データでは、薬指、小指は少し欠けているものの、指の形状までかなり詳細に復元できている。この実験により、カメラワークプランニング後に撮影した画像は、全身を撮影した画像よりも空間解像度が高いため、これまでよりも詳細な形状の復元が可能であることがわかった。

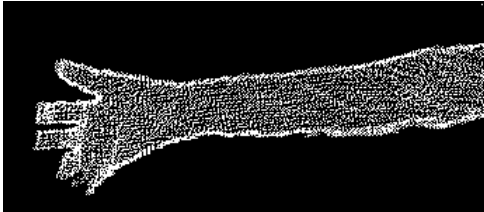
5 結論と今後の展開

本研究では、視体積交差法で細部の形状まで復元するためのアイデアとして、対象をパーツという単位に分割し、それぞれに注目することで空間解像度の高いシルエット像を取得することを考えた。

まず、カメラワークプランニングを、複数のカメラのうち、どのカメラがどのパーツを注目するのかを決定し、そのパーツに注目するためのカメラコン



(a) プランニング前



(b) プランニング後

図 7: マネキンの右手部分のボクセルデータ

トロールを行う問題であると定義した。次に、この実現に関わる問題に対し、対象の 3 次元形状モデルを使用した解決法を提案した。

そして、マネキンを対象として、カメラワークプランニングを適用し、得られたカメラワークで撮影した画像を使って、視体積交差法による 3 次元形状復元実験を行った。その結果、これまでよりも詳細な形状が取得できることを示した。

本研究の結果をふまえ、今後考えていくべき問題としては次のようなものがある。

カメラ群のグループ分けの評価の改善: 本研究で提案したカメラ群のグループ分けは、視体積交差法による形状復元で一般的に成り立つ事柄に基づいていて、形状復元の対象そのものの特徴を考慮に入れていない。例を挙げると、実験で行ったように指を開いた手の形状を復元する場合、5 本の指がきちんと撮影されている画像が取得できなければ、指の形状まで復元するのは困難である。パーツをどの方向から撮影するのが形状復元にとって良いのかを表す情報を、形状モデルに付加する必要があると考えられる。

動的環境下でのカメラワークプランニング: 本研究では静的環境下でのカメラワークプランニング

について述べてきたが、動きのある物体の形状を細部まで復元するには、状況に応じた動的なカメラワークプランニングが必要になってくる。動的なプランニングには、本論文で述べた処理を実時間で行う必要があることに加えて、カメラのコントロールなどに生じる遅延を考慮して、パーツの動きを事前に予測した上でカメラワークプランニングを行わなければならない。

なお、本研究は科学研究費補助金 基盤研究 A 13308017 の補助を受けて行った。

参考文献

- [1] A. Bottino and A. Laurentini. A silhouette based technique for the reconstruction of human movement. *Computer Vision and Image Understanding No. 83*, pp. 79–95, 2001.
- [2] A. Laurentini. How far 3d shapes can be understood from 2d silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17(2), pp. 188–195, 1995.
- [3] B. G. Baumgart. Geometric modeling for computer vision. *Technical Report AIM-249, Artificial Intelligence Laboratory*, 1974.
- [4] P. Srivasan, P. Liang, and S. Hackwood. Computational geometric methods in volumetric intersections for 3d reconstruction. *Pattern Recognition*, Vol. 23(8), pp. 843–857, 1990.
- [5] W. N. Martin and J.K. Aggarwal. Volumetric description of objects from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 5(2), pp. 150–158, 1987.
- [6] ウ小軍, 和田俊和, 東海彰吾, 松山隆司. 平面透視投影を用いた並列視体積交差法. *情報処理学会論文誌*, Vol. 42, No. SIG6(CVIM2), pp. 33–43, 2000.