

## 視体積交差におけるカメラ位置ずれが 復元形状に及ぼす影響

### Influence of Camera Position Displacement on Reconstructed Shapes in Volume Intersection

高橋 将<sup>†</sup> 東海林 健二<sup>†</sup> 外山 史<sup>†</sup> 宮道 壽一<sup>†</sup>  
Masaru Takahashi Kenji Shoji Fubito Toyama Juichi Miyamichi

#### 1 はじめに

画像からの3次元形状復元は、服飾や健康・医療の分野や3Dコンテンツ作成などへの応用があり、盛んに研究が行われている。しかし、人のように形状が変化していく対象物体の3次元形状復元には特殊な環境や装置が必要となり、手軽に利用できるものではない。画像からの3次元形状復元が手軽に利用できれば、その応用分野はさらに広がると予想される。そのため、撮影段階での準備が容易な手法が求められる。3次元形状復元手法の一つとして視体積交差法が知られている。視体積交差法とは、図1に示すように、複数のカメラによって復元対象物体を多方向から撮影したシルエット画像を用い、視点となるそれぞれのカメラ位置を頂点としシルエットを断面とする錐体(視体積)を3次元空間に逆投影し、それらの共通部分として3次元形状を復元するものである。この手法は原理が簡単であり高速化が容易なため、広く用いられている<sup>[1]~[3]</sup>。

視体積交差法を用いて3次元形状を正しく復元するためには、カメラパラメータが正確でなければならない。カメラパラメータは、焦点距離やレンズの歪み係数、画像の歪み中心といった内部パラメータと、カメラの空間位置と姿勢からなる外部パラメータに分けられる。このうち内部パラメータは単焦点レンズを用いる場合、値がほぼ変化しない。一方、外部パラメータはカメラを設置する度に異なる値となる。外部パラメータは、撮影空間に設置したチェス盤のような基準物体を撮影・処理することで得られるが、撮影対象と同程度のサイズの基準物体が必要となり、その準備と設置は簡単とは言えない。

基準物体を使わず外部パラメータを得ることを考えると、外部パラメータのうち空間位置は比較的容易に得ることができる。たとえば、床や地面の既存のマーキングや、床や地面に適当な半径で円を描いて円周上等分割した点を基準にして指定した高さにカメラを設置すれば、それらのカメラの空間位置が決まる。一方、外部パラメータのうち姿勢は、ジャイロセンサー等を用いれば計測可能であるが、一般的な機器ではないため、利用は難しい。そこで我々は、外部パラメータのうち比較的容易に得ることができる空間位置を既知とし、視体積交差結果を元のシルエットにより評価して姿勢調整を繰返す方法を提案している<sup>[4]</sup>。これは、画像の特徴点を使う代わりに、元のシルエットと視体積交差結果のシルエットの差分を利用する一種のバンドル調整である。本稿では、カメラの空間位置を既知としカメラの姿勢推定を行い形状復元を行う提案手法において、カメラの空間位置に含まれる誤差がどの程度復元結果に影響がどうかをシミュレーション実験により検討し、提案手法がカメラの位置ずれに比較的頑健であることを示す。

#### 2 カメラ姿勢推定法

カメラの空間位置を既知とし、カメラの初期姿勢を与え、視体積交差を行って得られる復元形状を元のシルエットにより評価して姿勢調整を繰返すという提案手法<sup>[4]</sup>を説明する。

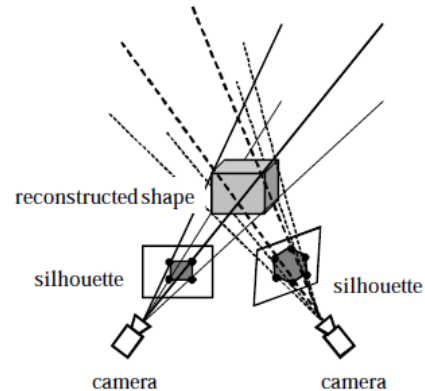


図1: 視体積交差の概念図

#### 2.1 シルエットの無効領域

視体積交差では、図1に示すように、それぞれのカメラ位置からカメラレンズの光軸方向にシルエットを逆投影することによって造られる錐体の交差部分として3次元形状が計算される。対象物体が存在する空間の範囲を定めた場合、カメラの姿勢が正しくないときは、復元された3次元形状は真の形状より小さくなる傾向がある。このとき、復元された3次元形状を各カメラの画像平面に再投影すると、再投影シルエットは元のシルエットの内部に存在し、元のシルエットをはみ出すことはない。図2に示すように2つのシルエットの差の領域は視体積交差結果の構成に寄与しない領域となる。この領域を無効領域と呼ぶ。シルエットと視体積交差結果を表す空間が連続という理想的な場合は、カメラの空間位置と姿勢が正確ならばシルエットの無効領域の総面積は0である。一般にその逆は真ではないが、対象物体が球のような特殊な形状でない限りは、シルエットの無効領域の総面積が小さいほど、より正確なカメラ姿勢であると考えられる。

#### 2.2 カメラ姿勢の最適化

シルエットの無効領域の総面積を評価値として、これを最小化するようにカメラ姿勢を最適化する手法を説明する。本研究で採用した視体積交差は文献[4]のボクセル空間でのラン表現に基づく手法である。ボクセルの集合で表された3次元形状を画像平面に再投影する計算は画素単位で行った。従って、評価値であるシルエットの無効領域の総面積は整数値となる。そのため、文献[4]では勾配や微分を用いない最適化手法の一つである滑降シンプレックス法(downhill simplex method: DSM)を利用している。しかし、DSMは大域的最適解を探索する手法ではないため、局所解に陥る可能性がある。そこで本研究のシミュレーション実験では、DSMに焼きなまし法(simulated annealing: SA)の考え方を取り入れた最適化手法<sup>[5]</sup>を用いる。以下、これをDSMSAと呼ぶ。

<sup>†</sup>宇都宮大学大学院工学研究科

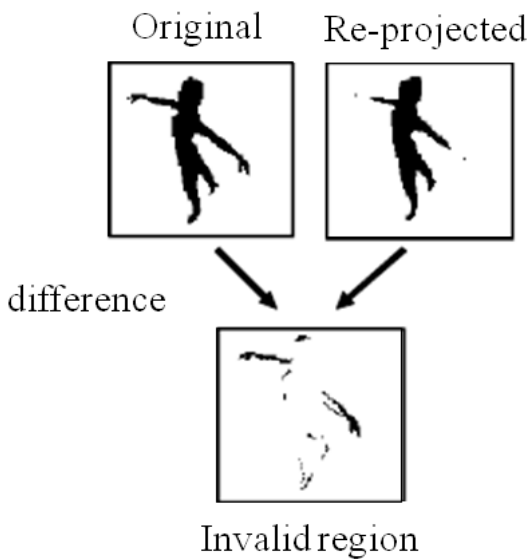


図 2: 無効領域の例

### 3 シミュレーション実験

我々は文献 [4] で、カメラの空間位置を既知とし、カメラの初期姿勢を与え、視体積交差を行って得られる復元形状をシルエットの無効領域の総面積で評価して姿勢を最適化するという手法を提案し、シミュレーション画像と実画像を用いた実験でその動作原理を確認した。しかし、既知であるとして与えるカメラの空間位置には必ず誤差が含まれているため、提案手法の有効性を示すためにはカメラ空間位置の誤差がどの程度復元結果に影響がでるかを明らかにする必要がある。本節では、カメラの空間位置に誤差を与えたときの復元形状が元の 3 次元物体形状とどの程度異なるかを調べたシミュレーション実験について報告する。

使用した対象物体とカメラの配置等は以下の通りである。図 3 に示すように、地面を  $x-z$  平面とし、原点の上に 3 次元対象物体として 3 次元人体形状作成支援を行うソフトウェア POSER で作成した任意姿勢の人体形状物体を置いた。人体形状物体のサイズは高さ約 150 である。このシミュレーション空間での 1 の長さは現実の空間で 1cm に相当する。9 台のカメラ ( $C_0 \sim C_8$ ) を  $y$  軸を中心とする半径 400 の円柱の側面上に円周上等間隔の空間位置に設置した。人がカメラを持って撮影することを想定しているため各カメラの高さは  $y = 135 \sim 160$  の間で、それぞれ異なる値を与えた。すべてのカメラは点  $(0, 100, 0)$  を向くような姿勢とした。すべてのカメラは、撮影画像サイズ  $640 \times 480$  画素、焦点距離 700 画素とした。以上のような設定で、9 台のカメラそれぞれからシーン画像を計算により得た。図 4 にそれらのシーン画像を示す。

シミュレーション実験では、図 4 に示すシルエットを用い、既知として与えるカメラの空間位置に誤差を加える一方、カメラの初期姿勢は正解姿勢を与え、DSMSA でカメラ姿勢 (yaw, pitch, roll) の最適化を行う。その結果、カメラの空間位置のずれをキャンセルする方向に、カメラ姿勢が初期姿勢 (正解姿勢) からずれ、結果として、視体積交差結果の復元形状は元の 3 次元形状から小さなずれに収まることが期待できる。

DSMSA のパラメータ設定として、初期温度は 1000 度、その更新係数は 0.9 とした。すなわち、温度が 1000 度、900 度、810 度、729 度のように減少していく。各温度でのシンプレックスの最大評価回数は 10 とし、終了条件は評価値の変化の割

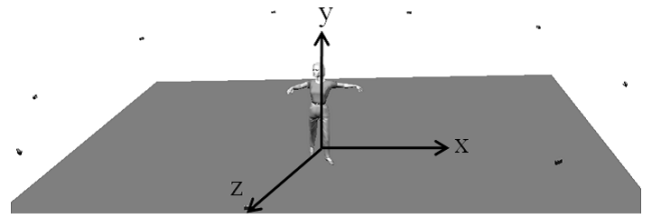


図 3: VRML 空間

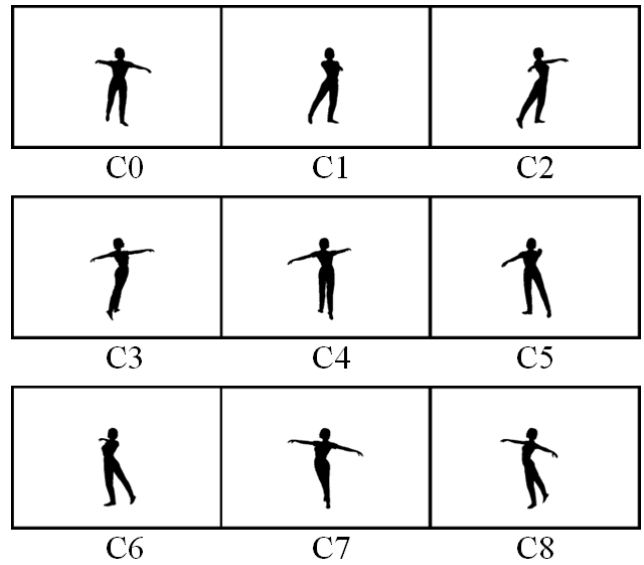


図 4: シルエット画像

合が 0.001 以下または温度更新回数が 1000 回行われたときとした。これらの値は予備の実験により得たものである。

以下、既知として与えるカメラの空間位置に加えた誤差と、DSMSA で最適化して得られた復元形状の元の 3 次元形状からのずれの評価方法、そして、実験結果を述べる。

#### 3.1 既知として与えるカメラ位置のずれ

既知として与えるカメラの空間位置は、真の位置から正規乱数に従い円柱の側面上で移動させることにより誤差を加える。具体的には、平均 0、標準偏差  $\sigma$  で発生させた正規乱数を円周方向の移動量  $e_c$  と  $y$  軸方向の移動量  $e_y$  とし、カメラの真の位置を  $(x_i, y_i, z_i)$ 、円柱の半径を  $r (= 400)$  としたとき、移動後のカメラの位置  $(x'_i, y'_i, z'_i)$  を

$$x'_i = \cos(e_c/r) \times x_i + \sin(e_c/r) \times z_i$$

$$y'_i = y_i + e_y$$

$$z'_i = -\sin(e_c/r) \times x_i + \cos(e_c/r) \times z_i$$

として誤差を加えた。

#### 3.2 復元形状の評価

元の 3 次元形状を基準として、DSMSA で最適化して得られた復元形状のずれを評価する方法について述べる。

元の 3 次元形状はサーフェースモデルで表されている一方、視体積交差により復元される形状はボクセル表現であるため、両者を直接に比較することはできない。そこで、本研究では、原点に対象物体を置き、中心が原点にある正 20 面体の各面の

中心に配置したカメラによる視体積交差で出来たボクセル表現による 3 次元形状を基準形状とした。この基準形状とカメラ姿勢の最適化によってできた復元形状とのずれの大きさをボクセル単位での両者の排他的論理和で残った“1”ボクセルの個数で評価する。このとき、復元形状はカメラ位置のずれにより形状だけでなく位置と姿勢も基準形状からずれている。そこで、基準形状と復元形状のずれの大きさが最小になるように基準形状に並行移動と回転を加え、最小化された基準形状と復元形状のずれの大きさを評価値とする。両者の形状間のずれの大きさの最小化にも 2.2 と同様、DSMSA 用いた。

### 3.3 実験結果と考察

図 5 に示す人体形状 arabesque を対象物体として、カメラに位置ずれを与える正規乱数の標準偏差  $\sigma$  を 0.0 から 25.0 まで 1.0 刻みで変化させ、各  $\sigma$  でカメラに位置ずれを与えてカメラ姿勢最適化と形状復元を行う処理を 30 回実施した。この対象物体の基準形状を図 6 に示す。

初期カメラ姿勢での復元形状の基準形状からのずれのボクセル数と、最適化されたカメラ姿勢での復元形状の基準形状からのずれのボクセル数について、30 回の平均値とカメラに位置ずれを与える正規乱数の標準偏差  $\sigma$  の関係を図 7 に示す。図 7 では、前者を initial、後者を optimized と表している。この実験で用いた基準形状の体積は約 48,000 ボクセルである。図 7 で、復元形状の基準形状からのずれのボクセル数が  $\sigma = 20 \sim 25$  で initial も optimized も約 48,000 ボクセルとなっているのは、初期カメラ姿勢での復元形状も、カメラ姿勢最適化後の復元形状もボクセル数が 0 であることを意味している。

初期カメラ姿勢での復元形状とカメラ姿勢最適化後の復元形状の評価値の差が一番の大きい  $\sigma = 8$  のときの両者の復元形状を図 8 に示す。また図 5 の人体形状 arabesque の他に 2 つの人体形状 walk と stand-up についても同様の実験を行った。それぞれの基準形状を図 9, 10 に、実験結果を図 11, 12 に示す。これらの基準形状の体積はそれぞれ約 34700 と約 35500 である。図 7, 11, 12 のグラフから  $\sigma = 8$ 、あるいは  $\sigma = 9$  程度まで初期視体積と最適化後視体積の評価の差が開いていることが分かる。

基準形状と復元形状を目視で比較した結果 3 種類のいずれの形状でも  $\sigma = 5$  以下ならばほとんど差がないと判断できた。これより、カメラ位置が平均的に  $\sigma = 5$  程度ずれていても復元形状への影響が少ないことが分かった。

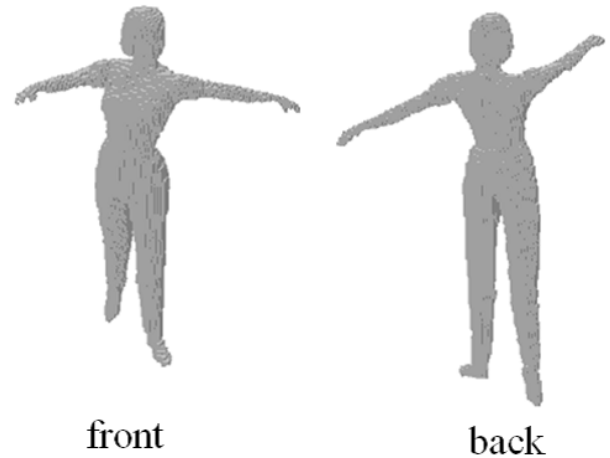


図 6: arabesque の基準形状

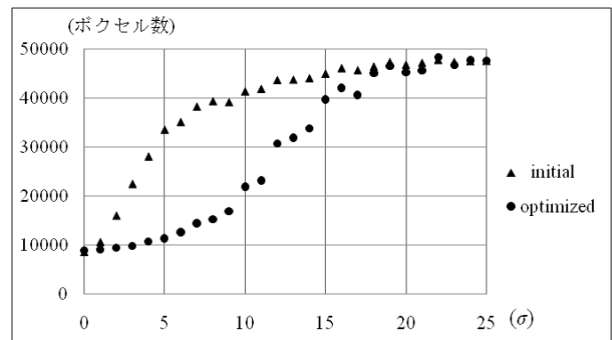


図 7: arabesque の評価結果

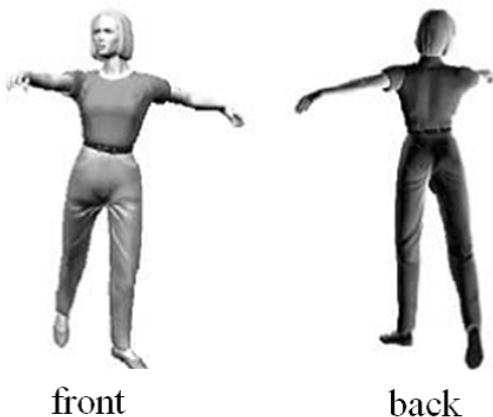


図 5: arabesque

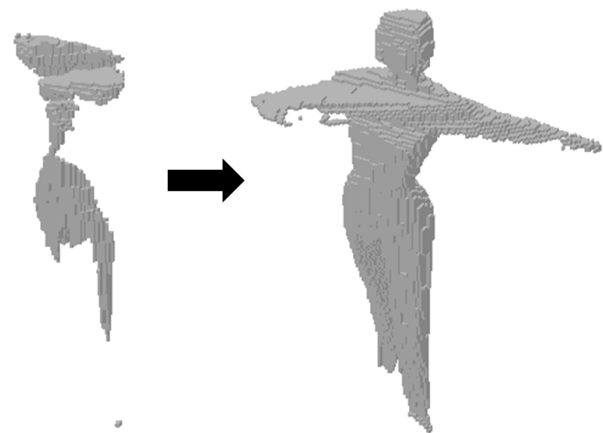


図 8: arabesque の復元形状

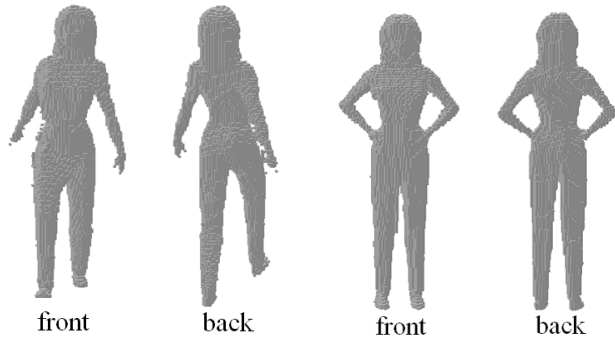


図 9: walk の基準形状

図 10: stand-up の基準形状

#### 4 結論

本稿では、カメラの空間位置を既知としカメラの姿勢推定を行い形状復元を行う提案手法において、カメラの空間位置に含まれる誤差がどの程度復元結果に影響がどうかをシミュレーション実験により検討し、提案手法がカメラの位置ずれに比較的頑健であることを示した。今後の課題としては、実画像に対して提案手法を適用することや対象物体を撮影する際の最適なカメラ配置の検討などがあげられる。

#### 参考文献

- [1] M. Li, M. Magnor, H. P. Seidel. "Hardware-Accelerate Visual Hull Rendering," The 8th Fall Workshop on Vision, Modeling, and Visualization 2003 (VMV 2003):151-158, 2003. 1-6056U-12, 2006.
- [2] J. M. Hasenfatz, M. Lapierre, J.-D. Gascuel, and, E. Boyer. "Real-time Cature, Reconstruction and Insertion into Virtual World of Human Actors," Vision, Video, and Graphics 2003 (VVG '03):49-56, 2003.
- [3] M. V. Uriol, M. Sain, F. Kuester, and, N. Bagherzadch, "Automatic Creation of Three-Dimensional Avatars," Proc. of SPIE-IS & T Electronic Imaging, SPIE 5013:14-25,2003.
- [4] Sayaka Sakamoto, Kenji Shoji, Hiroki Iwase, and Fubito Toyama, "Volume intersection with imprecise camera parameters," Proceedings of SPIE, vol.6056, pp60560U-1-60560U-12, 2006.
- [5] William H.Press, Saul A.Teukolsky, William T.Vetterling, Brian P.Flannery "Numerical Recipies in C 2nd" 444-455, 1992.

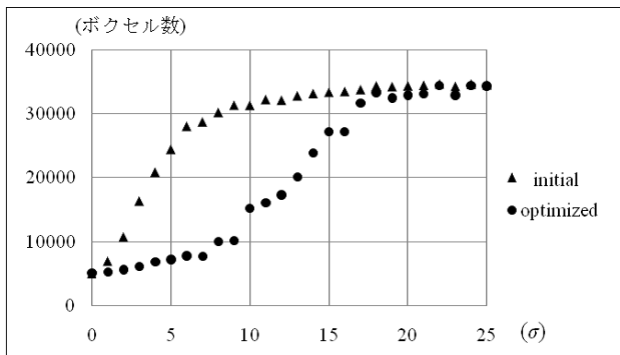


図 11: walk の評価結果

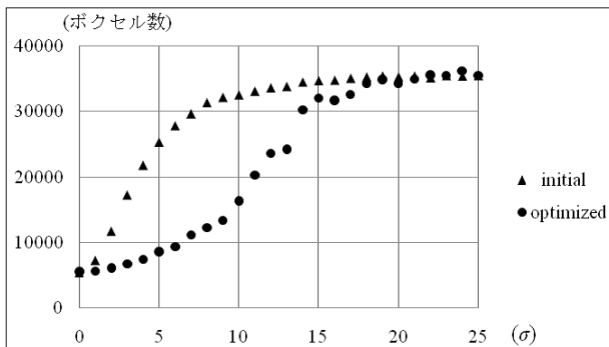


図 12: stand-up の評価結果