Direct Shape Carving: 表面形状復元のための 3 次元点群と法線の推定法

松田　一樹†　　浮田　宗伯†

† 奈良先端科学技術大学院大学
E-mail: †{kazuki-ma,ukita}@is.naist.jp

あらまし　本稿では，滑らかで正確な 3 次元表面形状を多視点画像から得るための手法を提案する．提案手法では，表面点群と法線から陰関数表現により 3 次元表面形状を得る手法を利用する．陰関数表現による復元表面形状の精度は，点群と法線の精度に加え，それらの間の整合性にも強く依存する．そこで，陰関数表現による表面形状復元に適するという特徴を重視して，多視点画像から点群と法線を獲得する．こうした特徴を備えた点群と法線の獲得のため，視体積交差法とステレオ視を独立に適用し，各手法の上記特徴に対する適正（例：点群の信頼性や法線の滑らかさ）を考慮しながら復元結果を統合する．実験により，提案手法が復元の精度や安定性において類似手法よりも優れていることを確認する．

キーワード　3 次元形状復元, 視体積交差法, ステレオ視, ステレオマッチング法

1. はじめに

人間が外界の環境を知る方法の一つが視覚である．我々は物体の形・色・位置といった基本的な情報の大部分を視覚から得ている．人間と同様，機械に視覚から物体の形状・位置・表面のテクスチャを理解させるというのは，コンピュータービジョンの大きな課題である．

特に，人体に関する理解，例えば多数のカメラを利用したモーションキャプチャは，映画などのメディアコンテンツの作成や，機械によるリハビリテーション支援などといった，様々な応用が可能である．このような応用のためには，3 次元面形状復元，テクスチャマッピング，姿勢推定など様々な技術が必要である．本論文はその中でも，3 次元形状復元を対象とする．画像を入力とする 3 次元点群復元、およびそれを元にした 3 次元表面形状復元は古くから研究されており，かつ今なお重要度の高い課題となっている．

本論文では，視体積交差法とステレオ視という二つの手法の利点を統合した結果を得る方法について述べる．視体積交差法はシルエット制約から，ステレオ視はフォトメトリック制約から物体の 3 次元点群情報を得る手法である．シルエット制約とは，画像から抽出した物体領域の中に，物体が必ず含まれるという制約であり，もう一方のフォトメトリック制約とは，ある点が異なる複数のカメラに観測されたとき，各画像におけるその点の周りのテクスチャは，投影変換の寄与を除いて同一であるという制約である．

本手法のように，シルエット制約とフォトメトリック制約を統合するアプローチは数多く研究されている．その多くは，視体積交差法で得られた結果を出発点として，フォトメトリック制約が満たされるように，結果を漸近的に変化させていくものである [1]〜[6].

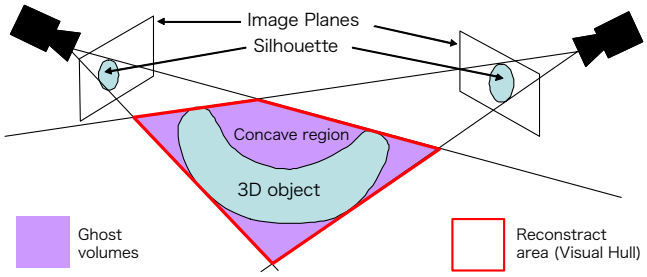


図 1　視体積交差法

これに対して本手法は，以下に示す特徴を備える．

（1）　繰り返し処理における局所最適解を避けるため，視体積交差法から得られた点群のうち，ゴーストボリュームの元となる点を直接，ステレオ視によって得られた点群と入れ替える．

（2）　視体積交差法から得られる稠密で滑らかな法線群と比較することにより，ステレオ視による疎な復元法線の中から滑らかさに欠けるものを抽出し，誤った復元結果と見なす．

2. 3 次元復元の基本的な手法とその特徴・問題

本章では，既存の代表的な 3 次元復元手法，その長所と既知の問題点について述べる．3 次元点群復元の代表的な手法としては，視体積交差法とステレオ視（ステレオマッチング法）がある．

提案手法を用いることで，これらが持つ問題点を解決することができる．

2.1　視体積交差法

図 1 は，視体積交差法 [7] を実行した結果である交差領域を示したものである．視体積交差法では，まずそれ

ぞれのカメラで撮影された物体のシルエットを３次元空間に射影して，カメラの投影中心を頂点とする錐体を作成する．各カメラ画像から作成された錐体の共通部分が，求めるべき結果である．視体積交差法によって得られた復元結果である交差領域は， Visual Hull と呼ばれる．

視体積交差法はステレオ視による３次元復元に比べ，ロバスト性，得られる点群の密度と連続性のどの点でも勝っている．しかし，物体にお椀のように窪んだ領域が存在する場合や，カメラの数が十分ではない場合，ゴーストボリューム (Ghost Volume) 等と呼ばれる偽りの体積が発生する問題がある．特に，窪んだ領域の周辺に於けるゴーストボリュームの発生は，カメラの数やその設置位置・方向に依らず，視体積交差法を用いる限り避けることができない問題である．

この問題があるにもかかわらず，視体積交差法は人体などの複雑な物体の形状復元に欠かせない存在である．なぜならば，スタジオなど既知の環境下に於けるシルエット抽出はステレオ視に於ける対応点探索より容易であり，対象表面のテクスチャによらず，高密度な点群が安定して復元できるからである．

実際の視体積交差法では，Visual Hull はボクセルの集合として復元される．Visuall Hull を構成するボクセルのうち，外部と接するボクセルを表面ボクセルとする．表面ボクセルの中心が，視体積交差法から求まる点群である．さらにその各点について，近傍のボクセルから接平面が推定でき，接平面に垂直な直線として法線を求めることができる．

2.2　ステレオ視

一方のステレオ視は，物体表面のテクスチャ相関性を元に，複数のカメラから三角測量の原理で物体各点までの距離を求める手法である．物体の輪郭情報しか使用しない視体積交差法に比べ，テクスチャの見え方を利用するステレオ視は，表面テクスチャが十分に特徴的である部分ではよい結果をもたらすが，そうでない部分では得られる点群は疎で，物体全体の形状を復元することができない．

ステレオ視に基づく初期の方法では，全ての点を独立したものとして扱い対応関係を求めていた．しかし，近年，滑らかさに関する制約を利用し，形状全体のフォトメトリック制約を用いて点群座標を求めている手法が開発された（例えば Level Sets [8] や EM による最適化 [9]）．さらに， Patch-based multiview stereo [10] などの新手法では，点群の位置と共に，各点に於ける法線も復元することが可能である．

ステレオ視によって精密な点群座標を求めることができるといっても，それを人体などの対象に適用しようとした場合には一つの問題に対処しなくてはならない．ほとんどの被服というのは対応点探索が容易な，特徴的なテクスチャというものを持っていないため，先に述べたように，間違った対応付けや疎な復元結果しか得られないといった問題である．

2.3　視体積交差法とステレオ視を改善する既存手法

視体積交差法を通して得られた Visual Hull からゴーストボリュームを削り取る手法として，Space Carving [1] がある．これは，初めに得られた形状を出発点として，フォトメトリック制約を満たさない部分を，内側へと削り込むように，漸近的に後退さて，制約に対して最適化する方法である．

最適化する対象は，フォトメトリック制約以外にも，形状の滑らかさに関する制約などを取り入れることもできる（例えば，連続的・局所的な最適化である Gradient Descent [2]，離散的・大域的な最適化である Graph-cuts [3], [4]，連続的・大域的な最適化 [5]）．他にも，自己遮蔽を明示的に考慮した対象形状の整合性 [6] などを利用する研究も存在する．

フォトメトリック制約に基づく， Visual Hull の削り取りは，対象表面のテクスチャをステレオ視のために設定できる場合は非常に強力なものとなる．例えば，人体の形状復元を行う場合，細かく独立な色模様がついた被服が役に立つ [11]．

また，フォトメトリック制約とは異なるが，特に人体など，構造が既知の形状の三次元形状復元には，形状モデル（例えば，Articulated Mesh [12]）が精度とノイズへのロバスト性向上のために有効である．

しかし，この章で述べた手法には，以下のような問題が存在する．

- 視体積交差法の結果を出発点とした漸近的最適化では，局所最適解に陥ることが多い．
- フォトメトリック制約を活用するために，対象に特別な表面テクスチャを設定することは，応用範囲を狭めることになる．逆に，そのような条件を強制しない，テクスチャの少ない一般的な対象においては，フォトメトリック制約を利用した最適化の効果が限定的になってしまう．
- 多関節剛体モデルを初めとした多くの形状モデルは，スカート，ワンピース，着物といった，非密着の衣服に対応できない．

2.4　ポワソン表面形状復元

表面点群から面を推定する３次元面形状復元は，コンピューターグラフィクス分野における一般的な問題であり，映画を初めとした産業応用の基礎技術となっている．いくつかの手法がある中で Marching Cube [13] による点群座標を利用した面付けが有名であるが，陰関数を利用した後発手法もあり，同じ入力点群に対してより滑らかに面形状を復元することができる．本論文では表面復

元手法としてポワソン表面形状復元 [14] を採用した．ポワソン表面形状復元では，点群の 3 次元位置だけでなく，各点について推定された法線も入力とし，それぞれの点が持つ法線が復元結果における面の法線と等しくなるよう面付けが行われる．

図 3 は，視体積交差法とステレオ視 [10] それぞれの手法で得られた点群に対して，ポワソン表面形状復元を行った結果である．使用したカメラは 8 台であり，各カメラからの画像を図 2 として掲載した．それぞれの手法の問題点をわかりやすくする目的で，対象の衣服は，ステレオ視による復元が容易な，特別に色分けされた着物とした．

まず，テクスチャの特徴が少ない，「顔」や「足」の部分に注目する．視体積交差法を元にした結果は，復元に失敗して削除された領域がなく，安定した結果を出している．一方で，ステレオ視を元にした復元結果では，一部の領域が失われてしまっていることがわかる．これは，テクスチャを元にした対応点探索が失敗したことを示している．

次に，テクスチャの特徴が豊富な衣服で，特に空間的に薄い「袖」の部分に注目する．ステレオ視を元にした結果では，形状が正しく復元されていることがわかる．対して，視体積交差法では，薄い形状の周りにゴーストボリュームがまとわりつき，正しい形状が失われている．

座標と法線が共に正しい点群を得ることが，ポワソン表面形状復元を使って，真の形状に対する過不足のない結果を得るために必要な条件である．本稿は，視体積交差法とステレオ視の結果を統合し，正確な法線付き点群を得ることを目的としている．

3. 視体積交差法とステレオ視の結果を利用した Direct Shape Carving

Space Carving といった既存の手法は，フォトメトリック制約が満たされるまで視体積交差法の結果を内側に縮小していくものであった．しかし，フォトメトリック制約を利用した Visual Hull の削り取りは，結果が真の形状に到達する前に局所最適解で停止してしまう場合がある．

本手法ではステレオ視 [10] によって得られた点群を直接採用し，Visual Hull から得られた点群のうち，ステレオ視によって得られた点群の外側にあるものを削除する．この手法の，技術的な特徴は以下の 2 つである

- ステレオ視によって得られた信頼のできる点群を，視体積交差法の結果とは無関係に，そのまま採用することで，
- 視体積交差法によって得られた点群のうち，各カメラとステレオ視によって得られた点群の間にあるものを直接に削除し，形状を削り込む．

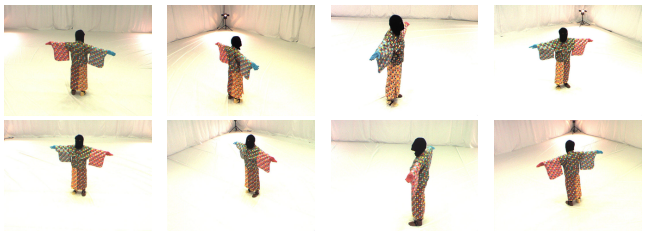
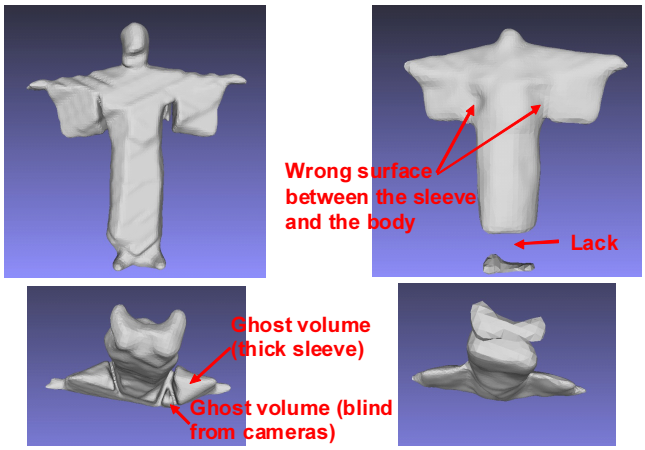


図 2　カメラセットから撮影された画像



視体積交差法　　ステレオ視

図 3　図 2 で示した画像の視体積交差法とステレオ視を利用した形状復元結果．点群からの表面復元にはポワソン表面形状復元を利用．

まずはじめに，得られた画像セットに対し，視体積交差法とステレオ視を独立に適用し，点群を得る．ステレオ視によって得られた点群は基本的に全て採用するが，近傍に於ける視体積交差法の結果と矛盾するような法線を持ったものだけは信頼性の低い点として取り除いた．これには以下の理由がある

- 周囲と比較して，あまりにも一貫性のない法線が復元された場合，その近傍におけるテクスチャの特徴が低く，対応付けに失敗している可能性が高い．
- 一貫性のない法線はポワソン方程式を適用することをより難しくさせる．

これらの判断に基づいて，ステレオ視による点群は

（1） 視体積交差法によって得られた点が十分に近い距離にあり，かつ

（2） 法線がその点の法線と著しく異なっている場合

に削除される．具体的には，

（1） 視体積交差法で得られた最近傍の点との距離が $w_l l$ より短く，かつ

（2） 法線のなす角が閾値 θ_a より大きい

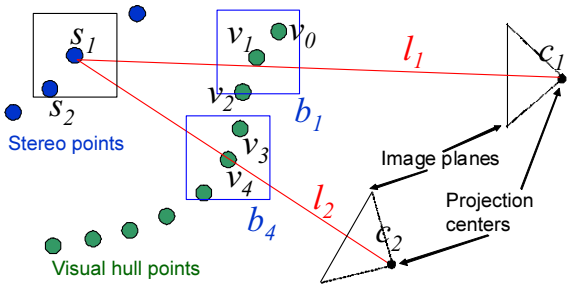


図 4　Visual Hull から得られた点群周辺のバウンディングボックスと，カメラ投影中心からステレオ視点群への視線の衝突判定

点を削除した．ここで l はボクセル一辺の長さであり，w_l はパラメータである．今回の我々の実験では，$w_l = 8$, $\theta_a = 30$ を使用した

次に，残りのステレオ視によって得られた各点を，CG レンダリングにおけるレイトレーシングのように，それぞれのカメラの像平面上に投影する．各点と，それを観測したカメラの投影中心を結ぶ線分が，視体積交差法によって得られた点に衝突したら，視体積交差法によって得られた点を，ゴーストボリュームをもたらすものであるとして削除する．実際には，視体積交差法の各点の周囲にバウンディングボックスを設定し，視線の衝突判定を行うことになる [15]．視線がバウンディングボックスに接触した場合，そのバウンディングボックスを備えた点は削除される．

この交差判定において，バウンディングボックスのサイズの決定が重要である．サイズが小さすぎ，あるいは大きすぎた場合，削除されるべき点が残り，あるいは残るべき点が削除されることになる．このサイズは，ステレオ視によって得られた点群のうち，対象となる点とその最近傍の点との距離から決める．

図 4 はこの交差判定を説明する図である．ステレオ視で復元された点 S_1 について考える．これを観測したカメラの投影中心を C_1, C_2 とする．まず，再近傍の点との距離に基づいて，バウンディングボックスのサイズを決定する．ここでは，S_1 から S_2 までの距離がバウンディングボックスのサイズであり，これを b とおく．1 番めのカメラについて考える．S_1 と C_1 を通る線分と距離 b の範囲内にある視体積交差法に基づく点を削除する．ここでは v_0, v_1, v_2 が該当する．2 番目のカメラに関しても同じように，線分 S_1C_2 との距離が b 未満である v_3, v_4 を削除する．

このようにして，視体積交差法から得られた点群から，ゴーストボリュームの原因となる点を取り除くことができた．その後，残った点群についてポワソン表面形状復元 [14] により表面形状を復元する．

なお，本手法の結果である点群には，視体積交差法で得られた点群と，ステレオ視を元に得られた点群との間に，空間的なギャップが発生している可能性がある．この様子を図 8 に示した．しかし，ポワソン表面復元手法は，多少空間的に距離が離れている二つの点群であっても，その間を滑らかに繋ぐ面を生成する．今回は，この性質を使用することで，生成された点群のギャップを埋める処理は特に行わなかった．

図 6 は，図 5 で示された映像を入力として復元され

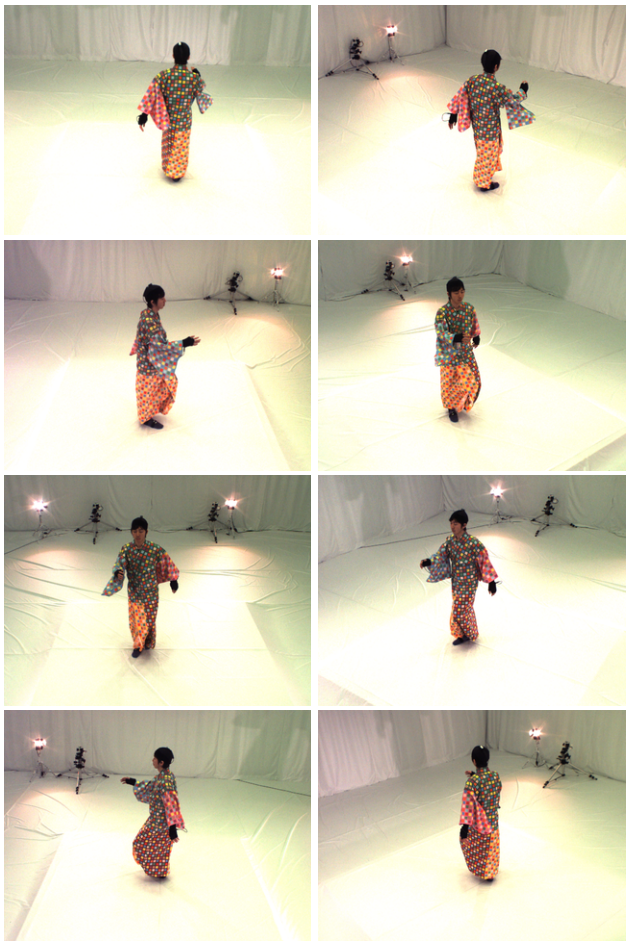
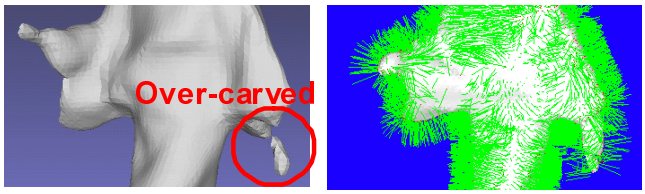
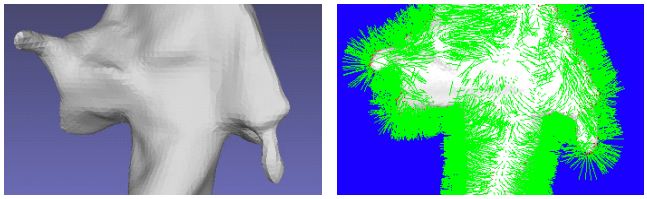


図 5　カメラセットから同時に撮影された画像



復元結果　　復元結果と点群の法線

図 6　ステレオ視による全点群と，削り取りによって残った視体積交差法による点群を合計したものを元にした表面形状復元の結果．推定された法線を緑色の線で示した．



復元結果　　復元結果と点群の法線

図 7　本手法の結果に対して表面形状復元を行った結果

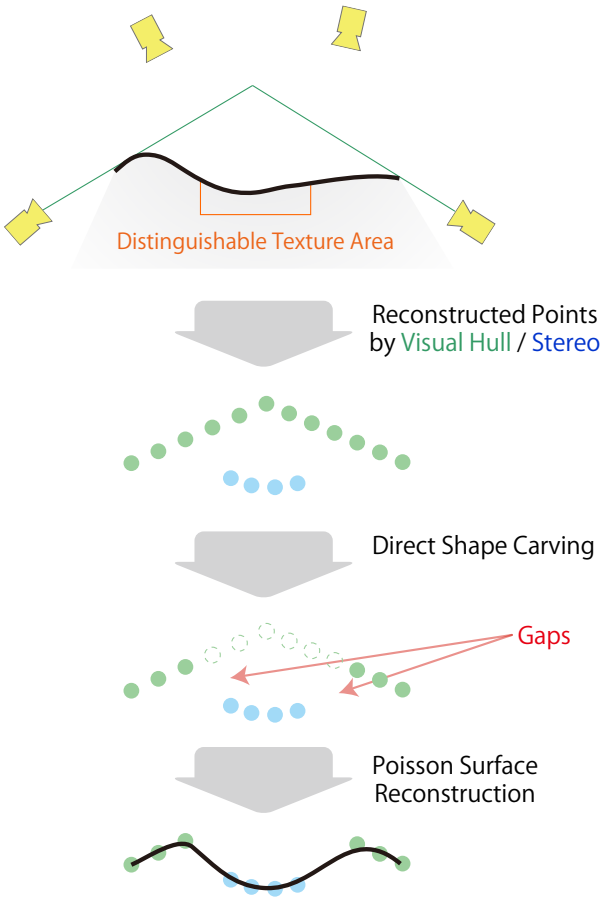
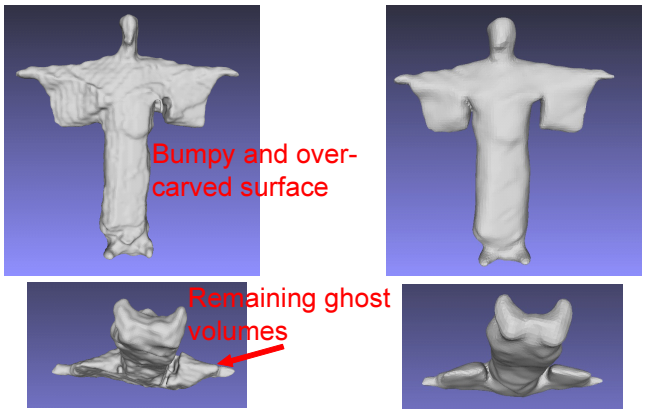


図 8　Direct Shape Carving 後に発生する点群のギャップと，ポワソン面復元によって生成される最終結果．



Space Carving　　本手法

図 9　図 2 の画像に Space Carving と本手法を適用し，表面形状復元を行った結果．

た表面形状を示しており，ゴーストボリュームの削り取りの必要性を示した典型的な例となっている．左側のイメージはステレオ視によって得られた点群と，残った視体積交差法の点群，それらの法線をもとに，そのままポワソン表面形状復元を施したものである．右の画像では法線を可視化して示した．左手に該当する部分の復元に失敗していることがわかる．これは，テクスチャが十分でなく，散らばった法線が，ポワソン方程式への当てはめを困難にしたためである．

図 6 とは対照的に，我々の結果である図 7 は左手の部分においてもよりよい結果を出していることがわかる．これは，表面復元の前に，乱雑な法線を持つ点を削除したためである．

本手法がゴーストボリュームを削減するのに有用であることを示すために，Graph-cuts を利用した Space Carving [4] と本手法の比較を行った．結果は図 9 に示した．Space Carving では ゴーストボリュームが残っていることが見える一方で，本手法ではそれらを取り除くことに成功していることがわかる．また，Space Carving による方法では，結果に本来は存在しない突起が認められる．これは削り取りが局所最適解で停止してしまったためである．

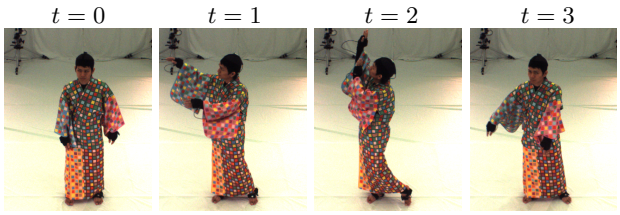


図 10　舞踊：正面カメラからの時系列画像

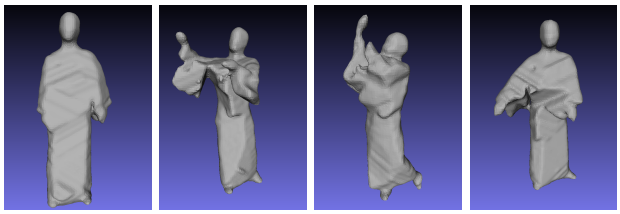


図 11　舞踊：視体積交差法のみを利用した復元結果

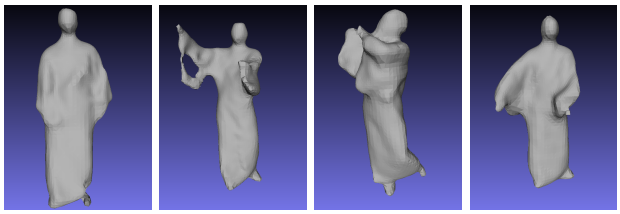


図 12　舞踊：ステレオ視のみを利用した復元結果

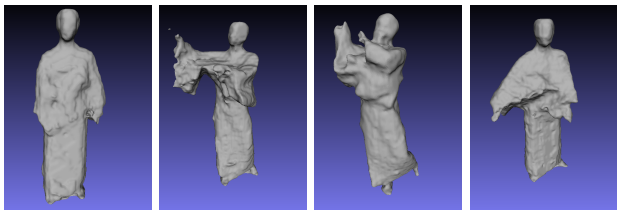


図 13　舞踊：Space Carving [4] を利用した復元結果

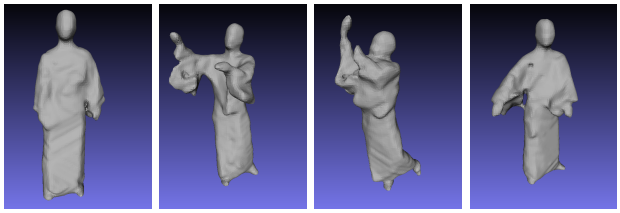


図 14　舞踊：本手法を利用した復元結果

4. 動画像を元にした実験

更に有用性を確かめるため，複数視点からの動画像に対して本手法を適用した．

図 10 – 24 に，実際に観測された動画像の一部と，それらに本手法を適用した結果を示している．ここでも Graph-cut を利用した Space Carving [4] による結果を比較として掲載した．シルエットとポワソン表面形状復元 [14] のパラメーターは，全ての実験で同一である．また，Space Carving のパラメーターは [4] 文中で人体復元に使用されているものと同一である．

実験には，物体を囲むように 8 台のカメラを使用した．視体積交差法で発生するゴーストボリュームの量は，物体の形状とカメラの配置に大きく依存する．図 1 からもわかるように，図 10 – 14 のような舞踊の場合では，平面に近い形をしている袖の裏表にゴーストボリュームが発生しやすい．このような場合，この平面を延長した空間，例えば真上から袖部分を見下ろすようにカメラを設置することができれば，視体積交差法によるゴーストボリュームを削減することに効果的である．しかし，カメラ配置と対象の形状・ポーズの関係の制限は，一般的な物体の形状復元に適用することはできない．

図 17 – 19 の下半身, 図 20 – 24 の全身は，ほとんど無特徴のテクスチャであり，ステレオ視による形状復元

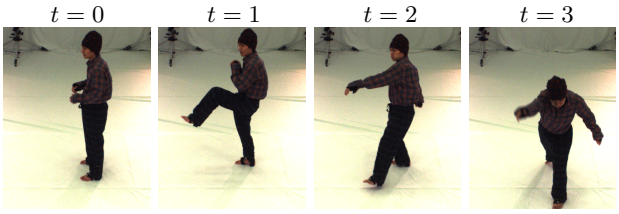


図 15　投球動作：正面カメラからの時系列画像

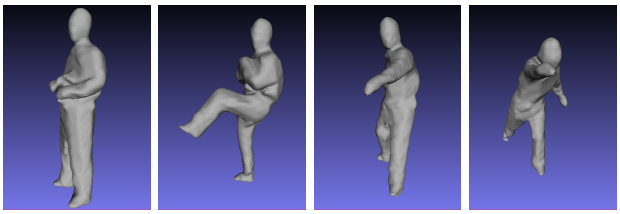


図 16　投球動作：視体積交差法のみを利用した復元結果

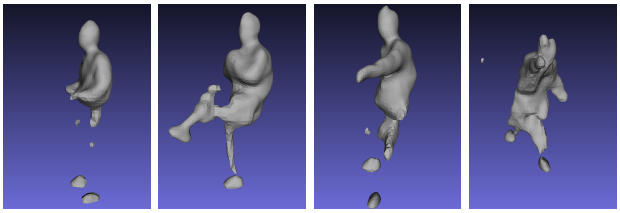


図 17　投球動作：ステレオ視のみを利用した復元結果

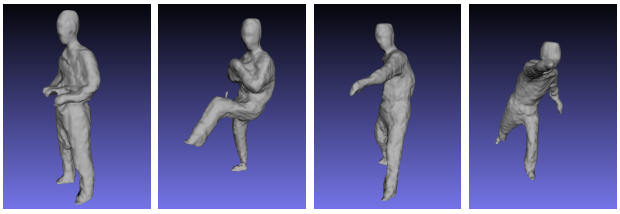


図 18　投球動作：Space Carving [4] を利用した復元結果

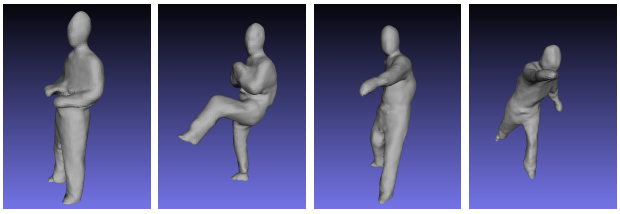


図 19　投球動作：本手法を利用した復元結果

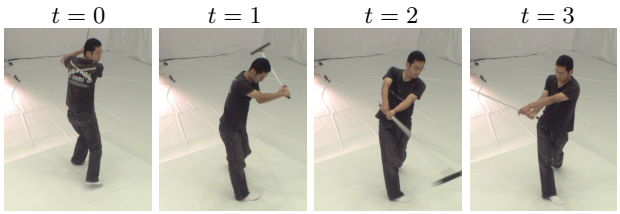


図 20　素振り：正面カメラからの時系列画像

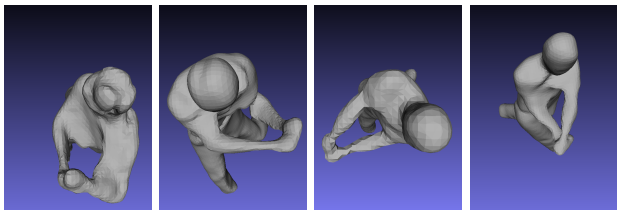


図 21　素振り：視体積交差法のみを利用した復元結果

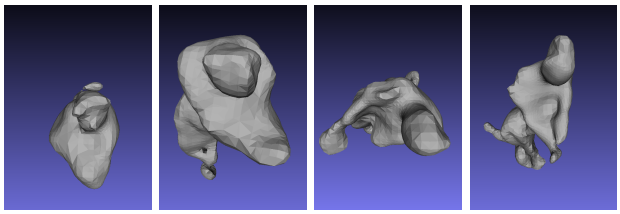


図 22　素振り：ステレオ視のみを利用した復元結果

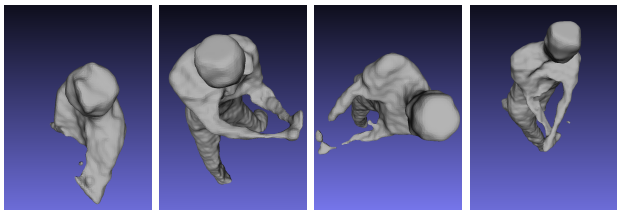


図 23　素振り：Space Carving [4] を利用した復元結果

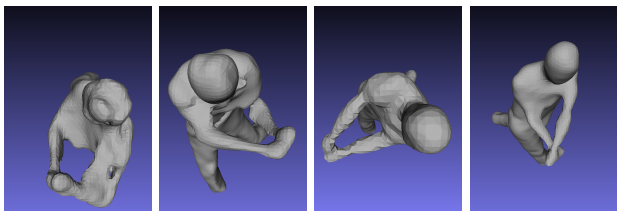


図 24　素振り：本手法を利用した復元結果

では非常に精度の低い結果しか得られていない．特に最後の例，図 20 – 24 では，閉じた両椀の領域の中が周囲に設置された 8 台のカメラからは見通しにくく，視体積交差法の結果に於いて，多くのゴーストボリュームが残存している．しかし，本手法を適用した結果，特に $t = 4$ 左腕内側で若干ながら結果が改善した．

以上いずれの実験においても，本手法が Space Carving と比較してゴーストボリュームがより少ない復元を可能とすることが確かめられた．

5. 結果と考察

本論文は，複数視点からの三次元法線付き点群の復元手法を提案した．本手法は視体積交差法とステレオ視によって得られた点群を用いるものである．これら 2 種類の点群を用いて，2 段階の点群縮減を行った

（1） ステレオ視によって得られた点群のうち，異常な法線を持つものの除去

（2） 視体積交差法によって得られた点群のうち，ゴーストボリュームに対応するものの除去

本手法は，点群を入力として，より妥当だと考えられる点群を出力するものであった．今後の発展として，次の 3 つを考えている．

- 得られた点群から面を再構成した結果を基にして，Space Carving のような，面ベースのフォトメトリック制約に対する最適化を行う．視体積交差法を出発点とした Space Carving と異なり，頂点が物体外側に移動する可能性も考慮に入れる必要があるが，最適化前の形状がより真の形状に近いという利点がある．
- 得られた点群そのものの中に，本手法を適用した結果としてギャップが発生している部分がある．この部分について，フォトメトリック制約や何らかの幾何学的制約を使用して点群を内挿したり，頂点座標・法線を修正する．

PMVS [10]，ポワソン表面形状復元 [14]，Space Carving [4] のソースコードはそれぞれ Y. Furukawa，M. Kazhdan および S. Nobuhara によって書かれたものを使用した．

文　献

[1] K. N. Kutulakos and S. M. Seitz: “A Theory of Shape by Space Carving”, International Journal of Computer Vision, **3**, 38, pp. 197–216 (2000).

[2] C. Hernandezesteban and F. Schmitt: “Silhouette and stereo fusion for 3D object modeling”, Computer Vision and Image Understanding, **96**, 3, pp. 367–392 (2004).

[3] G. Vogiatzis, P. H. S. Torr and R. Cipolla: “Multi-View Stereo via Volumetric Graph-Cuts”, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, pp. 391–398 (2005).

[4] T. Tony and N. Shohei: “Simultaneous super-resolution and 3D video using graph-cuts”, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society (2008).

[5] K. Kolev, M. Klodt, T. Brox and D. Cremers: “Continuous Global Optimization in Multiview 3D Reconstruction”, International Journal of Computer Vision, **84**, 1, pp. 80–96 (2009).

[6] A. Delaunoy and E. Prados: “Gradient Flows for Optimizing Triangular Mesh-based Surfaces: Applications to 3D Reconstruction Problems Dealing with Visibility”, International Journal of Computer Vision, pp. 1–24–24 (2010).

[7] G. Cheung, T. Kanade, J.-Y. Bouguet and M. Holler: “A real time system for robust 3D voxel reconstruction of human motions”, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, pp. 714–720 (2000).

[8] J.-P. Pons, R. Keriven and O. Faugeras: “Modelling Dynamic Scenes by Registering Multi-View Image Sequences”, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, pp. 822–827 (2005).

[9] C. Strecha, R. Fransens and L. Van Gool: “Combined Depth and Outlier Estimation in Multi-View Stereo”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, pp. 2394–2401 (2006).

[10] Y. Furukawa and J. Ponce: “Accurate, dense, and robust multiview stereopsis.”, IEEE Transactions on Pattern Analysis and Machine Intelligence, **32**, 8, pp. 1362–1376 (2010).

[11] R. White, K. Crane and D. A. Forsyth: “Capturing and animating occluded cloth”, ACM Transactions on Graphics, **26**, 3 (2007).

[12] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel and S. Thrun: “Performance capture from sparse multi-view video”, ACM Transactions on Graphics, **27**, 3 (2008).

[13] W. E. Lorensen and H. E. Cline: “Marching cubes: A high resolution 3D surface construction algorithm”, ACM Transactions on Graphics, **21**, 4, pp. 163–169 (1987).

[14] M. Kazhdan, M. Bolitho and H. Hoppe: “Poisson surface reconstruction”, Proceedings of the fourth Eurographics symposium on Geometry, pp. 61–70 (2006).

[15] T. Möller and B. Trumbore: “Fast, minimum storage ray/triangle intersection”, ACM Press, New York, NY, USA (2005).