

多視点画像を用いた対象および影領域抽出法

津田 佳行 延原 章平 松山 隆司

概要 本論文では多視点情報を用いた高精度シルエット抽出法を提案する。視体積交差法による三次元形状復元は、少ない計算量で安定して対象物体の三次元形状の概形が得られるため広く用いられている。しかし、シルエット抽出には対象自身によるキャストシャドウ、背景との色の類似など、一視点からのみの情報では対処し難い問題が存在する。そこで我々は多視点情報と撮影環境の事前知識を用いて三次元空間上で影の存在しうる領域を特定することにより、これらの問題を同時に解決する手法を提案する。また、実画像を用いた実験によりその有効性を検証する。

Shadow Carving: Accurate Object Silhouette Extraction from Multi-viewpoint Images

YOSHIYUKI TSUDA, SHOHEI NOBUHARA and TAKASHI MATSUYAMA

Abstract In this paper, we propose a method for accurate object silhouette extraction from multi-viewpoint images. Shape-from-silhouette is one of the most widely used methods of 3D shape reconstruction since it reconstructs rough 3D shape stably with a little calculation. However, it is difficult to obtain accurate silhouette using each single-viewpoint information because of cast shadow, similarity of the color between object and background, and so on. To solve this problem, we detect shadow region using multi-viewpoint information and geometry of studio. We show some experimental results with real examples.

1. はじめに

1.1 3D ビデオ

実世界の情報を記録するという行為は人間社会において必要不可欠であり、その方法や媒体も、壁画、文字、彫刻、写真など文明の発達にともなってさまざまなものが発明されてきた。そして現代においては、ビデオカメラを用いて時々刻々と変化する実世界の様子を記録することが当たり前のようにできるようになった。しかし、記録している情報はあくまでも3次元的世界の情報を2次元の撮像面に投影したものであるため、撮影対象が持つ3次元的位置・形状などの情報が大幅に欠落している。

そこで近年、実世界の情報を2次元的に記録するのではなく、3次元的な情報をそのまま記録する3Dビデオというメディアの研究が盛んに行われている。このメディアは、撮影対象の3次元情報をそのまま保持しており、ユーザーからの要求に応じて自由な視点からの映像を生成・提供することはもちろん、立体のまま表示することもできる。このため、プロスポーツ選手のフォーム解析や伝統技能の保存、高臨場感遠隔会議システムなど、様々な活用方法が期待される。

1.2 3次元形状復元法と問題点

多視点画像を用いた3次元形状復元の手法としては、

スペースカービング法¹⁾²⁾ や視体積交差法³⁾ などの手法が研究されており、近年では、両者を統合することによって形状復元の正確さと安定性を両立させた手法が提案されている⁴⁾⁵⁾⁶⁾⁷⁾⁸⁾⁹⁾。ここで、視体積交差法に代表される多くの手法は、カメラで得られた画像中から対象のシルエットを正確に抽出できることを仮定しているが、シルエット抽出は、たとえば広く用いられる手法である背景差分法では、事前に撮影された、対象の存在しない画像(背景画像)と、シルエットを抽出したい対象を含む画像(前景)の差をとって抽出結果とするが、

- 対象が存在することによって環境に生じる(影などの)変化
- 対象のテクスチャと背景のテクスチャが近い場合のあいまいさ

などから、今なお困難な問題であるといえる。

1.3 本論文の焦点と構成

そこで本研究では、対象物体が作る影がシルエット抽出の精度を低下させる要因になると考え、撮影環境の情報を用いることにより影を含んだシルエット(およびそれを元に復元した3次元形状)から影領域を削除する手法を提案する。

第2節では、従来のシルエット抽出法における問題点を明らかにする。第3節では、撮影環境の情報を用いた対象および影領域の抽出について説明する。第4

節では、アルゴリズムを実装し実験を行う。第5節では、本研究の結論と課題を述べる。

2. 関連研究

本節ではシルエット抽出における関連手法とその問題点について説明する。従来は、撮影したそれぞれの視点において独立にシルエットを抽出し、それらを使用して視体積交差法などにより3次元形状を復元するという手法が一般的であったが、近年では多視点からの情報を積極的にシルエット抽出に利用し、3次元形状復元とシルエット抽出を同時に行う研究も盛んに行われている。このため、単視点情報を用いたものと多視点情報を用いたもの分けて説明する。

2.1 単視点情報を用いたシルエット抽出法

単視点情報を用いた手法の代表的なものに背景差分法がある。この方法では、あらかじめ同じ視点から撮影しておいた背景画像との差分を取り、ある閾値よりも差分量が大きい画素を抽出するが、この方法では、

- (1) 照明の揺らぎにより本来背景である領域を対象として検出する
- (2) 対象物体と背景の色度類似性により本来対象である領域を背景として検出する
- (3) 対象物体がシーン中に作る影を対象として検出する

などの原因により、シルエットと背景を正しく分ける閾値が存在しないような状況が多く発生する。このため、これらの問題点を解決する方法が研究されている。1については、波部ら¹⁰⁾が画像類似度の統計的性質を用いて解決する手法を提案している。2については、次節で述べるように大浜らが対象領域の見かけ上の連結性により解決する手法を提案している。そして3については Horprasert ら¹¹⁾が明度と色度の類似度を利用して影を取り除く手法を提案している。

Horprasert らの手法は、画像上のある点 (x, y) における対象画像および背景画像の明度と色度を計算し、以下のように画素を対象 (F)、背景 (B)、背景中の影 (S)、照らされた背景 (H) に分類する。

- F 色度のずれが大きい、もしくは対象画像の明度がある程度より小さい
- B 明度、色度ともにずれが小さい
- S 色度のずれが小さく、対象画像の明度が背景画像より小さい
- H 色度のずれが小さく、対象画像の明度が背景画像より大きい

また、複数枚撮影した背景画像を用いた閾値の決定や、色の量子化間隔のために色度の誤差が拡大する現象への対処も同時に行っている。この手法によってある程度影の分類が可能になるが、あくまでも各ピクセルを独立に処理し、周囲との空間的連結性や3次元位置を考慮していないため、人物の髪など本来は影であるはず

のない部分まで影としてしまう可能性がある。

2.2 多視点情報を用いたシルエット抽出法

Zeng ら¹²⁾ は、シルエットの見かけ上の連結性に着目し、以下に示す **intersection consistency constraint** と **projection consistency constraint** という二つの制約を導入することによって未知背景の多視点画像からシルエットを抽出する手法を提案した。

Intersection consistency constraint 全ての視点のシルエットから復元した3次元形状をそれぞれの視点に投影した像は、それぞれの視点におけるシルエットと完全に一致する。

Projection consistency constraint 対象物体と背景の境界にエッジがあれば、その境界はそれぞれの視点における撮影画像でもエッジとなって現れる。即ち、輝度値の近い画素を統合して領域分割すれば、シルエット内部のセグメントの集合とシルエット外部のセグメントの集合に完全に分けることができる。

大浜¹³⁾¹⁴⁾ は Zeng らの手法を改良し、背景差分画像を用いて連結関係の誤り検出・訂正を行いながらシルエットを抽出する手法を提案した。いずれの手法も各撮影画像中で intersection consistency constraint と projection consistency constraint の両方を満たす領域を抽出し、それをシルエットとしている。このため複雑な背景に対しても安定してシルエットを抽出することができるが、対象物体と対象物体が床に作る影を合わせた領域もまた intersection consistency constraint と projection consistency constraint を満たしているため、画像中に影が写っていた場合には抽出された領域に影が含まれてしまい、対象そのものの3次元形状を復元することができない。

2.3 従来手法のまとめ

単視点情報を用いた手法は処理が単純で計算量が少ないという利点があるが、

- そのピクセルが対応する3次元位置を考慮しないので影ができないはずの部分も処理の対象にする
- 他の視点から抽出の妥当性を確認できない

という理由のために、影領域の抽出には限界がある。また、従来の多視点情報を用いた手法は抽出の妥当性を確認するため安定してシルエットを抽出できるが、影領域も妥当性の条件を満たすので本質的に影領域の抽出能力がない。

本研究ではこの問題を解決するため、撮影環境(撮影空間の形状、材質および光源)と多視点からの情報を利用して、従来の手法(大浜の手法)により抽出されたシルエット(対象そのもののシルエット+影)から対象そのもののシルエットを抽出する手法を提案する。

3. 撮影環境の情報を用いた対象および影領域の抽出

3.1 本研究の焦点

対象物体が作る影には以下の二つのものがある．

- (1) 対象物体が対象物体自身に作る影 (セルフシャドウ)
 - (2) 対象物体が背景に作る影 (キャストシャドウ)
- 各カメラの観測画像で考えると、1のセルフシャドウは対象物体のテクスチャを変化させるが、シルエットの形状に与える影響は小さい．一方、2のキャストシャドウは対象物体として誤って抽出されることによってシルエット形状に影響を与える．このため本研究では、二種類の影のうち2のキャストシャドウのみを扱い、大浜の手法によって抽出されたシルエット (対象そのもののシルエット + 影領域) から対象のシルエットを抽出する手法を提案する．

このとき、従来手法における問題点を解決するためには

- 画像上のピクセルと対応する3次元位置を考慮した処理
- 他の視点からの妥当性の確認

を行う必要がある．そこで本研究では、キャリブレーション済みのカメラ群で、対象を異なる複数の視点から撮影する状況を想定する．このような状況において撮影環境の情報を用いて影の存在し得る領域を限定し、その領域内で各画像を参照することにより対象および影領域の抽出を実現する．

3.2 領域抽出のアプローチ

本研究における対象および影領域抽出処理の流れは次のようになる．

- (1) まず、大浜らの手法により影を含んだシルエット $\{SIL_{est}\}$ と、 $\{SIL_{est}\}$ を用いて視体積交差法で復元した3次元形状 (visual hull) VH_{est} を得る．
- (2) そして、多視点からの情報を用いて VH_{est} の表面を、影を表す領域 SDW と対象物体を表す領域 OBJ とに分類する．
- (3) 最後に、それぞれの視点 i に OBJ を投影することにより影を除いたシルエット $\{SIL\}$ を得る．また、 SDW を投影すれば視点 i における影領域を得ることができる．

3.3 設定

3.3.1 データ構造

各画像におけるカラーモデルとしては、RGB表色系を用いる．各画素の輝度を赤 (R) 緑 (G) 青 (B) の量子化された三要素で表現する．

また、3次元形状を表現するためのデータ構造として、頂点の連結関係や表面の法線の情報が容易に利用できることから、メッシュモデルを採用する．ここで

は、まず視体積交差法で復元したボリュームデータの表面を Marching Cube 法によって三角形パッチで近似し、パッチを構成する頂点と辺の情報を保持する．そして、メッシュで表された Visual hull VH_{est} 上の頂点を処理単位とし、 SDW および OBJ は VH_{est} を構成する頂点の部分集合として、それぞれの頂点が SDW に含まれるかどうかを判定することにより対象表面のセグメンテーションを実現することになる．そして、セグメンテーションされた領域を各画像上に投影する際には、 OBJ に含まれる点のみで構成された多角形のうち、投影する視点から見えている領域のみを投影することにより対象のシルエット $\{SIL\}$ を得る．また、 SDW に含まれる点を1つでも構成要素に持つ多角形を投影することにより影領域を得ることができる．

各視点からメッシュ表面の点が見えているかどうかの判定には z バッファ法を使用する． z バッファ法では、画像上の各画素について RGB 値に加え、その視点から見た奥行き情報である z 値を z バッファに保持しておく．そして、 z バッファに格納されている値よりも遠くにある点は物体に隠蔽されているとして可視判定を行う．

3.3.2 撮影環境

本研究では、床は世界座標で $z = 0$ で表される平面とし、対象、カメラ、照明が図1のように $z \geq 0$ の領域に存在する．視体積交差法においては、 $z \geq 0$ の空間についてのみ visual hull を求めるものとする．また、光源は天井に分布した白色光源、材質は完全拡散反射を仮定する．

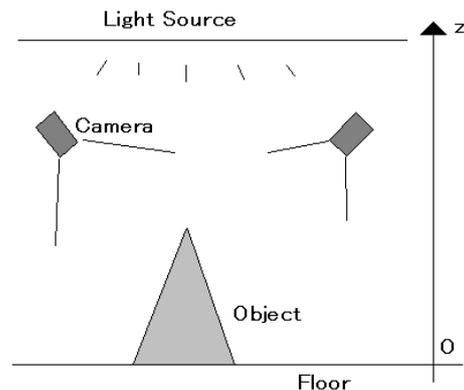


図1 撮影環境

3.4 影が持つ条件

本節では、上のように撮影環境を設定したときにメッシュ上の点が SDW に含まれるための条件を、幾何学的な条件と各画像上での条件とに分けて説明する．

3.4.1 幾何学的な条件

点 $p \in SDW$ が持つ幾何学的な条件は以下の2つである．これらの条件は visual hull 中で影が存在する

領域を限定するための条件である。

条件 1: SDW は床平面 $z = 0$ と連結している。

条件 2: 全ての視点 i について、カメラの投影中心 C_i と点 p を結ぶ直線と床面 $z = 0$ との交点を q_i とすると、点 q_i は VH_{est} の内部にある。

これらは visual hull の持つ性質による条件である。実空間において影は床平面 $z = 0$ にのみ存在するので、視体積交差法により 3 次元形状を復元した場合、復元された visual hull において影領域は床面 $z = 0$ を含む立体となる。従って、その表面上の頂点集合である SDW もまた床面 $z = 0$ と連結である。

また、床面 $z = 0$ 上で VH_{est} の外側の領域には対象物体の影は存在しないので、点 q_i が VH_{est} の外側にあるとき、点 q_i は影のない床面に存在することになる。ゆえに、視点 i からは床上の点 q_i ではなくその手前にある対象物体そのものが見えていることがわかる。

以上のことから、 SDW に含まれる全ての頂点は条件 1 と条件 2 を満たす。

3.4.2 各画像上での条件

点 $p \in SDW$ が持つ各画像上での条件は以下のようになる。この条件は影の持つ色の条件である。

条件 3: 点 p が複数視点から見えており、点 p を観測する全ての視点 i において、点 p を画像平面上に投影した点の、撮影画像上の画素値 $I_{c=R,G,B}(p, i)$ と背景画像上の画素値 $J_{c=R,G,B}(p, i)$ について以下の関係式が成り立つ。ただし T_1, T_2 および T_3 は適当に設定した閾値である。

$$0 \leq \sum_c (J_c(p, i) - I_c(p, i)) \leq T_1 \quad (1)$$

$$or \left\{ \begin{array}{l} T_1 \leq \sum_c (J_c(p, i) - I_c(p, i)) \leq T_2 \\ \frac{\sum_c I_c(p, i) J_c(p, i)}{\sqrt{\sum_c I_c(p, i)^2 \sum_c J_c(p, i)^2}} \geq T_3 \end{array} \right. \quad (2)$$

式 2 は、影領域の明度は背景画像の明度よりも小さく、色度がほぼ同じであるという性質を表している。また、RGB の量子化間隔の影響により、 $J_{R,G,B}(p, i)$ が 0 に近い値を持つ場合に明度のずれに対して色度のずれが大きくなる現象があるため、閾値 T_1 を設定し式 1 の条件を当てはめることによって対処する。同様の理由により、 $I_{R,G,B}(p, i)$ が 0 に近い値の場合も明度のずれに対して色度のずれが大きくなるため、明度のずれが閾値 T_2 を越えた場合は対象と見なすことによって対処する。

また、点 p は影によって作られた偽のボリュームの表面上にあるため、点 p を観測する視点 i は実際にはその先にある床面上の点 q_i を観測している。従って、点 p を観測できる全ての視点において明度と色度につ

いての条件式 1 または式 2 を満たす。従って、この条件 3 は Horprasert らの条件¹¹⁾ を多視点に拡張したものであるといえる。

3.5 アルゴリズム

以上が SDW に含まれる点が高満たす条件であり、それぞれの頂点について条件 1~3 を全て満たすかどうかを判定すれば対象表面の分類が行える。しかし、条件 1 を判定するためにはその周囲に SDW に含まれるかが存在するかどうか事前に判明していなければならない。

そこで、まず VH_{est} の中で平面 $z = 0$ 上にある点 (このような点は必ず条件 1 を満たす) について条件 2 と条件 3 の判定を行い、条件を満たす点があれば SDW に追加する。このとき、追加した点に隣接する頂点は条件 1 を満たすことが保証されるので、条件 2 と条件 3 の判定をすれば SDW に含まれるかどうかわかる。

このように、 SDW に含まれる点に隣接する頂点を順次調べていけば、条件 1 を満たしている頂点のみを処理対象として扱うことができるため、実際には処理対象となった各頂点が条件 2 と条件 3 を満たすかどうかを判定すればよいことになる。

このことを踏まえて、アルゴリズムを以下のように設計する。

多視点情報を用いた対象シルエット抽出アルゴリズム

Step 1 キャリブレーション済みのカメラ群から撮影した対象画像、背景画像および背景差分画像を入力として、大浜のアルゴリズムにより影 (キャストシャドウ) を含むシルエット $\{SIL_{est}\}$ と、 $\{SIL_{est}\}$ から視体積交差法および Marching Cube 法により復元した visual hull VH_{est} を得る。また、各視点 i において VH_{est} から z バッファを作成する。

Step 2 V を VH_{est} の底面 $z = 0$ 上の点に隣接する頂点の集合、 OBJ 、 SDW を空の頂点集合とする。

Step 3 $p \in V$ について、 $IsShadow(p)$ が真なら **Step 3-true** を、偽なら **Step 3-false** を行う。

Step 3-true p に隣接する頂点のうち SDW にも OBJ にも含まれない頂点の集合を U とすると、 $V = V + U - p$ 、 $SDW = SDW + p$ とする。

Step 3-false $OBJ = OBJ + p$ 、 $V = V - p$ とする。

Step 4 $V = \phi$ でなければ **Step 3** に戻る。 $V = \phi$ なら全ての視点 i について、 OBJ 中の頂点のみで構成された VH_{est} 上の多角形のうち、視点 i から見えている領域を i に投影して SIL_i を求め、 $\{SIL_i : \text{for all } i\}$ を出力して終了する。

ただし、 $IsShadow(p)$ は点 p が条件 2 と条件 3 を同時に満たしたときに真を返す関数であり、また、このアルゴリズムでは対象形状は変化しないので z バッファ

の更新は必要ない。

このアルゴリズムでは、図 2 に示すように、まず $z = 0$ の頂点を処理対象としてセグメンテーション処理を開始し、影領域に含まれる頂点が見付かったら、その頂点に隣接する頂点へを処理対象に追加していく。つまり VH_{est} を下から上へと探索しながら領域のセグメンテーションを行い、処理対象 (SDW に分類される可能性のある未処理の頂点) がなくなった時点で終了する。ただし、図 2 では SDW に分類された頂点を灰色、その他の頂点を黒色で表している。

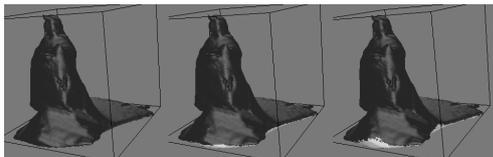


図 2 処理の進行

これにより、従来手法で問題になった

- 各画像上のピクセルと 3 次元空間上の位置とを対応づけることで、選択的に影の判定を行うこと
- 他の視点からの妥当性の確認

を解決した。

4. 評価実験

4.1 実験環境

高さ約 2m、半径約 3m の円筒状の部屋に 15 台のカメラを配置した (ただし各カメラは正確にキャリブレーション済みとする)。それぞれのカメラで撮影した対象画像および背景画像から大浜のアルゴリズムによってシルエットを抽出し、それをを用いて影領域を取り除いた。

4.2 評価方法

実装したアルゴリズムのシルエット抽出精度を定量的に評価するために、再現率 (Recall) および適合率 (Precision) に基づいて求められる F 値 (F-measure) を評価値として用いる。

まず、シルエット抽出結果に含まれる画素を以下の四種類に分類する。

True positive (TP) 本来シルエットである画素を抽出したもの

True negative (TN) 本来シルエットでない画素を抽出しなかったもの

False positive (FP) 本来シルエットでない画素を抽出したもの

False negative (FN) 本来シルエットである画素を抽出しなかったもの

そして、全ての画像に含まれる TP, TN, FP, FN の面積の合計 $S(TP)$, $S(TN)$, $S(FP)$, $S(FN)$ を用いて Precision, Recall, および F-measure を以下の式により求

める。

$$\text{Recall} = \frac{S(TP)}{S(TP) + S(FN)} \quad (3)$$

$$\text{Precision} = \frac{S(TP)}{S(TP) + S(FP)} \quad (4)$$

$$F\text{-measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

Recall は真のシルエットのうち実際に抽出できた割合を示し、Precision は抽出したシルエットに含まれる真のシルエットの割合を示す指標であり、F-measure はそれらを総合的に判断する指標である。これらはいずれも値が大きいくほど正確であることを示すが、Recall, Precision および F-measure を計算するためには本来のシルエットが既知でなければならないが、この実験では実物体を撮影しているため正しいシルエット情報を得ることができない。そこで、人手によって切り抜いたシルエットを基準として使用した。

提案したアルゴリズムによって、いくつかの閾値について求めた OBJ 領域からそれぞれの視点におけるシルエットを作成し、式 5 によって F-measure を計算する。また、比較対象として、

- (1) 大浜の手法により抽出したシルエット
- (2) 背景差分法により抽出したシルエット
- (3) Horprasert らの手法により抽出したシルエットについても同様の評価を行う。

4.3 実験

4.3.1 データセット 1

対象の撮影画像を図 3(a)~(d) に、大浜のアルゴリズムによって抽出されたシルエットを図 4(a)~(d) に示す。このデータでは対象物体が複雑な形状をしており、裾が広がったところにある影がシルエットとして抽出されていることがわかる。

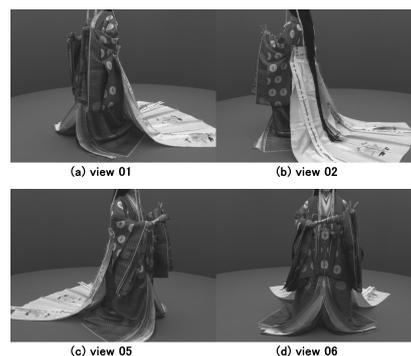


図 3 対象画像

4.3.2 結果

本手法によるシルエット抽出結果を図 5(a)~(d) に、シルエットを対象画像に重ね合わせた画像を図 6(a)~(d) に示す (閾値 $T_1 = 3$, $T_2 = 150$, $T_3 = 0.995$)。ただし、図 6 では本手法で抽出した対象領域を薄く表

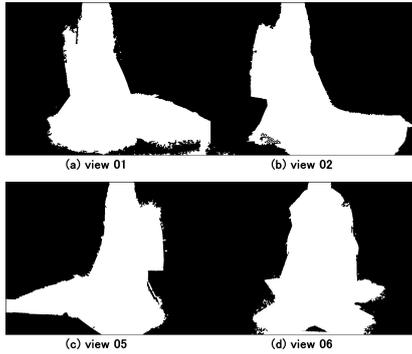


図 4 大浜の手法による抽出結果

した。また、背景差分法によるシルエット抽出結果を図 7(a)~(d)に(閾値 $T = 5, 10, 20, 30 \cdot \text{view 01}$), Horprasert らの手法によるシルエット抽出結果を 8(a)~(d)に示す。

さらに、本手法、背景差分法によるシルエット抽出結果の F-measure をそれぞれ表 1, 2 に示す。Horprasert らの手法では 0.972, 大浜の手法では 0.969 であった。

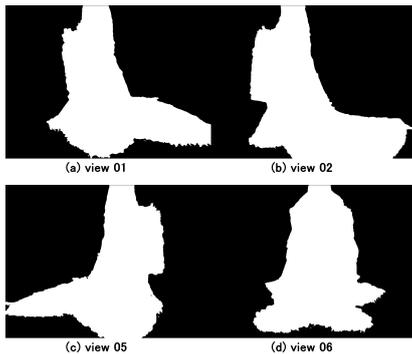


図 5 本手法による抽出結果

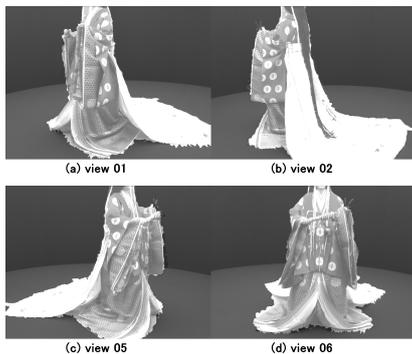


図 6 本手法による抽出結果+対象画像

4.3.3 データセット 2

対象の撮影画像を図 9(a)~(d)に、大浜のアルゴリ

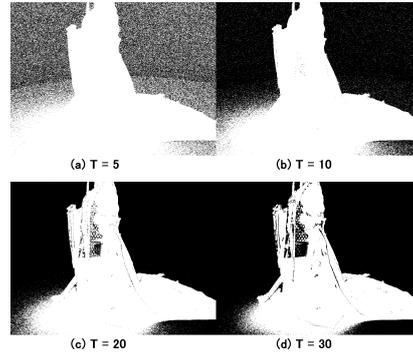


図 7 背景差分法による抽出結果 (view 01)

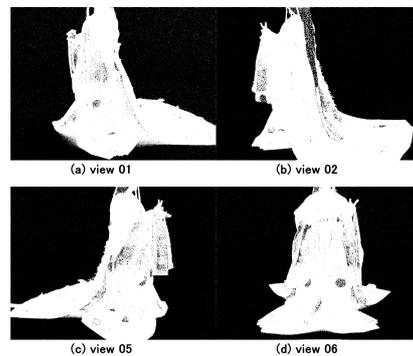


図 8 Horprasert らの手法による抽出結果

閾値 T_1	3	3	3	3	3	3
閾値 T_2	150	200	250	150	200	250
閾値 T_3	0.990	0.990	0.990	0.995	0.995	0.995
F-measure	0.978	0.978	0.978	0.977	0.978	0.978

表 1 本手法による F-measure

閾値	5	10	15	20	25	30
F-measure	0.619	0.804	0.855	0.869	0.872	0.869

表 2 背景差分法による F-measure

ズムによって抽出されたシルエットを図 10(a)~(d)に示す。このデータでは左脚の下側と右足首の周辺にある影がシルエットとして抽出されている。

4.3.4 結果

本手法によるシルエット抽出結果を図 11(a)~(d)に、シルエットを対象画像に重ね合わせた画像を図 12(a)~(d)に示す(閾値 $T_1 = 3, T_2 = 150, T_3 = 0.995$)。ただし、図 12 では本手法で抽出した対象領域を薄く表した。また、背景差分法によるシルエット抽出結果を図 13(a)~(d)に(閾値 $T = 5, 10, 20, 30 \cdot \text{view 01}$), Horprasert らの手法によるシルエット抽出結果を 14(a)~(d)に示す。

さらに、本手法、背景差分法によるシルエット抽出結果の F-measure をそれぞれ表 1, 2 に示す。Horprasert らの手法では 0.984, 大浜の手法では 0.845 であった。

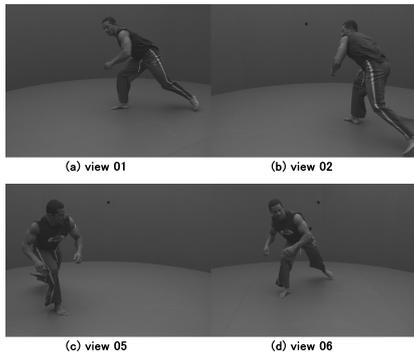


図 9 対象画像

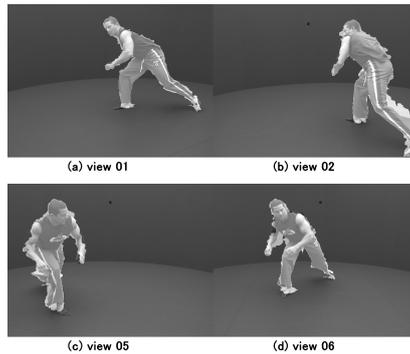


図 12 本手法による抽出結果+対象画像

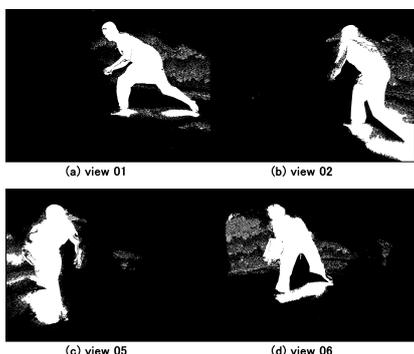


図 10 大浜のアルゴリズムによる抽出結果

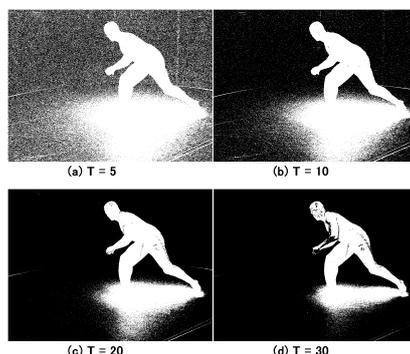


図 13 背景差分法による抽出結果 (view 01)

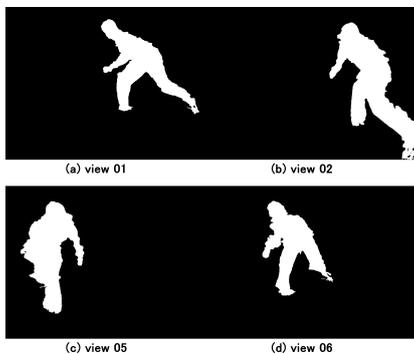


図 11 本手法による抽出結果

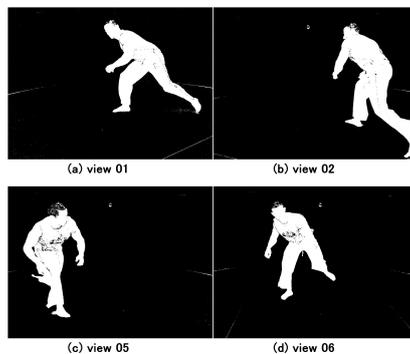


図 14 Horprasert らの手法による抽出結果

閾値 T_1	3	3	3	3	3	3
閾値 T_2	150	200	250	150	200	250
閾値 T_3	0.990	0.990	0.990	0.995	0.995	0.995
F-measure	0.945	0.945	0.945	0.945	0.945	0.945

表 3 本手法による F-measure

閾値	5	10	15	20	25	30
F-measure	0.325	0.553	0.669	0.732	0.774	0.799

表 4 背景差分法による F-measure

4.4 考 察

以上の実験結果から、大浜の手法による抽出結果から本手法により影が削除され、精度が改善されていることがわかる。単視点からの手法と比較すると、背景差分法ではどのように閾値を設定しても対象そのもののシルエットの欠損 (false negative error) と影領域などの背景の過検出 (false positive error) のいずれか、も

しくは両方の影響が顕著に現れているが、本手法を用いるとこれらのエラーが抑制されている。また、Horprasert らの手法ではデータセット 1 において人物の髪が影とみなして削除しているが、本手法では正しく抽出できている。しかし、裾の広がった部分の影が一部削除できずに残った。さらに、データセット 2 では本手法で影領域が削除されているが、腕や背中など人物の体が正しく抽出されておらず、Horprasert らの手法

よりも精度が悪くなった。

この原因としては、人物のテクスチャが背景色と近いために大浜の手法で影以外の部分で多くの欠損と過検出があったことが考えられる。本手法では、対象そのもののシルエットと影領域のみが抽出されていることを前提としているため、このような場合には効果が小さい。また、その他にも visual hull の持つ情報の不完全性が原因として考えられる。visual hull は真の 3 次元形状を包含する凸立体であるため、ある視点からは影と観測されるが別の視点からは対象と観測されるような点が存在する。このため、対象領域と影領域の境界付近では影領域が正しく抽出されない可能性がある。これらの問題を解決するには、本手法で用いた条件をシルエット抽出の段階で適用し、3 次元形状復元と 3 次元の意味の解釈を行いながらシルエットを抽出する必要がある。

その他の原因として、複数視点からの確認が困難な位置に影ができていないことが挙げられる。このような部分に対しては、カメラの数や姿勢を工夫して全ての領域を複数視点から撮影するか、単視点のみの情報を用いて影領域の抽出を行う必要がある。しかし、このような場合であっても、Horprasert らの手法のように全ピクセルに対して画一的に影かどうかの判断を単視点で行う方法と比較した場合、床、すなわち影領域になりえる部分に絞って処理を行うことができる。

5. 結 論

本研究では、視体積交差法で復元した 3 次元形状の表面を分類することによって影を含んだシルエットから影領域を削除する手法を提案し、実画像による実験を行った。その結果、概ね正しいシルエットを抽出することができたが、対象物体の凹面の領域に影ができていた場合に影の検出能力が下がることも確認された。

その原因としては、visual hull の表面を分類するという本手法の処理に起因する部分と、多視点からの撮影画像の不完全さに起因する部分がある。この問題を解決する方法としては、前者に関しては予め視体積交差法で復元した 3 次元形状を用いるのではなく、3 次元形状の復元と影領域の抽出を同時に行うことが考えられる。後者に関しては使用するカメラの台数と復元精度とのトレードオフとなる。

6. 謝 辞

本研究は文部科学省科学技術振興調整費リーディングプロジェクト「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」および総務省戦略的情報通信研究開発推進制度 (SCOPE) プロジェクト「和装文化振興を目的とした実写 3 次元映像のインターネット配信技術の研究開発」の支援を受けて行った。

参 考 文 献

- 1) K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," In Proceedings of ICCV 1999, pp. 307-314
- 2) S. M. Seitz and C. Dyer, "Photorealistic scene reconstruction by voxel coloring," IJCV, 25(3), November 1999
- 3) A. Laurentini, "How far 3d shapes can be understood from 2d silhouettes," IEEE TPAMI, 17(2):188-195, 1995.
- 4) T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara, "Real-time 3d shape reconstruction, dynamic 3d mesh deformation and high fidelity visualization for 3d video," CVIU, 96:393-434, December 2004
- 5) P. Fua and Yvan G. Leclerc, "Using 3-dimensional meshes to combine image-based and geometry-based constraints," Proceedings of ECCV, pages 281-291, 1994
- 6) P. Fua and Yvan G. Leclerc, "Object-centered surface reconstruction: combining multi-image stereo and shading," IJCV, 16(1):35-55, 1995
- 7) P. Fua, "Reconstructing complex surfaces from multiple stereo views," Proceedings of the ICCV, pages 1078-1085, Jun 1995
- 8) G. Cross and A. Zisserman, "Surface reconstruction from multiple views using apparent contours and surface texture," Proceedings of NATO Advanced Research Workshop on Confluence of Computer Vision and Computer Graphics, pages 25-47. Kluwer Academic Publishers, 2000
- 9) C. Hernández and F. Schmitt, "Multi-stereo 3d object reconstruction," Proceedings of the International Symposium on 3D Data Processing, Visualization and Transmission, pages 159-167, Padova, Italy, June 2002
- 10) T. Matsuyama, T. Wada, H. Habe, and K. Tanahashi, "Background subtraction under varying illumination," SCJ, Vol.37, Issue.4, pp.77-88, 2006.4.
- 11) T. Horprasert, D. Harwood, L. S. Davis, "A robust background subtraction and shadow detection" In Proceedings of the 4th ACCV, Vol. 1, January 2000, pp. 983-988
- 12) G. Zeng and L. Quan, "Silhouette Extraction from Multiple Images of an Unknown Background," ACCV(2004)
- 13) 大浜 郁, "多視点画像を用いた誤り検出・訂正機能を持つ高精度対象領域抽出法", 京都大学情報学研究科 修士論文, 2006
- 14) 大浜 郁, 延原 章平, 松山 隆司, "多視点画像を用いた誤り検出・訂正機能を持つ高精度対象領域抽出法", 画像の認識・理解シンポジウム, MIRU2006, pp.688-693, 2006.7.