

## 基準平面への平行投影を用いた多視点画像からの物体領域抽出

Object Extraction from Multi-view Images  
using Parallel Projection to Base Plane

茨木 雅典<sup>†</sup>      坂本 雄児<sup>†</sup>  
Masanori Ibaraki    Yuji Sakamoto

## 1. まえがき

多視点画像から物体の形状を復元する手法の一つとして、視体積交差法 [1](SFS: Shape From Silhouette) がある。これは画像中の物体のシルエットと、撮影したカメラの位置や角度・レンズの歪み係数などのカメラパラメータを用いて、視体積と呼ばれる物体の存在する空間を求め、その共通部分を計算することで形状復元を行う手法である。

SFS ではシルエットを利用して処理を行うため、撮影した画像から物体領域のみを抽出した画像を作成する必要がある。シルエット画像と呼ばれるこの画像は、一般的に背景差分法 [2] によって自動生成されるが、差分を行うためには物体を撮影した画像と同数の背景画像が必要になるほか、時間変化する背景に対して適用できないことや、物体による影の部分で誤抽出が発生することなど、いくつかの問題点がある。

本稿では、物体の置かれている床面への投影変換を利用することで、物体を撮影した画像のみからシルエット画像を生成する手法を提案し、背景差分法との比較を含めた検討を行う。

## 2. 背景差分法

背景差分法とは、あらかじめ背景画像を撮影しておき、得られた画像との差分を求め閾値によって 2 値化し、シルエット画像を作成する手法である。これには様々な応用や改良法が報告されているが、本稿では前述したごく一般的なアルゴリズムによって実装したものと比較する。

## 3. 手法の概要

提案手法の説明をする。本手法は、大きく投影変換による画像変換と投票処理による領域判定の 2 つに分けられる。

## 3.1 投影変換

物体は水平面上に置かれて撮影されていると仮定し、この平面を基準平面と呼ぶ。一般的な撮影環境を想定したときには、この仮定は多くの場合満たされるものであると考えられる。カメラパラメータを用いて各入力画像を基準平面へと投影変換し、投影画像を得る (図 1)。投影変換は次の式によって表される。

$$\mathbf{m}' = \mathbf{A}[\mathbf{R}|\mathbf{t}]\mathbf{M}' \quad (1)$$

ここで  $\mathbf{m}'$  は入力画像における座標、 $\mathbf{M}'$  は基準平面における座標、 $\mathbf{A}$  はカメラの内部パラメータ、 $[\mathbf{R}|\mathbf{t}]$  はカメラの外部パラメータを表す。投影変換を行うことで、基準平面と同一平面上にある点は全ての投影画像上で同じ座標に投影されるが、同一平面上にない点は各投影画像ごとに異なった座標へと投影される。物体は基準平面に対し高さを持つため、各投影画像ごとに異なった座標へと投影されることになる。こうして得られた投影画像群を比較することで、物体領域と背景領域の判別を行う。判別の方法については次節で述べる。なお、カメラパラメータは既知であるとする。

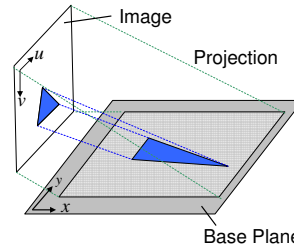


図 1 投影変換

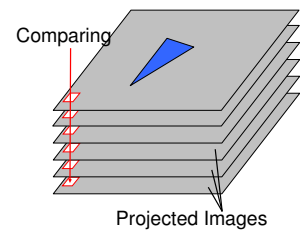


図 2 比較・投票処理

## 3.2 投票処理

物体領域の判別方法について述べる。各投影画像において同じ座標を持つ画素ごとに、特徴量を用いた比較を行い、次の式で表される  $P_k$  と  $V_l$  を求める (図 2)。

$$P_k = \sum_{l=0}^n V_l \quad (2)$$

$$V_l = \begin{cases} 1 & (|I_k(x, y) - I_l(x, y)| < T) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

ここで、 $n$  は入力画像の数、 $I_k(x, y)$ 、 $I_l(x, y)$  はそれぞれ、注目している投影画像および比較する投影画像の  $(x, y)$  座標における特徴量、 $T$  は閾値である。

$V_l$  は  $l$  番目の投影画像との比較の結果、類似した特徴を持つ画素であると判断されたとき  $P_k$  に対し 1 を投票する値であり、 $V_l$  を集計した  $P_k$  はその画素の総得票数となる。ある画素において  $P_k$  が大きくなることは、他の投影画像の同じ座標において類似画素が多く存在することを表し、このような画素は投影によって座標に差が生じない点、すなわち基準平面上の点であり、物体領域である可能性は低いと考えられる。よって、 $P_k$  の値が大きい画素は背景領域であると判定する。以上の処理を全ての投影画像の全ての画素に対して行うことで、物体領域と背景領域を判別する。なお、使用する特徴量には RGB 値、HSV 値を用い、それぞれ独立に処理を行った後、得られた物体領域について共通部分を計算することで最終的な物体領域とする。

この投票処理の後、再度カメラパラメータを用いて平行投影の逆処理を行うことにより、シルエット画像を得る。

## 4. 実験

本手法の有効性を確認するために実験を行った。SFS によって得られるモデルはボクセルモデルであり、ボクセル数やボクセルのサイズによって精度が変化する。本実験のパラメータを表 1 に示す。本実験では、物体形状、物体表面の材質、物体表面のテクスチャ、床面のテクスチャなどを様々に変更して本手法を適用し、物体の抽出が行われていることを確認する。

## 4.1 実験 1 物体形状の抽出

実験 1 では被写体として布製のぬいぐるみを用い、物体の抽出が行われることを確認した。結果を図 3 に示す。(a) は入

<sup>†</sup> 北海道大学大学院情報科学研究科  
Graduate School of Information Science and  
Technology, Hokkaido University

表1 実験パラメタ

Number of Images	8
Image size	1280 × 960 [pixels]
Number of Voxels	400 × 400 × 400 [voxels]
Voxel size	2.0 × 2.0 × 2.0 [mm]

力画像, (b) は得られたシルエット画像である。物体の形状が抽出されていることが確認できる。さらに, 得られたシルエット画像を用いて SFS を行い, 物体形状を復元した結果が (c) である。ぬいぐるみは複雑な形状をしており, 手足の部分では手前にある部分が後方の部分を隠してしまう自己オクルージョンが発生している。得られるモデルは SFS のアルゴリズムに依存するので, オクルージョンによって隠れた部分や凹部などを復元することはできない。

#### 4.2 実験2 背景の変化および影の影響

実験2では背景が変化する場合について検討した。背景差分法では, 背景の撮影と物体の撮影には時間的な隔たりが存在するため, 図4の(a)と(b)のように, 小さなものが移動するなどの変化があった場合, 差分の結果図4(c)のように目的の物体以外にも抽出が行われてしまう。また, 被写体が落とす影による誤抽出が確認できる。提案手法では, 背景画像を用いないため, 背景の変化や物体による影の影響を受けず, (d)に示すように目的の物体のみが抽出される。(e)には実験1と同様, 形状復元を行った結果を示した。

#### 4.3 実験3 床面にテクスチャが存在する場合

画像中の物体抽出手法には, 背景差分法のみならず, 色やエッジ情報を用いて画像の分割を行う手法が数多く存在する。しかし, 図5(a)のように, 床面に様々な色や形のテクスチャが存在する場合, 一般的な領域分割の手法では物体領域のみを正しく抽出することは非常に困難である。本手法では, 物体の三次元位置情報を用いるため, このような場合に対しても物体抽出・形状復元を行うことができる(図5(b))。

#### 4.4 実験4 物体面と床面にテクスチャが存在する場合

図6(a)のように物体・床面ともに様々な色のテクスチャが存在する場合, 実験3よりもさらに物体領域の抽出は困難であるが, 本手法を用いることで物体表面や床面のテクスチャにかかわらず, 物体領域の抽出および形状復元を行うことができることを確認した(図6(b))。

### 5. まとめ

本手法は物体を撮影した画像のみを用いて画像中の物体領域を認識し, 抽出を行う手法であり, 背景画像を必要としないため, 物体や背景が時間変化する場合においても適用が可能である。また, 物体による影の影響や, 物体面や床面に簡単なテクスチャが存在するような場合でも適用が可能であることを確認した。今後の課題として, 床面に近い部分の抽出や, 撮影画像上で比較対象となる画素が少なくなってしまうような場合における処理などについて検討が必要である。

### 参考文献

- [1] H. Baker, "Three-Dimensional Modeling", Proc. IJCAI, Vol. 2, pp.649-655, 1997.
- [2] 波部齊, 和田俊和, 松山隆司, "照明変化に対して頑健な背景差分法", 情報処理学会研究報告, CVIM, [コンピュータビジョンとイメージメディア] 99(29), 17-24, 1999

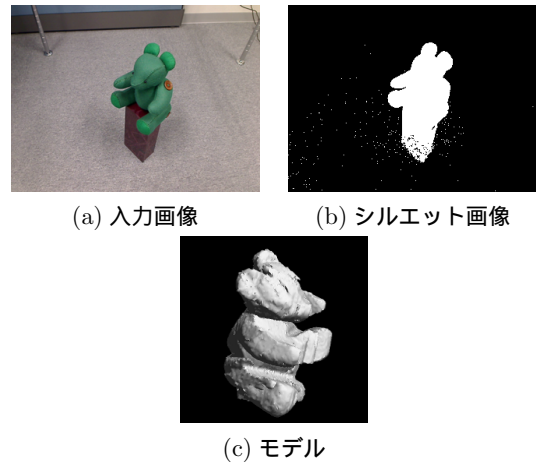


図3 実験1 結果画像

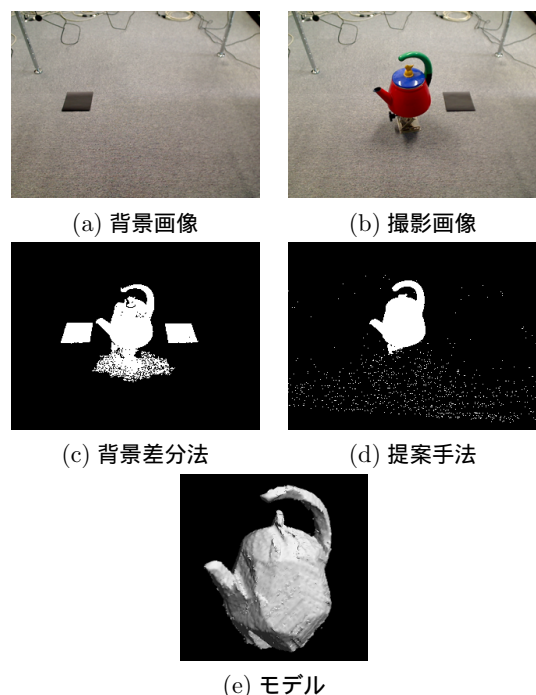


図4 実験2 結果画像

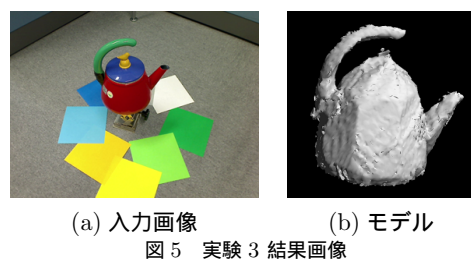


図5 実験3 結果画像

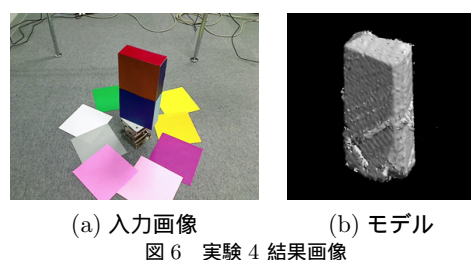


図6 実験4 結果画像