

# X-means 法による Depth 値のクラスタリングと InfiniTAM を用いた隠消現実感

竹嶋 渉太<sup>†</sup> 小池 崇文<sup>†</sup>

法政大学情報科学部<sup>†</sup>

## 1. はじめに

映像中から不要な物体を視覚的に除去する技術は、隠消現実感(DR)と呼ばれる。DR では、除去対象物体を表す除去領域の推定と補完を行う必要がある。除去対象の背景(隠背景)が複雑な形状を持つ場合、その復元が困難であるため、従来のDRシステムでは、背景が壁のような平面を想定するものが多く存在した。本研究では、RGB-Dカメラを用いて作成された3Dメッシュモデルを利用することで、隠背景が複雑な形状を持つ場合に対応したDRシステムを提案する。

## 2. 関連研究

Zokai らの研究では、大きく離れた3つの視点から撮影された静止画像の投影モデルを用いて適切な背景を生成する手法を提案している[1]。この手法では、隠背景が平面の組み合わせによって構成されていると仮定している。Cosco らは、Image Based Rendering を利用して事前にシーンのモデルを作成し、DR を実装している[2]。この研究では、隠背景が数面で構成された既知の幾何形状であると仮定することで、立体的な隠背景であっても復元を可能にした。

このように、隠背景が平面であることを想定するDRの研究が多く存在する。そこで、本研究では Simultaneous Localization and Mapping (SLAM)の1つである InfiniTAM を用いて作成した除去対象周辺の3Dメッシュモデルを利用することで複雑な形状を持つ隠背景に対応する。また、本研究では Depth 値のクラスタリングを用いて Image Inpainting の高画質化手法を提案した海野らの研究[4]を参考に、除去領域の推定を行う。

## 3. 提案手法

図1にシステムの概要を示す。RGB-Dカメラを用いて Kahler らの手法[3]を基に作成したSLAMである InfiniTAM[4]を使用し、除去対象を消した3Dメッシュモデルを作成する。本研究では、これを利用してDRを実現する。

図2に提案手法の流れを示す。まず、1フレーム目取得したDepth画像に対してDepth値のクラスタリングを行い、ユーザーが指定したクラスタを除去領域とする。そこで指定された除去領域内のDepth値をNaNとすること

で、マスク領域を設定する。2フレーム目以降は、InfiniTAMを用いて推定されたカメラの位置姿勢のパラメータを用いて、フレーム毎にカメラの動きに合わせてDepth画像上でマスク領域を再計算する。そして、マスク領域を設定したDepth画像に基づいて除去対象となる物体が存在しないかのような3Dメッシュモデルを作成する。映像中から除去領域に対応する部分を取り除き、3Dメッシュモデルの隠背景部分の色情報をマスク領域と同じ座標に埋めあわせることでDRを実現する。

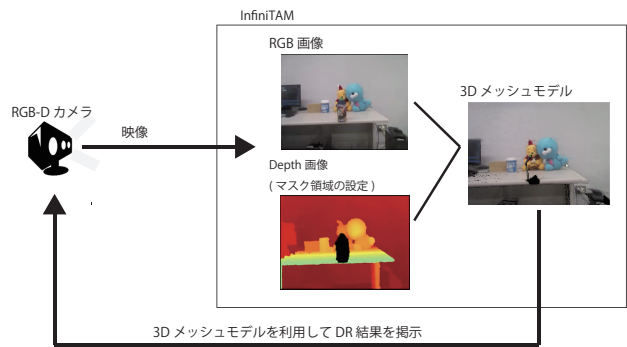


図1:システムの概要

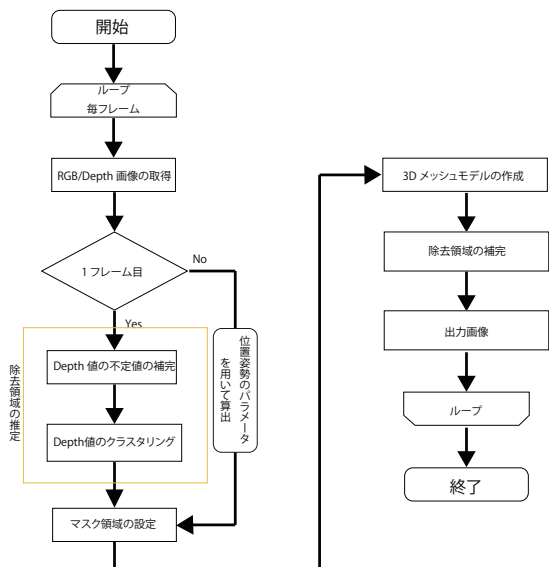


図2:提案手法の流れ

### Diminished Reality with InfiniTAM and Clustering Depth Image by X-means

<sup>†</sup> Shota TAKESHIMA, Takafumi KOIKE

<sup>†</sup> Faculty of Computer and Information Sciences, Hosei University Tokyo, Koganei-shi, 3-7-2 Kajino-chou

### 3.1. 除去領域の推定

除去領域の推定には、Depth 値のクラスタリングを用いる。Depth 画像では、RGB 画像と比べて奥行きの異なる物体の Depth 値が明確に異なるため、物体間のセグメンテーションの精度が向上する。Depth 値のクラスタリングにおいてクラスタ数の事前設定というユーザーの行動をなくすために、クラスタ数が自動で決定可能な x-means 法 [5] を用いる。Depth 値を取得する際、光の反射や RGB-D カメラとの角度が原因で Depth 値が不定値を持つ領域が存在する。不定値が存在してしまうと除去領域を正確に推定できなくなってしまうため、事前に不定値の領域の補完を行う。クラスタリング後、除去対象に対応するクラスタを指定することで除去領域を設定する。

### 3.2. 3D メッシュモデルの作成(マスク領域の設定)

InfiniTAM では、Depth 画像の情報を基に 3D メッシュモデルを作成している。そのため、Depth 画像上の指定された除去領域内の Depth 値を NaN にすることで除去領域のモデルは作成されなくなる。

InfiniTAM は、3D メッシュモデルの作成と同時にカメラの位置姿勢の推定も行っている。この位置姿勢のパラメータを用いて 2 フレーム目以降毎フレーム、カメラの動きに合わせた除去領域の位置姿勢を求め、その部分の Depth 値を NaN にすることで除去対象となる物体が存在しない 3D メッシュモデルを作成する。

### 3.3. 除去領域の補完

RGB 画像上の除去領域部分は、毎フレーム推定される位置姿勢のパラメータを用いて算出されたマスク領域と同じ座標になる。算出された除去領域に隠背景の 3D メッシュモデルの色情報を代入することで、隠背景を重畳した出力画像を得ることができる。

## 4. 実験

### 4.1. 実験内容

除去対象の隠背景が複雑な形状を持つ場合でも本手法が機能するかを確認する実験を行った。背景が複雑な形状になるように、机の上にぬいぐるみなど 3 次元形状を持つものを配置し、除去対象としてペットボトルケースに入ったペットボトルを用意した。RGB-D カメラは、Microsoft Kinect v1 を使用している。RGB-D カメラは、除去対象の正面を向くようにして机から約 70cm 離れた場所に配置し、除去領域の推定を行った。除去領域の推定を行った後は、RGB-D カメラを動かし、さまざまな方向から除去対象を撮影する。

### 4.2. 実験結果

本手法による除去結果を図 3 に示す。図 3(a) と図 3(b) を比較すると映像からペットボトルが除去され、背後にあるぬいぐるみが確認できる。図 3(c) と図 3(d) では、除去対象との距離が近くなるなど視点位置が大きく変化しているにもかかわらず、除去領域にぬいぐるみなどが正しく表示されていることが確認できる。実験により、事前処理などを用いて隠背景の 3 次元形状を取得する必要なく、また、カメラが自由に移動可能な状況で DR を実現できることが示された。

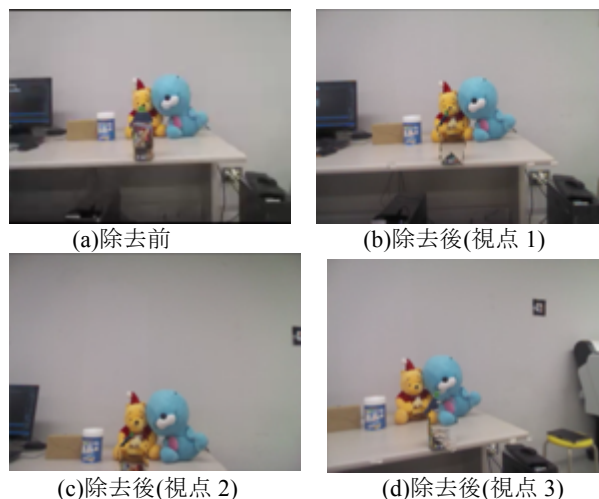


図 3: 除去結果

図 3(d) のように除去領域周辺で除去対象が除去できていない部分が存在する。これは、床など除去対象と接している部分の 3D メッシュモデルが取得できていない、位置姿勢の推定誤差によって除去対象の 3D メッシュモデルが作成できてしまっていることが原因であると考えられる。

## 5. 結論

本研究では、InfiniTAM を用いて作成した 3D メッシュモデルを利用することで、隠背景が複雑な形状を持つ場合でも、事前処理や複数台のカメラを必要としない DR を実現することができた。今後の課題としては、除去領域の推定精度、補完精度の向上が挙げられる。

## 文 献

- [1]. S. Zokai, J. Esteve, Y. Genc, and N. Navab: "Multiview paraperspective projection model for diminished reality," Proc. ISMAR 2003, pp. 217 - 226, 2003.
- [2]. F. I. Cosco, C. Garre, F. Bruno, M. Muzzupappa, and M. A. Otaduy: "Augmented touch without visual obtrusion," Proc. ISMAR 2009, pp. 99 - 102, 2009.
- [3]. 海野彩, 小池崇文, "X-meansによるDepth値のクラスタリングを用いたImage Inpainting," 映像情報メディア学会技術報告, vol. 40, No. 20, pp11-14, 2016.
- [4]. O. Kahler, Victor A. Prisacariu, Carl Y. Ren, Xin Sun, P. Torr, and D. Murray, "Very High Frame Rate Volumetric Integration of Depth Images on Devices", *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, 2015.
- [5]. victorprad/InfiniTAM · GitHub, <https://github.com/victorprad/InfiniTAM>, 2017/1/6
- [6]. 石岡恒憲, "x-means 法改良の一提案:k-means 法の逐次的繰り返しとクラスターの併合," 計算機統計学, vol. 18, no. 1, pp. 3-13, 2006.