

ボロブドゥール寺院壁面レリーフの3次元復元のための 学習データセット作成支援法

Ji Shenyu (立命館大学 大学院 情報理工学研究科)

Li Liang · 長谷川 恭子 · 田中 覚 (立命館大学 情報理工学部)

概要 : ボロブドゥール寺院遺跡のレリーフの一部が補強工事により保存されており、見えない状態にあるため、レーザや多視点写真による3次元計測は大変困難である。各レリーフについては単眼写真が1枚ずつ残されており、この1枚の写真から壁面レリーフの3次元形状を復元するための「深層学習による深度推定手法」が提案されている。しかし、この深度推定手法は深層学習のための学習データセットを必要とし、そのデータセットの作成に多大な時間を浪費する。本研究では、「ボロブドゥール寺院遺跡レリーフ」を対象とした、レリーフの3次元点群復元支援のための高効率な学習データセットの作成支援ソフトウェアを開発した。この支援ソフトウェアを用いることで学習データセットの作成効率が向上した。

キーワード : ボロブドゥール寺院壁面レリーフ, 3次元点群, 学習データセット, デジタルアーカイブ

A Training Dataset Preparation Method for Deep Learning-based 3D Reconstruction of the Borobudur Reliefs

Shenyu Ji (Graduate School of Information Science and Engineering, Ritsumeikan University)
Liang Li, Kyoko Hasegawa, Satoshi Tanaka (College of Information Science and Engineering,
Ritsumeikan University)

Abstract: The Borobudur temple, a UNESCO World Heritage Site, has the most complete collection of Buddhist reliefs. However, some of the reliefs have been buried behind stone walls due to reinforcement. Those hidden reliefs are no longer available for 3D measurement by laser scanning or photogrammetry. In our previous work, we have proposed a deep learning-based depth estimation method to reconstruct 3D point clouds of the hidden reliefs from single monocular photos. However, preparation of the training data requires a lot of labor and time. In this study, we propose an efficient training data preparation tool for supporting the 3D point cloud reconstruction of the hidden Borobudur reliefs. The proposed method has greatly improved the efficiency of training data preparation.

Keywords: Digital archive, Borobudur, Heritage relief, 3D point cloud, Deep-learning Dataset

1. 序論

近年、3次元点群計測技術の発展と共に、文化財のデジタルアーカイブ化が効率的に行えるようになった。文化財の実物が存在する場合には、レーザ計測や写真計測などの3次元計測が用いられる機会が増えている。そのため、3次元計測によって取得された大規模な3次元点群がデジタルアーカイブに用いられるようになり、点群の可視化の必要性が高まっている。一方、ボロブドゥール寺院壁面レリーフの一部のような、補強工事のため石垣の中に隠された場合、3次元計測により実物を計測することは困難である。しかし、実物が破壊したり、紛失したり、消失している場合でも、写真のみが残されている場合がある。このような場合では、複数枚の写真が存在すれば、多視点による3次元デジタル復元が可能となる。しかし、写真が一枚のみ残され

た記録も多い。これらに対し有効な3次元復元の手法が必要である。

本研究はボロブドゥール寺院遺跡のデジタルアーカイブを目的としている。ボロブドゥール寺院遺跡の保有しているデータとして、レーザや写真による3次元計測データ、地下の基礎工事部分に関する土台の2次元CAD設計図、レリーフの単眼古写真などがある。レリーフは補強工事のため石垣の中に隠された部分があるため、古写真から深層学習を用いて3次元データを復元する研究が近年行われている[1,2]。この復元手法には学習データセットが必要であり、学習データセットは計測された点群から学習データ(RGB写真とデプスマップ)を作成する。学習データの作成は、3次元点群編集ソフトウェアを用いて各レリーフ部分のみ抽出した後、画像生成作業のために手作業による視点調整が必要になるため、効率が非常に低いという問題点がある。寺院壁面の3次元データは非常に巨大であり、手動で点群を編集す



図 1 ボロブドゥール寺院遺跡



図 2 ボロブドゥール寺院レリーフ



図 3 隠されているレリーフ



図 4 レリーフ (22 番) の投影写真 (上) と
デプスマップ (下)

ることは容易ではない。また、抽出したレリーフ点群は向きも様々であるため、画像生成のために、投影方向がレリーフに対して正確に真正面を向いている必要がある。このように、学習のためのデータセットを作成する事は効率が非常に低い作業を必要とする。

本研究では学習データセットの作成効率を向上するために、自動的にデータセットを作成するソフトウェアを開発する。このソフトウェアは計測点群データから必要なレリーフ領域をユーザが指定し、正面の判定や画像の生成は自動で行う、半自動の支援ソフトを目指す、これにより、学習データセット作成の時間短縮を目指す。

2. ボロブドゥール寺院遺跡

本研究の実験対象は、インドネシアジャワ島の中中部にある「ボロブドゥール寺院遺跡」(図 1)である。ボロブドゥール寺院遺跡は世界最大級の仏教寺院遺跡で、ユネスコ世界文化遺産にも登録されている。この寺院の一階から五階までの回廊には、建造物の壁面に彫られた、1,460 面の仏教説話に基づくレリーフ(図 2)がある。これも今世界中最大級の仏教説話レリーフのコレクションである。

総延長 5km におよぶ方形壇の回廊には、仏教説話にもとづいた 1460 面におよぶ浮彫彫刻レリーフが時計回りにつづいており、登場人物は 1 万人におよぶとされている。同様に 1212 面の装飾浮彫には、天人や羅刹、鳥獣、植物文様およびインド神話に登場する伝説上の鳥獣などがみられ

る。なお、外層、内層ともに四方に階段をもち、各面いずれも全く同形同構造で、どれを正面とするかわからない、幾何学的に均質な構造となっている[3]。しかし、総計 1,460 面の仏教説話に基づくレリーフの中で、第 1 階層の一部である Karmawibhanga レリーフ(156 面)が補強工事により石垣の中に埋まって見えない状態になる(図 3)。ただし、埋まった各面のレリーフに対しては、古写真一枚ずつのみ残されている状態になっている。

3. 関連研究

3.1.3 次元復元の関連研究

レーザー計測[4]技術の発展と共に、無傷または欠陥のある文化遺産のデジタル保存が効率的に行えるようになった[5]。しかし、ボロブドゥール寺院壁面レリーフのような、補強工事のため石垣の中に隠された場合、レーザー計測により復元するのは不可能である。3次元情報を取るため、複数の画像を使用し、写真計測[6]により3次元モデルを復元する手法もある[7]。しかし、ボロブドゥールレリーフのような複数の画像が残っていないような、単眼写真のみが残る例も多くある。そのため、単眼写真から文化遺産を3次元復元するための効率と精度に優れた手法が提案された[1]。この論文では、ボロブドゥールレリーフを対象として、「単眼写真のみ残された文化財」に対し、効率的な3次元デジタル化と可視化手法が提案され、ボロブドゥール寺院遺跡を対象として提案手法の有用性が証明された。

提案手法で使ったレリーフデータセットは、レリーフの単眼写真とその写真と対応するデプスマップのペアで構成される。デプスマップは単眼写真と同じ解像度であり、各点の画素値を深度値に替えた画像データである。学習データセットに関しては次節で説明する。

Panらによって提案された手法は「エッジ強調ネットワーク」と「残差ネットワーク」の2つのニューラルネットワークモデルを使って古写真からデプスマップを推定している。[1,2]. 深度推定モデルを学習させるには、「ボロブドゥール寺院遺跡」のレリーフに基づいて、データ量の影響を判明するため、比較実験を行い、データ量の拡大より、精度が向上できることを明らかにした。また、二つのネットワークモデルを使用して深度推定および3次元復元を行った。その結果、「エッジ強調ネットワーク」の精度は90%で、画像のローカル情報、つまりレリーフの人物のような細部の微細な変化を捉えることができる。「残差ネットワーク」の深度推定の性能は精度としては95%の結果を得ている。しかし、ローカル情報は「エッジ強調ネットワーク」より少なく、レリーフの細部の深度が再現できていないことが判明した。上記の比較実験展望を通して、データ量の拡大とネットワーク層数の増加により復元精度の向上が期待されることが分かった[8-19].

3.2 深層学習のデータセット

「深層学習によるボロブドゥールレリーフの深度推定手法」[6, 7]は、教師あり学習であり、2次元写真の画素値と深度値（点群データの奥行き値）の対応関係を学習する。学習したこのマッピング関係はデプスマップで表現される[6].

マッピング関係の深度値とは、2次元写真の各ピクセルに対応する計測点群の各点の深度について、レリーフの最小深度～最大深度を0～255に正規化したものである。深度値を推定したあと、深度値の線形変換によって3次元Z軸の座標値を算出し、単眼写真の3次元モデルが復元できる。つまり、本研究の教師あり深層学習の手法に必要な「教師」といえるグラウンドトゥルースは各点の深度値である[7].

「深層学習によるボロブドゥールレリーフの深度推定手法」は、レリーフデータの学習データ

セットとして、レリーフの単眼写真とその写真に対応するデプスマップのペアが用いられる。デプスマップは単眼写真と同じ解像度の各点の画素値を深度値に替えた画像データである。図4にレリーフの単眼写真と対応のデプスマップの一例を示す。デプスマップの濃淡はレリーフの凹凸を表す。色が白いほどカメラとの距離が遠くなり、黒いほど距離が近くなる。例えば、図4で示されているように、背景部分は、人物が描かれた部分よりもカメラから遠いため、白く描かれていることがわかる。

4. 学習データセット作成支援ソフトウェアの開発

本研究では、OpenGLを主要なコンピュータグラフィックスライブラリとして、OpenCVやEigenなどのライブラリを利用する。デジタルアーカイブされた「ボロブドゥール寺院遺跡」の3次元点群データを対象として、深層学習による深度推定手法におけるボロブドゥール寺院遺跡レリーフの3次元点群復元のために効率的なデータセット作成支援ソフトウェアを開発する。ソフトウェアは図5に示すフローチャートに従って以下のように処理を行う：

- ステップ1：レリーフ点群データの読み込み
- ステップ2：注目領域の選択
- ステップ3：注目領域の抽出
- ステップ4：レリーフの正面判定
- ステップ5：レリーフの正規化
- ステップ6：写真化とデプスマップ化(投影)

ここで、以上のステップにおいては、ステップ2のみ手動作業が必要となる以外は自動で処理を行う。以下に各ステップで提案する手法について詳述する。

ステップ1：レリーフ点群データの読み込み

本ソフトウェアで取り扱う点群データのファイルフォーマットはCG分野で取り扱うことが多いPLYファイルフォーマットを採用する。PLY形式では、オブジェクトを頂点、面、およびその他の要素の集まりとして記述し、これらの要素に付加できる色や法線方向などのプロパティも含めることが可能である。本研究では、点群データを対象としている。点群データのPLYファイルの場合は、基本的にレーザ計測や写真計測な

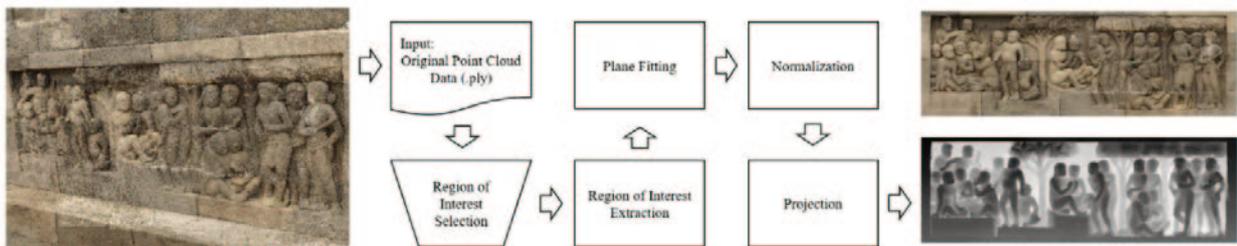


図5 フローチャート

どの3次元計測が用いられ、点の位置情報(x, y, z)と色情報(R, G, B)以外はないため、読み込むデータは点の位置情報と色情報だけを対象とする。PLY ファイルから点群の位置情報と色情報を読み込んだ後、OpenGL に点を一つずつレンダリングする。

ステップ2：注目領域の選択

本論文では、抽出や処理したいレリーフを注目レリーフと呼び、注目レリーフの周りの情報(点群)を注目領域と呼ぶこととする。注目領域を選択するため、まず四つの点をユーザーが選択する。選択した四つの点から構成された四角形でレリーフを囲む。レリーフの形状把握と正面判定のため、必ずレリーフの「左上」「右上」「右下」「左下」の順番で点を選択する。

注目領域は基本的に選択された四つの点を基準とする。注目レリーフを抽出するために、注目レリーフの形状や位置が必要であり、選択された四つの点は注目レリーフを全体的に把握するために必要である。また、レリーフの点群データに対して、レリーフの正面か裏面を判別するのは困難である。本ソフトウェアでは、正面を判別するために、四つの点の選択順を定義することで、正面かどうかを確定している。「左上」「右上」「右下」「左下」の順番で点を選択すれば、レリーフの正面方向は選択順によって自動的に判別することができる。

ステップ3：注目領域の抽出

注目領域の抽出では、レリーフの周りの点群データを収集する。注目領域の抽出が必要な理由としては、入力した点群データはデータサイズが大きくなるため、不要な点を除外することで処理の軽量化を図るためである。注目領域の抽出とは、入力データから、ユーザーが選択した四つの点によって一定の領域を確保して点の情報を集める、つまり領域以外の点が絶対要らないということであり、計算量の減少とレリーフ形状把握可能のため、領域以外の点を削除するのは注目領域の抽出という過程である。

提案手法での注目領域の抽出方法について説明する。図6のようにまず選択した四つの点(左上1番目の点(x₁, y₁, z₁), 右上2番目の点(x₂, y₂, z₂), 右下3番目の点(x₃, y₃, z₃), 左下4番目の点(x₄, y₄, z₄))の座標を平均して注目領域の中心(x_c, y_c, z_c)を

$$\begin{cases} x_c = \frac{x_1 + x_2 + x_3 + x_4}{4} \\ y_c = \frac{y_1 + y_2 + y_3 + y_4}{4} \\ z_c = \frac{z_1 + z_2 + z_3 + z_4}{4} \end{cases} \quad (1)$$

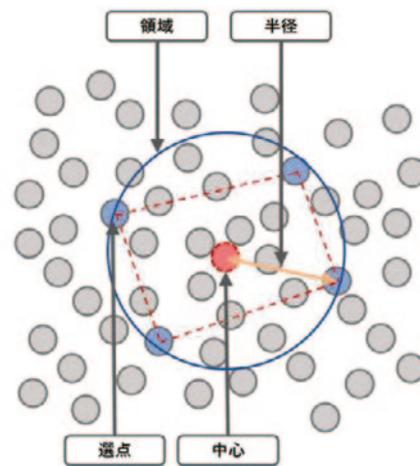


図6 注目領域の抽出

で計算する。四つの点の中で中心から一番遠い距離を計算して、この距離を領域の半径とする。入力データの全点に対して、中心までの距離が領域半径より小さければ、この点を保留、中心までの距離が領域半径より大きい点は削除する。これにより、保留した点が球体になり、球体領域の情報が得られる。

ステップ4：レリーフの正面判定

「注目領域の抽出」で抽出したのは注目レリーフを含んだ球体領域の点群であり、レリーフがどの方向を向いているかはわからない状態である。レリーフの向きを把握のために、まずレリーフの正面判定が必要になる。ここで、正面判定とは、レリーフに平面を当てはめて、その平面の法線ベクトルを計算することである。

本研究では、特異値分解を用いてレリーフ点群の正面判定を行う。まず、「注目領域の抽出」で収集した点をn×3の行列として、各点から平均値(x̄, ȳ, z̄)を引いた行列Aを:

$$A = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} & z_1 - \bar{z} \\ x_2 - \bar{x} & y_2 - \bar{y} & z_2 - \bar{z} \\ \vdots & \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} & z_n - \bar{z} \end{pmatrix}$$

のように定義する。行列Aを特異値分解すると、

$$A = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{pmatrix}$$

のようになる。ここで、当てはめられた平面の法線ベクトル n は最小特異値σ₃に対応するベクトルであり、v₃が法線ベクトルとして採用される。

ステップ5：レリーフの正規化

「正面判定」ではレリーフ点群平面の法線ベクトルが得られた。つまりレリーフ点群の位置と向き方向が確定したことになる。入力データから抽出したレリーフ点群は入力データの世界座標系をもち、位置や向きはレリーフの元々の位置に依存する。「正規化」はレリーフの位置と向き方向を一致する過程である。

本ソフトウェアにおいて「正面」とはZ軸に対して垂直な向きと定義する。そのため、ステップ4で求めたレリーフ点群の法線ベクトルをZ軸に水平になるように以下の操作を行う。

まず、レリーフ点群の中心を原点とするように、式(1)で計算した中心座標を利用して平行移動の変換行列を

$$T = \begin{pmatrix} 1 & 0 & 0 & -x_c \\ 0 & 1 & 0 & -y_c \\ 0 & 0 & 1 & -z_c \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

で求める。次に、原点を中心にして法線ベクトルがZ軸に水平になるように、回転行列を求める。回転行列は法線ベクトル \mathbf{v}_3 を利用して回転行列を算出する。まずは法線ベクトルをX-Y平面への投影してX軸との角度を φ とすれば、回転行列

$$R_z = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 & 0 \\ -\sin \varphi & \cos \varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

が求められる。行列 R_z を作用すれば、Z軸を中心としてZ-X平面に回転させる事が出来る。次に、法線ベクトルとZ軸の角度を θ とせば、

$$R_y = \begin{pmatrix} \cos \theta & 0 & -\sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

が得られる。行列 R_y を利用してY軸を中心として回転すれば、法線ベクトルがZ軸と同一になる。行列 R_z と行列 R_y を合成するとした行列：

$$R' = R_y R_z$$

により、レリーフの向きはZ軸に調整できる。

続いて、レリーフの角度再調整を行う。ここまでアフィン変換によって、レリーフの向きはZ軸に大して垂直になったが、X-Y平面上においては、X軸またはY軸に対して傾いている可能性がある。そのため、ユーザが選択した四つの点を利用して再調整する。4点の位置情報からX軸との傾きを

$$\alpha = \tan^{-1} \frac{\frac{(y_2 - y_1) + (y_3 - y_4)}{2}}{\frac{(x_2 - x_1) + (x_3 - x_4)}{2}}$$

で求めれば、回転行列は、

$$R_{adj} = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 & 0 \\ -\sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

で求められる。従って、最終的な回転行列は、

$$R = R_{adj} R_y R_z$$

で定義される。

以上より、正規化のための計算としては、正規化後の点の位置 \mathbf{x}_p は各レリーフ点群位置 \mathbf{x}_i に対して、

$$\mathbf{x}_p = R T \mathbf{x}_i$$

を計算すればよい。以上がレリーフの正規化である。

ステップ6：写真化とデプスマップ化(投影)

正規化したレリーフの3次元点群を回転や移動させ、レリーフの近似平面はすでにX-Y平面に当てはめている。つまり、レリーフの点群にある点の深度値はZ軸の値となり、2次元のデプスマップに投影することが容易にできる。

点群には3次元 (x, y, z) 座標があり、写真には2次元 (x, y) 座標しかない。しかし、投影する前にレリーフの点群はすでに正規化し、Z軸の値が深度値である。つまりここで点の (x, y) 座標を抽出すれば投影した後の2次元座標に対応できる。そのため、点のRGB値や深度値に対応して写真化とデプスマップ化ができる。

つまり、点 (x_p, y_p, z_p) に対してデプスマップに投影後の位置 (x_i, y_i) を以下のように算出できる。

$$x_i = \frac{x_p - x_{min}}{x_{max} - x_{min}} W$$

$$y_i = \frac{y_p - y_{min}}{y_{max} - y_{min}} H$$

$$D_i = \frac{z_p - z_{min}}{z_{max} - z_{min}}$$

ここで、投影画像の横幅はW、縦幅はHである。

「深層学習によるボロブドゥールレリーフの深度推定手法」に適用するため、投影した2次元写真とデプスマップの解像度は $3,200 \times 1,024$ が要求される。また、一面のレリーフ点群は約 5×10^6 点であり、 5×10^6 点から約 3×10^6 の画素



図 7 実験 1 の入力データ



図 9 実験 2 の入力データ

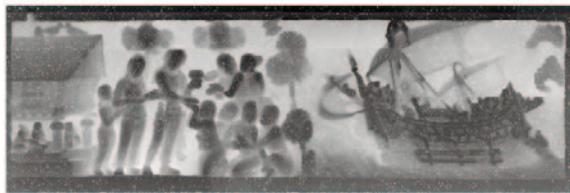


図 8 実験 1 の生成データセット



図 10 実験 2 の生成データセット

値に投影するにはダウンサンプリングの必要がある。つまり複数の点を一つの画素に投影する場合がある。つまり複数の点の (x, y) 座標値が近く、 z 座標値が異なるという状況である。「深層学習によるボロブドゥールレリーフの深度推定手法」への適用のために、手前にある点は後ろにある点より色や深度の表現が顕著であるため、複数の点を一つの画素に投影する場合は一番手前にある点の RGB 値と深度値を採用する。以上が、本研究で考案したダウンサンプリングの手法である。

5. 適用実験

本研究では、効率的なレリーフ学習データセットの作成手法を提案した。提案手法の実用性を証明するため、データセット作成支援ソフトウェアを作成した。支援ソフトウェアで実行して生成した結果画像を示し、また、点数が違う点群データから 1 つのレリーフを抽出するために必要な時間を提示する。最後は提案手法として作成した学習データセットの正確性を証明するため、検証実験を行った。

5.1 単眼写真とデプスマップの生成

本実験は壁面レリーフの入力点群データから、一つのレリーフを抽出し、単眼写真とデプスマップを生成する実験である。

実験 1

実験 1 として用いる点群データを図 7 に示す。図 7 の右側で示したように実験 1 で用いられるデータは二つのレリーフが上下に並んでいる。図 8 は壁面の下部分のレリーフを抽出して写真化(図 8 の上部分) とデプスマップ(図 8 の下部分) 化した結果である。

実験 2

実験 2 として用いる点群データを図 9 で示す。図 9 の左側で示したように実験 2 で用いられるデータは壁の一角であり、角の両側に複数のレリーフが並んでいることがわかる。

図 10 は図 9 の右側の中心に示しているレリーフを抽出し写真化(図 10 の上部分) とデプスマップ(図 10 の下部分) 化した結果である。

5.2 処理時間テスト

本研究の実験で使用している実行環境を表 1 に示す。

表 1 実行環境

OS	MacOS
CPU	Intel Core i7 3.1GHz
RAM	16 GB
GPU	NVIDIA Geforce GT 750M 1024M

規模が異なる 6 種類の点群データに対して実験を行った。レリーフの 1 つ(約 3 百万～5 百万点)を抽出して、学習データセットを作成するのに要した時間を表 2 に示す。

表2 実行時間 (単位: 秒)

Data	Number of Points	Relief Processing		Projection	Total
		Fitting	Norm.		
A	5,138,403	1.151	16.427	0.167	17.745
B	51,860,878	5.680	59.327	0.276	65.283
C	55,411,392	5.735	60.107	0.229	66.071
D	65,519,391	3.927	32.378	0.158	36.463
E	167,015,996	13.296	99.710	0.516	113.522
F	357,670,619	46.416	42.932	0.164	89.512

従来のデータセット作成手段では、3次元点群を編集するソフトウェアを用いて手動作業に多大な時間を要し、レリーフの学習データセットを作成するために、作業効率が非常に低いという問題点があった。このような手作業を主とする従来の手法に対して、効率テストをするのは困難なため、従来の学習データセット作成する時間としては、我々の作業時間を比較のために記載し、提案手法との比較を表3に示す。

表3 実行時間 (比較)

Data	Number of Points	Conventional Method	Proposed method
A	5.1×10^6	3 hours	17 s
B	5.1×10^7	1 day	65 s
C	5.5×10^7	1 day	66 s
D	6.5×10^7	1 day	36 s
E	1.6×10^8	2 days	113 s
F	3.5×10^8	>3 days	89 s

5.3 正確性の検証

提案手法として作成した学習データセットの正確性を証明するため、検証実験を行った。検証実験とは、処理時間テストで作成した6つのデプスマップを用いて再現した点群データを生成する。生成した点群データ(以下で生成点群と呼ぶ)と壁面から直接抽出した点群データ(以下で抽出点群と呼ぶ)のC2C distanceを計算す。図11は1つのレリーフについてC2C distanceの大きさを色で示した結果である。正確性の判断基準は生成点群と抽出点群を比較し、点と点の距離を計算し、距離の平均値を誤差として正確性を検証する。結果は表4に示す。

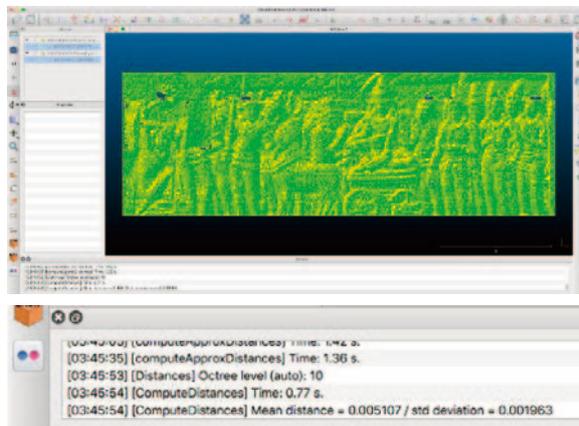


図11 入力点群と生成点群間のC2Cdistance

表4 検証実験結果 (平均距離)

Data	A	B	C	D	E	F
Dis.	0.0031	0.0052	0.0051	0.0067	0.0079	0.0048

5.4 考察

本章では学習データセット作成支援ソフトウェアの作成効率と正確性を検証するため、6つの実験を行い、それぞれ規模が異なる入力点群データに対して実験した。最大のデータ(実験F)は3.5億点を超え、約60GBのファイルであり、非常に巨大なデータである。本実験によって提案手法の有用性が実証され、非常に巨大なデータに対しても2分以内で計算完了できるということも分かった。

また、処理時間について考察すると、計算時間はオリジナルデータの大きさに影響しているが、完全に大きさに関係しているわけではない。表2の実験CとDを見よう。実験DのデータはCより大きいですが、処理の合計時間はDの方が早い。その原因として、計算時間はレリーフの点密度に関係すると考えている。

最後は作成したデータセットの正確性を検証するための検証実験を行った。結果として、生成した点群データの平均誤差は0.05程度であり、レリーフ1つあたりの横幅は約24である。つまり、例えば一面のレリーフの横幅は24メートルだったら、平均誤差は5ミリのレベルであり、誤差が非常に低いということが分かった。

6. 結論

本研究では、デジタルアーカイブされた「ボロブドゥール寺院遺跡」の3次元点群データを対象として、「深層学習によるボロブドゥールレリーフの深度推定手法」のための効率的な学習データセット作成支援手法を提案した。さらに、OpenGLを主要なコンピュータグラフィックスライブラリとして、OpenCVやEigenなどのライブラリを利用し、ボロブドゥール寺院遺跡レリーフの3次元点群復元のための学習データセット作成支援ソフトウェアを作成し、提案手法の実行可能性を証明した。学習データセット作成支援ソフトウェアの作成効率と正確性を検証するため、6種類のそれぞれ規模が異なる入力点群データを用いて実験した。処理時間については、非常に巨大なデータに対しても短時間で計算完了ができ、学習データセットの作成効率が向上した。また、生成した学習データセットの平均誤差が微小であり、学習データセットの作成精度と正確性が検証した。

今後の展望としては、学習データセット作成支援として、学習データセットの作成効率と作成制度を更に向上していくために、更に巨大な入力点群データに対応できるように改善する。また、支援ソフトウェアとして、簡明なインターフェースを使用し、ユーザに対して使いやすいソフトウェアに改善することである。

参考文献

- [1] Jiao Pan, Liang Li, Hiroshi Yamaguchi, Kyoko Hasegawa, Fadjat I. Thufail, Brahmantara, Satoshi Tanaka, Fused 3D Transparent Visualization for Large-Scale cultural Heritage Using Deep Learning-Based Monocular Reconstruction. XXIV ISPRS Congress, 2020.
- [2] Jiao Pan, Liang Li, Hiroshi Yamaguchi, Kyoko Hasegawa, Fadjat I. Thufail, Brahmantara, Satoshi Tanaka, 3D Transparent Visualization of Relief-Type Cultural Heritage Assets based on Depth Reconstruction Of Old Monocular Photos, 19th Asia Simulation Conference(AsiaSim2019), Singapore, October 30-November 1 (October 31), 2019. CCIS, volume 1094, pp 187-198, 2019.
- [3] NHK 取材班ほか『NHK 美の回廊をゆく 東南アジア至宝の旅』日本放送出版協会, 1991年. ISBN 4-14-009156-8.
- [4] Robert Gregor, Ivan Sipiran, Georgios Papaioannou, Tobias Schreck, Anthousis Andreadis, and Pavlos Mavridis. Towards Automated 3D Reconstruction of Defective Cultural Heritage Objects. Eurographics Work. Graph. Cult. Herit., pages 135–144, 2014.
- [5] Renato Hermoza and Ivan Sipiran. 3D Reconstruction of Incomplete Archaeological Objects Using a Generative Adversarial Network. nov 2017.
- [6] Ioannides, A. Hadjiprocopi, N. Doulamis, A. Doulamis, E. Protopapadakis, K. Makantasis, P. Santos, D. Fellner, A. Stork, O. Balet, M. Julien, G. Weinlinger, P. S. Johnson, M. Klein, and D. Fritsch. Online 4d reconstruction using multi-images available under open access. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., 2(5/W1):169–174, 2013.
- [7] Georgia Kyriakaki, Anastasios Doulamis, Nikolaos Doulamis, Marinos Ioannides, Konstantinos Makantasis, Eftichios Protopapadakis, Andreas Hadjiprocopis, Konrad Wenzel, Dieter Fritsch, Michael Klein, and Guenther Weinlinger. 4D Reconstruction of Tangible Cultural Heritage Objects from Web-Retrieved Images. Int.J.Herit.Digit.Era, 3(2):431–451, jun 2014.
- [8] Salem A, Williams S, Fairhead J D, et al. Tilt-depth method: A simple depth estimation method using first-order magnetic derivatives[J]. The leading edge, 2007, 26(12): 1502-1505.
- [9] Renato Hermoza and Ivan Sipiran. 3D Reconstruction of Incomplete Archaeological Objects Using a Generative Adversarial Network. nov 2017.
- [10] Min Lu, Bo Zheng, Jun Takamatsu, Ko Nishino, and Katsushi Ikeuchi. Preserving the Khmer smile: classifying and restoring the faces of Bayon. Proc. 12th Int. Conf. Virtual Reality, Archaeol. Cult. Herit., pages 161–168, 2011.
- [11] Jin Ho Park, Tufail Muhammad, and Ahn Jae-Hong. The 3D reconstruction and visualization of Seokguram Grotto World Heritage Site. Proc. 2014 Int. Conf. Virtual Syst. Multimedia, VSMM 2014, pages 180–183, 2014.
- [12] Beyang Liu, Stephen Gould, and Daphne Koller. Single image depth estimation from predicted semantic labels. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pages 1253–1260, 2010.
- [13] Miaomiao Liu, Mathieu Salzmann, and Xuming He. Discrete-Continuous Depth Estimation from a Single Image. In 2014 IEEE Conf. Comput. Vis. Pattern Recognit., pages 716–723. IEEE, jun 2014.
- [14] Thomas P. Kersten and Maren Lindstaedt. Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services Automatische 3D-Objektrekonstruktion aus digitalen Bilddaten für Anwendungen in Archit. Photogramm. - Fernerkundung - Geoinf., 2012(6):727–740, 2013.
- [15] David Eigen, Christian Puhrsch, and Rob Fergus. Depth Map Prediction from a Single Image using a Multi-Scale Deep Network, 2014.
- [16] Anirban Roy and Sinisa Todorovic. Monocular Depth Estimation Using Neural Regression Forest. In 2016 IEEE Conf. Comput. Vis. Pattern Recognit., pages 5506–5514. IEEE, jun 2016.
- [17] Guizilini V, Ambrus R, Pillai S, et al. 3d packing for self-supervised monocular depth estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 2485-2494.
- [18] Thurston J B, Smith R S, Guillon J C. A multimodel method for depth estimation from magnetic data[J]. Geophysics, 2002, 67(2): 555-561.
- [19] Zhuo W, Salzmann M, He X, et al. Indoor scene structure analysis for single image depth estimation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern recognition. 2015: 614-622.