



Article

Integrated High-Definition Visualization of Digital Archives for Borobudur Temple

Jiao Pan ^{1,*}, Liang Li ² , Hiroshi Yamaguchi ³, Kyoko Hasegawa ², Fadjar I. Thufail ⁴, Brahmantara ⁵ and Satoshi Tanaka ²

- ¹ Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan
² College of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan; liliang@fc.ritsumei.ac.jp (L.L.); hasegawa@media.ritsumei.ac.jp (K.H.); stanaka@media.ritsumei.ac.jp (S.T.)
³ Nara National Research Institute for Cultural Properties, Nara 630-8577, Japan; yamaguchi-h62@nich.go.jp
⁴ Research Center for Area Studies, National Research and Innovation Agency, Jakarta 12710, Indonesia; fadj001@lipi.go.id
⁵ Borobudur Conservation Office, Magelang 56553, Indonesia; brahmantara@kemdikbud.go.id
* Correspondence: gr0342ir@ed.ritsumei.ac.jp

Abstract: The preservation and analysis of tangible cultural heritage sites have attracted enormous interest worldwide. Recently, establishing three-dimensional (3D) digital archives has emerged as a critical strategy for the permanent preservation and digital analysis of cultural sites. For extant parts of cultural sites, 3D scanning is widely used for efficient and accurate digitization. However, in many historical sites, many parts that have been damaged or lost by natural or artificial disasters are unavailable for 3D scanning. The remaining available data sources for these destroyed parts are photos, computer-aided design (CAD) drawings, written descriptions, etc. In this paper, we achieve an integrated digital archive of a UNESCO World Heritage site, namely, the Borobudur temple, in which buried reliefs and internal foundations are not available for 3D scanning. We introduce a digitizing framework to integrate three different kinds of data sources and to create a unified point-cloud-type digital archive. This point-based integration enables us to digitally record the entire 3D structure of the target cultural heritage site. Then, the whole site is visualized by stochastic point-based rendering (SPBR) precisely and comprehensibly. The proposed framework is widely applicable to other large-scale cultural sites.

Keywords: cultural heritage; digital archive; 3D reconstruction; transparent visualization; the Borobudur temple



Citation: Pan, J.; Li, L.; Yamaguchi, H.; Hasegawa, K.; Thufail, F.I.; Brahmantara; Tanaka, S. Integrated High-Definition Visualization of Digital Archives for Borobudur Temple. *Remote Sens.* **2021**, *13*, 5024. <https://doi.org/10.3390/rs13245024>

Academic Editors: Nikos Grammalidis, Kosmas Dimitropoulos and Anastasios Doulamis

Received: 14 November 2021

Accepted: 7 December 2021

Published: 10 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recently, the three-dimensional (3D) recording of cultural heritage (CH) has attracted enormous interest worldwide [1]. The complete 3D recording process of CH usually contains five steps: 3D digitization, 3D data processing, archiving, visualization, and reproduction [2–4]. The steps up to archiving address the problem of permanent preservation of CH. Visualization is an important way to analyze CH, and reproduction involves virtual representation, application, post-disaster reconstruction, etc. In this study, we focus on the 3D digitization and visualization of large-scale CH assets and use the Borobudur temple in Indonesia, a UNESCO World Heritage site, as our experimental subject.

Herein, we consider three main requirements in digitally recording large-scale CH assets. First, due to the huge workload, the digitization method and the visualization method of the large-scale CH should be as efficient as possible because the efficiency determines the feasibility of the entire project. Second, many historical CH assets can be partially or even completely damaged, destroyed, or lost by natural and man-made disasters [5]. Thus, it is necessary to execute digital reconstruction from traditional documenting artifacts, including manual drawings, photographs from several different angles, and written materials. Third, for complete visualization, we need to integrate (1) the 3D scanned data of the

extant parts and (2) the 3D reconstructed data of the missing parts. The latter includes 3D data reconstructed from photographic images and computer-aided design (CAD) drawings modeled based on academic surveys of the target CH asset.

For the first requirement, laser scanning and photogrammetry [6–11] are available and widely used techniques that can generate realistic 3D models with high geometric accuracy. In the case of extant complete large-scale CH sites, a laser system can extract the geometry of the site surfaces. However, laser scanning usually requires expensive and large equipment. In contrast, photogrammetry can implement high-quality and reliable models of large-scale targets with a good-quality camera and standard personal computer. Developments in software for photogrammetry assistance also lead to increased automation and a higher performance [12]. Thus, we adopt photogrammetry in this study due to the site-specific constraints (refer to Section 3.2). However, only using photogrammetry is not sufficient in our case because scanning techniques are only applicable when the target object is extant and complete. The destroyed or lost parts cannot be handled by scanning, thus leading to the need for the second requirement.

For the second requirement, 3D reconstruction of the parts unavailable for scanning from preserved traditional 2D archiving materials such as manual drawings, photographs, and documents is needed. Many studies use image-based methods to digitize CH using multiple photos from different directions [13–15]. However, in many cases, no good photos for such methods remain. Sometimes, only one single monocular photo of each object is preserved. When no photo remains, the target parts need to be reconstructed from manual drawings or documentation material if available. Our target, the Borobudur temple, is a typical example of such a case.

For the third requirement, we should use a high-quality transparent (see-through) visualization method to recognize and investigate every integrated part. The method should be able to deal with large-scale 3D scanned point cloud data. In our recent work, we proposed an appropriate method called stochastic point-based rendering (SPBR) [16,17].

The Borobudur temple is one of the UNESCO World Heritage sites and is known as the largest Buddhist temple ruin in the world (see Figure 1). The temple is a stupa built with stones extracted from the volcanic rocks outcropping nearby. It is built on a hill that serves as the foundation of the temple. The hill is artificially built from three smaller naturally formed hills. According to the results of the UNESCO boring survey, the hill is divided into four layers based on soil qualities [18]. This four-layer foundation forms the internal structure of the temple. For the external structure, the temple building consists of nine stacked platforms, six square and three circular, topped by a central dome [19]. On the stone walls of the six square platforms, 2672 panels of bas-reliefs (sculptural reliefs in which forms extend only slightly from the background), containing 1460 narrative panels and 1212 decorative panels, are carved. These reliefs can be divided into five sections, each of which describes a different Buddhism story. Unfortunately, due to safety concerns, the foot encasement of the temple was reinstalled, and the section carved on the foot encasement named “Karmawibhanga” was buried behind the stone walls. This section has 160 relief panels, and only four panels at the southeast corner are visible to visitors. The remaining 156 relief panels are hidden by the stone walls. According to the abovementioned situation, the Borobudur temple can be divided into three parts, as shown in Figure 2: (1) the temple building, (2) the foundation, and (3) the hidden reliefs.

Among the above three parts, only the first part, i.e., the temple building, including the currently visible reliefs on the stone walls, can be archived using 3D scanning. We executed a 3D scanning of these extant buildings and reliefs based on the structure from motion - multiview stereo (SfM-MVS), which is a photogrammetry measurement method. We created two 3D scanned point datasets of the temple building. One dataset is created by remote and sparse photogrammetry executed with the help of an unmanned aerial vehicle (UAV). The other dataset is created by close-range and detailed photographic scanning. photogrammetry is a collaboration project between the Borobudur Conservation Office, the Indonesian Institute of Sciences, and Ritsumeikan University. Currently, 25% of detailed

photographic scanning has been completed. The complete detailed point dataset will be available by 2024.



Figure 1. The Borobudur Temple in Indonesia (photograph). The entire temple consists of a series of concentric terraces of decreasing size that rise like steps to a central peak. The first and lowest part is a square base measuring 113 m on each side [20]. The upper platforms have diminishing height and the highest point of the monument is 35 m above ground level.

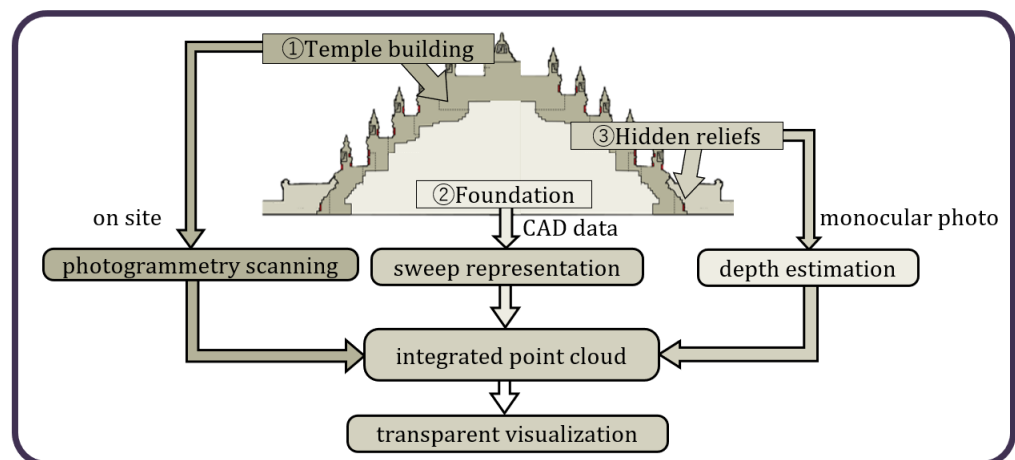


Figure 2. The overall strategy of the integrated transparent visualization.

The foundation and the hidden reliefs cannot be recorded by 3D scanning because the foundation is inaccessible and the hidden reliefs carved on the foot encasement are buried behind the stone walls. Fortunately, both parts have preserved 2D archiving materials left. First, the 2D CAD drawings of the 4-layer foundation based on the UNESCO boring survey remain. In this paper, we recover the 3D shape of the 4-layer formed foundation based on the sweep representation. Second, for each panel of hidden reliefs, a single monocular old photo remains in the museums. We propose a deep learning-based 3D reconstruction method to reconstruct the 3D model of the hidden reliefs from these monocular photos and achieve good results with high accuracy and realistic visual effects [21,22]. Then, we convert these two reconstructed datasets of invisible Parts (2) and (3) into 3D points and merge them with the 3D scanned points of visible Part (1). Then, we execute a visualization of the integrated 3D point data to show the whole 3D structure of the temple. As a result, we have achieved a complete point-cloud-type archive of the Borobudur temple.

The organization of this paper is as follows. After a brief review of related works in Section 2, we explain our proposed strategy in detail in Section 3. Then, in Section 4, we present the details of our experiment settings as well as the discussion on our results. In Section 5, we summarize the achievements of this paper.

2. Related Work

2.1. Digitization of Extant Cultural Heritage

3D Scanning technologies are the most widely adopted methods for the 3D digitization of CH sites. Tallon [23] performed laser scanning of the entire structure of the Notre Dame Cathedral, which helped professionals rebuild the cathedral after the tragic fire in 2019. Moreover, photogrammetry is an inexpensive solution that can be performed with a good-quality camera and a personal computer [24]. Therefore, many works adopt a hybrid approach by combining photogrammetry and other scanning methods to digitize CH assets [3,7,9,11,25]. In our work, as the structure of the Borobudur temple is complex and the trails are too narrow to place large devices, we consider photogrammetry to be the most suitable solution at present.

2.2. Reconstruction of Destroyed/Inaccessible Cultural Heritage

For the destroyed or inaccessible parts of CH sites, 3D reconstruction from previously preserved 2D archiving materials is widely studied. Many works [8,13,14,26,27] have created 3D models of CH sites from historical/open-source photographs and videos. Bruha et al. [28] combined diverse 3D data sources as well as referential terrain models to reconstruct the vanished medieval town of Sekanka and the surrounding landscape. There are also many studies that have digitized CH assets from other documentation such as CAD drawings. Li et al. [29] created 3D models of the festival floats in the Virtual Yamahoko Parade based on CAD drawings. Kuler et al. [30] reconstructed old Safranbolu houses using 3D modeling software from CAD drawings. Moreover, with the rapid development of machine learning, many studies have applied machine learning to the reconstruction of CH assets. Hermoza et al. [31] proposed a generative adversarial network to reconstruct complete 3D points from incomplete archaeological objects. Belhi et al. [32] proposed a learning-based image inpainting method to reconstruct incomplete artworks. In addition, in our previous work, we proposed a depth estimation network to reconstruct 3D points of relief from a single monocular photo [21,22].

2.3. Transparent Visualization

To visualize point datasets, point-based rendering is the most straightforward and widely used strategy, where 3D points are used as rendering primitives [33]. Based on point-based rendering, many studies have presented 3D opaque visualizations of CH assets [24,34,35]. However, these opaque visualization methods cannot achieve see-through imaging, which helps to understand the internal 3D structures of CH assets. Zwicker et al. [36] proposed elliptical weighted average (EWA) splatting to achieve point-based transparent visualization. Seemann et al. [37] used transparent points in a conventional point-based rendering technique, such as point sprites, to achieve transparency. However, these methods require depth sorting of many 3D points, which leads to long computation times and is thus not practical for dealing with large-scale 3D scanned point cloud data. Zhang et al. [38] proposed a point-based transparent rendering method that divides the visualized point dataset into many layers to avoid rendering artifacts. However, it is difficult to divide very dense points into proper layers, and the frame rate tends to degrade when visualizing a large-scale point dataset. In our previous work [16,17], we proposed a stochastic algorithm-based transparent visualization method called stochastic point-based rendering (SPBR) to achieve an accurate depth perception without any depth sorting steps.

3. Methods

This section presents our strategy and framework to digitally archive the Borobudur temple and visualize its entire 3D structure transparently. We first explain the outline of our digitization framework in Section 3.1. Then, we explain our digitization methods for the three different temple parts one by one: the extant temple building in Section 3.2, the inaccessible foundation in Section 3.3, and the hidden reliefs in Section 3.4. The point-based transparent visualization method, which we recently proposed and adopted in this study, is explained in Section 3.5.

3.1. Overview

As we have mentioned in the introduction, the Borobudur temple can be divided into three parts. We apply a different digitization method for each part, depending on its available data source. Figure 2 shows the overall strategy of our digitization: (1) photogrammetry is adopted to create the 3D point dataset of the surface of the temple building, including the extant reliefs. (2) The sweep representation is adopted to create 3D points from the CAD drawing of the inaccessible foundation. (3) Deep learning-based 3D reconstruction is executed to restore the 3D shape of the hidden reliefs from a single monocular photo, and then the restored 3D shape is converted to 3D point data. After executing the above three ways of digitization, the obtained 3D point datasets are merged to create an integrated point dataset. Then we transparently visualize the integrated point dataset, which describes the entire 3D structure of the Borobudur temple. This visualization enables us to observe both the surficial and internal structure of the temple.

3.2. Digitizing the Extant Temple Building from Photogrammetric Data

The temple building of Borobudur is a large-scale stupa with nine stacked platforms, six square and three circular, topped by a central dome. One aspect that adds to the cultural value of the Borobudur temple is the rich and vivid sculptures, including the Buddhist reliefs carved on the wall of the six square platforms and Buddha statues placed on the top platform. Although laser scanning and photogrammetry both seem to be optimal solutions, the narrow corridors of the temple make laser scanning difficult. Therefore, in this work, we adopt photogrammetry as the digitization method. Figure 3 shows examples of the Buddhist reliefs, Buddha statues, and narrow corridors of the temple.



Figure 3. From left to right: photograph of the Buddha statue, the relief, and the narrow corridor.

Moreover, as our goal is to achieve high-definition visualization results of the temple, both the building structure and the valuable sculptures should be visualized in detail. We adopt two photography strategies in this paper: remote photography to capture the building structure and close-range photography to capture the full details of the building, including the sculptures. The remote photography strategy performs quickly but results in sparse photogrammetry data, which lack building details. On the other hand, close-range photography captures the full details of the building but requires a long time and a large amount of manpower. The details of both photography strategies are presented as follows.

- **Remote scanning:** Remote scanning of the temple is performed with a UAV in the sky over the temple building. Sixty shots with a resolution of 4000×3000 pixels were

taken by the UAV carrying a digital camera (DJI FC300S). The vertical distance from the camera to the highest point of the temple is about 20 m and the overlap of each photo is about 60%.

- **Close-range scanning:** The close-range scanning is performed on the narrow corridor of the temple. To capture the high place of the temple wall of each platform, a monopod is used to support the camera. The photos with a resolution of 6000×4000 pixels are captured by a digital camera (RICOH GR III). The distance from the camera to the temple building is about 2 m and the overlap of each photo is about 60%.

The camera positions of the two proposed strategies are shown in Figure 4. The structure from motion-multi-view stereo (SFM-MVS) method is adopted as the photogrammetry method. In our work, we use the commercial multiview stereo software, Agisoft Metashape, to generate dense 3D points from the photogrammetry data. Moreover, during the close-range scanning, the operation is performed separately by several staff members. Therefore, the total station is used to record the coordinates to integrate the separate photogrammetry data. Besides, to reduce the impact of different light conditions, the color correction is performed with a color card taken together with the photos.

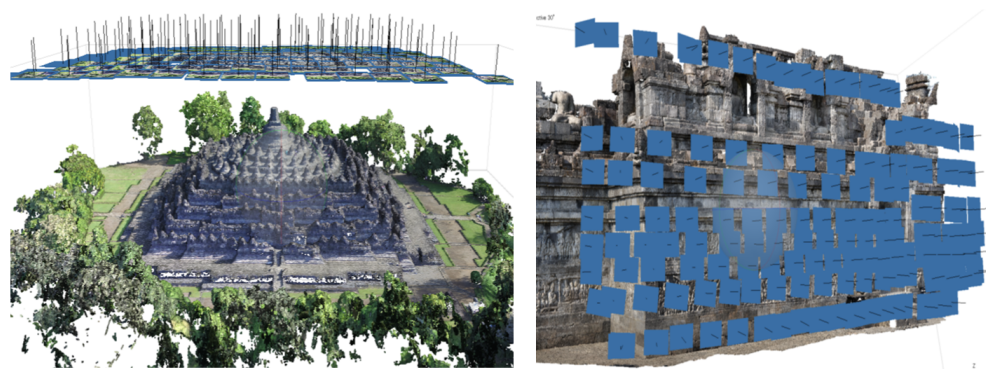


Figure 4. The camera positions of the two proposed strategies: the remote photography is on the left and the close-range photography is on the right. The blue box in the picture shows the direction and the position of the camera while taking the target photo.

3.3. Digitizing the Inaccessible Foundation from CAD Drawings

The foundation of the temple is a hill that is artificially built with stones carried from surrounding volcanoes on three smaller naturally formed hills. As shown in Figure 5, the layers from A to C are naturally formed, and layer D is artificially built. Based on the result of the UNESCO boring survey, the naturally formed hills can be divided into three layers by soil types: from bottom to top: soft volcanic tuff, foam with andesite, and andesite. The Borobudur temple can be imagined as a huge stone sculpture built around the surface of the foundation without entries or an internal space. Therefore, the foundation is not available for the scanning methods, and the only information remaining is the CAD drawing based on the boring survey. The CAD drawing was created based on a boring survey conducted by UNESCO in 1968. It is archived as an AutoCAD-format file (see Figure 5).

The available information on the foundation is restricted to the CAD drawings. Only the shape of the cross-section of the foundation is presented without any information on the perspective relationship or color. The vertical cross-sectional shape of the foundation is presented as 2D points with continuous coordinates that form a line. The sweep representation can construct symmetrical 3D objects by specifying a 2D shape and a sweep that moves the shape through a region of space. However, only layer A of the foundation is symmetrical, and the others are asymmetric; thus, it is insufficient to construct the foundation using only the sweep representation. In this paper, we propose a method to construct the 3D shape of this 4-layer foundation from the CAD drawings. The proposed method can be divided into three steps (see Figure 6): (1) 2D point interpolation, (2) 3D polygon creation, and (3) 3D point generation.

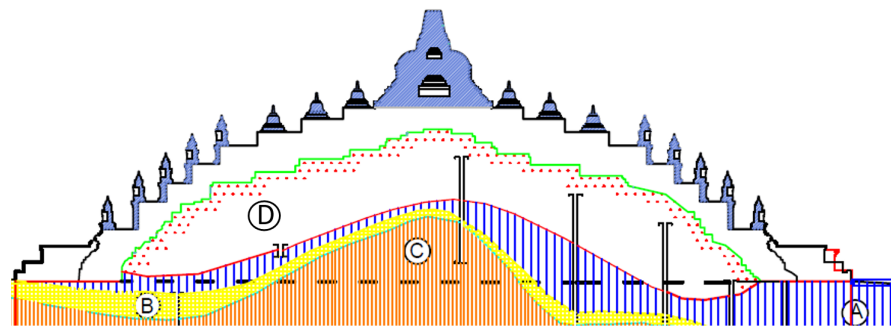


Figure 5. The CAD drawing of the foundation of Borobudur Temple is based on the UNESCO boring survey. From A to D in the photo: (A) andesite, (B) foam with andesite, (C) soft volcanic tuff, and (D) andesite with stone chips.

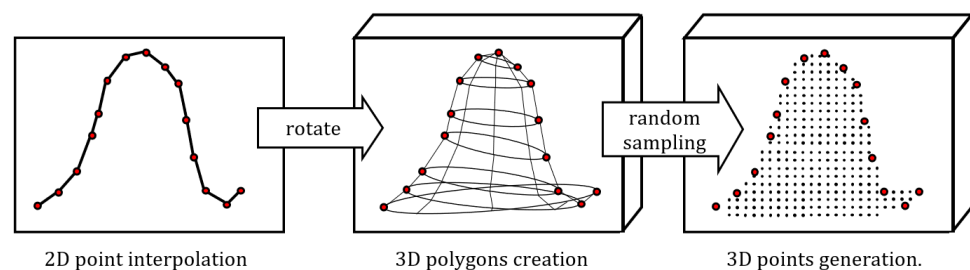


Figure 6. The flow chart of the proposed method.

- **STEP 1:** The first step of the proposed method is point generation and interpolation from the CAD drawing. The CAD drawing presents the shape of the vertical cross-section as vertex points with 2D coordinates. The point with the largest value on the y -axis is defined as a reference point that splits the points into two groups, that is, the left side and the right side. Then, we use cubic Hermite interpolation on the points to make the two sides have equal points. In our work, the number of interpolated points is 1000 per side. Figure 7 shows the point interpolation results of layer A.
- **STEP 2:** The second step is to rotate the results in Step 1 and create 3D polygons which present the shape of each layer. First, the points from the left side and right side are matched one by one. As Figure 8 shows, for point $A(x_1, y_1, 0)$ from the left side, there is a corresponding point $B(x_2, y_2, 0)$ on the right side. Then we rotate each point to its corresponding point and select sampling points from the transition zone. By connecting all the sampling points, the 3D polygon mesh which presents the shape of the foundation can be created.

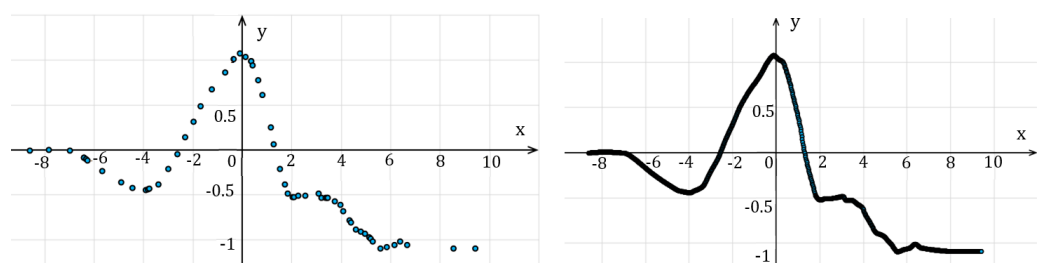


Figure 7. The points of layer A before interpolation on the right side and the points of layer A after interpolation on the right.

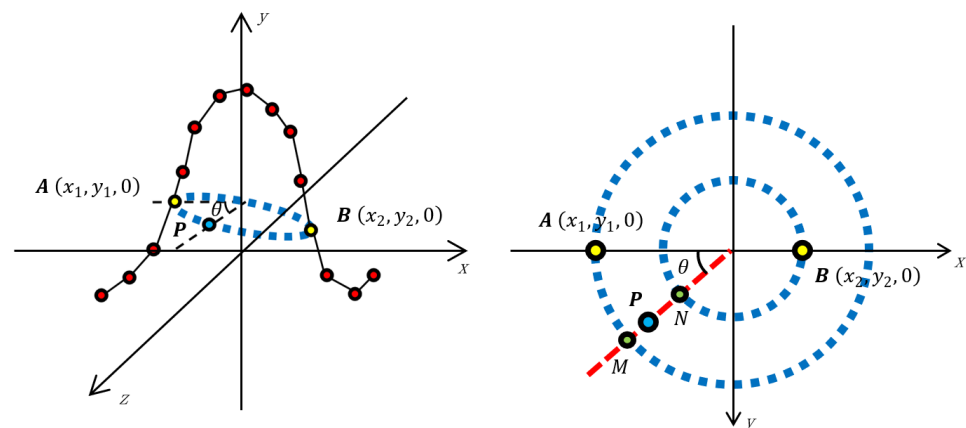


Figure 8. The method to calculate the coordinate of sampling point P . The left and right picture represents the 3D and 2D sketch map of the calculation method, respectively.

For each sampling point P , the θ is defined as the angle between the x -axis and the line formed by origin and P . By defining the distance between the rotation track of A and B as 1, t is the distance between point P to the rotation track of A (see Figure 8). The equation to calculate the coordinate of P is as follows:

$$P = \begin{pmatrix} \cos \theta x_1 & -\cos \theta x_1 \\ y_1 & y_2 \\ \sin \theta x_2 & -\sin \theta x_2 \end{pmatrix} \begin{pmatrix} t \\ 1-t \end{pmatrix}, t = \frac{\theta}{\pi}, \theta \in (0, \pi). \quad (1)$$

- **STEP 3:** The final step is to generate 3D points from the 3D polygon data of Step 2. Generating points directly from 3D polygon data will lead to a huge calculation cost, thus it is necessary to split the 3D polygon into an amount of 2D triangles. For a polygon with n vertex points, by connecting each vertex point to any two other points, $n - 2$ sub-triangles are created. Then sampling points can be randomly generated from each triangle based on the coordinates of its three vertex points. The number of the points generated from each triangle is fixed to make the output points have a uniform distribution.

Please note that layer D is symmetrical, thus we directly apply the sweep representation to create the 3D polygon and then generate 3D points. Layer A, B, and C follow Steps 1 to 3 because the layers are asymmetric.

3.4. Digitizing the Hidden Reliefs from Single Monocular Photo

The Buddhist reliefs carved on the walls of the six square platforms of the Borobudur temple are the largest collection in the world. They are sculptural reliefs in which forms extend only slightly from the background. These reliefs can be divided into five sections, each of which describes a different Buddhism story. Unfortunately, one section named “Karmawibhangga” was buried behind the stone walls after the reinstatement of the foot encasement due to safety concerns. This section has 160 relief panels, and only four panels at the southeast corner are visible to visitors. For each hidden panel, there remains one grayscale photo taken directly in front of the reliefs in 1890. Figure 9 shows an example of the old photo (left top), the stone wall covering the hidden reliefs (left bottom), and the southeast corner with the remaining four “Karmawibhangga” reliefs (right).

As there is only one monocular photo available for each panel, the traditional multi-view image-based method is not applicable. Thus, we proposed a deep learning-based method in our previous work to reconstruct the 3D models of the hidden reliefs. The 2D monocular photo contains the color information and value of the x and y -axis for each point. Thus, to reconstruct the 3D model from a single monocular photo, the only information left to find is the value of the z -axis (depth) in 3D coordinates. Therefore, we apply a monocular depth estimation network to estimate the corresponding depth map

of the monocular photo. The value of each pixel in the depth map represents the distance between the point and the camera when the photo is taken. Thus the value of the z-axis (depth) can be calculated by a linear transformation from the depth map. The proposed method is a supervised learning method, the training and validation data are obtained from the photogrammetry scanning data of the extant relief panels. Once the model is trained, the corresponding depth map of the old photos of the hidden reliefs can be predicted by the trained network. Then the 3D points can be generated by combining the information in the predicted depth map and the old photo.



Figure 9. The example of the old photo (left top), the stone wall covering the hidden reliefs (left, bottom), and the southeast corner with the remaining four “Karmawibhanga” reliefs (right).

In this paper, we apply a monocular depth estimate network [39] to estimate the depth of relief data. The proposed network structure follows an encoding-decoding structure. The encoder utilizes DenseNet [40] to extract dense feature maps and a denser version [41] of the atrous spatial pyramid pooling layer to extract the contextual information. The decoder recovers the resolution of feature maps gradually with a factor of 2. A novel upsampling layer, local planar guidance (LPG) layer is applied to locate geometric guidance to improve the depth estimation. For finer scales with high resolution, the LPG layer learns regions with sharp curvatures, such as object boundaries. For coarser scales with low resolution, the LPG layer learns the major structures. Thus, by placing the LPG layer at different scales of the decoder part, both the major structures and the object boundaries can be extracted properly.

3.5. Transparent Visualization

As we hope to provide transparent imaging of the entire Borobudur temple, including the inaccessible foundation and hidden reliefs, we should adopt an efficient transparent visualization method applicable to large-scale point datasets. Most of the traditional transparent visualization methods realize transparency based on the depth sorting of rendering primitives. However, the depth sorting requires computation time proportional to $N \log_N$ with N the number of rendering primitives, i.e., the number of points. Therefore, traditional transparent visualization methods are not suitable for large-scale point datasets. Thus, we adopt stochastic point-based rendering (SPBR), which we recently proposed [16,17]. This rendering method eliminates the need for depth sorting by adopting a noble stochastic algorithm. Below, we briefly explain the algorithm of SPBR. The procedure of SPBR can be divided into three steps as shown in Figure 10.

- **Step 1:** Create multiple subgroups of points by randomly dividing the original point dataset. Each subgroup should have the same point density and be statistically independent. Below, the number of subgroups is denoted as L , which is usually set to a few hundred.
- **Step 2:** For each point subgroup in Step 1, execute the standard point-based rendering by projecting its constituent 3D points onto the image plane, which creates an inter-

mediate image. In the projection process, the point occlusion is considered per pixel. A total of L intermediate images are obtained.

- **Step 3:** Create an average image of the L intermediate images created in Step 2. This average image becomes the final transparent image, in which the measurement noise is automatically eliminated per pixel based on the statistical effect [17].

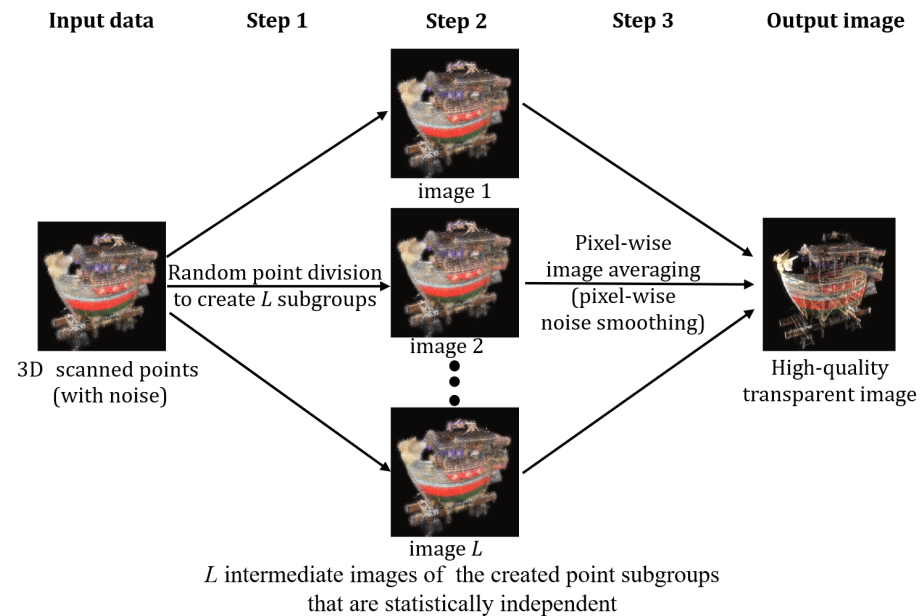


Figure 10. The stochastic point-based rendering (SPBR). The method consists of three steps to create a high-quality transparent image from a group of 3D scanned points: Step (1) random point division, Step (2) intermediate images creation, and Step (3) image averaging.

In the above steps, L represents the number of averaged intermediate images and thus controls the image quality. The local surface opacity α takes the following value, depending on L :

$$\alpha = 1 - \left(1 - \frac{s}{S}\right)^{n/L} \quad (2)$$

where s is the point sectional area, whose image overlaps only one pixel and S is the area of the local surface segment that contains n points. The opacity α can be controlled by point-number adjustment, i.e., by tuning the local number of points, n . Increase of n from the raw number is made by copying selected points on S . Decrease of n is made by eliminating randomly selected points on S .

The above algorithm of SPBR provides a straightforward solution to achieve the integrated transparent visualization of multiple point datasets. What we should do is: (1) Execute the point-number adjustment to tune the opacity of each point dataset. (2) Unify all the adjusted point datasets. (3) Apply SPBR to the unified point dataset. We can highlight a selected portion of the visualized 3D shape by assigning a higher opacity to selected point datasets in action (1).

4. Experimental Results

In this section, we present the results of the abovementioned digitization and visualization. In Section 4.1, we present digitization results of the three parts of the Borobudur temple. In Section 4.2, we present the transparent visualization of the integrated point dataset merging the three datasets. In Section 4.3, we describe the implementation details of our experiments.

4.1. Digitization Results

4.1.1. Extant Temple Building

Here, we present the digitization results of the extant temple building by the proposed remote and close-range photography strategies. Figure 11 shows the point dataset obtained by UAV-based remote scanning. This remotely scanned point dataset (Figure 11) contains more than 2.7×10^7 points including the entire temple building, paths, and surrounding forests. The nine platforms and the narrow corridors are presented clearly. However, as the numbers and the resolution of the photos taken by the UAV are not large enough, the quantity of the remote photogrammetry scanned dataset is limited. For example, in the zoom-in figures of Figure 11, details such as reliefs carved on the temple wall are blurry.

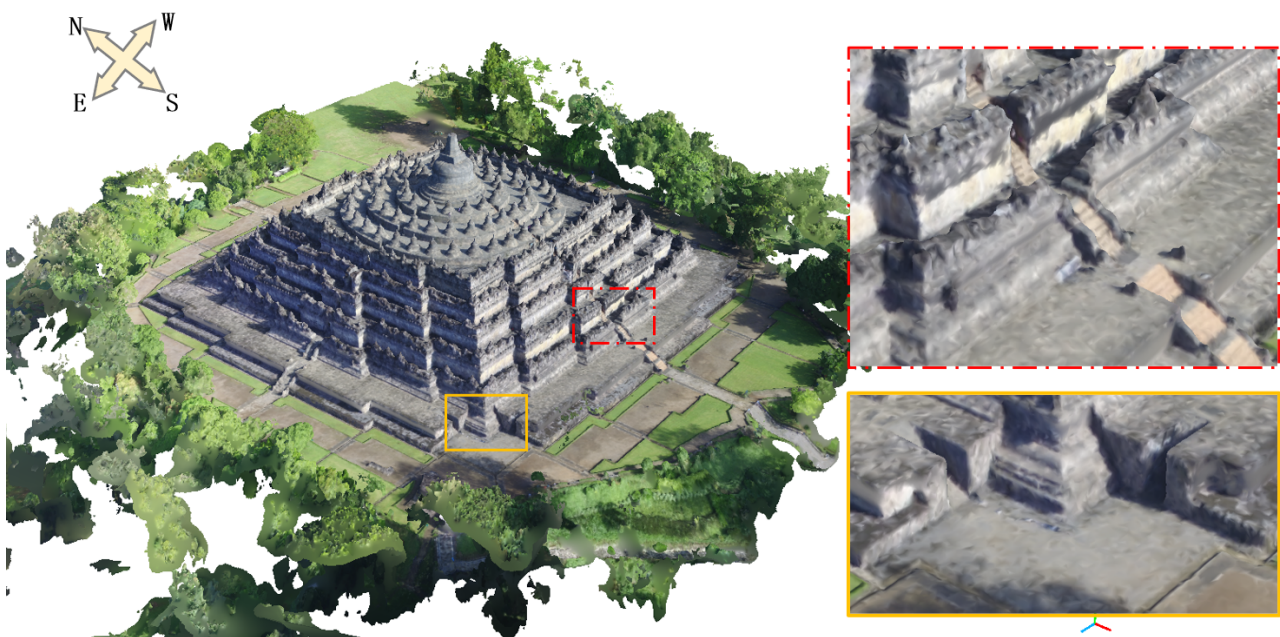


Figure 11. Remote scanned points of the Borobudur temple.

The close-range photogrammetry results of these two parts (images on the right of Figure 11) are shown in Figure 12. The left image of Figure 12 shows the digitization result of the southern parts near the stairs on the first level that corresponds to the top-right image in Figure 11. The right image of Figure 12 shows the digitization result of the southeast corner of the temple, on which the remaining four panels of the Kamavibangga reliefs are carved, that corresponds to the top-right image in Figure 11. Both datasets contain over 10^8 points. The Buddha statues, the Buddhist reliefs, and the decorations on the corridors can be seen clearly, owing to the high quality of the photos provided by the close-range photogrammetry.

This close-range photogrammetry project is currently being performed and we present 25% of the point dataset representing 75% of the second platform of the temple in Figure 13, which was photographed in 2020. The complete detailed point dataset will be available by 2024.

4.1.2. Inaccessible Foundation

In this section, we present the reconstructed 3D model of the foundation of the Borobudur temple. The foundation consists of four layers, as shown in Figure 5, and each layer is converted into a point dataset based on the rotating extrapolation described in Section 3.3. Figure 14 shows the reconstructed points of each layer and the number of points of each layer is summarized in Table 1. The results achieved smooth curved surfaces describing the shape of each layer. Because the available CAD drawing is only for one direction of the temple, the reconstructed 3D shapes of the asymmetric layers are

not entirely accurate. However, the reconstructed layers by the proposed method are still advantageous for understanding the structure of the Borobudur foundation.

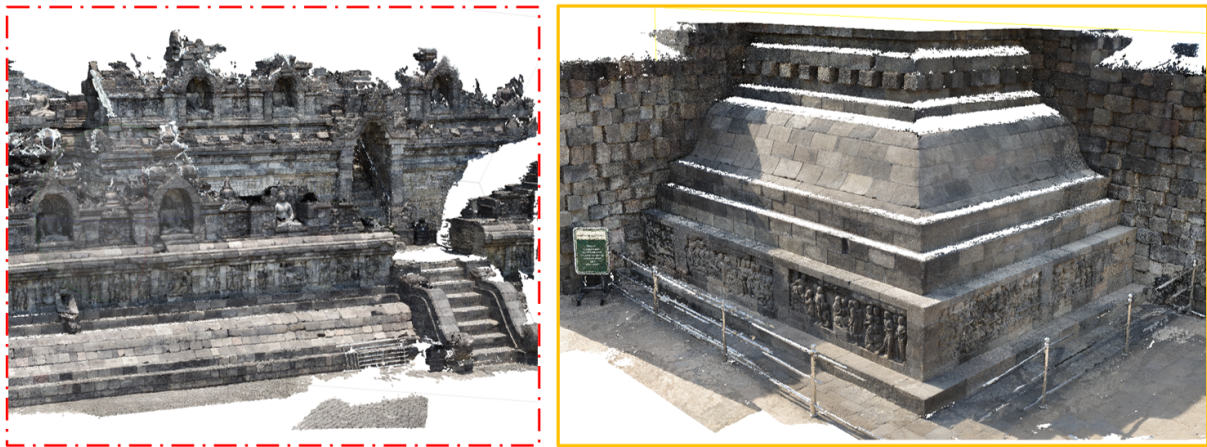


Figure 12. Close-range scanned points of the Borobudur temple.

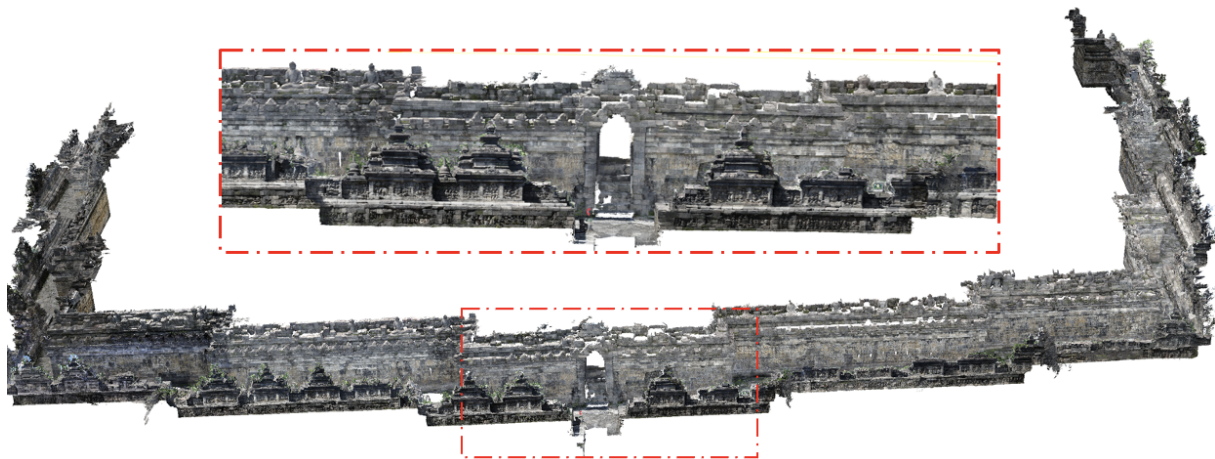


Figure 13. Digitization result of 75% of the first level of the Borobudur temple.

Table 1. Number of points of each reconstructed point dataset.

Index	Color	Number of Points	Symmetry
D	green	716,043	Symmetry
A	blue	661,552	Asymmetry
B	yellow	1,266,911	Asymmetry
C	red	1,823,015	Asymmetry

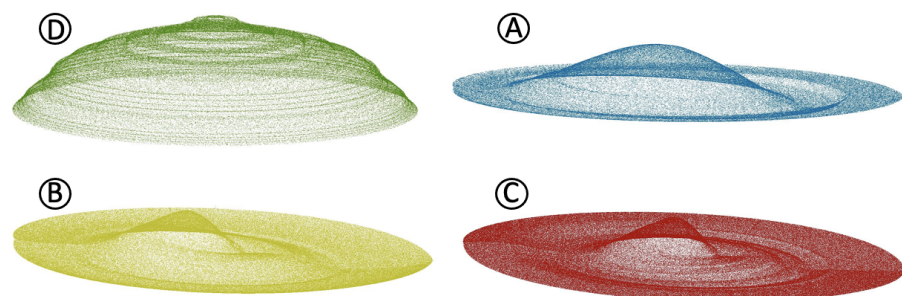


Figure 14. Reconstructed 3D points of each layer in the foundation from the corresponding 2D vertex points extracted from CAD drawing. Subfigures (A–D) represent the four layers shown in Figure 5.

Moreover, we merged the four layers together to describe the complete shape of the foundation. Figure 15 shows the fused transparent results of the merged dataset of the foundation. The shape of the foundation, which is formed from the three smaller natural hills (layers A, B, and C in Figure 5) and the artificially built hill (layer D in Figure 5), is visualized clearly.

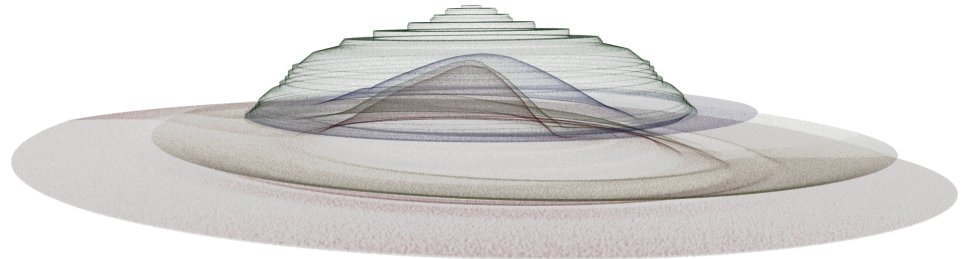


Figure 15. The fused transparent results of the reconstructed points of the four foundation layers of the unreachable foundation.

4.1.3. The Hidden Reliefs

In this section, we present the reconstructed 3D model of the reliefs in the Borobudur temple obtained by the proposed deep learning-based method. As the hidden reliefs are buried by the stones with only monocular photos remaining, the ground truth of the hidden reliefs is not available for quantitative or qualitative comparison. Thus, we chose one visible Karmawibhangga relief as test data to evaluate the proposed method. The qualitative comparison results are shown in Figure 16. The proposed method is compared with three models, a multiscale CNN-based model [21], a ResNet-based model [22], and a DenseNet-based model [42].

Figure 16 shows that the results of the proposed method predict more accurate depth values compared to the ground truth. The shape of the carved figure is clearer and the noise which happens in other models is reduced. Here, for noise, we mean that the unstable depth changes in the background. Moreover, the proposed method can extract some details like human faces which have only slight depth change comparing to the surrounding areas. Note that the artifacts between the patches are produced because the path-wise training ignores the context relationship between neighbourhood patches.

Besides, seven error metrics are utilized to evaluate the proposed method as shown in Table 2. The ratio of the pixels correctly labeled in the predicted depth map is calculated by three thresholds: α_1 , α_2 and α_3 . As the depth value is always positive or zero (in the range of 0 to 255) with a highly skewed distribution, which makes the symmetric loss function such as RMSE not applicable enough, a logarithmic transformation (RMSE_log) is applied to obtain a less skewed distribution. The absolute and squared relative differences are also applied. The quantitative results show that the proposed 3D monocular reconstruction method achieves more accurate depth estimation results than the other models.

Table 2. Results of the comparison experiment on the relief dataset.

	Higher Is Better			Lower Is Better			
	$\theta_1 \leq 1.25$	$\theta_2 \leq 1.25^2$	$\theta_3 \leq 1.25^3$	RMSE	RMSE_log	abs_rel	sq_rel
CNN [21]	0.306376	0.597804	0.777289	10.28911	0.780762	3.067239	2.000889
ResNet-50 [22]	0.34431	0.607586	0.778217	10.17354	0.589355	3.029139	1.770056
DenseDepth [42]	0.377522	0.641598	0.790830	9.995814	0.633301	3.872130	2.193610
Ours [39]	0.440557	0.76995	0.920572	9.840669	0.455078	4.07383	2.127823

The reconstructed 3D point datasets of the hidden reliefs are shown with the corresponding old photos and the estimated depth maps in Figure 17. Note that the value of

each pixel in the depth map represents the value of the axis z of the corresponding point in the reconstructed point. The value of the z -axis was set in the range of 0 to 255. To recover the real depth range of the reliefs, a linear transformation is applied. The real Borobudur relief was 2.7 m wide, 0.92 m high, and 0.15 m deep, thus the range is transformed into 0 m to 0.15 m by the linear transformation to create the coordinate of each point.

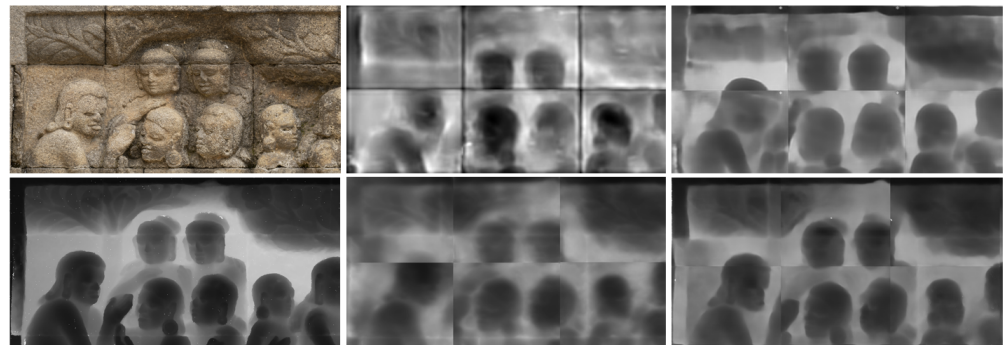


Figure 16. The comparison results of the depth estimation results between the proposed method and the other models. From top left to bottom right: the monocular photo, the result of CNN [21], the result of ResNet-50 [22], the ground truth, the result of DenseDepth [42], and the result of the proposed method [39].

The proposed results provide distinct boundaries between the figures and the background. The 3D points provide natural visual effects from different angles and the depth perception is correctly reconstructed. As the point dataset is reconstructed from a single gray-scale monocular photo, the number of points is equal to the image resolution and the color of each point follows the intensity in the old monocular photo. The image resolution is 3200×1024 and the number of the reconstructed points is 3,276,800.

4.2. Integrated Visualization

Based on the abovementioned digitization results of the extant temple building, the inaccessible foundation, and the hidden reliefs, the integrated complete digital archive of the Borobudur temple is established. The 3D models of the three parts are all digitized/reconstructed into point datasets to achieve a high-definition transparent visualization. Figure 18 shows the transparent visualization results of the complete Borobudur temple. For the extant temple buildings, the remote photogrammetry results are merged with parts of the close-range photogrammetry results. Besides, the four reconstructed 3D points of each layer of the inaccessible foundation are placed below the temple building in their correct location. Furthermore, the hidden reliefs are reconstructed by the proposed deep-learning based reconstruction method from the monocular old photo and placed into their correct position according to official documents of the Borobudur conversation office [19]. Note that the reconstructed hidden reliefs are aligned manually. Their real locations may slightly differ from the visualization result. From this single transparent image, the external, internal, and details of the temple can be discerned clearly.

Figure 19 shows the zoom-in transparent visualization results of the hidden reliefs. The image shows the southeast corner of Borobudur, where there are four visible Karmawibhanga reliefs and 16 reconstructed hidden relief panels. The first row is opaque visualization results, which have the same visual effect as in the real world, and the hidden reliefs behind the stone wall are invisible. The second row is the result of transparent visualization, which enables us to see through the stone wall and recognize the 3D appearance of the hidden reliefs. The third row is the focused view of the relief panels, in which the 3D appearance of the hidden reliefs is clearly visualized. The visualization is executed at an interactive speed, which is quick enough for various visual analyses.



Figure 17. Depth estimation and reconstruction results from two examples: The top four pictures and the bottom four pictures represent the results of two old photos, respectively. From top right to bottom left in each group: the old monocular photo, the depth estimation result, and the screenshot of the 3D reconstructed points from the left and right sides, respectively.

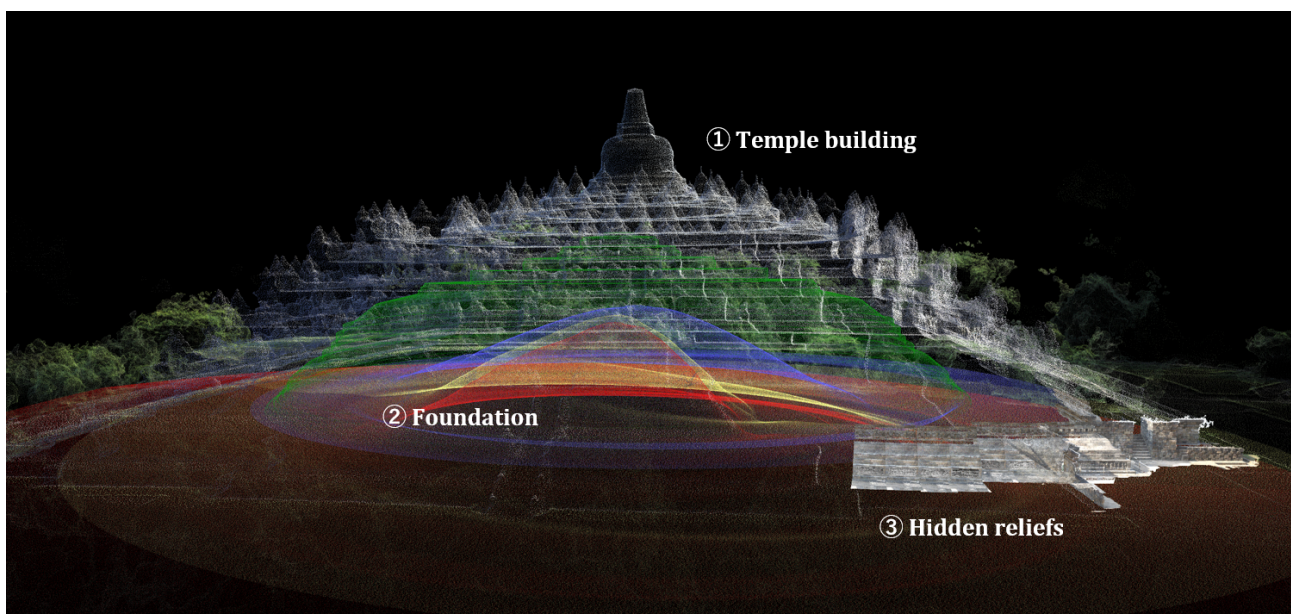


Figure 18. Integrated visualization result.

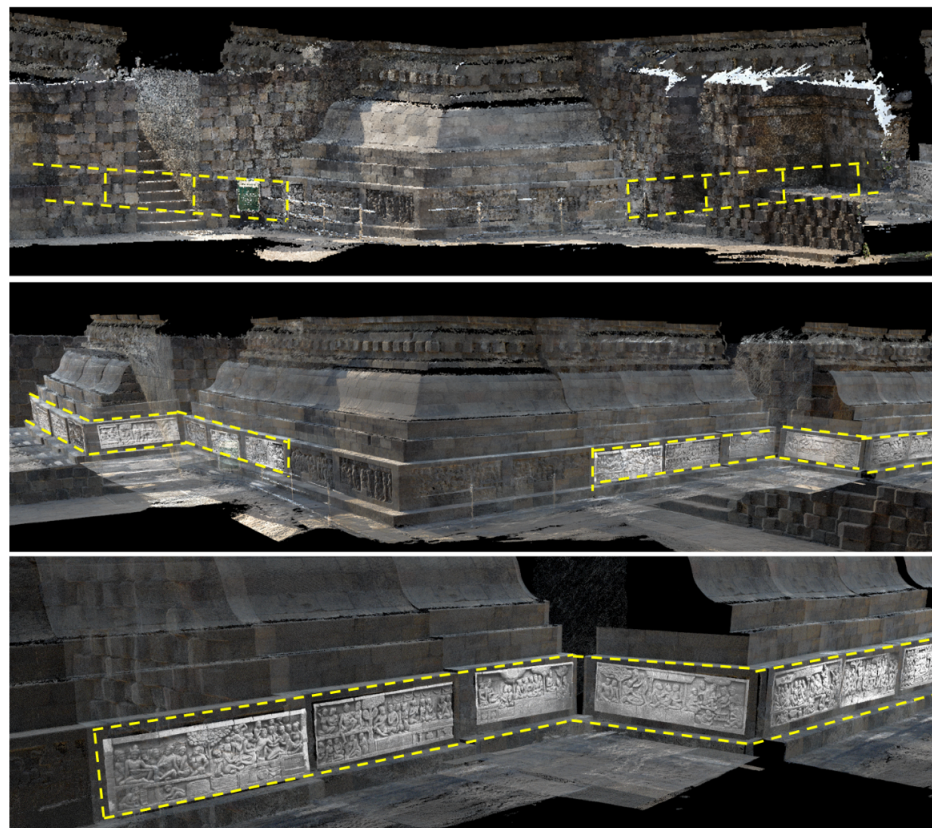


Figure 19. The zoom-in transparent visualization results of the hidden reliefs.

4.3. Implementation Details

For the deep learning-based 3D relief reconstruction demonstrated in this paper, PyTorch [43] was utilized to implement the proposed network. The network was trained on a single NVIDIA GeForce GTX 1080Ti with 12 GB of GPU memory. The weights of the encoder part in the proposed network were initialized by DenseNet-161 [40] pretrained on the image classification dataset. For training, we use the Adam optimizer [44] with polynomial decay from a base learning rate of 10^{-4} with power $p = 0.9$. The total number of epochs is set to 50 with a batch size of 16 for all experiments in this work. To avoid overfitting, we augment images before input to the network using random rotation, horizontal flipping, and random brightness adjustment in a range of [0.9, 1.1], with 50% of chance.

For the visualizations demonstrated in this paper, the computations were executed on a LINUX PC with an Intel Xeon(R) W-2245 (3.90 GHz \times 16) CPU and an NVIDIA Corporation TU104GL (Quadro RTX 4000) GPU. We confirmed that this PC could handle (several) $\times 10^9$ 3D points (10^8 3D points could be rendered using a laptop PC with a 3.07 GHz Intel Core i7 processor, 8 GB of memory and an NVIDIA GeForce GT 480M GPU).

5. Conclusions and Discussion

In this paper, we proposed a technical framework for the enrichment phase of the digital preservation process for large-scale CH sites. We successfully applied the proposed framework to a UNESCO World Heritage site, the Borobudur temple, offering a method for 3D recoveries of incomplete or inaccessible portions of the site. The paradigm policy of the framework is that we unify different types of archived data from various sources as an integrated point dataset. The entire 3D structure described by the integrated point dataset can be visualized by stochastic point-based rendering (SPBR) precisely and comprehensively. The framework is widely applicable to other cultural sites for which the efficiency of digitization, 3D reconstruction of destroyed/missing parts, and integrated visualization of the entire site are required.

There are three future plans for the continuation of this research. First, we will complete the Borobudur digital archive by adding the close-range photogrammetry data for the remaining parts. Moreover, recently developed high-precision portable laser scanners should also be available in our digitization project. We plan to use them in combination with photogrammetry in future. Second, we will improve the accuracy of the 3D reconstruction of the hidden Karmawibhangga reliefs. The edges of the carved patterns on the reliefs contain essential features of both semantic information and the geometric nature of the reliefs. Therefore, edge information should be available to improve accuracy. Third, we will apply our integrated visualization to immersive VR scenes such that our digital archives can become available for applications such as digital museums, mobile guidance purposes, and other applications.

Author Contributions: Conceptualization, J.P., L.L., K.H. and S.T.; methodology, J.P., K.H., L.L. and S.T.; software, J.P.; validation, J.P.; formal analysis, J.P.; investigation, J.P.; resources, H.Y., F.I.T. and B.; data curation, J.P. and L.L.; writing—original draft preparation, J.P.; writing—review and editing, J.P., L.L., K.H. and S.T.; visualization, J.P.; supervision, L.L. and S.T.; project administration, S.T.; funding acquisition, S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partially supported by JSPS KAKENHI Grant Number 19KK0256, 21H04903, and the Program for Asia-Japan Research Development (Ritsumeikan University, Kyoto, Japan).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: In this paper, the images of the Borobudur temple are presented and with the permission of the Borobudur Conservation Office and Research Center for Area Studies, National Research and Innovation Agency, Indonesia. Additionally, photogrammetry scanning point dataset was provided by the Nara National Research Institute for Cultural Properties. We deeply thank these three institutions for their generous cooperation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bachi, V.; Fresa, A.; Pierotti, C.; Prandoni, C. *The Digitization Age: Mass Culture Is Quality Culture. Challenges for Cultural Heritage and Society*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 786–801. [\[CrossRef\]](#)
2. Pavlidis, G.; Koutsoudis, A.; Arnaoutoglou, F.; Tsioukas, V.; Chamzas, C. Methods for 3D digitization of Cultural Heritage. *J. Cult. Herit.* **2007**, *8*, 93–98. [\[CrossRef\]](#)
3. Bitelli, G.; Balletti, C.; Brumana, R.; Barazzetti, L.; D’Urso, M.G.; Rinaudo, F.; Tucci, G. The gamher research project for metric documentation of cultural heritage: Current developments. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 239–246. [\[CrossRef\]](#)
4. Adane, A.; Chekole, A.; Gedamu, G. Cultural Heritage Digitization: Challenges and Opportunities. *Int. J. Comput. Appl.* **2019**, *178*, 1–5. [\[CrossRef\]](#)
5. Long, J.; Shelhamer, E.; Darrell, T. Research Challenges for Digital Archives of 3D Cultural Heritage Models. *J. Comput. Cult. Herit.* **2010**, *2*, 1–17.
6. Yastikli, N. Documentation of cultural heritage using digital photogrammetry and laser scanning. *J. Cult. Herit.* **2007**, *8*, 423–427. [\[CrossRef\]](#)
7. Nuttens, T.; De Maeyer, P.; De Wulf, A.; Goossens, R.; Stal, C. Terrestrial Laser Scanning and Digital Photogrammetry for Cultural Heritage: An Accuracy Assessment. In Proceedings of the 4th International Workshop on 3D Geo-Information, Marrakesh, Morocco, 18–22 May 2011; pp. 18–22.
8. Dhonju, H.K.; Xiao, W.; Sarhosis, V.; Mills, J.P.; Wilkinson, S.; Wang, Z.; Thapa, L.; Panday, U.S. Feasibility study of low-cost image-based heritage documentation in Nepal. *ISPRS Arch.* **2017**, *42*, 237–242. [\[CrossRef\]](#)
9. Remondino, F. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sens.* **2011**, *3*, 1104–1138. [\[CrossRef\]](#)
10. Lerma, J.L.; Cabrelles, M.; Navarro, S.; Fabado, S. From Digital Photography to Photogrammetry for Cultural Heritage Documentation and Dissemination. *Disegnarecon* **2013**, *6*, 1–8. [\[CrossRef\]](#)
11. Girelli, V.A.; Tini, M.A.; D’Apuzzo, M.G.; Bitelli, G. 3D digitisation in cultural heritage knowledge and preservation: The case of the neptune statue in bologna and its archetype. *ISPRS Arch.* **2020**, *43*, 1403–1408. [\[CrossRef\]](#)

12. Alidoost, F.; Arefi, H. Comparison of uas-based photogrammetry software for 3D point cloud generation: A survey over A historical site. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 55–61. [[CrossRef](#)]
13. Kersten, T.P.; Lindstaedt, M. 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. Fernerkund. Geoinf.* **2013**, *2012*, 727–740. [[CrossRef](#)]
14. Kyriakaki, G.; Doulamis, A.; Doulamis, N.; Ioannides, M.; Makantasis, K.; Protopapadakis, E.; Hadjiprocopis, A.; Wenzel, K.; Fritsch, D.; Klein, M.; et al. 4D Reconstruction of Tangible Cultural Heritage Objects from Web-Retrieved Images. *Int. J. Herit.* **2014**, *3*, 431–451. [[CrossRef](#)]
15. Grün, A.; Remondino, F.; Zhang, L.I. Photogrammetric reconstruction of the great buddha of Bamiyan, Afghanistan. *Photogramm. Rec.* **2004**, *19*, 177–199. [[CrossRef](#)]
16. Tanaka, S.; Hasegawa, K.; Okamoto, N.; Umegaki, R.; Wang, S.; Uemura, M.; Okamoto, A.; Koyamada, K. See-through Imaging of Laser-scanned 3D Cultural Heritage Objects based on Stochastic Rendering of Large-scale Point Clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III-5*, 73–80. [[CrossRef](#)]
17. Uchida, T.; Hasegawa, K.; Li, L.; Adachi, M.; Yamaguchi, H.; Thufail, F.I.; Riyanto, S.; Okamoto, A.; Tanaka, S. Noise-robust transparent visualization of large-scale point clouds acquired by laser scanning. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 124–134. [[CrossRef](#)]
18. Tokyo National Museum. *Borobudur dan Seni Purbakala Indonesia*; Tokyo National Museum: Tokyo, Japan, 1980.
19. Balai Konservasi Borobudur. *Pada Relief Pada Relief Karmawibhangga*; Balai Konservasi Borobudur. 2012. Available online: <http://borobudurpedia.id/book/adegan-dan-ajaran-hukum-karma-pada-relief-karmawibangga/> (accessed on 5 December 2021).
20. Miksic, J.; Tranchini Marcello, T.A. *Borobudur: Golden Tales of the Buddhas*; Tuttle Publishing: Clarendon, VT, USA, 2012.
21. Pan, J.; Li, L.; Yamaguchi, H.; Hasegawa, K.; Thufail, F.I.; Bramantara, K.; Tanaka, S. *3D Transparent Visualization of Relief-Type Cultural Heritage Assets Based on Depth Reconstruction of Old Monocular Photos*; Springer: Singapore, 2019; Volume 1, pp. 187–198. [[CrossRef](#)]
22. Pan, J.; Li, L.; Yamaguchi, H.; Hasegawa, K.; Thufail, F.I.; Brahmantara, K.; Tanaka, S. Fused 3D Transparent Visualization for Large-Scale Cultural Heritage Using Deep Learning-Based Monocular Reconstruction. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *5*, 989–996. [[CrossRef](#)]
23. Tallon, A. Divining Proportions in the Information Age. *Archit. Hist.* **2014**, *2*, 15. [[CrossRef](#)]
24. Aicardi, I.; Chiabrande, F.; Maria Lingua, A.; Noardo, F. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *J. Cult. Herit.* **2018**, *32*, 257–266. [[CrossRef](#)]
25. Alshawabkeh, Y.; El-Khalili, M.; Almasri, E.; Bala'awi, F.; Al-Massarweh, A. Heritage documentation using laser scanner and photogrammetry. The case study of Qasr Al-Abidit, Jordan. *Digit. Appl. Archaeol. Cult. Herit.* **2020**, *16*. [[CrossRef](#)]
26. Themistocleous, K. Model reconstruction for 3d vizualization of cultural heritage sites using open data from social media: The case study of Soli, Cyprus. *J. Archaeol. Sci.* **2017**, *14*, 774–781. [[CrossRef](#)]
27. Condorelli, F.; Rinaudo, F. Cultural heritage reconstruction from historical photographs and videos. *ISPRS Arch.* **2018**, *42*, 259–265. [[CrossRef](#)]
28. Brůha, L.; Laštovička, J.; Palatý, T.; Štefanová, E.; Štych, P. Reconstruction of lost cultural heritage sites and landscapes: Context of ancient objects in time and space. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 604. [[CrossRef](#)]
29. Li, L.; Choi, W.; Hachimura, K.; Yano, K.; Nishiura, T.; Tanaka, H.T. [Paper] Virtual Yamahoko Parade Experience System with Vibration Simulation. *ITE Trans. Media Technol. Appl.* **2014**, *2*, 248–255. [[CrossRef](#)]
30. Külür, S.; Şahin, H. 3D Cultural heritage documentation using data from different sources. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *XXXVII*, 353–356.
31. Hermoza, R.; Sipiran, I. 3D reconstruction of incomplete archaeological objects using a generative adversarial network. *ACM Int. Conf. Proceeding Ser.* **2018**, *2018*, 5–11.
32. Belhi, A.; Al-Ali, A.K.; Bouras, A.; Foufou, S.; Yu, X.; Zhang, H. Investigating low-delay deep learning-based cultural image reconstruction. *J. Real-Time Image Process.* **2020**, *17*, 1911–1926. [[CrossRef](#)]
33. Gross, M.; Pfister, H. *Point-Based Graphics (The Morgan Kaufmann Series in Computer Graphics)*; Morgan Kaufmann Publishers: Burlington, MA, USA, 2007; p. 552.
34. Kersten, T.P.; Keller, F.; Saenger, J.; Schiewe, J. *Automated Generation of an Historic 4D City Model of Hamburg and Its Visualisation with the GE Engine*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 55–65. [[CrossRef](#)]
35. Dylla, K.; Frischer, B.; Mueller, P.; Ulmer, A.; Haegler, S. Rome Reborn 2.0: A Case Study of Virtual City Reconstruction Using Procedural Modeling Techniques. *Comput. Graph. World* **2008**, *16*, 62–66.
36. Zwicker, M.; Pfister, H.; Van Baar, J.; Gross, M. EWA splatting. *IEEE Trans. Vis. Comput. Graph.* **2002**, *8*, 223–238. [[CrossRef](#)]
37. Seemann, P.; Palma, G.; Dellepiane, M.; Cignoni, P.; Goesele, M. Soft Transparency for Point Cloud Rendering. *EGSR* **2018**. [[CrossRef](#)]
38. Zhang, Y.; Pajarola, R. Deferred blending: Image composition for single-pass point rendering. *Comput. Graph.* **2007**, *31*, 175–189. [[CrossRef](#)]
39. Lee, J.H.; Han, M.K.; Ko, D.W.; Suh, I.H. From big to small: Multi-scale local planar guidance for monocular depth estimation. *arXiv* **2019**, arXiv:1907.10326.

40. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
41. Yang, M.; Yu, K.; Zhang, C.; Li, Z.; Yang, K. DenseASPP for Semantic Segmentation in Street Scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [[CrossRef](#)]
42. Alhashim, I.; Wonka, P. High Quality Monocular Depth Estimation via Transfer Learning. *arXiv* **2018**, arXiv:1812.11941
43. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An imperative style, high-performance deep learning library. *arXiv* **2019**, arXiv:1912.01703.
44. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015; pp. 1–15.