

A DIGITAL ARCHIVE OF BOROBUDUR BASED ON 3D POINT CLOUDS

M. Kawato¹*, L. Li², K. Hasegawa², M. Adachi³, H. Yamaguchi⁴, F. I. Thufail⁵, S. Riyanto⁶, Brahmantara⁷, S. Tanaka²

¹ Graduate of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan
- is0351sr@ed.ritsumeikan.ac.jp

² College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan
- liliang@fc.ritsumeikan.ac.jp, hasegawa@media.ritsumeikan.ac.jp, stanaka@is.ritsumeikan.ac.jp

³ Shrewd Design Co. Ltd, Fushimi-ku, Kyoto, Kyoto 612-8362, Japan - motoaki.adachi@shrewd-sekkei.com

⁴ Nara National Research Institute for Cultural Properties, Nijo-cho, Nara City, 630-8577, Japan - yamaguchi-h62@nichi.go.jp

⁵ Research Center for Area Studies, Indonesian Institute of Sciences (P2W-LIPI), Jakarta, Indonesia
- fadjarthufail@gmail.com

⁶ Yogyakarta Archaeology Office (Balai Arkeologi DIY), Yogyakarta, Indonesia - sugeng66riyanto@gmail.com

⁷ Borobudur Conservation Office, Magelang, Jawa Tengah, Indonesia - bramantarayk@gmail.com

Commission II, Photogrammetry

KEY WORDS: Virtual Reality, 3D Scanned Point Clouds, Digital Archive, Cultural Heritage.

ABSTRACT:

Three-dimensional point clouds are becoming popular representations for digital archives of cultural heritage sites. The Borobudur Temple, located in Central Java, Indonesia, was built in the 8th century. Borobudur is considered one of the greatest Buddhist monuments in the world and was listed as a UNESCO World Heritage site. We are developing a virtual reality system as a digital archive of the Borobudur Temple. This research is a collaboration between Ritsumeikan University, Japan, the Indonesian Institute of Sciences (LIPI), and the Borobudur Conservation Office, Indonesia. In our VR system, the following three data sources are integrated to form a 3D point cloud: (1) a 3D point cloud of the overall shape of the temple acquired by photogrammetry using a camera carried by a UAV, (2) a 3D point cloud obtained from precise photogrammetric measurements of selected parts of the temple building, and (3) 3D data of the hidden relief panels recovered from the archived 2D monocular photos using deep learning. Our VR system supports both the first-person view and the bird's eye view. The first-person view allows immersive observation and appreciation of the cultural heritage. The bird's eye view is useful for understanding the whole picture. A user can easily switch between the two views by using a user-friendly VR user interface constructed by a 3D game engine.

1. INTRODUCTION

Three-dimensional (3D) point clouds are popular representations for digital archives of cultural heritage sites and objects. A point cloud is represented as a set of points having coordinate values and color values for an object based on laser scanning or photogrammetry. The Borobudur Temple, located in Central Java, Indonesia, was built in the 8th century. Borobudur is considered one of the greatest Buddhist monuments in the world and was listed as a UNESCO World Heritage site (see Figures 1 and 2). The temple has the world's largest and most complete collections of Buddhist reliefs (John Miksic, 2012). However, some of the reliefs were covered by stone walls and were hidden following the reinforcements during the Dutch rule. For the hidden parts, a 3D reconstruction method based on deep learning is applied to reconstruct the invisible parts into point cloud data from old monocular photos taken by Kassian Cephas in 1890 (Pan et al., 2020). The point clouds used in this study were created from multiple data sources using different methods. This research is an international collaborative project of Ritsumeikan University, Japan, Indonesian Institute of Sciences (LIPI), and Borobudur Conservation Office, Indonesia. In this paper, we present a virtual reality (VR) system as a digital archive of the Borobudur Temple. VR is suitable for viewing and studying archived cultural heritage sites and objects. In Section 2, we introduce some examples of digital preservation of cultural prop-

erties using VR.



Figure 1. Borobudur Temple in Indonesia (photograph).

* Corresponding author



Figure 2. A photograph of our selected part of the Borobudur Temple (the area around the Karmavibhangga relief at the ground level).

2. RELATED WORK

A digital archive is a method of digitizing and preserving materials and cultural heritage. For example, by digitizing valuable historical and cultural heritages that are deteriorating or damaged and generating high-precision reproductions, it is possible to protect valuable cultural heritages from loss, destruction, and deterioration and to pass them on to future generations in their original form. In addition, by sharing digitized cultural properties through networks, it is possible to access them without being restricted by location.

Currently, such digital archives are used not only to digitize two-dimensional materials such as documents and paintings but also to digitize and preserve tangible cultural properties such as archeological sites and 3D objects, as well as intangible cultural properties formed by sound, acoustics, and human body movements. Recently, the methods using VR and AR have become popular as digital archiving methods of cultural properties. Digital archiving of cultural assets using VR and AR can enhance how culture is experienced. The benefit is in terms of both the number of people who can have access to knowledge and the quality of the knowledge diffusion. The effectiveness of VR for virtually experiencing cultural heritage sites has been proven in many studies.

Gonzalez Vargas et al. surveyed the latest trends in AR, VR, and MR systems from the perspective of cultural heritage preservation (Gonzalez Vargas et al., 2020). They showed that VR technology is most suitable for virtual museums, which is the goal of our study. For example, Ch'ng et al. showed agreement in virtual reality as environments for learning and experiencing cultural heritage (Ch'ng et al., 2020). In another study, Paladini et al. developed a virtual museum for the Bagan archeological sites in Myanmar to deepen the understanding of cultural heritage values. In the virtual museum, the users are able to increase their interest in the temples and awareness of the importance of their preservation through VR experiences of real temples (Paladini et al., 2019).

When constructing a VR system for cultural heritage, how to create 3D models of cultural properties, especially 3D models of large-scale complex objects, is a problem that needs to be considered according to the available resources and the characteristics of different target objects. For example, 3D models of the floats in the Virtual Yamahoko Parade were created based

on CAD data of the floats (Li et al., 2014).

Tallon made laser scanning of the entire structure of the Notre Dame Cathedral (Tallon, 2014). The high-precision point clouds of the cathedral are not only important for preserving the iconic piece of Gothic architecture in France but can also help professionals rebuild the cathedral after the tragic fire in 2019.

Photogrammetry is another popular technology to capture the 3D point clouds of cultural heritages. With the development of computer vision algorithms and software tools, photogrammetry can achieve good accuracy with simple equipment (Aicardi et al., 2018). Alshawabkeh et al. adopted a hybrid approach combining photogrammetry and laser scanning for 3D documentation of the Qasr Al-Abidit palace in Jordan (Alshawabkeh et al., 2020). In our study, we adopt photogrammetry for 3D scanning of the surficial structure of the Borobudur Temple and integrate materials from multiple data sources to create 3D point clouds for the VR system.

3. ACQUISITION AND CREATION OF THE ARCHIVED POINT CLOUD DATA

We are working on a digital archive of the Borobudur Temple. Each part of the temple is measured and recorded using a method appropriate for the section. Therefore, the archived datasets are not uniform and have different forms. However, by converting each dataset into point cloud data, we can integrate them to create a large integrated point-based dataset.

3.1 Overall scanning of the whole temple building

3D measurement is becoming an increasingly popular method for obtaining point cloud data of buildings. The main methods for 3D measurement of buildings are laser measurement and photogrammetric measurement. In this study, the SfM-MVS method, which is one of the photogrammetric measurement methods, was used to generate point cloud data of the entire temple. In the SfM-MVS method, photogrammetric data taken with a digital camera is imported into Agisoft Metashape, a multiview stereo photogrammetry software, and 3D data are generated by merging multiple photogrammetric data.

UAV (unmanned aerial vehicles) can be used to acquire photogrammetric data of areas that are difficult to photograph, such as roofs. It is possible to capture a complete image by combining data taken by a UAV.

By using 60 shots of the entire temple taken by a UAV and a digital camera (DJI FC300S) with a resolution of $4,000 \times 3,000$ pixels, 3D data of the entire temple are generated using the photogrammetric method.

When the photogrammetric data are acquired, it should be assumed that the data will be converted into 3D data. The photogrammetric measurement method requires overlapping images to be combined.

Therefore, by being conscious of taking a series of photos, accurate 3D data can be generated. The equipment required for the photogrammetric measurement method is basically only a digital camera, and compared to laser measurement, it requires no extensive preparation and has the advantage of being inexpensive.

3.2 Detailed scanning of a focused area

Point clouds of the details of select parts of the temple were also created using photogrammetry from multiple photos taken with

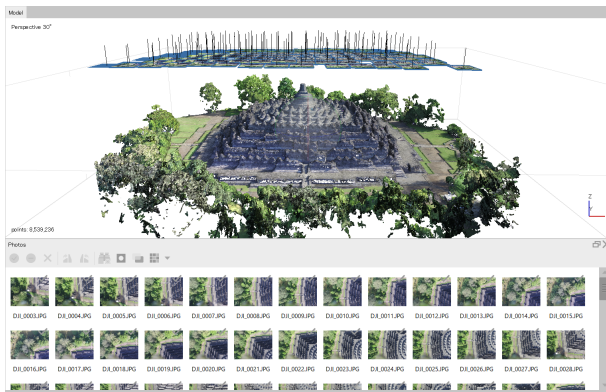


Figure 3. Example of the camera positions for reconstructing the whole temple building by means of photogrammetry.

a high-resolution digital camera. Details of a focused area can be captured using close-range digital photogrammetry. In our work, RICOH GR III with a resolution of $6,000 \times 4,000$ pixels was used for image collection. A monopod was used to help support the camera (Figure 4). Figure 5 shows an example of the camera positions for reconstructing the detailed point cloud of a focused area.



Figure 4. A snapshot photo showing our 3D measurement activity.

3.3 3D recovery of hidden reliefs from 2D photo images

The Borobudur Temple features a collection of the world's largest Buddhist reliefs. However, some temple reliefs are hidden behind stone walls and are not currently visible. The only remaining visual data about the 160 panels of the hidden relief are grayscale monocular photographs (Figure 6) taken in 1890. Pan et al. proposed a deep learning-based method to reconstruct the 3D point cloud of the hidden reliefs from grayscale monocular photographs (Pan et al., 2020).

The idea of 3D reconstruction was based on a depth estimation neural network. The network was first trained with a set of monocular photographs and their corresponding depth maps. The training data were generated from the photogrammetric point clouds of the visible reliefs. Then, it was possible to efficiently estimate its corresponding depth map using the trained neural network for an input monocular photograph with unknown depth information. The 3D point clouds were generated from the estimated depth maps and were imported into our VR system.

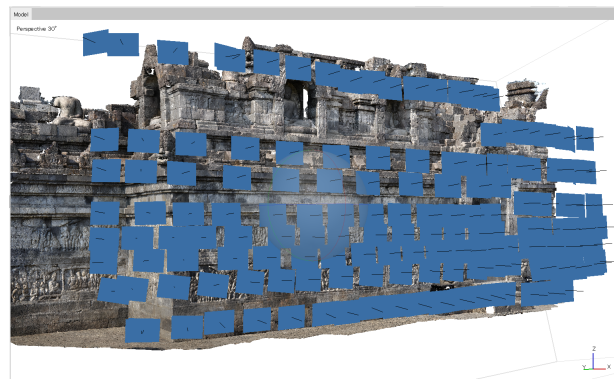


Figure 5. Example of the camera positions for acquiring the detailed point cloud of a focused area by means of photogrammetry.



Figure 6. An example of the 2D monocular photos taken in 1890s that record the reliefs currently covered by the stone walls and made invisible.

4. 3D VIEWING WITH A POINT-BASED VR SYSTEM

In our work, we used a 3D game engine that enabled us to develop the required VR functions. Recent game engines support point rendering in addition to conventional polygon rendering, as well as the combination of point and polygon renderings. The limitation of the point rendering by the 3D game engine is that we need to reduce the data size by thinning the original points, which is crucial to realize real-time rendering. In addition, uniform thinning is required to guarantee good rendering quality.

4.1 VR platform

A game engine is a software library or an integrated development environment developed for creating computer games. Originally, game engines were designed for computer graphics (CG) using polygons. Recently, however, point graphics using point clouds as input data have become possible, allowing the use of 3D point clouds. The game engine uses the most advanced computer graphics technology and is equipped with tools for advanced rendering, animation, and collision detection. In addition, there are many functions for creating VR spaces and user interface functions to control VR spaces. Therefore, beyond the original purpose of game production, it is widely used for video production, scientific and technological visualization, and VR development for tourism, architecture, and medicine.

We aim to apply the VR space constructed by the game engine to the digital archiving of cultural properties. There are several game engines that are currently available. We imported, integrated, and rendered the point clouds using Unity (developed and provided by Unity Technologies, Inc.), which has the largest

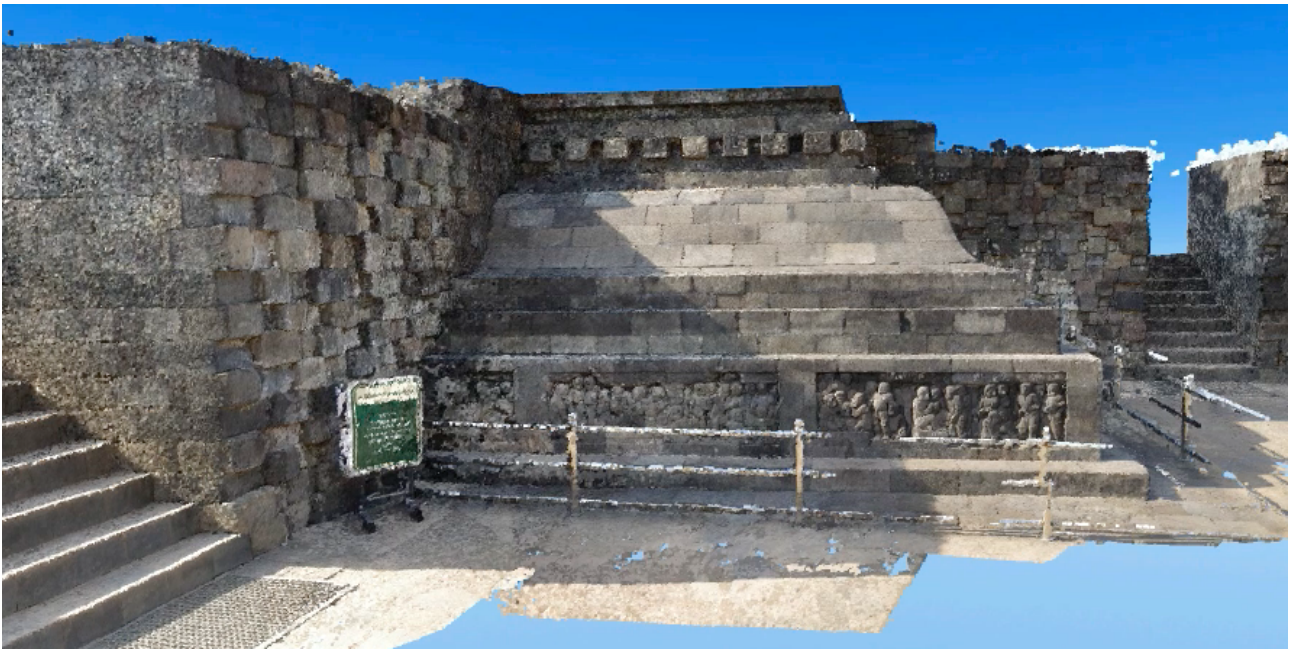


Figure 8. First-person view of the Borobudur Temple viewed in our VR system.

market. Unity is also an excellent VR development environment and enabled us to easily develop VR content compatible with many VR devices.

Three types of point clouds introduced in Section 2 were integrated and incorporated into our VR system. To load point cloud data, we converted the point clouds into PLY format and imported them into Unity using an asset called Pcx (Pcx - Point Cloud Importer/Renderer for Unity, n.d.).

A low frame rate can cause motion sickness in a VR space. However, the integrated point cloud of the original data contains more than 100 million points. Rendering such a large-scale point cloud in VR space may significantly reduce the frame rate. Therefore, we downsample the three original point clouds using the data-thinning method based on Poisson disc sampling (PDS) (Bridson, 2007). PDS is an algorithm that generates a uniform point cloud by downsampling the point cloud with a predefined point distance. In this study, we reduced the number of points in each point cloud to 4 million using PDS. By this data reduction, we realized a rendering speed of 30 fps in our VR system. In addition, the point size was enlarged to avoid possible visibility degradation after downsampling.

4.2 Bird's-eye view and first-person view

There are two major types of viewpoints that can be set in the VR space. One is the bird's-eye view, and the other is the first-person view. The bird's-eye view is convenient to perceive the whole scene.

We can also move and rotate the temple building by the mouse operation and set the viewpoint to the user's favorite angle. The first-person view takes advantage of the VR features. This viewpoint allows us to observe and appreciate the cultural site with a sense of immersion. We implemented a function to switch one viewpoint to another with a simple mouse operation (see Figure 7).

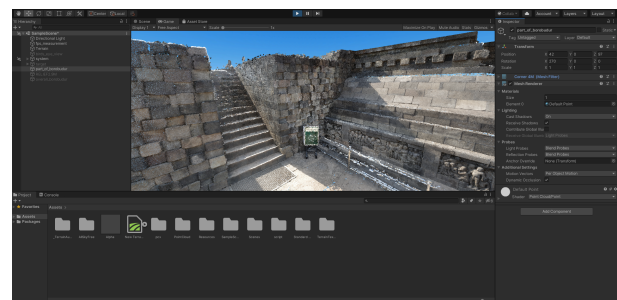


Figure 7. Switching of the viewing styles using the Unity function.

4.2.1 VR with first-person view A first-person view is the viewpoint of the user who is operating the system. Fig. 8 shows an example of the image viewed by the user from a subjective viewpoint. The first-person view enables an immersive virtual tour of Borobudur with mouse and keyboard controls.

From a subjective point of view, our VR system allows people worldwide to virtually experience the World Heritage Site of Borobudur without the restraints of time and space. What is important in the virtual experience is a natural walkthrough. To achieve this, it is important to determine the collision between the camera and the ground or the camera and the temple building. To implement the collision detection function between a camera and a point cloud, we use transparent polygon meshes to represent the ground and the stairs of the temple. By integrating these meshes into the point cloud, a natural walkthrough is enabled in the VR environment.

4.2.2 VR with bird's-eye view The bird's-eye view is a viewpoint that looks down at the whole temple from above. Figure 9 shows an example of the image viewed by the user from the bird's-eye view. The bird's-eye view provides an overall perspective of the entire temple as well as its surroundings from

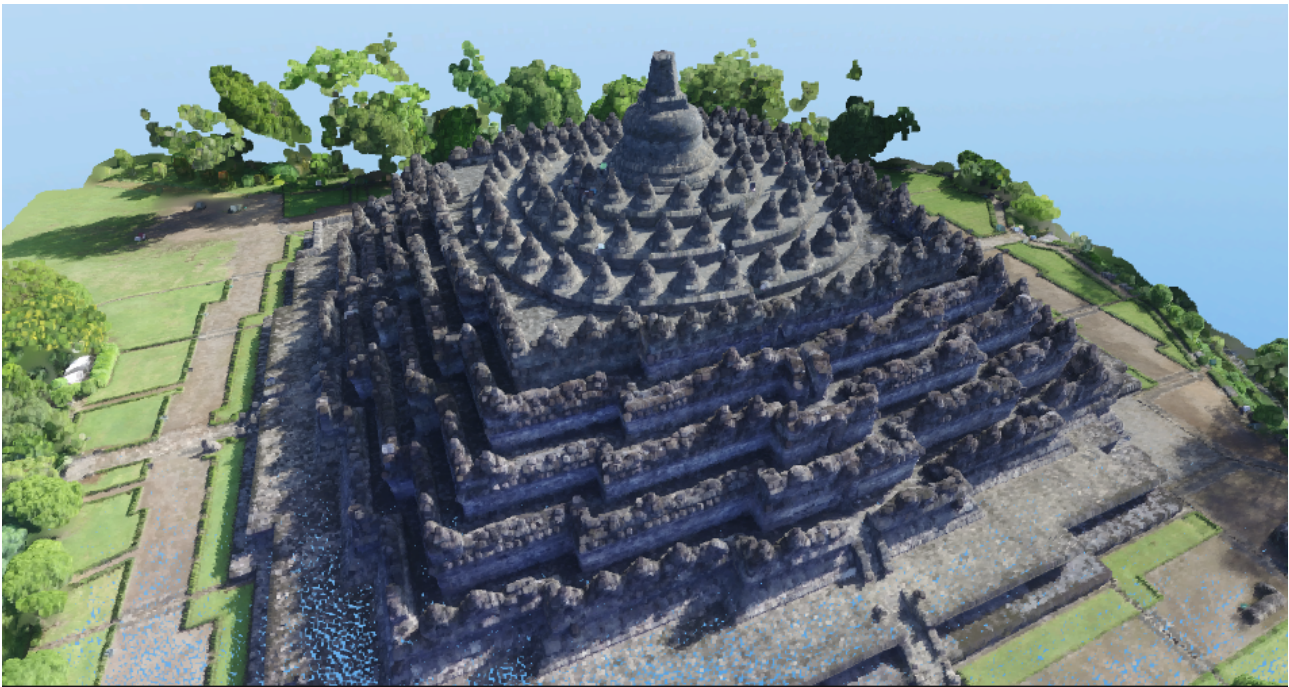


Figure 9. Bird's eye view of the Borobudur Temple viewed in our VR system.



Figure 10. Point cloud of a hidden relief panel reconstructed from a single monocular photo using our (pan et al., 2020) deep learning-based depth estimation method.

various angles. The bird's-eye view is advantageous for understanding the whole cultural property, which is useful for academic analysis and school education. The Borobudur Temple is well known as a school excursion destination for elementary and junior high school students in Indonesia. It can be used to prepare for such school trips and to review after trips. We have been working with local community centers and Indonesian archeologists to make our VR system available to the general public. Our VR system can be used for lifelong learning for the general public.

4.3 3D viewing of the recovered reliefs

These reliefs of the Borobudur Temple, which have great cultural value, depict Indonesian customs and lifestyle at the time of the Borobudur Temple's construction. In addition to the zero-level reliefs of the hidden foot that have been reconstructed into

3D using deep learning as described in Section 3.3, we are also working on measuring and visualizing the visible relief panels of the higher levels. The results will be incorporated into our VR system. Figure 10 shows an example of the reconstructed 3D relief panel.

5. CONCLUSION AND DISCUSSION

This paper presented a virtual reality system as a digital archive of the Borobudur Temple using a 3D point cloud. In our VR system, we integrated a 3D point cloud of the overall shape of the temple acquired by photogrammetry using a camera carried by a UAV, a 3D point cloud obtained from precise photogrammetric measurements of selected parts of the temple building, and 3D data of the hidden relief panels recovered from archived 2D monocular photos using deep learning. Our VR system supports both the first-person view and the bird's eye view. The first-person view allows immersive observation and appreciation of cultural heritage. The bird's eye view is useful for understanding the whole picture. In the next few years, we plan to carry out detailed 3D measurements of all the levels of the Borobudur Temple and reconstruct the high-definition VR scene of the entire temple. Currently, we are working on measuring the first level. See-through viewing (Tanaka et al., 2016, Uchida et al., 2020) and other ingenious attempts to visualize the entire Borobudur site (visible and invisible portions) are also planned. Our system is developed based on Unity. Therefore, it is easy to apply to various platforms and environments. For example, various applications, such as web-based remote VR systems and stereoscopic vision by holography, can be implemented in the future.

ACKNOWLEDGEMENT

In this paper, images of the Borobudur Temple are presented with the permission of the Borobudur Conservation Office and

Research Center for Area Studies (P2W) of the Indonesia Institute of Sciences (LIPI). Additionally, photogrammetry scanning point cloud data were provided by the Nara National Research Institute for Cultural Properties. We would like to thank Pan Jiao (Ritsumeikan University) for providing the point cloud data of the reconstructed relief panels. We would like to express our deepest gratitude to all of them for their cooperation. This research is supported by JSPS KAKENHI Grant No. 19KK0256 and the AY2020 Program for Asia-Japan Research Development, Ritsumeikan University.

Uchida, T., Hasegawa, K., Li, L., Adachi, M., Yamaguchi, H., Thufail, F. I., Riyanto, S., Okamoto, A., Tanaka, S., 2020. Noise-robust transparent visualization of large-scale point clouds acquired by laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 161, 124–134.

REFERENCES

Aicardi, I., Chiabrando, F., Lingua, A. M., Noardo, F., 2018. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *Journal of Cultural Heritage*, 32, 257–266.

Alshawabkeh, Y., El-Khalili, M., Almasri, E., Bala 'awi, F., Al-Massarweh, A., 2020. Heritage documentation using laser scanner and photogrammetry. The case study of Qasr Al-Abidit, Jordan. *Digital Applications in Archaeology and Cultural Heritage*, 16, e00133.

Ch ' ng, E., Li, Y., Cai, S., Leow, F.-T., 2020. The Effects of VR Environments on the Acceptance, Experience, and Expectations of Cultural Heritage Learning. *J. Comput. Cult. Herit.*, 13(1).

Gonzalez Vargas, J. C., Fabregat, R., Carrillo-Ramos, A., Jové, T., 2020. Survey: Using Augmented Reality to Improve Learning Motivation in Cultural Heritage Studies. *Applied Sciences*, 10(3).

John Miksic, Anita Tranchini, M. T., 2012. *Borobudur: Golden Tales of the Buddhas (Periplus Travel Guides) (English Edition)*. Periplus Editions.

Li, L., Choi, W., Hachimura, K., Yano, K., Nishiura, T., Tanaka, H. T., 2014. [Paper] Virtual Yamahoko Parade Experience System with Vibration Simulation. *ITE Transactions on Media Technology and Applications*, 2(3), 248-255.

Paladini, A., Dhanda, A., Reina Ortiz, M., Weigert, A., Nofal, E., Min, A., Gyi, M., Su, S., Van Balen, K., Santana Quintero, M., 2019. IMPACT OF VIRTUAL REALITY EXPERIENCE ON ACCESSIBILITY OF CULTURAL HERITAGE. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W11, 929–936.

Pan, J., Li, L., Yamaguchi, H., Hasegawa, K., Thufail, F., Brahmanantara, Tanaka, S., 2020. Fused 3D transparent visualization for large-scale cultural heritage using deep learning-based monocular reconstruction. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 5(2).

Tallon, A., 2014. Divining Proportions in the Information Age. *Architectural Histories*, 2(1).

Tanaka, S., Hasegawa, K., Okamoto, N., Umegaki, R., Wang, S., Uemura, M., Okamoto, A., Koyamada, K., 2016. SEE-THROUGH IMAGING OF LASER-SCANNED 3D CULTURAL HERITAGE OBJECTS BASED ON STOCHASTIC RENDERING OF LARGE-SCALE POINT CLOUDS. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, III-5, 73–80.