DEVELOPING A MULTIMODAL DATABASE OF DIGITAL ARCHIVES FOR CULTURAL HERITAGE SITES – A CASE OF DIGITALLY PRESERVING THE BOROBUDUR TEMPLE OF INDONESIA

B. Batjargal [1*], J. Pan [1 **], Sh. Ji [1], L. Li [2], H. Yamaguchi [3], K. Hasegawa [1], T. Nishibayashi [4], A. Maeda [2], U. Sarjiati [5], F. I. Thufail [5], Brahmantara [6], S. Tanaka [2]

[1] Research Organization of Science and Technology, Ritsumeikan University, Shiga 525-8577, Japan - biligsaikhan@gmail.com, (panjiao, jishenyu)@gst.ritsumei.ac.jp, hasegawa@media.ritsumei.ac.jp
[2] College of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan - liliang@fc.ritsumei.ac.jp, (amaeda, stanaka)@is.ritsumei.ac.jp
[3] Nara National Research Institute for Cultural Properties, Nara 630-8577, Japan - yamaguchi-h62@nich.go.jp
[4] College of Letters, Ritsumeikan University, Kyoto 603-8577, Japan - xilin@lt.ritsumei.ac.jp
[5] Research Center for Area Studies, National Research and Innovation Agency, Jakarta 12710, Indonesia - upik.sarjiati@gmail.com, fadj001@lipi.go.id
[6] Borobudur Conservation Office, Magelang 56553, Indonesia - brahmantara@kemdikbud.go.id

KEY WORDS: Cultural Heritage, Digital Archive, Digital Humanities, Database, 3D Reconstruction, Information Retrieval, Borobudur Temple.

ABSTRACT:

Digital archives that support the three-dimensional (3D) digitization of cultural heritage sites are gaining popularity and attracting enormous interest as a critical strategy for digital preservation. This paper investigates the construction of a multimodal database for cultural heritage sites, which aims to preserve 3D data, images, videos and other materials along with multilingual text data. 3D scanning is widely used for efficient and accurate digitization of cultural heritages. However, some sites have been damaged entirely or partially, or sometimes, it is impossible to conduct 3D scanning of a portion of a site that is hidden from sight. At the Borobudur temple, a UNESCO World Heritage site, there are buried reliefs that are hidden at the basement and inaccessible for 3D scanning. Currently available data for these inaccessible portions of the site include old photos, computer-aided design (CAD) drawings and historical written descriptions. To digitally reconstruct 3D models of hidden parts, we introduce a method that use a single monocular photo. We constructed "Borobudur Visual Archives" to store, access, view and retrieve Borobudur-related digital assets, including the 3D data that we created.

1. INTRODUCTION

1.1 Background

As a result of worldwide digitization over the last decades, many cultural institutions (libraries, galleries, museums or archives) have started to digitize their collections and expose digital objects on the Internet in a variety of formats and languages through digital archives. Such digital archives offer global users access to a variety of resources through the internet from all over the world. There have been increasing efforts from libraries, galleries, museums and archives to preserve their collections using 3D technologies and to make them available online for users who want to take a close look. Similarly, the need for digital preservation of cultural heritage sites is growing using recent advancements in the 3D digitization of cultural heritage sites (Bachi et al., 2014). With increasingly more digital objects of cultural heritage sites becoming available on the Internet, it is essential to develop methods to access these vast and valuable collections of cultural heritage easily and thoroughly. Users require an effective and efficient way to search and view valuable collections of cultural heritage sites. Thus, the goal of this study is to develop a multimodal database of multilingual digital archives, which allows searching and browsing 3D data, images, videos and other related materials of cultural heritage sites.

1.2 Primary Target Cultural Heritage Site

The primary target of this research is the Borobudur Temple, a UNESCO World Heritage Site in Indonesia. This paper summarizes the approaches with regard to 1) reconstructing the 3D models of hidden parts of the Borobudur temple, and 2) developing a multimodal database that allows searching and browsing Borobudur-related digital assets, including images, videos, and 3D data that we have created.

To capture the 3D structure of the Borobudur temple digitally, we must understand the architecture of the temple. 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 parts are square bases measuring 113 meters on each side (Miksic et al., 2012). The upper platforms have diminished height, and the highest point of the monument is 35 meters above ground level. The upper platforms include a large stupa built of volcanic rock from a nearby volcano that is built on a hill that serves as the foundation of the temple, which is artificially built from three smaller naturally formed hills. The exterior of the temple consists of nine stacked platforms, of which the lower six are square and the upper three are circular. The upper three circular platforms contain seventy-two surrounding small stupas and are topped by

* Corresponding author
** This author is currently affiliated with School of Intelligence Science and Technology, University of Science and Technology Beijing, Beijing 100083, China. Email: panjiao@ustb.edu.cn

https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-713-2023

one large central dome (Santiko and Nugrahani, 2012). On the stone walls of the lower six square platforms, 2672 panels of bas-reliefs containing 1460 narrative panels and 1212 decorative panels, are carved. These reliefs can be divided into five sections, each of which describes a different Buddhist story. Unfortunately, the section carved on the basement called "Karmawibhangga" was buried behind the stone walls after the reinstalling the base encasement due to safety concerns. Karmawibhangga has 160 relief panels, and only four panels at the southeast corner are visible. The remaining 156 panels are hidden by the stone walls.

We propose a 3D reconstruction method for the hidden reliefs, which will be explained in Section 3.3.3.1. This paper is organized as follows. In Section 2, some related studies are introduced. The proposed approach is then introduced in Section 3. Some concluding remarks are given in Section 4.

2. RELATED WORK

Laser scanning and photogrammetry are widely used to generate 3D models. Although laser scanning requires expensive and large equipment, it can extract the geometry of the site surfaces. For example, Tallon performed laser scanning of the entire structure of the Notre Dame Cathedral (Tallon, 2014), which helped professionals rebuild the cathedral after the tragic fire in 2019. Photogrammetry is an inexpensive solution that can be performed with a good-quality camera and a computer (Aicardi et al., 2018). Software developments in photogrammetry scanning have also led to increased automation and higher performance (Alidoost and Arefi, 2017). Some studies (Nuttens et al., 2011; Remondino, 2011; Girelli et al., 2020; Bitelli et al., 2019; Alshawabkeh et al., 2020) have used hybrid approaches that combine photogrammetry and other scanning methods to digitize cultural heritages. However, it is impossible to conduct 3D scanning of a portion of a site that is hidden from sight.

Instead, 3D reconstruction from previously preserved 2D archiving materials is widely used to reconstruct destroyed or inaccessible parts of cultural heritage sites. Many studies (Dhonju et al., 2017; Kersten and Lindstaedt, 2012; Kyriakaki et al., 2014; Themistocleous, 2017; Condorelli and Rinaudo, 2018) have created 3D models of cultural heritage sites from historical or open-source photographs and videos. Brůha et al. combined diverse 3D data sources and referential terrain models for reconstructing the vanished medieval town of Sekanka and the surrounding landscape (Brůha et al., 2020). Li et al. created 3D models of festival floats in the Virtual Yamahoko Parade from CAD drawings (Li et al., 2014). Külür and Şahin reconstructed old Safranbolu houses from CAD drawings using 3D modelling software (Külür and Şahin, 2008).

Moreover, with the rapid development of machine learning, many studies have also applied machine learning techniques to the reconstruction of cultural heritage assets. Hermoza and Sipiran proposed a generative adversarial network to reconstruct complete 3D points from incomplete archaeological objects (Hermoza and Sipiran, 2018). Belhi et al. proposed a deep-learning-based image inpainting method to reconstruct incomplete artworks (Belhi et al., 2020).

Zheng et al. made high-quality digital archives of the Bas-reliefs of a historical temple building of Angkor Wat (Zheng et al., 2015). Overall, creation of digital archives that support the 3D models of large-scale cultural heritage sites that have a portion hidden from sight remains challenging. Many studies only focus on generating datasets, such as acquisition methods of measurement data and classification methods of acquired data.

3. PROPOSED APPROACH

In this paper, we introduce an approach to construct a multimodal database for cultural heritage sites, which aims to preserve 3D data, images, videos and other related materials along with multilingual text data. We also introduce a method to reconstruct 3D models from a 2D photo. Using our achievements, we have constructed a multimodal database named "Borobudur Visual Archives" to store, share, access, view and retrieve Borobudur temple-related digital assets.

3.1 Borobudur Visual Archives

We consider Borobudur Visual Archives (BoVA) as a multimodal database that is capable of storing, browsing and retrieving Borobudur temple-related digital assets, including 3D data, images, videos and multilingual text data. The primary features of the BoVA let users 1) aggregate various digital objects of multiple digital archives; 2) review multilingual data that contain detailed information and links to the original objects; and 3) search and browse various different digital objects such as 3D data, images and videos in a unified database. The BoVA allows the registration of multiple objects or different variants of the same digital objects. If text data such as explanations or metadata written in different languages are available, the proposed system can also store text data in its database. Such a feature is useful because some digital archives are primarily available in their native language; thus, users who do not understand a certain language may not find the desired information.

3.2 Design and Structure of the BoVA

We adopted a database framework with a concept named the "portal database" (Yamaji and Akama, 2021) after reviewing the designs and structures of databases capable of handling 3D data, images, videos and multilingual text data. The portal database framework is capable of integrating various databases based on the similarity of their contents.

3.2.1 The Portal Database Framework: It supports multi-users or multi-projects, where the different projects can combine their databases into a single database system, or each group, project or collector can also operate their database independently. While the data are stored in a unified database system, access levels for a database can be set by access permissions such as which projects can use which data. By using the portal database framework, many traditional or contemporary cultural heritages of Japan are being digitized and made publicly available as searchable databases, which serve as catalogues of cultural heritage materials. Also, the portal database framework's disclosure feature can control whether the data shall be open to the public or for limited use. Combining the security layers of the databases and content security, the portal database framework has been taking a multi-layered approach to manage data. The portal database framework can be used for a variety of purposes, depending on the needs, such as accumulating, managing, publishing and doing research. We used the portal database framework for constructing the BoVA, which can store, share, access, view and retrieve Borobudur temple-related digital assets. The BoVA can also be used for various cultural resources, which can accommodate any kind of cultural resource, including multimedia, images, audio videos and 3D data. The accumulation and management of cultural resources and materials (i.e., the archive) is a particularly essential function for cultural organizations.

3.2.2 Multilingual Support: The BoVA have basic multilingual support to 1) translate user interface, 2) translate metadata element names, and 3) store multilingual text data separately. To translate and change the display language, we use language files, which correspond to the terms and phrases used in the user interface and metadata element names. Language configuration files can also be set at the individual database level for customized translation over the system-wide translation. Although all the language files available on the web can be installed in the BoVA, we have deployed English and Japanese language only, where English is the default language. To store multilingual text data, the BoVA use separate metadata fields for different languages. Text data including metadata and explanations of the Borobudur-related digital assets in the BoVA are written in Indonesian, Dutch, English, and Japanese.

3.3 Current Content in the BoVA

The BoVA aim to preserve Borobudur-related 3D data, images, videos and other related materials along with multilingual text data. As of April 2023, the BoVA contain the following contents: 1) 872 photos of visible relief panels; 2) 464 photos of the hidden reliefs Karmawibhangga; and 3) 160 reconstructed 3D models of hidden Karmawibhangga reliefs, which will be explained in Section 3.3.1, Section 3.3.2 and Section 3.3.3, respectively. The BoVA uses the reliefs' numbering according to the Krom's archaeological description (Krom, N.J., 1920). Thus, in the BoVA, each relief panel has own identifier, and the position of every single panel can be identified by using its identifier.

3.3.1 Photos of Visible Relief Panels: As mentioned in Section 1.1, there are 2672 relief panels in the Borobudur temple. We have been photographing and conducting analysis of each panel. 872 photos of visible relief panels, including a) 120 photos of 120 reliefs of the main walls of the first gallery; b) 151 photos of 122 reliefs of the inner walls of the first gallery; c) 130 photos of 128 reliefs of the main walls of the second gallery; d) 98 photos of 87 reliefs of the main walls of the third gallery; e) 91 photos of 88 reliefs of the inner walls of the third gallery; f) 74 photos of 70 reliefs of the main walls of the fourth gallery; g) 137 photos of 82 reliefs of the inner walls of the fourth gallery; and h) other 10 photos are now viewable and searchable in the BoVA.

As shown in Figure 2, text data and metadata of visible relief panels are in English and Japanese. Multilingual metadata are used to describe objects (e.g., relief panels) and consist of several descriptive metadata elements (e.g., title, creator, owner). These metadata provide information on digital objects and assist users in searching for them. Particularly in collections of images where textual contents are not available, metadata provide major access points. Users can see search results as a list of multilingual data holding basic metadata and links of the original records along with detailed information (see Figure 2). Users can also see a detailed page with a bigger image and detailed metadata (see Figure 3) by clicking the thumbnails in the list of search results.

Figure 1. Thumbnail view of the Borobudur-related images in the BoVA.

Figure 2. Displaying metadata and thumbnails as a list in the BoVA.

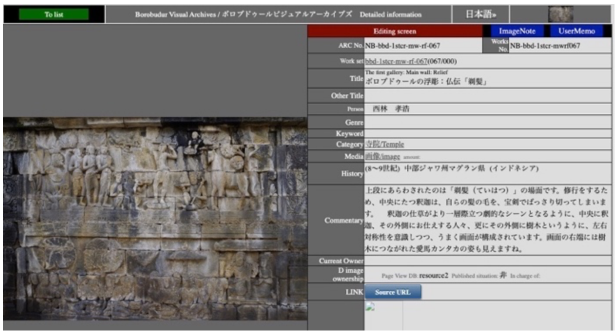

Figure 3. Detailed view of an object in the BoVA.

3.3.2 Photos of the Hidden Reliefs Karmawibhangga: As mentioned before, 156 relief panels of Karmawibhangga are buried behind the stone walls and are no longer visible. Only four panels at the southeast corner are visible. Fortunately, for each hidden panel, there remains a single photo taken by Kassian Cephas (1845-1912) in 1890 (Adams, 1990). Digitized versions of these photos of the hidden relief panels of Karmawibhangga can be downloaded from the collection of "Southeast Asian and Caribbean Images" of the Leiden University Libraries (KITLV) under a Creative Commons CC BY License. There are two versions available: 1) KITLV A762 - Bas-reliëf op de Boroboedoer bij Magelang (Leiden University Libraries, 2014a) and 2) KITLV A765 - Bas-reliëf op de Boroboedoer bij Magelang (Leiden University Libraries, 2014b). As shown in Figure 4, the KITLV A762 collection includes a single photo of each panel taken from the front of the reliefs. In addition, as shown in Figure 5, the KITLV A765 collection includes cropped and edited versions of these photos. By considering the archaeological value and preventing the loss of information on cropped parts, both versions of the original photos and cropped versions are registered in the BoVA. As of April 2023, 156 Karmawibhangga relief photos of the KITLV A762 collection (reliefs 69, 70, 71, and 72 are missing) and 154 cropped photos of the KITLV A765 collection (123 and 160 are missing; 57, 58, 59 and 60 have different images) are now searchable and browsable through the BoVA (see Figure 6). As shown in Figure 7, text data of the KITLV collections are in Dutch and English, which show a list of bilingual data holding basic metadata, and links to the original records, IIIF manifest along with detailed information. As shown in Figure 8, a detailed page with a bigger image and detailed metadata can be shown by clicking the thumbnails in the list of search results.

The online collection database of the Nationaal Museum van Wereldculturen (NMVW) also hosts digitized photos of the hidden relief panels of Karmawibhangga (see Figure 9) under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license (Nationaal Museum van Wereldculturen, 2014). As shown in Figure 10, the NMVW collection also includes grayscale photos of the hidden relief panels. Both versions are registered in the BoVA. As of April 2023, 159 Karmawibhangga relief photos of the NMVW collection (60 are missing) and 158 grayscale photos of the NMVW collection (68 and 160 are missing) are now searchable and browsable through the BoVA (see Figure 11 and Figure 12). As shown in Figure 13 and Figure 14, text data and metadata of the NMVW collections are shown in Dutch and English with the links to the original records and detailed information. As shown in Figure 15, a detailed page with a bigger image and detailed metadata can be shown by clicking the thumbnails in the list of search results.

Figure 4. Example of an old photo of a hidden relief at the KITLV A762 collection.

Figure 5. Cropped and edited example of an old photo of a hidden relief.

Figure 6. Thumbnail view of KITLV collections in the BoVA.

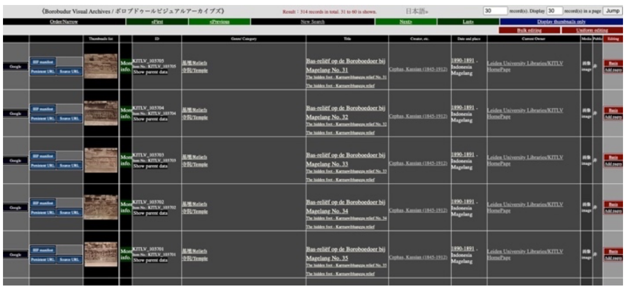

Figure 7. Displaying metadata and thumbnails of the KITLV collections in the BoVA.

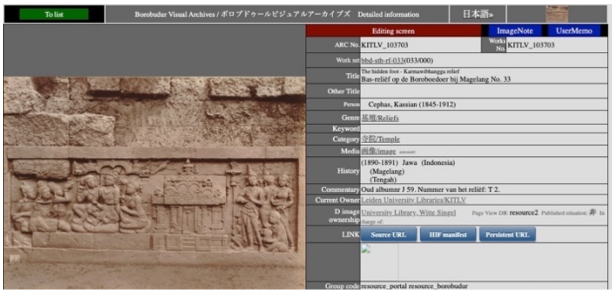

Figure 8. Detailed view of the KITLV collections in the BoVA.

Figure 9. Example of an old photo of a hidden relief at the NMVW collection.

Figure 10. Grayscale example of an old photo of a hidden relief at the NMVW collection.

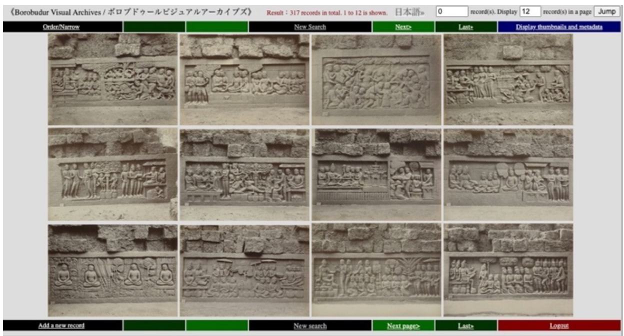

Figure 11. Thumbnail view of the NMVW collections in the BoVA.

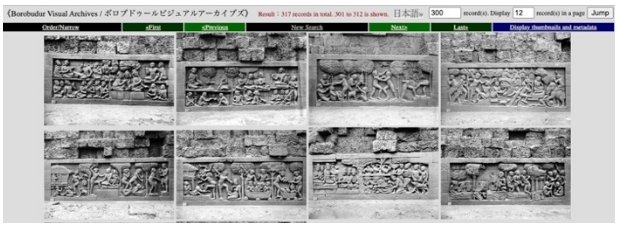

Figure 12. Thumbnail view of grayscale photos of the NMVW collections in the BoVA.

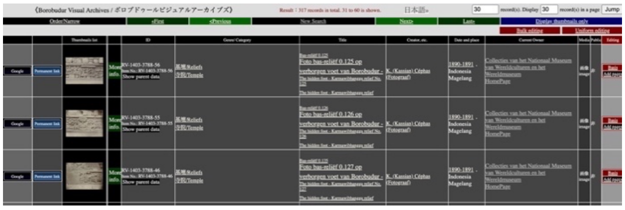

Figure 13. Displaying metadata and thumbnails of the NMVW collections in the BoVA.

Figure 14. Displaying metadata and thumbnails of grayscale photos of the NMVW collection in the BoVA.

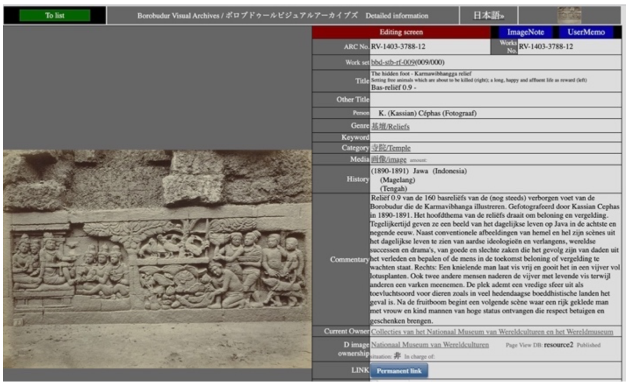

Figure 15. Detailed view of the NMVW collection in the BoVA.

In 2012, Balai Konservasi Borobudur published a book titled "Adegan dan Ajaran Hukum Karma pada Relief Karmawibhangga" (Santiko and Nugrahani, 2012), which also includes photos of all 160 Karmawibhangga reliefs of the hidden foot. As shown in Figure 16, this book includes cropped and edited grayscale photos. In this book, explanations are provided in Indonesian and English. These photos, along with the explanations in Indonesian and English, can also be searched and viewed in the BoVA.

Figure 16. Cropped and edited example of a hidden relief at the Balai Konservasi Borobudur.

3.3.3 Reconstructed 3D Models of 160 Hidden Karmawibhangga Reliefs: 3D scanning is widely used for digitizing cultural heritage objects and creating 3D digital archives of cultural heritage sites. However, Karmawibhangga relief panels are hidden from sight and are not accessible for 3D scanning. Because there is a single monocular photo available for each panel, the traditional multi-view image-based method is not applicable and is insufficient for 3D reconstruction. Thus, we used a deep-learning-based method to reconstruct 3D models of the hidden Karmawibhangga reliefs. We reconstructed all 160 Karmawibhangga hidden relief panels as 3D point clouds by relating and combining information of the old 2D photos and the predicted depth map of the trained deep monocular depth estimation network. Figure 17 shows an example of an old monocular photo, the depth estimation result, and the 3D reconstructed points from the left and right sides.

Figure 17. Example of depth estimation and 3D reconstruction results. From top right to bottom left: an old monocular photo, the depth estimation result, and the 3D reconstructed models.

We used an open-source framework 3D Heritage Online Presenter (3DHOP) (Potenziani et al., 2015) to present the reconstructed 3D models in the BoVA. All the reconstructed 3D models of 160 hidden Karmawibhangga reliefs are now interactively viewable through the BoVA (see Figure 18) along with text data in English and Japanese. To show the reconstructed 3D models on the web using 3DHOP, we converted the point cloud data to the Nexus multiresolution format (Ponchio and Dellepiane, 2015). The next section discusses the proposed method to reconstruct a 3D model of the Karmawibhangga relief from a single monocular old photo.

3.3.3.1 3D Reconstruction of the Karmawibhangga Reliefs from a Single Monocular Photo: To reconstruct 3D models of the hidden relief panels from old monocular 2D photos, we use monocular deep-learning-based methods to estimate the depth values from the 2D photos. If the depth of each point is known, we can reconstruct a 3D model from a single monocular photo.

We use a depth estimation network with an encoder-decoder structure (Pan et al., 2021) to reconstruct hidden reliefs, and trained with the photogrammetry scanning data of the visible reliefs. The depth estimation network uses edge information as guidance in different stages of the decoder. DenseNet (Huang et al., 2017) used as the feature extractor to output a dense feature map. Before the decoder, a denser version (Yang et al., 2018) of the atrous spatial pyramid pooling layer is placed to extract the contextual information. The decoder then restores the resolutions of the feature maps to the original resolution. At different scales in the decoder, we place the edge guidance layer to locate geometric guidance to improve the depth estimation. It improved the accuracy around the soft edges in the reliefs that form boundaries between the carved items and the wall background.

Figure 18. Reconstructed 3D model of a hidden Karmawibhangga relief in the BoVA.

The proposed network employs prior edge images as guidance information, and the corresponding edge images are required for both training and testing. We use the "opacity-based feature-highlighting method" (Kawakami et al., 2020) as the edge extraction method for extracting edge images from 3D point clouds. To extract 3D edges as feature regions on the surface of a 3D scanned point-cloud object, we examine the directions of the local spread of the point distribution. We calculate the 3D structure tensor (i.e., the co-variance matrix of position vectors of points) from local points using a statistical method (Weinmann et al., 2015). Then, we define feature values such as "change of curvature", "linearity" and "aplanarity" using the three eigenvalues of the tensor. Using the opacity-based feature-highlighting method, the centerlines of the 3D edges are drawn more clearly, resulting in thinning the drawn 3D edges and realizing sharper edge lines to express the 3D structure.

This method can also express the soft-edge area in relief as a gradation of surface opacity. If we set the feature-value threshold as a finite width, it is also effective for highlighting soft edges (i.e., rounded edges). Such soft edges often appear in the cultural properties of artworks and also appear when the original clear-cut edges deteriorate over time. The idea of this process is to use the degree of opacity to highlight 3D edges. The extracted soft edges (see Figure 19) can be used in the edge guidance layer of the depth estimation network to enhance more details during upsampling. Figure 19 shows two different edge highlights of a portion of the Karmawibhangga relief of the Borobudur temple. The change in curvature is used as the feature value. Many human characters are drawn, and each has soft 3D edges as its outline.

Figure 19. Highlighting the soft edges of a Karmawibhangga relief of the Borobudur temple.

The upper image of Figure 19 is created by simply assigning a red colour to points with a feature value greater than a fixed threshold value. Because the soft edges tend to have larger widths than the clear-cut edges, the visualized red regions form wide bands rather than thin lines; thus, the outlines of human objects are unclear. Conversely, the bottom image of Figure 19 shows the edge highlighting of the opacity-based feature-highlighting method. Red gradation appears, and the brightness is higher around the centerline, which makes the outlines recognizable more clearly. Generally, the opacity-based feature-highlighting method, which can highlight both the clear-cut and soft edges of point cloud data, contributes to improve the accuracy of the 3D reconstruction of the hidden reliefs from the old 2D photos.

The corresponding depth map of the old 2D photos of the hidden reliefs can be predicted by the trained depth estimation networks. The training dataset is composed of pairs of monocular photos, edge images, and depth maps. This method can extract some details, such as human faces, which have only marginal depth changes compared to the surrounding areas. The artifacts between the patches are produced because the pathwise training ignores the context relationship between neighbourhood patches. This method provides 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. Generally, the shape of the carved figure is clearer in the results of this approach, and noise such as unstable depth changes in the background, which occurs in other models, is reduced. This method also provided better results with detailed tree figures at the middle of the depth maps, distinct decoration objects at the bottom sides of the reliefs, and the correct depth values of the human faces. The proposed 3D reconstruction method was evaluated quantitatively using the TLS data of visible reliefs in the southeast corner. The accuracy of the reconstructed point clouds reached 97%. Refer to our previous study (Pan et al., 2021) for details of a quantitative evaluation and algorithm implementation.

3.3.4 Integrating Multiple 3D Models: To digitally preserve and visualize the entire 3D structure of Borobudur temple, we 1) integrated the 3D scanned data of the visible parts; 2) reconstructed 3D models of the underground foundation; and 3) reconstructed 3D models of the hidden Karmawibhangga reliefs into a unified 3D point cloud dataset. For the visible parts, accurate connections are made by using a total station survey. For hidden relief panels and underground foundations that are hidden from sight, data integration was performed under the supervision of the Borobudur Conservation Office to ensure the depth criteria of multiple 3D models. Refer to our previous study (Pan et al., 2021) for details.

3.3.5 Merging Multiple Results of the Same Object: The BoVA can register items from different organizations and integrate various objects based on the similarity. Multiple identical results of the same object can also be merged and shown as a list. For example, as shown in Figure 20, seven different types of digital assets of Karmawibhangga relief No. 20 can be shown under a unified title: "The hidden foot - Karmawibhangga relief No. 20." The following seven digital assets are registered as the identical object in the BoVA: 1) a photo from the KITLV A762 collection; 2) a photo from the KITLV A765 collection; 3) a photo of the NMVW collections; 4) a grayscale photo of the NMVW collections; 5) a photo from the Balai Konservasi Borobudur; 6) a reconstructed 3D model; and 7) a photo taken by us. Various different digital assets of all 160 Karmawibhangga reliefs of the multiple cultural institutions shown by clicking the "click here to view each relief" button shown in Figure 21.

Figure 20. Unified view of the Karmawibhangga relief No. 20 of the hidden foot.

Figure 21. List of the Karmawibhangga reliefs of the hidden foot.

4. CONCLUSIONS

In this paper, a multimodal database for cultural heritage sites that can preserve high-quality 3D data, images, videos and other related materials along with multilingual text data is introduced. We proposed a 3D digitization method to reconstruct 3D models from 2D photos. We successfully applied the proposed method to the Borobudur temple, a UNESCO World Heritage site. Using these methods, we constructed the multimodal database BoVA to store, share, access, view and retrieve Borobudur temple-related digital assets. Such a database along with multilingual text data is useful for users and allows searching and browsing all Borobudur temple-related digital assets in a single place so that it helps users save time.

In addition, several technologies for 3D visualization were also developed, including 1) a feature-highlighting visualization that can highlight 3D structures of complex shapes in cultural heritage objects; and 2) a method for 3D reconstruction of incomplete or inaccessible parts of cultural heritage sites. The proposed framework can be applied to other large-scale cultural heritage sites. The proposed digital archiving and visualization methods are accurate and efficient and can be deployed to a wide range of tangible cultural heritages.

For future digitization research, we plan to use recent high-precision portable laser scanners in combination with photogrammetry. The edges of the carved patterns on the reliefs contain essential features of both semantic information and the geometric nature of the reliefs. Therefore, effective usage of edge information could improve the reconstruction accuracy. We also plan to apply the proposed integrated visualization to immersive virtual reality scenes such that the proposed digital archives can become available for applications such as digital museums, mobile guidance purposes, and other applications. Moreover, in the future, by linking the 3D visualization system and database system, users will be able to interactively specify the scale of the 3D models. Since all our 3D models are point cloud data, it is possible to specify the 3D scale by referring to them. The scale of the restored point clouds of the 3D reconstructed models of the hidden reliefs was created by using the scale of the visible reliefs.

ACKNOWLEDGEMENTS

In this paper, images of the Borobudur temple are presented with the permission of the Borobudur Conservation Office and Research Center for Area Studies, National Research and Innovation Agency, Indonesia. Additionally, a photogrammetry scanning point dataset was provided by the Nara National Research Institute for Cultural Properties. The Art Research Center and the Asia Japan Institute, Ritsumeikan University, provided research funding, server infrastructure, data storage, and a digital archiving framework. We would like to thank these institutions for their generous cooperation.
This study was partially supported by JSPS KAKENHI Grant Numbers 17K02331, 19KK0256, and 21H04903.

REFERENCES

Adams, R., 1990: *Borobudur in photographs: past and present: photographs in Dutch collections = Borobudur dalam foto: di masa lampau dan kini : foto dari berbagai koleksi Belanda / Roy Adams*. University of Leiden.

Aicardi, I., Chiabrando, F., Maria Lingua, A., Noardo, F., 2018. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *J. Cult. Herit.*, 32, 257–266. doi.org/10.1016/j.culher.2017.11.006.

Alidoost, F., Arefi, H., 2017. Comparison of uas-based photogrammetry software for 3D point cloud generation: A survey over A historical site. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-4/W4, 55–61. doi.org/10.5194/isprs-annals-IV-4-W4-55-2017.

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. *Digit. Appl. Archaeol. Cult. Herit.*, 16 (2020), e00133. doi.org/10.1016/j.daach.2019.e00133.

Bachi, V., Fresa, A., Pierotti, C., Prandoni, C., 2014. The Digitization Age: Mass Culture Is Quality Culture. Challenges for Cultural Heritage and Society. In *Proceedings of the 5th International Conference, EuroMed 2014*, 786–801. doi.org/10.1007/978-3-319-13695-0_81.

Belhi, A., Al-Ali, A.K., Bouras, A., Foufou, S., Yu, X., Zhang, H., 2020. Investigating low-delay deep learning-based cultural image reconstruction. *J. Real-Time Image Process*., *17*, 1911–1926, doi.org/10.1007/s11554-020-00975-y.

Bitelli, G., Balletti, C., Brumana, R., Barazzetti, L., D'Urso, M.G., Rinaudo, F., Tucci, G., 2019. The gamher research project for metric documentation of cultural heritage: Current developments. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W11, 239–246. doi.org/10.5194/isprs-archives-XLII-2-W11-239-2019.

Brůha, L., Laštovička, J., Palatý, T., Štefanová, E., Štych, P., 2020. Reconstruction of Lost Cultural Heritage Sites and Landscapes: Context of Ancient Objects in Time and Space. *ISPRS Int. J. Geo-Inf.*, 9(10), 604. doi.org/10.3390/ijgi9100604.

Condorelli, F., Rinaudo, F., 2018. Cultural heritage reconstruction from historical photographs and videos. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 259–265. doi.org/10.5194/isprs-archives-XLII-2-259-2018.

Dhonju, H.K., Xiao, W., Sarhosis, V., Mills, J.P., Wilkinson, S., Wang, Z., Thapa, L., Panday, U.S., 2017. Feasibility study of low-cost image-based heritage documentation in Nepal. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W3, 237–242. doi.org/10.5194/isprs-archives-XLII-2-W3-237-2017.

Girelli, V.A., Tini, M.A., D'Apuzzo, M.G., Bitelli, G., 2020. 3D digitisation in cultural heritage knowledge and preservation: The case of the neptune statue in bologna and its archetype. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B2-2020, 1403–1408. doi.org/10.5194/isprs-archives-XLIII-B2-2020-1403-2020.

Hermoza, R., Sipiran, I., 2018. 3D reconstruction of incomplete archaeological objects using a generative adversarial network. In *Proceedings of the Computer Graphics International 2018*, 5–11, doi.org/10.1145/3208159.3208173.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2017)*, 2261–2269, doi.org/10.1109/CVPR.2017.243.

Kawakami, K., Hasegawa, K., Li, L., Nagata, H., Adachi, M., Yamaguchi, H., Thufail, F. I., Riyanto, S., Brahmantara, Tanaka, S., 2020. Opacity-based Edge Highlighting for Transparent Visualization of 3d Scanned Point Clouds. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci,* V-2-2020, 373–380. doi.org/10.5194/isprs-annals-V-2-2020-373-2020.

Kersten, T.P., Lindstaedt, M., 2012. Automatic 3D Object Reconstruction from Multiple Images for Architectural, Cultural Heritage and Archaeological Applications Using Open-Source Software and Web Services. *Photogramm. Fernerkund. Geoinf.*, 6, 727–740. doi.org/10.1127/1432-8364/2012/0152.

Krom, N.J., 1920: *Beschrijving van Barabudur. I: Archeologische Beschrijving*. The Hague.

Külür, S., Şahin, H., 2008. 3D Cultural heritage documentation using data from different sources. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, XXXVII, 353–356.

Kyriakaki, G., Doulamis, A., Doulamis, N., Ioannides, M., Makantasis, K., Protopapadakis, E., Hadjiprocopis, A., Wenzel, K., Fritsch, D., Klein, M., Weinlinger, G., 2014. 4D Reconstruction of Tangible Cultural Heritage Objects from Web-Retrieved Images. *Int. J. Herit*., 3, 431–451. doi.org/10.1260/2047-4970.3.2.431.

Leiden University Libraries, 2014a. KITLV A762 - Bas-reliëf op de Boroboedoer bij Magelang. hdl.handle.net/1887.1/item:936610 (20 April 2023).

Leiden University Libraries, 2014b. KITLV A765 - Bas-reliëf op de Boroboedoer bij Magelang. hdl.handle.net/1887.1/item:936613 (20 April 2023).

Li, L., Choi, W., Hachimura, K., Yano, K., Nishiura, T., Tanaka, H.T., 2014. Virtual Yamahoko Parade Experience System with Vibration Simulation. *ITE Trans. Media Technol. Appl*., 2, 248–255. doi.org/10.3169/mta.2.248.

Miksic, J., Tranchini, A., Tranchini, M., 2012: *Borobudur: Golden Tales of the Buddhas*. Tuttle Publishing.

Nationaal Museum van Wereldculturen, 2014. Online collection database. collectie.wereldculturen.nl/ (20 April 2023).

Nuttens, T., De Maeyer, P., De Wulf, A., Goossens, R., Stal, C., 2011. Terrestrial Laser Scanning and Digital Photogrammetry for Cultural Heritage: An Accuracy Assessment. In *Proceedings of the 4th International Workshop on 3D Geo-Information*, 18–22.

Pan, J., Li, L., Yamaguchi, H., Hasegawa, K., Thufail, F. I., Brahmantara; Tanaka, S., 2022. 3D reconstruction of Borobudur Reliefs from 2D Monocular Photographs Based on Soft-Edge Enhanced Deep Learning. *ISPRS J. Photogramm. Remote Sens,* 183*,* 439–450, doi.org/10.1016/j.isprsjprs.2021.11.007.

Ponchio, F., Dellepiane, M., 2015. Fast decompression for web-based view-dependent 3D rendering. In *Proceedings of the 20th International Conference on 3D Web Technology*, 199–207, doi.org/10.1145/2775292.2775308.

Potenziani, M., Callieri, M., Dellepiane, M., Corsini, M., Ponchio, F., Scopigno, R., 2015. 3DHOP: 3D heritage online presenter, *Computers & Graphics,* 52, 129–141, doi.org/10.1016/j.cag.2015.07.001.

Remondino, F., 2011. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sens*., 3, 1104–1138. doi.org/10.3390/rs3061104.

Santiko, H., Nugrahani, D.S., 2012: *Adegan dan ajaran hukum karma pada relief Karmawibhangga*. Balai konservasi Borobudur, Borobudur, Indonesia.

Tallon, A., 2014. Divining Proportions in the Information Age. *Archit. Hist.*, 2, 15. doi.org/10.5334/ah.bo.

Themistocleous, K., 2017. Model reconstruction for 3d vizualization of cultural heritage sites using open data from social media: The case study of Soli, Cyprus. *J. Archaeol. Sci.*, 14**,** 774–781. doi.org/10.1016/j.jasrep.2016.08.045.

Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015. Semantic Point Cloud Interpretation based on Optimal Neighborhoods, Relevant Features and Efficient Classifiers. *ISPRS J. Photogramm. Remote Sens,* 105*,* 286–304. doi.org/10.1016/j.isprsjprs.2015.01.016.

Yang, M., Yu, K., Zhang, C., Li, Z., Yang, K., 2018. DenseASPP for Semantic Segmentation in Street Scenes. In *Proceedings of the 31st IEEE Conference on Computer Vision and Pattern Recognition (CVPR2018)*, 3684–3692. doi.org/10.1109/CVPR.2018.00388.

Yamaji, M., Akama, R., 2021. On the ARC Portal Database System: How to Handle its Multilingualization and Specialized Terminology. *ART RESEARCH, Technical support Newsletter*, 21, 105–114. doi.org/10.34382/00014482.

Zheng, B., Oishi, T., Ikeuchi, K., 2015. Rail Sensor: A Mobile Lidar System for 3D Archiving the Bas-reliefs in Angkor Wat, *IPSJ Transactions on Computer Vision and Applications*, 7, 59–63. doi.org/10.2197/ipsjtcva.7.59.