

Proactive Preservation of World Heritage by Crowdsourcing and 3D Reconstruction Technology

Hidehiko Shishido
University of Tsukuba
Tsukuba, Japan
shishido@ccs.tsukuba.ac.jp

Yutaka Ito
Akita University
Akita, Japan
utaka.ito@gipc.akita-u.ac.jp

Youhei Kawamura
Akita University
Akita, Japan
y.kawamura@gipc.akita-u.ac.jp

Toshiya Matsui
University of Tsukuba
Tsukuba, Japan
t-matsui@heritage.tsukuba.ac.jp

Atsuyuki Morishima
University of Tsukuba
Tsukuba, Japan
mori@slis.tsukuba.ac.jp

Itaru Kitahara
University of Tsukuba
Tsukuba, Japan
kitahara@ccs.tsukuba.ac.jp

Abstract—Since over one million tourists annually visit the Angkor ruins, the effect on the buildings from the vibrations caused by these tourists is a huge problem for maintaining them. Such organisms as bryophytes, which adhere to the surface of the stones of the ruins, is another factor that damages them. Using crowdsourcing and 3D reconstruction technology, we are organizing a proactive preservation project for the Angkor Thom Bayon Temple, which is a world cultural heritage site. We evaluated its damaged parts and visualized the damaged state.

Keywords: *Crowdsourcing; 3D reconstruction; World heritage; Angkor ruins*

I. INTRODUCTION

Many world heritage sites registered by UNESCO are annually visited by millions of tourists. Due to the high historical value of such sites, activities that preserve them are critical. We are engaged in research to preserve Cambodia's Angkor ruins [1], which were registered in 1992 by UNESCO as an endangered site, and our current repair efforts are focusing on ruins. Restoration reproduces the buildings/structures and maintains their appearances to prevent collapses and beautify the outer walls. Since over one million tourists annually visit the Angkor ruins, the influence on them from the walking vibrations of tourists is severe. One major task is to maintain the appearance of the ruins. In addition, such microorganisms as bryophytes, which adhere to the stone surfaces of the ruins, are a major cause of the deterioration of the shape of the relief on the ruin surfaces. For our investigations (Fig. 1), we applied different chemicals to the stones made of the same material as the ruins, observed the breeding process of bryophytes, and studied a conservation method. In this follow-up observation, based on visual confirmation by researchers, we introduce digital technology in studies that measure with instruments equipped with a distance sensor on a balloon for the 3D measurement of world cultural sites [2].

We compared sensor images to minimize the distortion of the distance data caused by such external forces as wind to estimate sensor motions and correct the distortion. Our technique successfully measured regions that could not be



Figure 1. Different chemicals were applied to stones made of same material as ruins to observe progress. (Left: Various stone materials used for ruins, Right: Stone pillar of ruins of observation target)

measured by conventional sensors. Other research movements applied Structure from Motion (SfM) to the multi-viewpoint images of unglazed earthenware, reconstructed 3D information, and generated archeology materials [3,4].

As part of the conservation activities of cultural heritage sites using such digital technology, we are organizing a preventive preservation project of the Angkor ruins using crowdsourcing and 3D reconstruction technology. For the Angkor Thom Bayon Temple, a Cambodian world cultural heritage site, we are evaluating its damaged areas and visualizing the damaged states by digital technology to generate 3D models of it and its damaged areas. For follow-up observations, 3D models with different observation stages are required. In the above method, generating a 3D model using a distance sensor is very time-consuming, and so measurement experiments with high frequency are not realistic. Therefore, in this research, we solve the problem by crowdsourcing, which is a sourcing model that gathers contributions from Internet users. To reconstruct the object's 3D shape using SfM, many images are necessary, which are obtained by capturing objects from various angles, which is also very time-consuming. Therefore, in this research, we

solve the problem by building a crowdsourcing platform that captures the damaged area as a microtask. Tourists and local residents become crowdsourcing workers (people who request to perform tasks). By capturing a large number of tourists, we can take pictures of the entire large-scale space and comprehensively observe the damaged areas scattered in the ruins. With the observation data, we discriminated the damaged areas and visualized them.

II. PROACTIVE PRESERVATION ACTIVITY PLAN OF ANGKOR RUINS UTILIZING CROWDSOURCING AND 3D RECONSTRUCTION TECHNOLOGY



Figure 2. Flow of proactive preservation activities of Angkor ruins

We sequentially explain the flow of the proactive preservation activity of the Angkor ruins using crowdsourcing and 3D reconstruction technology (Fig. 2).

1. We capture target objects from various angles. We manually acquired in advance the initial captured images of our research target, the Angkor Thom Bayon Temple.
2. 3D models were generated using the acquired images of the ruins by applying SfM to the captured image groups. The position and orientation of the camera and a sparse point group were generated by searching for corresponding points between images. Next we minimized the re-projection error to generate a dense point group. After the application, we deleted unnecessary point clouds and attached a surface model to complete a 3D model with surface information.
3. A damaged area is distinguished based on archaeologists' opinions and vibration sensors. Crowdsourcing was not used to identify the damaged areas of the ruins. We improved the efficiency using expert opinions or vibration sensor information to identify the contents to be processed.
4. Tasks were outsourced to an unspecified number of crowds (workers) over the Internet. We requested them to capture the target based on the information of the damaged areas determined from the viewpoint of the vibration sensors and experts. When the captured images are diverse, the 3D model's accuracy increased.

5. The workers captured the damaged areas using cameras, including cameras mounted on portable terminals, from various angles for comprehensive observations.
6. We reconstructed a 3D model of the subject using images of the damaged areas. We examined the existence of the target image and excluded error images. Experts received visualization data of the damaged state.

We evaluated the damaged areas of the cultural sites and visualized the damaged states by repeating the above flow to promote proactive preservation activities.

III. 3D MODEL GENERATION OF STONE PILLAR OF ANGKOR THOM BAYON TEMPLE

In this section, we introduce our approach and focus on the bryophytes that are attached to the stone surface of the ruins. As shown on the right of Fig. 1, we applied various chemicals to stones made of the same material as the ruins and generated a 3D model to observe the breeding process of the bryophytes. We installed a virtual camera in the space of the generated 3D model and generated a rendering image of the 3D model in the position and the orientation captured by the workers. We conducted a demonstration experiment of our processing and confirmed that the images captured by the workers are accurately reflected in the 3D model. The experiment's flow is explained below.

Based on the above processing flow, we captured images from various angles of the stone pillar of our target, Angkor Thom Bayon Temple. Furthermore, assuming the worker capture them with Step 5 in Section 2, we captured the stone pillar in advance using a mobile camera that is different from what was captured by the camera. We applied SfM to each captured image to calculate the position and orientation of the multi-view camera and generated a sparse 3D point group of the subject. Next we applied a patch-based multi-view stereo (PMVS) approach from a sparse 3D point group and estimated a dense 3D point group. A mesh was generated using a poisson surface and pasting the texture. We acquired the position and orientation information of the multi-view camera in the generated 3D model and set the virtual camera at the same position as the acquired position. 3D models were rendered from virtual cameras installed at mobile camera positions that were different from previously captured cameras.

As shown in Fig. 3, to generate sparse 3D points, we used the images captured from various angles and confirmed by SfM that the captured camera positions and postures were estimated. As shown in Fig. 4, we installed a virtual camera at the same position as the estimated camera position in the 3D model space to which the texture is attached. An orange rectangle indicates a virtual camera at a position captured by different cameras. Fig. 5(a) shows an image rendered from a virtual camera at a position captured with different cameras. Fig. 5(b) shows a live-action image captured by different cameras. By comparing Figs. 5(a) and (b), we confirmed that the appearance of the image of the rendered 3D model is the same as the live-action image.

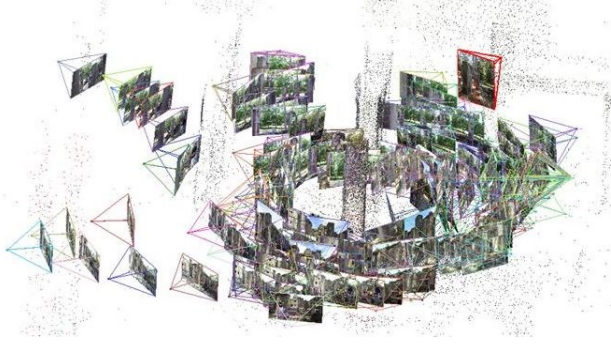


Figure 3. Position and posture of multi-view camera and a sparse 3D point cloud as a base of a 3D model



Figure 4. Virtual camera is installed at identical position as estimated camera position in generated 3D model space

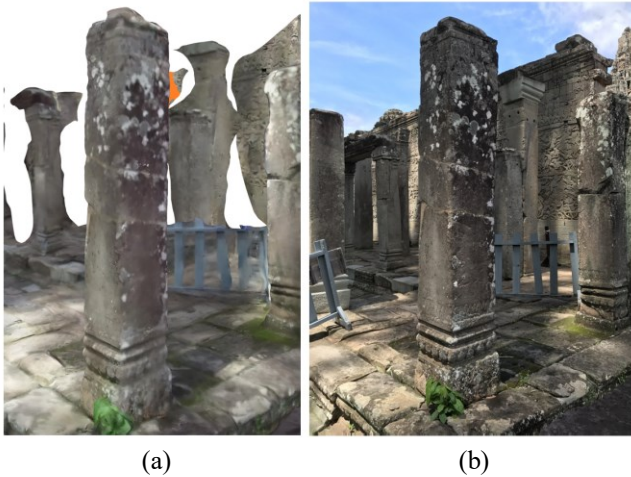


Figure 5. Various chemicals were applied to stone made of the same material as ruins and their stone pillar to observe the progress. (a) Image rendered from a virtual camera at a position captured with different cameras. (b) Live-action image captured by different cameras.

From the above demonstration experiments, we confirmed that a highly accurate 3D model can be generated even when a mobile camera different from the camera shot in advance was used (Verification of Section 2, Flow 1 to 4). Moreover, we confirmed that the images captured by the workers are

accurately reflected on the 3D model (Verification of Section 2, Flows 5 to 6). Therefore, we can compare the follow-up observations of bryophytes using the images of the rendered 3D model.

We will implement the following items to achieve this project: determination of the details of microtask using crowdsourcing, problem solving for worker's task remuneration, development of an interface that requests the captured damaged area, classification of the captured images of worker, a judgment algorithm for a damaged area, and the development of a visualization interface of damaged states.

IV. CONCLUSION

We described our proactive preservation project of the Angkor ruins using crowdsourcing and 3D reconstruction technology. We attempted such 3D reconstruction of a stone pillar of ruins by observing the progress and applying various chemicals to stones made of the same material as the ruins, confirmed that the stone pillar was captured from various angles using multiple kinds of cameras, and generated a 3D model with high accuracy. We confirmed that the images captured by the workers were accurately reflected on the 3D model and clarified that we can compare the breeding process of bryophytes using rendered 3D model images.

ACKNOWLEDGMENT

This work was supported by JST CREST Grant Number JPMJCR16E3, Japan.

REFERENCES

- [1] Toshiya Matsui, Masaaki Sawada, Saihachi Inoue, Takao Ebisawa, Emi Kawasaki, Yosuke Atomi, and Takeshi Nakagawa, "A Study on Conservative Material for the Bas-Reliefs of the Bayon Temple in Angkor Monuments," *Proceedings of the International Conference on Conservation of Stone and Earthen Architectural Heritage*, pp. 138-144, 2014.
- [2] A. Banno, T. Masuda, T. Oishi, and K. Ikeuchi, "Flying Laser Range Sensor for Large-Scale Site-Modeling and Its Applications in Bayon Digital Archival Project," *International Journal of Computer Vision*, 2008, Volume 78, Issue 2-3, pp. 207-222.
- [3] Yukinori Kawae, Yoshihiro Yasumuro, Ichiroh Kanaya, and Fumito Chiba, "3D Reconstruction of the "Cave" of the Great Pyramid from Video Footage," *Digital Heritage International Congress 2013*, pp. 227-230.
- [4] C. Wu, "Towards Linear-Time Incremental Structure from Motion," *2013 International Conference on 3DVision-3DV 2013*, pp. 127-134, 2013.