デジタルバイヨン
− 大型有形文化の e−monument 化の実際 −

Digital Bayon Temple
- e-monumentalization of large-scale cultural-heritage objects -

池内克史/東京大学大学院情報学環,　大石岳史/東京大学大学院情報学環,　高松淳/東京大学生産技術研究所
Ikeuchi Katsushi/Univ. of Tokyo, Oishi Takeshi/Univ. of Tokyo, Takamatsu Jun/Univ. of Tokyo
{ki,oishi,j-taka}@cvl.iis.u-tokyo.ac.jp

近年の電気コンピュータ技術の進展により、文化遺産をデジタル化し、保存、展示、解析等に使用することが可能となってきた。本稿では、このデジタル化された文化遺産のことを e−monument と呼ぶ。我々のグループは、カンボジアアンコール遺跡内バイヨン寺院を題材として、e−monument 化の技術を開発してきた。本稿では、まず、なぜこのバイヨン寺院を選んだかの理由について述べる。その後、バイヨン寺院の複雑さと大きさのために e−monument する際にどのような問題が発生したか、これらをどの様に解決したかを述べる。最後に、e−monument 化されたバイヨン寺院の映像を紹介する。

Recent advancement in Electrical Engineering and Computer Science technologies enables us to digitize cultural heritage objects for their preservation, exhibition, and utilization. In this paper, we refer to these digitized cultural heritage objects as “e-monuments.” As the target object for conversion to an e-monument, we chose the Bayon temple in Angkor, Cambodia. First, we will explain why we chose the temple as the target object, and then we will describe the challenges we encountered in the conversion process and how we overcame these challenges. Finally, we will show the obtained e-monument of the Bayon Temple.

1. Introduction

Recent advancement in Electrical Engineering and Computer Science (EECS) technologies enables us to have PCs with a large memory area, cameras with high resolution, and range sensors with high accuracy. We can digitize irreplaceable heritage objects by using these advanced EECS technologies. In this paper, we refer to these digitized heritage objects as “e-monuments.” Our group has been working on developing software to e-monumentalize large heritage objects, such as the great Buddha of Nara [1] and the Angkor ruins in Cambodia [2]. This project was sponsored by JST and MEXT of the Japanese government.

E-monuments can be used for preservation of the objects, utilization of digitization results for multimedia displays, and analysis of the data for archeological research. Irreplaceable heritage objects can be lost forever due to natural disasters such fires and earthquake, human disasters such as war, and vandalism such as that which occurred under the Taliban or during the Crusades. Without continuous effort to maintain e-monuments, we may lose these human treasures forever. It is worthwhile to obtain e-monuments because these treasures are irreplaceable, they attract public interest, and they are highly valuable when displayed. We can construct a virtual museum, using a collection of e-monuments, to be available from anywhere at any time. Further, such e-monuments provide important scientific data. Thus, it is urgent that we begin the process of converting cultural heritage objects into e-monuments as soon as possible.

Figure 1 the Bayon temple in Angkor ruin

This paper first explains the reason why we chose the Bayon Temple as the example of e-monumentation. Then we describe the challenges are encountered due to the choice and how we overcame these challenges. Finally, we will show some of the rendered images of the e-monumentalized Bayon Temple.

2. Why the Bayon Temple?

The Bayon Temple, shown in Figure 1, was constructed by the order of Khmer ruler Jayavarman VII in the twelfth century. The temple has a central tower with fifty-one surrounding towers. This temple, made of stone, spans 150 meters by 160 meters by 40 meters. The first reason for choosing the Bayon Temple was its size. Usually, the larger the site, the more difficult it is to e-monumentalize due to the memory and computational time required.

Second, the structure is complex and beautiful. The temple has a central tower representing Mt. Sumer, the meeting place of the gods above the world, and, together with the 51 surrounding towers with calm, smiling faces, exhibits the Khmer's idea of the Universe. The temple has two layers of terraces and double corridors with many hidden passages along with bas-relief depicting the victory of Jayavarman. Due to this complex structure, there are no reliable floor plans existing. Though the size is smaller than the Angkor Wat Temple, the Bayon Temple is considered to be one of the masterpieces among heritage objects in the Angkor ruin. This complexity and beauty was the second reason for us to be attracted to digitizing the Bayon temple.

Third, the Japanese Government team for safeguarding Angkor (JSA) has been conducting a preservation project. It is observed, however, that the central tower has been inclining year by year and there is a possibility that it may fall down in the near future. Before we lose this irreplaceable object, it is important to have its digital copy. For these three reasons, we decided to e-monumentalize the Bayon Temple.

3. Techniques for E-monumentalization

. E-monumentalization, the conversion of real heritage objects into e-monuments, consists of three steps: 1) **sensing step** to obtain partial geometric shapes using range sensors, 2) **alignment step** to determine relative relations among these obtained partial geometric shapes, and 3) **merging step** to connect these partial geometric shapes into a mesh model in a unified representation. For small indoor objects, techniques are well developed and there is even commercially available software to do this. However, large and/or complex objects, such as the Bayon Temple, provide various technical challenges in all aspects of these modeling steps. This section will explain these challenges and how we solved these issues.

3.1 Sensing for obtaining partial geometric data

Obtaining e-monuments begins with the acquisition step of obtaining partial data from various viewing directions. We usually obtain both color and range images from a single viewing direction using a color camera and a range sensor, respectively. One color image consists of a 2-dimensional array of red, green, and blue values. Similarly, one range image consists of a 2-dimensional array of range values, which are distance values from the image plane to corresponding physical points along their lines of sight.

A range image can be obtained by using a range sensor. Various types of range sensors are commercially available. A Cyrax sensor made by Leica Geosystems, which measures the distance by the flight time of laser light, can obtain a range image up to 100m with 5 mm accuracy. A Vivid sensor made by Konica-Minolta, which uses the triangular method to measure the distance, can measure an object within a 3m range with 0.1 mm accuracy. An Imager, made by Z+F, can obtain an omnidirectional range image.

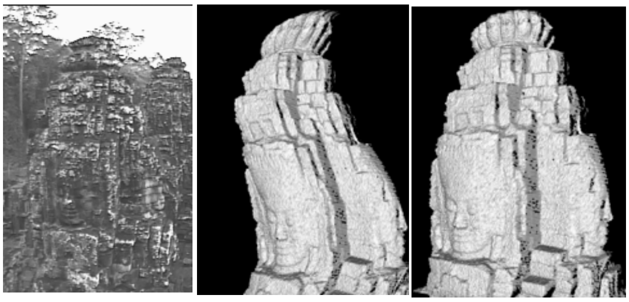
One of the technical challenges in applying these commercially available sensors to a large structure, such as the Bayon Temple, is that they are ground-based sensors. Some parts of the Bayon Temple are occluded and invisible from the ground. When a target object is small, we can build scaffolds around the object and bring a sensor to the top of the scaffolds. The size of the Bayon Temple, 150 meters by 160 meters by 40 meters, prohibits the use of this scaffold method because it would be a time-consuming effort to construct, impractically expensive, and ugly looking to visitors.

We designed a balloon sensor in order to remedy this issue. A high-speed range sensor hangs under a balloon. The balloon floats in the air and stays at any position; the balloon sensor provides freedom to choose any viewing position by simply maneuvering the balloon to a desired position. This balloon method, however, poses a new challenge: the sensor swings during range acquisition due to the wind, and thus the resulting range data are distorted.

Our researchers Banno and Ikeuchi proposed to rectify this distorted range data by combining motion estimation from an image sensor mounted on the range sensor so that sensor motion, predicted from an image sensor, is consistent with the distortion of the range data [3]. They introduced three constraints: range data, bundle, and smoothness constraints, and derived a minimization formula to determine the sensor motion and to rectify the range data. Figure 2 shows examples of the corrections; the leftmost images depict a scene image, the central images present a distorted range image, and the rightmost images present the rectified images using the method.

Another method, proposed by our researchers Masuda et al., assumes that the balloon-based sensor is an augmentation of the ground-based sensor, and there should be overlapping regions between data from ground-based and balloon-based sources [4]. They developed a rectification algorithm to minimize the difference between these two regions. This method is handy because it does not need a TV camera mounted on the range sensor, while the first method is more accurate and robust than this method.

The Bayon Temple has double corridors around the building, narrow passages connecting fifty-one towers, and high surrounding walls. A range sensor prefers orthogonal views on the target surface for accurate measurement. In narrow areas, one view only provides limited orthogonal areas due to the distance between the sensor and



(a) (b) (c)

Figure 2 Rectification of range image from the balloon sensor: (a) Scene, (b) distorted range image, (c) rectified range image using the Bannno-Ikeuchi

the wall; other areas are visible, but heavily inclined. A regular range sensor only provides unreliable data at these inclined regions. In order to overcome this issue, our researchers Ono et al. developed a moving 1D range sensor that we called "the climbing sensor" to obtain orthogonal data along a moving direction as shown in Figure 3 [5].

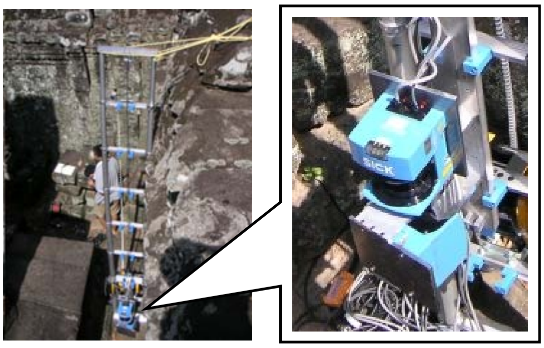


Figure 3 Climbing sensor

This type of sensor requires calculating the speed of the sensor to have evenly distributed sampling intervals among orthogonal data. The researchers developed an algorithm to determine the speed of the sensor using another 1D range sensor mounted on the orthogonal direction to the first 1D range sensor. The range images obtained by this senor are referred to as epipolar range images, following from epipolar-image analysis proposed by Bolles in the late 1980s [6].

Using newly developed sensors, such as the balloon sensor and the climbing sensor as well as commercially available sensors, such as Cyrax, Imager, and Vivid, we have obtained more than 20000 range images with 200GB to cover the entire surface of the Bayon Temple.

3.2 Alignment to determine relative relations among partial geometric data

Each range image is represented with respect to the sensor coordinate system at the time of the data collection. The alignment algorithm determines the relative relation among these coordinate systems. Traditional alignment algorithms, such as Iterative Closest Point algorithm (ICP), can handle a smaller number of range images [7-9]. These alignment algorithms determine rotation and translation parameters between two sensor coordinate systems by minimizing the distance of corresponding points between a pair of range images. The alignment process iteratively continues these pairwise alignments, one by one, along all the range images. If the object were small enough to be covered by a dozen range images, this pairwise alignment should be able to obtain all the relative relations correctly. However, when applying this iterative alignment to more than 20000 range images, a huge error is accumulated along the alignment chain and a large discrepancy occurs between the first and the last sensor coordinate systems. In order to remedy the problem of sequential pairwise alignment, simultaneous alignment algorithms, which load in all the data and determine all the relations simultaneously, have been proposed. However, 200GB data cannot be loaded into a single computer due to the overflow of the memory space.

Our researchers Oishi et al. proposed a two-step algorithm to remedy this issue. The first step is a rapid alignment algorithm for onsite alignment. When work is being done onsite, a rapid alignment algorithm is desirable for data debugging and sensor planning, though this involves the sacrifice of some robustness. One of the most time-consuming processes in alignment is to establish corresponding pairs of data points. The researchers derived an algorithm to establish these pairs by utilizing a graphic-processing unit, commonly available on recent PCs [10]. The resulting algorithm is 100 times faster than those based on the ICP technique [11].

The second step is a parallel alignment algorithm that runs on a PC cluster. For this simultaneous algorithm, we have to consider all the relations. In our case, the first algorithm already provides a fairly accurate initial configuration of all the data. Using these known relations, the algorithm divides the entire data set into sub-groups depending on the number of available PCs in the PC cluster, and distributes those sub-data groups over the cluster so that computational load and memory requirements are evenly distributed among the PCs [12].

3.3 Merging partial geometric data into a unified model

Each range data set from a single scanning covers only one small part over a whole object's surface. The previous alignment process determines relative relations among those partial data sets. Usually, 3D range data are represented as a mesh structure, connecting 3D data points with arcs, and forming triangular patches. It is necessary to connect those partial mesh representations into a uniform representation of a whole object's surface. The procedure of interconnecting these mesh structures is referred to as merging.

It is important to make the merging process robust against any noise that may be in the scanned range images and can also be inherited from the alignment procedure. Our researchers Sagawa and Ikeuchi represented a set of range images in a volumetric implicit-surface representation of signed distance distributions, which is then converted to a surface mesh by using a variant of the Marching Cubes algorithm [13]. Unlike previous techniques based on implicit-surface representations [14], this method estimates the signed distance to the object surface by determining a consensus of locally coherent observations of the surface. First, the researchers considered a method to increase computational and memory efficiency using parallel computing techniques. Later, they extended the method for considering not only range data but also color information.

Merging is also considered as a process to reduce noise for a smoother surface. In particular, noise in range data is inevitable on measuring outdoor objects from a distant position under severe sunlight. Sagawa et al. considered methods to generate a smooth surface by considering two issues: data error and data lacking. Previously described methods assume error distributions are evenly distributed in space. Some range sensors, such as Cyrax, have less accuracy in the depth dimension than in the spanning dimensions, because the depth measurement is obtained by a laser returning a signal while the spanning dimensions are measured with encoders for mirror motions. Sagawa et al. derived a method to reduce noise in range data by considering this anisotropic error distribution.

Another issue in data error is how to fill small holes due to small occlusions. We have to interpolate over such holes for water-tight surfaces. The method fills such data holes by considering nearby signed distance fields and flipping signs of signed distance fields under certain conditions. This algorithm is also parallelized for a large data set [16].

4. Digital Bayon Temple obtained

This section contains digital pictures obtained from rendering 3D data.

4.1 Entire structure of the Bayon temple

Figure 4 shows the obtained mesh model of the Bayon temple. Due to technical limitations, the current mesh model is 2cm resolution. We are currently working to develop software to obtain 1cm resolution. We have also determined various cross-sections and obtained floor plans of the temple, which provide the basis data for a future restoration project.

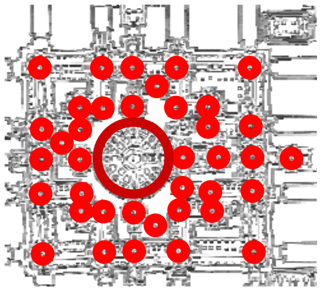
It turned out that the east-west axis of the Bayon Temple rotates 0.94 degrees in a counter-clockwise direction from the real east-west line. Note that we scanned not only the outside surface of the temple, but also the inner surfaces of the towers, and thus, we obtained accurate cross-sections and could measure the width of the walls of the towers.

Figure 4 Entire structure of the Bayon Temple

4.2 Deity faces on towers

The Bayon Temple has 51 towers beside the central tower, on which four noble, smiling faces of deities are engraved. In total, 173 faces of deities are currently visible from outside of the temple. We have acquired range images of all the faces, using Cyrax, Vivid and balloon sensors. Figure 5 shows the rendered images of some of them.

According to JSA art researchers, those faces are classified into three classes: Dava, Devata, and Ashora. This classification is, however, based on their subjective judgments. Some ambiguous classifications do exist. Thus, currently, we are working to classify these faces into a small number of classes using more accurate computer analysis.



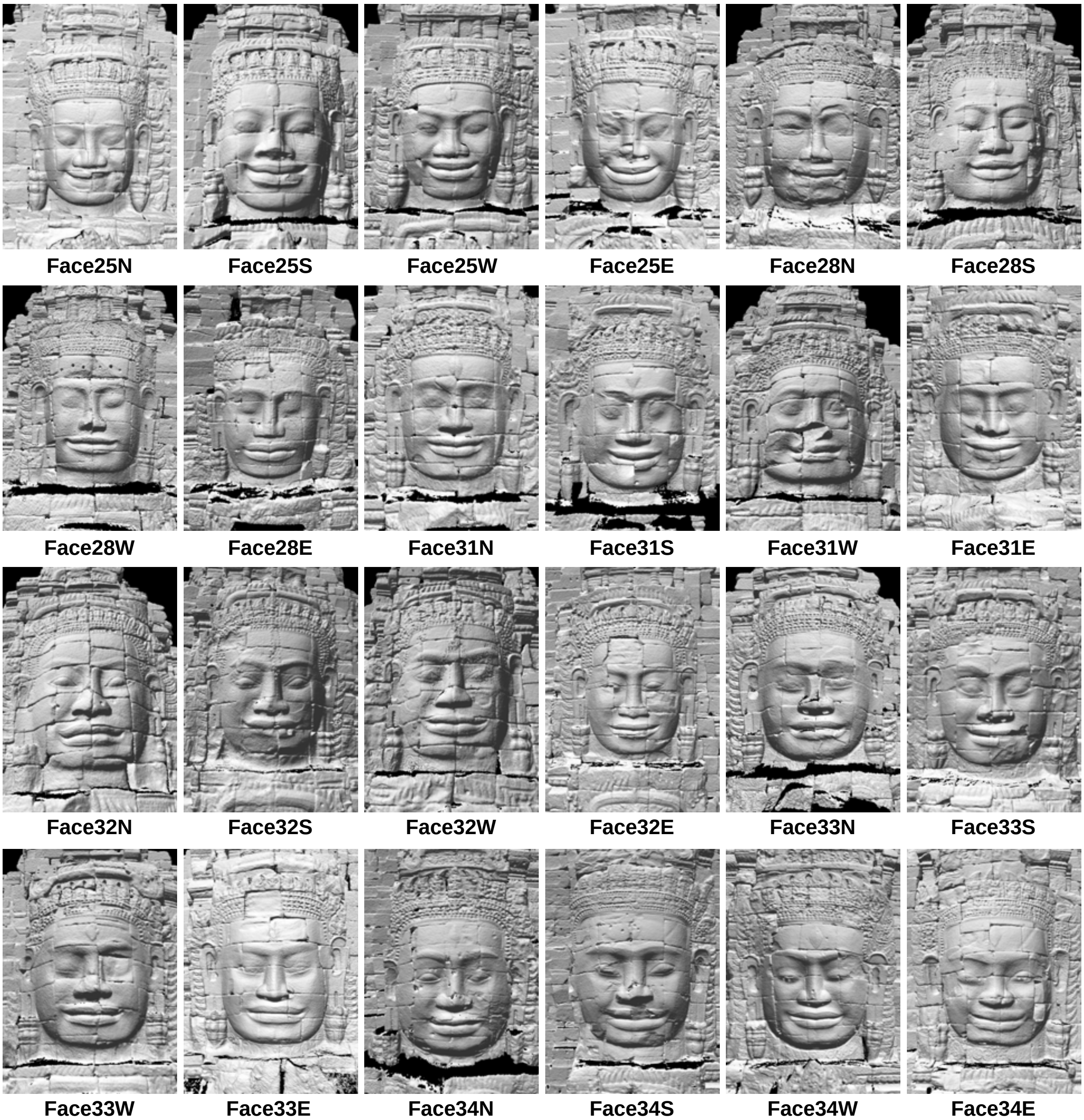


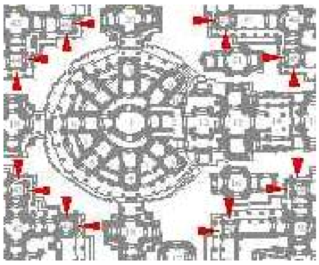
Figure 5 Deity face on the tower

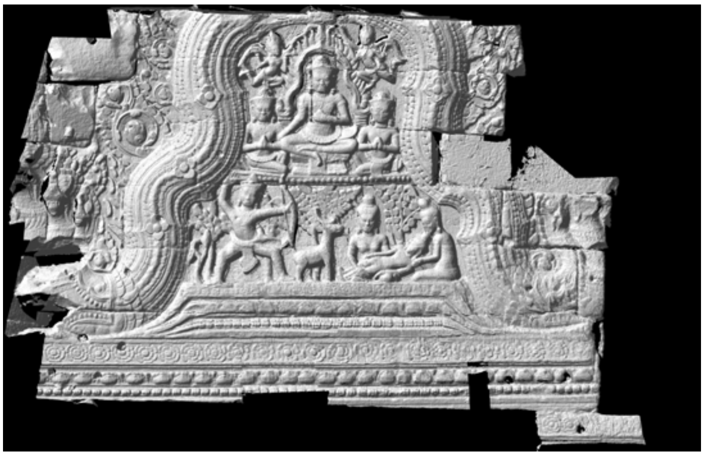
4.3 Pediments hidden under the terrace

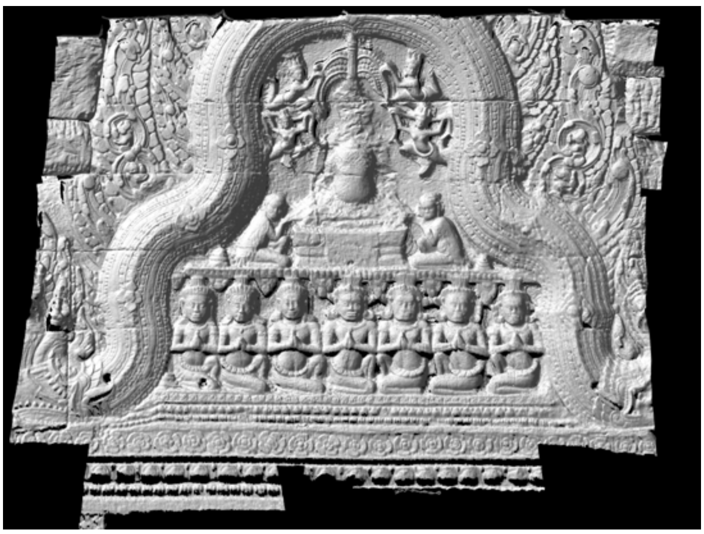
The Bayon Temple had been modified several times due to the religious change from the Hinduism to Buddhism. Due to this re-modification, most of the pediments, though they are beautiful and important, are hidden. Even visitors to the Bayon Temple today cannot observe those pediments. In fact, most of them had never been observed and pictures of them did not exist.

We digitized those hidden pediments using a Vivid sensor and a surface mirror. The clearance between the pediments and the adjacent wall is so small that we could not mount a Vivid sensor in front of the pediments. By putting a surface mirror in front of the pediments for re-orienting laser light from the Vivid sensor on the ground, and observing the pediment illuminated by the re-oriented light in the mirror, we scanned 16 hidden pediments.

Figure 6 shows the rendered images obtained from the range data. In one sense, this is the “world premiere” of the appearance of the pediments.







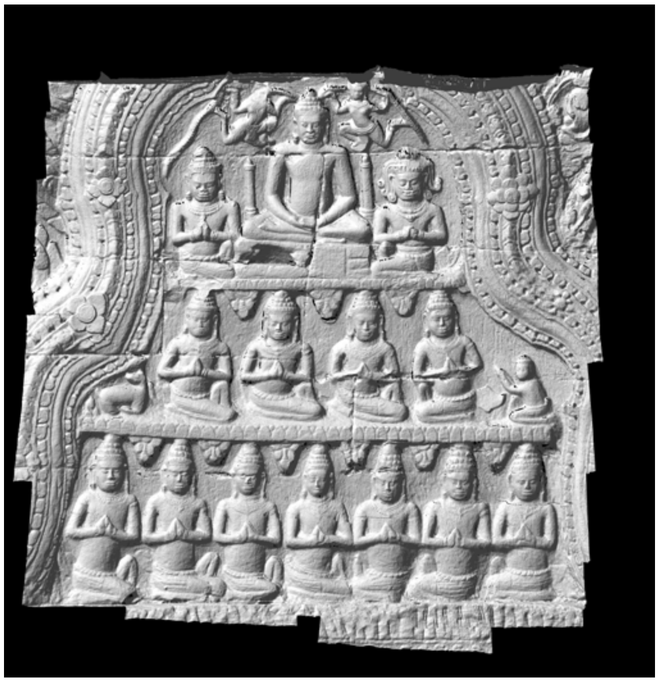
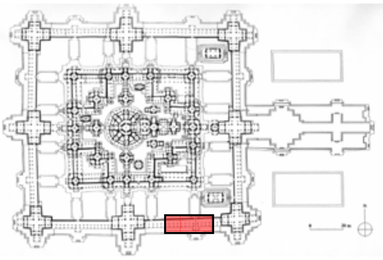


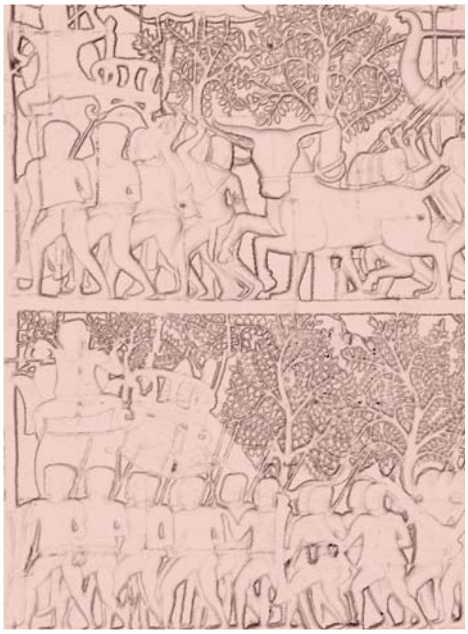
Figure 6 Pediment hidden under the terrace

4.4 Bas-relief on walls along the double corridors

The Bayon Temple has double corridors. Along the wall of these corridors, bas-reliefs are inscribed to depict the victory story of Jayavarman VII. We scanned the entire wall using Vivid sensors mounted on cranes. Since the Vivid sensor is prone to errors that distort a planar surface to a curved surface due to slight inaccuracy of factory-setting laser-camera calibration, we rectified these errors using a planar calibration object and aligned all the range images over a range image obtained by a Z+F range sensor.

Figure 7 shows some of the bas-relief. In particular, the lower right corner depicts the scene of a famous fighting cock. While the picture shows moth damage and traces of running water, and is thus difficult to see, the rendered image from range data provides a much more clear appearance. This effect becomes even clearer when we generate a movie of the bas-relief under varying illuminations. We plan to analyze the scene in detail with the JSA artists for further interpretation.





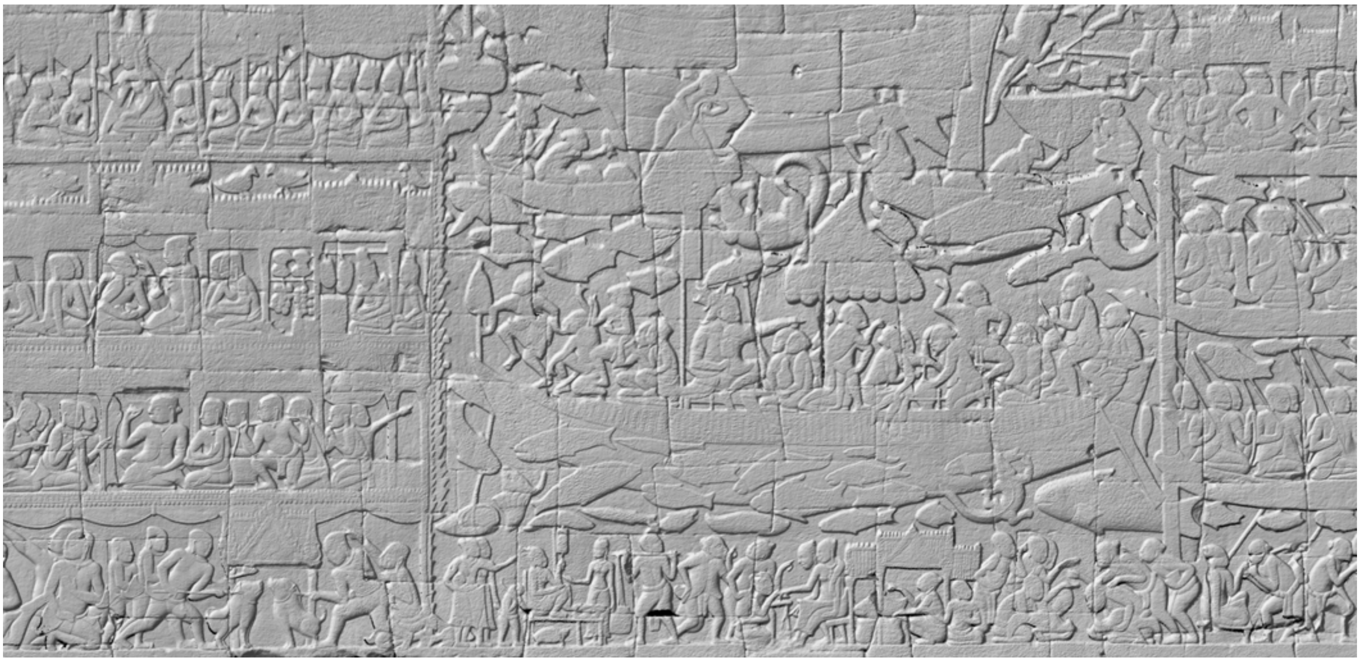


Figure 7 Bas-relief on walls

5. Conclusion

This paper describes the geometric pipeline designed for e-monumentalizing large-scale cultural heritage objects. We chose the Bayon Temple in the Angkor ruin as the target object to develop such e-monumentalizing techniques. First, we explained the reason why we chose the Bayon Temple as the target object. Then we described the challenges that were encountered in e-monumentalizing the Bayon Temple, and how we overcame these challenges. Finally, we showed the obtained e-monument of the Bayon Temple. In a future work, we will consider how to utilize these 3D geometric models for web-retrieval as well as display.

Acknowledgement

This research was conducted not only by the authors listed in the heading section but also by ex-staff members of our laboratory, Prof. Ryo Kurazume of Kyushu Univ., Prof. Atsushi Nakazawa of Osaka Univ., Prof. Koichi Ogawara of Kyusyu Univ., and Dr. Kazuhide Hasegawa of Fuji Film. Our current staff members, Dr. Masataka Kagesawa, Dr. Daisuke Miyazaki, Dr. Shunsuke Kudoh, Dr. Shintaro Ono, and Dr. Atsuhiko Banno, and ex-graduate students, Prof. Ko Nishino of Dexel Univ., Prof. Hiroshi Kawasaki of Saitama Univ., Dr. Yasuyuki Matsushita of MSRA, Dr. Imari Sato of NII, Dr. Ryusuke Sagawa of Osaka Univ., and Dr. Robby Tan of ANU. Dr. Shinichiro Nakaoka of AIST, Dr. Hiroki Unten of Toppan, and Dr. Tomohito Masuda of Toppan also contributed to this project. Without their efforts, we could not have achieved this result.

Research results presented in this paper are supported by three major research projects. We are deeply indebted for their guidance and management to the three leading program managers of the projects: Prof. Masao Sakauchi of NII for the Sin Program, Prof. Makoto Nagao of National Library of Japanese Diets for the CREST program, and Prof. Hiroshi Harashima of the Univ. of Tokyo for the LP program. This research was financially supported, in part by the Ministry of Education, Culture, Sports, Science, and Technology of the Japanese government, in part by the Japan Science and Technology Agency, and in part by Microsoft Research IJARC-MIC.

Reference

[1] K. Ikeuchi and Y. Sato, *Modeling from Reality*, Springer, Boston, 2001.

[2] K. Ikeuchi and D. Miyazaki, *Digitally Archiving Cultural Objects,* Springer, Boston, 2007.

[3] A. Banno and K. Ikeuchi, “Shape recovery of 3D data obtained from a moving range sensor by using image sequences,” *ICCV*, 2005.10.

[4] T. Masuda, Y. Hirota, K. Nishino and K. Ikeuchi, "Simultaneous determination of registration and deformation parameters among 3D range images," *3DIM,* pp.369-376, .2005.1.

[5] K. Matsui, S. Ono, and K. Ikeuchi, “The Climbing Sensor: 3-D Modeling of a Narrow and Vertically Stalky Space by Using Spatio-Temporal Range Image,” *IROS*, pp. 3997-4002, Aug.2005.8.

[6] R. C. Bolles, H. H. Baker and D. H. Marimont, “Epipolar-plane image analysis,” *IJCV*, 1:7-55, 1987.

[7] P.J. Besl and N.D. McKay, "A method for registration of 3-d shapes," *IEEE PAMI.,* 14(2):239-256, 1992.2.

[8] Y. Chen and G. Medioni, “Object modeling by registeration of multiple range images, *Image and Vision Computing*, 10(3):145-155, 1992.4

[9] S. Rusinkiewicz and M. Levoy, “Efficient variants of the IPC algorithm,” *3DIM, pp.145-152,* 2001.5.

[10] T. Oishi, A. Nakazawa, R. Kurazume, and K. Ikeuchi, “Fast simultaneous alignment of multiple range data,” *3DIM*, pp. 476-483, 2005.6

[11] M. D. Wheeler and K. Ikeuchi, “Sensor modeling probabilistic hypothesis generation, and robust localization for object recognition,” *IEEE PAMI*, 17(3): 252-265, 1995.3.

[12] T. Oishi, R. Sagawa, A. Nakazawa, R. Kurazume, and K. Ikeuchi, “Parallel alignment of a large number of range images,” *3DIM,* pp. 195-202, 2003

[13] R. Sagawa, K. Nishino, K. Ikeuchi, “Adaptively Merging Large-Scale Range Data with Reflectance Properties,” *IEEE PAMI*, 27(3): 392-405, 2005.3

[14] B. Curless and M. Levoy, “A volumetric method for building complex models from range images,” *SIGGRAPH* 1996.

[15] R. Sagawa, T. Oishi, A. Nakazawa, and K. Ikeuchi, “Iterative refinement of range images with anisotropic error distribution,” *IROS*, pp. 79-85, 2002.

[16] R. Sagawa and K. Ikeuchi, “Taking consensus of signed distance field for complementing unobservable surface,” *3DIM*, pp.410-417, 2003.