

Representing Cultural Heritage in Digital Forms for VR Systems through Computer Vision Techniques

[#]Katsushi Ikeuchi, ^{*}Atsushi Nakazawa, [#]Kazuhide Hasegawa, [#]Takeshi Ohishi
[#] Institute of Industrial Science, The University of Tokyo
^{*}Cybermedia Center, Osaka University
{ki, k-hase, oishi}@cvt.iis.u-tokyo.ac.jp
nakazawa@ime.cmc.osaka-u.ac.jp

Abstract

This paper overviews our research on digital preservation of cultural assets and digital restoration of their original appearance. Geometric models are digitally achieved through a pipeline consisting of scanning, registering and merging multiple range images. We have developed a robust simultaneous registration method and an efficient and robust voxel-based integration method. On the geometric models created, we have to align texture images acquired from a color camera. We have developed a texture mapping method to utilize laser reflectance. In an attempt to restore the original appearance of historical heritage objects, we have synthesized several buildings and statues using scanned data and literature survey with advice from experts.

1. Introduction

Currently, a large number of cultural heritage objects around the world are deteriorating or being destroyed because of natural weathering, disasters and civil wars. Among them, Japanese cultural heritage objects are quite vulnerable to fires and other natural disasters because most of them were constructed of wood and paper.

One of the best ways to prevent them from loss and deterioration is to digitally preserve them. Digital data of heritage objects can be obtained by using computer vision techniques. Once these data have been acquired, they can be preserved permanently, and then safely passed down to future generations. In addition, such digital data is suitable for many applications, including simulation, restoration, and creating multi-media contents. Such digital contents can be viewed through the internet from anywhere in the world, without moving the objects nor visiting the sites.

We have been working to develop such digital archival methods by using computer vision and computer graphics technologies [1]. Similar projects include: Stanford's Michelangelo Project [2], IBM's Pieta Project [3], and

Columbia's project [4], to name a few. Our project has a number of unique features; among them is its ability to digitize relatively large objects outdoor such as the Kamakura great Buddha, and Cambodia's Bayon. This presents several challenges. Also, our project consists not only of geometric modeling but also of photometric and

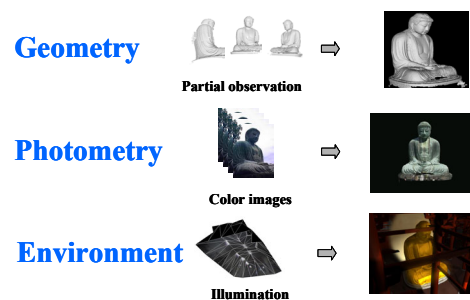


Figure 1. Three components of our project.

environmental modeling, as shown in Figure 1.

The remainder of this paper is organized as follows. Section 2 briefly describes the outline of the geometric pipeline developed, a parallel simultaneous alignment algorithm and a parallel voxel-based merging algorithm. Section 3 describes a method to align observed textures from a digital camera with a range data for texture mapping. Section 4 reports our efforts to restore the original appearance of these objects using acquired digital data and the literature survey. Section 5 summarizes this paper.

2. Geometric Modeling

Several computer vision techniques, such as traditional shape-from-X and binocular-stereo, or modern range sensors, provide cloud of point information. The cloud of point information certainly carries three-dimensional information pertaining to observed objects. However, there is no structural information among these points. Namely,

there is no information to represent adjacency among the points. The first step of geometric modeling is to convert the cloud of points into a surface representation such as a mesh model.

Since each observation provides only partial information, we have to combine these partial mesh representations into a whole geometric mesh representation. Thus, the second step in geometric modeling is to align these meshes so that their corresponding parts overlap one another (alignment).

To accomplish this step, we have developed several simultaneous alignment methods [11, 12]. Traditional alignment employs pair-wise alignment. By repeating this pair-wise alignment, all the relation among range images in a data set is determined. When handling a large number of data set, error gradually accumulates along the path, and a large gap exists locally in the final result. For avoiding error accumulation in a certain part, this algorithm employs a simultaneous alignment method, which determines all alignment relations among a data set at once. These simultaneous alignments evenly distribute error among all the relations.

Further, we also extended the algorithm into a parallel implementation so as to be able to align a very large data set [12]. The simultaneous algorithm, originally designed, requires all range images to be read into memory; even when the computation is distributed over a PC cluster, the amount of memory used on each PC is not reduced. For the parallel implementation, both time and memory performance have to be considered.

We remove redundant or weak data dependency among a data set for efficient memory usage. A pair of range image images that does not satisfies all these three conditions will be removed as redundant relations.

The bounding-boxes of two range images overlap each other.

The angle, θ , between ray directions of two range images, is less than a threshold value. Here a ray direction is the optical axis of a range sensor for scanning the data.

Two range images are adjacent and overlapping each other.

Since these assumptions work, the initial positions of range images are accurately estimated. For this purpose, as the initial step for this parallel alignment algorithm, we execute a pair-wise alignment algorithm for all adjacent overlapping range images.

We also make even distribution of the computational load over a PC cluster. The computational cost of a PC is estimated to be proportional to the number of vertices on the PC. By using this assumption, we assign range images to each PC. See [13] in details.

Table 1. Computational time.

	1Processor(ms)	16Processor(ms)
Hand Model	11723	1391
Kamakura Buddha	51715	5799

In Table 1, the computation times with 1 processor and 16 processors are described. The computation time with 16 processors is approximately 8.4 times faster than that with 1 processor for the hand model, and 8.9 times for the Great Buddha of Kamakura. As for the memory usage, we can reduce the amount of memory used decreases as the number of processor increases.

As shown in Table 2, our method can reduce the amount of memory used in approximately 60% for the hand model and in approximately 43% for the model of the Great Buddha of Kamakura. See [12] in more details.

Table 2. Amount of memory ratio.

	1Proc(MB)	16Proc(MB)	Ratio
Hand Model	49.2	29.3	0.5961
Kamakura Buddha	258.2	110.8	0.4293

The third geometric modeling step is to integrate the pre-registered multiple range images and to compose one complete geometric model, a step usually called 'merging'. The procedure can be considered as extracting one surface from multiple overlapped surfaces. In the merging procedure, it is important to make the integration framework robust against any noise which may be in the scanned range images and can also be inherited from the registration procedure [16,17,18].

Our method merges a set of range images into a volumetric implicit-surface representation, which is converted to a surface mesh by using a variant of the marching-cubes algorithm [14]. Unlike previous techniques based on implicit-surface representations, our method estimates the signed distance to the object surface by determining a consensus of locally coherent observations of the surface [15,16,17,18].

We utilize octrees to represent volumetric implicit surfaces, thereby effectively reducing the computation and memory requirements of the volumetric representation

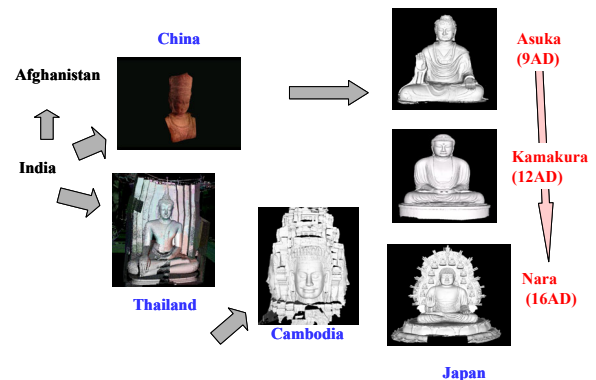


Figure 2. The World Buddha Library

without sacrificing accuracy of the resulting surface. We originally design software that merges a relatively small data set. However, our target requires to merge a huge set, we decide to design our algorithm to run on a PC cluster; the cluster parallel-processes the merging algorithm for saving the computation time and utilizing a large memory space of many PCs. We produced one integrated digital Great Buddha with this software. The whole data set consists of 3.3 M points, and 5 M polygons that can be merged in approximately 20 minutes on the PC cluster.

By using this geometric pipeline, we have digitally archived Japanese Buddhas, including Asuka, Kamakura, Nara, and foreign ones, including Thailand's Wat Si Chum and Cambodia's Biyon. We are continuing this effort toward completing the world Buddha library, as shown in Figure 2, which digitally display transitions of Buddha shapes in time and region.

3. Texture Mapping and Rendering

The geometric model is vital information regarding the cultural heritage objects because it enables us to analyze object in details. In addition to the geometric model, surface color distribution (texture) is also very important for some kinds of cultural properties. We have developed a method for mapping texture based on laser reflectance.

When a short-distance range sensors can be used, the most promising method is to calibrate the geometrical relationship between the image sensor and the range sensor before scanning using a calibration object. However, this method requires that the range and color sensors be fixed on the fixture once the relationship is calibrated. Further, the calibration-based method is accurate only around the position occupied by the calibration fixture. When a target object is very large, this method becomes unreliable due to the lens distortion. Thus, we need a method that does not rely on calibration.

Generally speaking, range sensors often provide reflectance images as side products of range images. The returned timing provides a depth measurement, while the

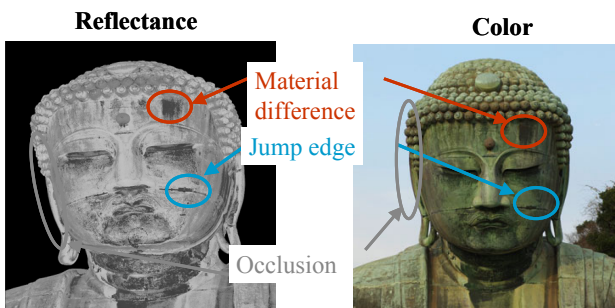


Figure 3. Reflectance and color edges. They share the similar characteristics.

returned strength provides a reflectance measurement. A reflectance image is a collection of the strength of returned laser energy at each pixel. This reflectance image is aligned with the range image because both images are obtained through the same optical receiving device. Commonly available range sensors, including ERIM, Preceptron, and our main sensor, CYRAX, provide this reflectance image.

We employ this reflectance image as a vehicle for the alignment of range images with color images [3,16]. Reflectance images share characteristics similar to color images due to the fact that both images are somehow related with surface roughness as shown in Figure 3. Since our CYRAX range scanner uses a green laser diode, reflectance edges can be observed along the boundary between two colors or material boundaries along difference reflectance ratios for this wavelength. Since different materials are of different colors, a discontinuity also appears in the color images. Jump edges along small ranges in a range image also appear as jump edges in a reflectance image as well as in a color image. Occluding boundaries are observed both in reflectance images and in color images.

Prior to the alignments, we paste the necessary reflectance edges onto the 3D geometric model. As mentioned above, since occluding boundaries vary depending on the viewing direction, edges along the occluding boundaries are first removed from the reflectance images. On the other hand, edges along the current occluding boundaries will be estimated from the 3D geometric model and the current viewing direction. Our algorithm extracts them automatically, and uses them for the alignment.

We align edges extracted from reflectance images with those in color images so that the 3D position error of those edges is minimized by iterative calculation as shown in Figure 4. Extracted edges are represented as a collection of points along them. The alignment is done between 3D

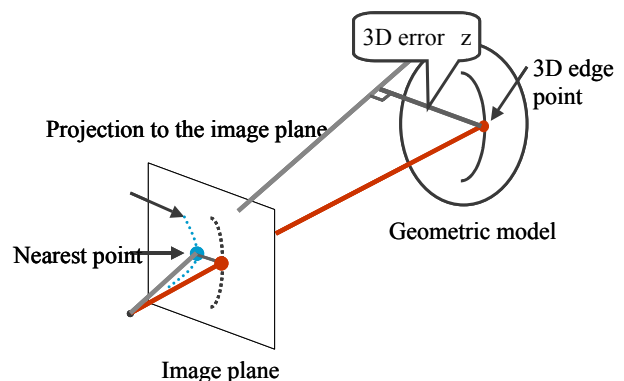


Figure 4. Correspondence between 3D reflectance and 2D color edges.

reflectance points on 3D geometric model projected on the image plane and 2D color edge points in the 2D image.

To establish correspondence, the system finds the color image points that are nearest to the projected reflectance points. This operation is similar to the ICP operation.

To determine the relative pose that coincides with the position of 2D color edges and projected 3D reflectance edges, we use the M-estimator.

First the distance between corresponding 2D color edge points and 3D reflectance edge points is evaluated as shown in Figure 4 : where z_i is a 3D error vector which is on a perpendicular line from a 3D reflectance edge point to the stretched line between the optical center and a 2D color edge point on the image plane.

$$\varepsilon_i = Z_i \sin \theta$$

where Z_i is the distance between the optical center and a 3D reflectance edge point, and θ is the angle between the color edge point and the reflectance edge point.

The system finds the configuration, P , which minimizes the total error, E , where ρ is an error function. The minimum of $E(p)$ can be obtained by:

$$\frac{\partial E}{\partial P} = \sum_i \frac{\partial \rho(\varepsilon_i)}{\partial \varepsilon_i} \frac{\partial \varepsilon_i}{\partial P} = 0$$

We can consider $\omega(\varepsilon)$ as a weight function to evaluate error terms.

$$\omega(\varepsilon) = \frac{1}{\varepsilon} \frac{\partial \rho}{\partial \varepsilon}$$

Thus, the minimization can be derived as the following least squared equation:

$$\frac{\partial E}{\partial \rho} = \sum_i \omega(\varepsilon_i) \varepsilon_i \frac{\partial \varepsilon_i}{\partial P} = 0$$

We choose the Lorentzian function for this function.

$$\omega(\varepsilon) = \left(1 + \frac{1}{2} \left(\frac{\varepsilon}{\sigma} \right)^2 \right)^{-1}$$

By solving this equation using the conjugate gradient method, we can obtain the configuration P that minimizes the error term and gives the relative relationship between the camera and the range sensor. Figure 5 shows the texture mapped Kamakura Buddha. Since this method minimizes a non-linear equation, we need an initial alignment. The initial alignment is given manually using our GUI. For the current implementation, relatively accurate alignment is necessary for rotation, but it is not the case for translation.



Figure 5. Texture Mapped Kamakura Buddha..

4. Restoring Hypothesized Original State

After we obtain the precise geometry and photometry information of the cultural assets in the current state, we can restore them to their hypothesized original state. In this section, we describe one of the examples: the restoration of the Nara Great Buddha and its main hall

4.1 Restoring Nara Great Buddha

Nara Great Buddha is the main statue of Toudaiji Temple. Unfortunately, the current statue is a rebuilt and repaired one because the original statue was burned and melted down due to a couple of civil wars. Accordingly, the shape of the current Great Buddha is different from that of the original one.

By using the geometrical modeling shown in Section 2, we have acquired the complete 3D geometrical model of Buddha in its current state. From this model, we have attempted to synthesize the original state by morphing the 3D geometry of the model.

From some literature inherited at the temple, we know the sizes of various face parts such as the nose and mouth. Using these data, we design a two-step morphing algorithm.

First, we globally change the scale of the whole portions



(a) Current (b) Hypothesized
Figure 6. Restored result of Nara Great Buddha.



Figure 7. Miniature and its 3D scanned models.

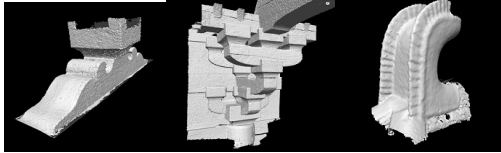
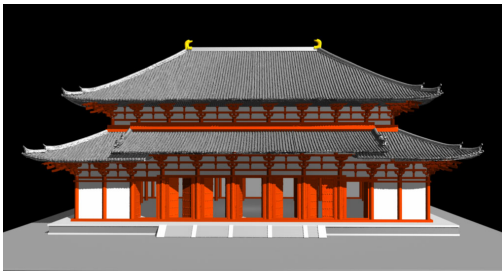


Figure 8. Partial models acquired at Tousyoudaiji.



(a) Current Toudaiji main hall built in 18AD.



(b) Hypothesize 8AD Toudaiji main hall.



(c) Hypothesized 8AD Nara Buddha in the hypothesized 8AD main hall.

Figure 9. Digitally restored Nara Great Buddha.

gradually modified. In the 2nd stage, vertices are moved one by one iteratively, similar to the constraint propagation algorithm, using smoothness and uniform constraints. The 2-stage morphing enables us to obtain the complete model of the original Great Buddha. Figure 6 shows the 3D models of the current (a) and the original Great Buddha (b). We can easily recognize that the original Buddha is larger and rather skinny.

4.2 Restoring Toudaiji Main Halle

The main hall of the Toudaiji Temple was built during the same decades as those of the Great Buddha (8th century). It was also rebuilt twice: in the 12th and 18th centuries. In the 12th century, Tenjiku architecture was imported from China and the main hall was rebuilt in a totally different architecture style. The rebuilding in the 18th century followed the same new style. As a result, the style of the current main hall is entirely different from that of the original building.

Fortunately, the Toudaiji temple has been displaying a miniature model of the original hall, constructed for the Paris Expo in 1900, as shown in Figure 7(a). We digitized it using the Pulsteck TDS-1500 and scaled it up to the original size as shown in Figure 7(b).

Due to the limitation of resolution, the detail parts cannot be obtained precisely. According to Prof. Keisuke Fujii, an architecture professor at the University of Tokyo, one of the experts on building style in the era, the Toudaiji and Touthou-daiji Temples share a similar format. Here, the main hall of Touthoudaiji Temple were also built during the same period (8th century). After scanning various key parts of the main hall at Touthoudaiji, as shown in Figure 14, we morphed these partial range data by expanding and shrinking the Touthoudaiji parts (Figure 8) to the scaled-up range data of the Toudaiji (Figure 7). The process was conducted by an extended alignment algorithm that allows scale change as well as configuration differences. Figure 9 shows the current Nara Great Buddha main hall and the original one digitally restored by our method.

5. Conclusion

In this paper, we introduced our project to digitally archive and restore cultural heritage objects. Our project's main goal is to develop a method of 'modeling from reality', in which the digital model of cultural properties is created by using various computer vision methods. For the observation of geometrical information, we used laser range finders and post process algorithms, including registering and merging the range images. For the texture information, we have developed several texture mapping methods. For the short distance range sensors, we

calibrated the relationship between the range sensor and the image sensor for complete texture alignment. For the long distance range sensors, we developed a non-calibrated texture alignment method by using laser reflectance features. Digital restoration of lost cultural heritage objects has a big advantage compared with other restoration methods such as physical construction of actual temples, because we can examine various hypotheses without any physical changes nor long building periods. We demonstrated the effectiveness of this method through the restoration of the Nara Great Buddha and its main hall. We are also conducting a project to create a digital library of the world great Buddhas, including three Japanese Buddhas, Sri Chum Buddha in Thailand, and Bion's in Cambodia. The models and restoration results constructed so far can be viewed at [21]

Acknowledgment

This research is sponsored, in part, by JST under Ikeuchi Crest program. The Bayon in Cambodia was digitized with the cooperation of Japanese Government Team for Safeguarding Angkor (JSA).

References

- [1] K. Ikeuchi and Y. Sato, *Modeling from Reality*, Kluwer Academic Press, 2001.
- [2] M. Levoy et. al., "The digital Michelangelo project," *SIGGRAPH 2000*, New Orleans.
- [3] J. Wasserman, *Michelangelo's Florence Pieta*, Princeton University Press 2003.
- [4] I. Stamos and P. Allen, "Automatic registration of 2-D with 3-D imagery in urban environments," *ICCV2001*, Vancouver.
- [5] P.J. Besl and N.D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Patt. Anal. Machine Intell.*, 14(2):239-256, 1992.
- [6] R. Benjemma and F. Schmitt, "Fast global registration of 3D shample surfaces using a multiple-z-buffer technique," *Int. Conf on Recent Advances in 3-D Digital Imaging and Modeling*, pp. 113-120, May 1997.
- [7] P. Neugebauer, "Geometrical cloning of 3D objects via simultaneous registration of multiple range images," *Int. Conf on Shape Modeling and Application*, pp.130-139, March 1997.
- [8] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, 10(3):145-155, April 1992.
- [9] S. Rusinkiewicz and M. Levoy, "Efficient variants of the IPC algorithm," *Int. Conf 3-D Digital Imaging and Modeling*, pp.145-152, May 2001.
- [10] H. Gagnon, M. Soucy, R. Bergevin, and D. Laurendeau, "Registration of multiple range views for automatic 3-D model building," *CVPR94*, pp.581-586.
- [11] K. Nishino and K. Ikeuchi, "Robust Simultaneous Registration of Multiple Range Images", *Fifth Asian Conference on Computer Vision ACCV '02*, pp454-461, 2002.
- [12] T. Oishi, R. Sagawa, A. Nakazawa, R. Kurazume, and K. Ikeuchi, "Parallel Alignment of a Large Number of Range Images on PC Cluster," *Int. Conf 3-D Digital Imaging and Modeling*, Oct 2003.
- [13] M. D. Wheeler and K. Ikeuchi, "Sensor Modeling, Probablistic Hypothesis Generation, and Robust Localization for Object Recognition", *IEEE PAMI*, 17(3): 252-265, 1995.
- [14] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," *SIGGRAPH 96*, New Orleans, LA.
- [15] M. Wheeler, Y. Sato, and K. Ikeuchi, "Consensus surfaces for modeling 3D object from multiple range images," *ICCV98*.
- [16] R. Sagawa, K. Nishino, M.D. Wheeler and K. Ikeuchi, "Parallel Processing of Range Data Merging", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1, pp577-583, 2001
- [17] R. Sagawa, T. Masuda, and K. Ikeuchi, "Effective Nearest Neighbor Search for Aligning and Merging Range Images," *Int. Conf 3-D Digital Imaging and Modeling*, Oct 2003
- [18] R. Sagawa and K. Ikeuchi, "Taking Consensus of Signed Distance Field for Complementing Unobservable Surface," *Int. Conf 3-D Digital Imaging and Modeling*, Oct 2003
- [19] R. Kurazume, M. D. Wheeler, and K. Ikeuchi, "Mapping textures on 3D geometric model using reflectance image," *Data Fusion Workshop in IEEE Int. Conf. on Robotics and Automation*, 2001.
- [20] R. Kurazume, K. Nishino, Z. Zhang, and K. Ikeuchi, "Simultaneous 2D images and 3D geometric model registration for texture mapping utilizing reflectance attribute," *Fifth Asian Conference on Computer Vision*, 2002.
- [21] http://www.cvl.iis.u-tokyo.ac.jp/galley_e/.