

CONSTRUCTION OF A SYSTEM FOR DEFINING AREAS WHICH ARE NOT OBTAINED DATA FROM STEREO IMAGES

Hideharu Yanagi ^{a,c}, Hirofumi Chikatsu ^b

^a Japan Association of Surveyors, 1-48-12, Itabashi, Itabashi-ku, Tokyo 173-0004, Japan - yanagi@geo.or.jp

^b Department of Civil and Environmental Engineering, Tokyo Denki University,
Hatoyama, Hiki-gun, Saitama 350-0394, Japan - chikatsu@g.dendai.ac.jp

^c School of Science and Engineering, Tokyo Denki University, Hatoyama, Hiki-gun, Saitama 350-0394, Japan

Commission V, WG V/1

KEY WORDS: Macro lens, Digital very close range photogrammetry, 3D modeling, Small object

ABSTRACT:

Recently, digital documentation and visualization of various cultural assets have been receiving attention. For example, a small Buddha with a height of approximately 4 cm is categorized as a cultural asset. Such a small object has to be documented. Generally, in order to perform 3D modeling of small objects from the viewpoint of digital very close range photogrammetry, multi view images are taken. However, it is important to confirm the occluded part and image quality after taking images. In order to confirm the defocusing of images and the areas which were not obtained data from stereo images, a system is proposed for confirming data with a macro lens and convenient 3D measurement software called 3DiVision. Finally, it is proposed that the data be supplemented in areas which are not obtained data.

1. INTRODUCTION

Recently, digital documentation of various cultural assets has been receiving attention. Small objects, such as an extremely small Buddha with a size of approximately 4 cm and an old coin that was unearthed from remains, are categorized as cultural assets. In order to document such small cultural assets, convenient 3D measurements are required.

On the other hand, convenient 3D measurements using consumer-grade digital cameras are eagerly awaited in various application fields. In these circumstances, the desire is not only for convenient 3D measurements using consumer-grade digital cameras but also for a convenient method to integrate multiple images without ground control points (GCPs). The authors have been attempting to develop convenient 3D measurement software called 3DiVision by using consumer-grade digital cameras (Chikatsu et al., 2006). 3DiVision has the capability to perform exterior and interior orientation simultaneously without GCPs.

Generally, in order to perform 3D modeling of a small object, multi view images are taken by digital very close range photogrammetry. However, it is important to confirm the occluded part after taking images. In 3D measurements using a terrestrial laser scanner, a problem due to occlusion has reportedly been solved by inserting a two-dimensional high resolution image into the missing location in the 3D model (Martínez et al., 2011). In digital very close range photogrammetry, the construction of an imaging system to confirm data in real time is needed in order to confirm as soon as possible the defocusing of images and the areas which were not obtained data from stereo image.

With this motivation, an imaging system with 3DiVision and a macro lens was proposed in the present study. Moreover, an investigation was conducted on the proposed method for supplementing the data in areas which were not obtained data from stereo images.

2. PROPOSED IMAGING SYSTEM

Table 1 lists the specifications of the macro lens that was used in this investigation. This lens was mounted on a Canon EOS 50D.

	Canon EOS 50D
Camera	15.1 megapixels, CMOS Sensor size: 22.3 × 14.9 mm
Lens	Canon EF100mm f/2.8 Macro USM Focal length: 100 mm, F: 2.8

Table 1. Specifications of camera and lens

In order to construct the proposed imaging system, the following are required:

- 1) Consider the working distance from the small object to the camera.
- 2) Take multi view images using one camera.
- 3) Construct a pseudo 3D model in real time.
- 4) Confirm the defocusing of images.
- 5) Make convenient 3D measurements using consumer grade digital cameras without GCPs.
- 6) Define the areas which were not obtained data from stereo images.
- 7) Supplement the data in areas which were not obtained data from stereo images.

The imaging system proposed in this study and that satisfies requirements 1 through 6 is shown in Figure 1. The size of the small object shown in Figure 1 is approximately 8.6 cm and the altitude is approximately 90 cm. The camera is connected directly to a personal computer (PC), and images are taken using the PC.

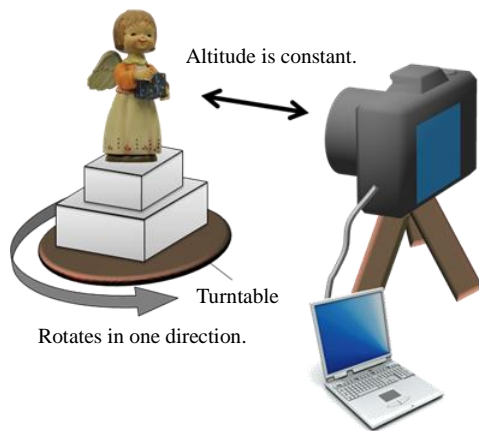


Figure 1. Proposed imaging system.

The depth of field depends on the principal distance and the aperture value of the lens. For 3D modeling of cultural assets in particular, it is important to adjust the depth of field for each image obtained. However, appropriate imaging modes can be used to achieve superior imaging. The camera listed in Table 1 offers a selection of imaging modes as follows:

- P mode (program automatic exposure):
Aperture value and shutter speed are set automatically depending on the brightness.
- AV mode (aperture value):
Aperture value can be changed manually.
- A-DEP mode (automatic depth of field):
Aperture value is set automatically depending on the auto focus (AF) frame.

In this study, the AV mode was selected for the ability to set an aperture value manually with respect to the depth of field. The focal length of the macro lens specified in Table 1 is fixed. Another characteristic feature of a macro lens is the working distance; i.e., the distance between the front of the lens and the object. Suitably adjusting the working distance enables the capture of sharp images without the influence of shadows caused by lighting. As shown in Figure 1, it is understood that requirements 1 and 2 are satisfied because the camera position is fixed and a turntable is used. Requirements 3 through 6 will be discussed in the sections below.

3. THE PSEUDO 3D MODEL

A characteristic of the proposed system is a fixed camera position. The small object is set at the center of a turntable. The turntable is rotated in one direction, and images of the small object are taken with one camera. Therefore, the small object is always in the center of the image. In addition, the images of the small object are saved in a PC, because the camera is connected directly to a PC. The images of the small object are taken by the PC while the turntable is rotated in one direction. After the turntable has rotated once, the images of the small object appear continuous like a 3D model rotated on a display. In this study, the continuous images that are shown in Figure 2 are called the pseudo 3D model.

Therefore, the image defocusing and the occlusion areas are confirmed from the pseudo 3D model. Thus, it is understood that requirements 3 and 4 are satisfied.

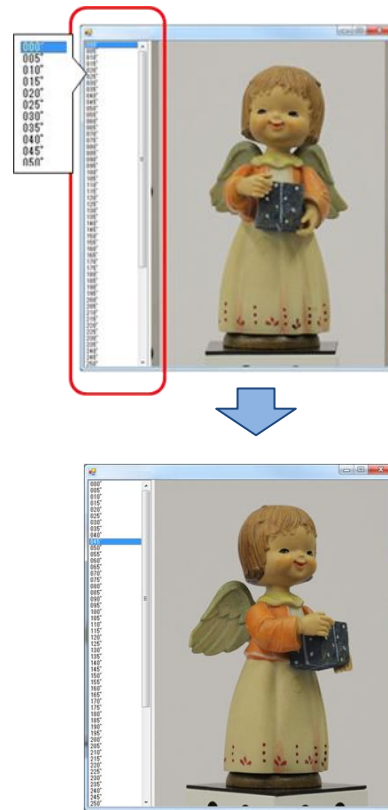


Figure 2. Pseudo 3D model.

4. 3D MODELING AND ORTHOPHOTOS

The proposed imaging system uses 3DiVision. Hence, it is understood that requirement 5 is satisfied. In this study, triplet images that are taken using the proposed imaging system (Figure 3) and 3DiVision are used for 3D models and orthophotos.

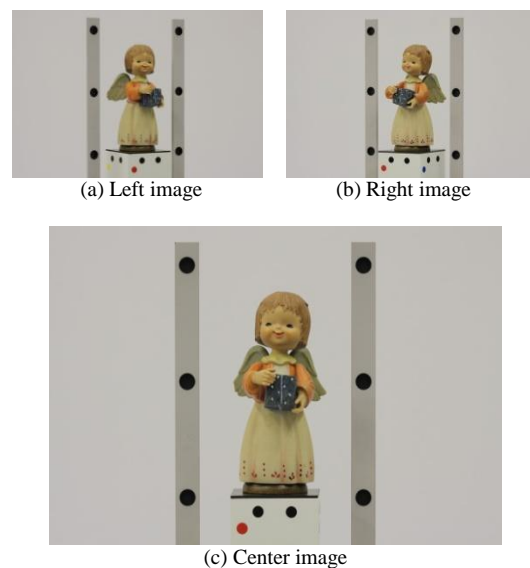


Figure 3. Triplet images.

4.1 Image processing

In this study, the feature points for the 3D modeling of the small object were obtained by stereo matching the left and right images of the triplet. However, it was difficult to obtain the feature points of the small object because there are many similar colors and patterns on its surface. Therefore, the feature points were increased by processing the images using a Wallis filter (Jazayeri & Fraser, 2008). The Wallis filter enhances the contrast of an image. Figure 4 shows an image processed by using the Wallis filter in equation (1) with a mask size of 51×51 . The parameter values applied to the Wallis filter were as follows: $m_f = 127$, $s_f = 80$, $c = 1.0$, and $b = 0.8$.

$$g^c(x, y) = [g(x, y) - m_g] (cs_g + s_f / c) + bm_f + (1 - b)m_g \quad (1)$$

where, $g^c(x, y)$: Highlighted pixel
 $g(x, y)$: Brightness value
 m_g : Original mean
 s_g : Original standard deviation
 m_f : Target mean
 s_f : Target standard deviation
 c : Contrast expansion constant
 b : Brightness forcing constant



(a) Original image (b) Image processed using a Wallis filter

Figure 4. Image processed using a Wallis filter.

4.2 Stereo matching

Generally, it seemed that the stereo matching needed much processing time. Therefore, in order to reduce the processing time, the image was divided into areas (Figure 5) by using an Otsu method (Otsu, 1980), and the stereo matching process was done in each area without a background.



(a) Original image (b) Classified image

Figure 5. Image processed using an Otsu method.

However, mismatching is an extremely important issue in stereo matching. In this study, in order to reduce the amount of mismatching, a reverse projection method (Sugiyama & Chikatsu, 2011) was used in the stereo matching process. The reverse projection method uses triplet images and each orientation element of the image. The reverse projection method is implemented as follows: first, the stereo matching is performed using the left and right images (Figure 3); second, the result point of the matching is checked with points on the center image that were projected using the interior and exterior orientation elements of the center image. In this study, the stereo matching method uses the area-based matching that is given by equation (2). The correlation value threshold and masking window size were set to 0.4 and 51×51 , respectively.

$$C_{ij} = \sum_{u=1}^m \sum_{v=1}^l \frac{(f(u, v) - f_{ave})(t(u, v) - t_{ave})}{\sqrt{f_s \cdot t_s}} \quad (2)$$

where, C_{ij} : Correlation value of pixel position (i, j) in the image,
 (i, j) is the center of the masking window

m : Height of masking window

l : Width of masking window

$f(u, v)$: Threshold of pixel position (u, v) in the target image

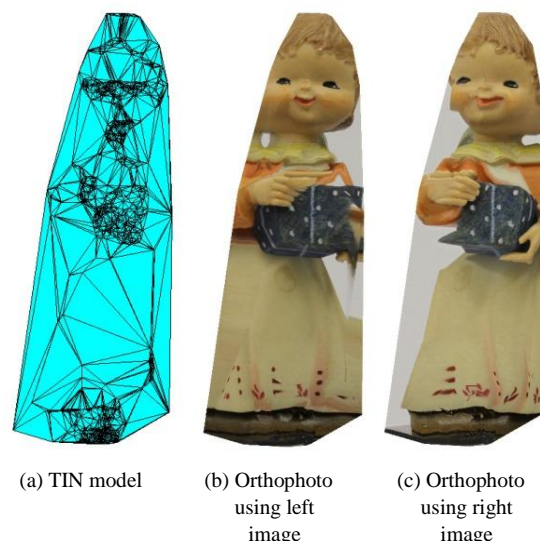
$t(u, v)$: Threshold of pixel position (u, v) in the template image

$$f_{ave} = \sum_{u=1}^m \sum_{v=1}^l \frac{f(u, v)}{l \times m}, t_{ave} = \sum_{u=1}^m \sum_{v=1}^l \frac{t(u, v)}{l \times m}$$

$$f_s = \sum_{u=1}^m \sum_{v=1}^l (f(u, v) - f_{ave})^2, t_s = \sum_{u=1}^m \sum_{v=1}^l (t(u, v) - t_{ave})^2$$

4.3 TIN modeling and orthophoto images

Figure 6(a) shows a triangulated irregular network (TIN) model using result points described in Section 4.2. Figures 6(b) and 6(c) show the orthophoto images obtained using the TIN with the left and right images, respectively. From these results, it is clear that each orthophoto image is a stretched texture image with textures other than those of the doll.



(a) TIN model (b) Orthophoto using left image (c) Orthophoto using right image

Figure 6. TIN model and orthophoto images.

4.4 Defining areas which are not obtained data from stereo images

Each orthophoto image is constructed from corresponding points on left and right images that have been extracted by stereo matching. Therefore, it seems that the texture images of the orthophoto images should each contain the same color information. However, Figure 7 shows that the texture images extracted from the left and right images are different. It is clear in Figure 7 that the color information within the red triangles is the same but the color information within the blue triangles is not.

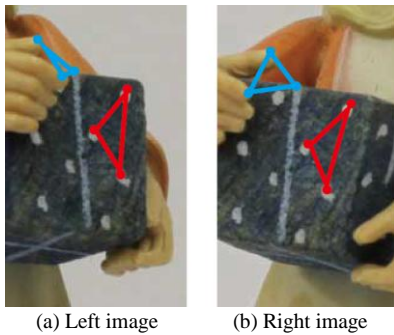


Figure 7. Difference in color information.

With this motivation, a difference image was constructed using Figures 6(b) and 6(c). Figure 8 shows the difference image that is extracted given the same color information at the same location in both images. The black areas shown in Figure 8 are the areas where the same color information cannot be extracted from both images (e.g., anywhere that color information recorded in the left image was not recorded at the same location in the right image). From these results, it is clear that the black areas in Figure 8 define areas which are not obtained data from stereo imagea.



Figure 8. Difference image.

5. DEFINING AREAS WHICH ARE NOT OBTAINED DATA FROM TRIPLET IMAGES

The triplet image is used in this proposed system. In Section 4.4, the left and right images are used for defining areas that are not

obtained data. Therefore, it can be inferred that additional data can be obtained from the pair of left and center images or the pair of center and right images. Figure 9 shows the difference image made from each pair. It is clear from this figure that different data were obtained in each image.



(a) Difference image using left and center images (b) Difference image using center and right images

Figure 9. Difference images.

However, the images shown in Figures 8 and 9 were calibrated using the same coordinate system. Therefore, all three images can be overlaid. Figure 10 shows the three images overlaid, and data have clearly been added to the black areas in Figure 8. However, there are still black areas in Figure 10. That is, the black areas in Figure 10 show the areas which are not obtained data from images.



Figure 10. Areas which are not obtained data from the triplet images.

Consequently, it is understood that requirement 6 of Section 2 is satisfied.

6. SEARCH FOR CORRESPONDING IMAGES

The orthophoto images constructed in this study are based on the TIN model. Therefore, in order to supplement the data in the areas which are not obtained data, the triangles forming the TIN are examined. The eye vector is a base vector that is defined between a triangle in the TIN and the camera position set in front of the small object. An image that corresponds to the area which is not obtained data is searched for by calculating the angle between the eye vector and normal vector of the triangle. Figure 11 shows the relationship between the eye vector and the normal vector. The angle between the eye vector and the normal vector is given by equation (3).

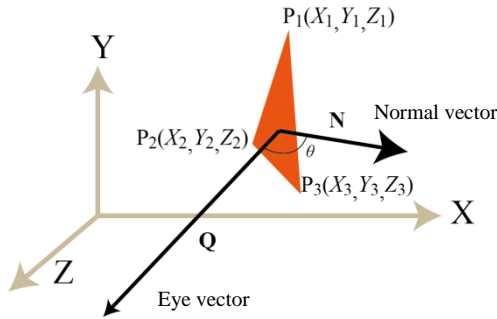


Figure 11. Relationship between eye vector and normal vector.

$$\cos \theta = \frac{\mathbf{N} \cdot \mathbf{Q}}{|\mathbf{N}| \cdot |\mathbf{Q}|}, \quad \mathbf{N} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \quad (3)$$

where, \mathbf{N} : Normal vector

\mathbf{Q} : Eye vector

$\mathbf{i}, \mathbf{j}, \mathbf{k}$: Unit vectors

a, b, c : Vector components

$$a = (Y_2 - Y_1)(Z_3 - Z_1) - (Z_2 - Z_1)(Y_3 - Y_1)$$

$$b = (Z_2 - Z_1)(X_3 - X_1) - (X_2 - X_1)(Z_3 - Z_1)$$

$$c = (X_2 - X_1)(Y_3 - Y_1) - (Y_2 - Y_1)(X_3 - X_1)$$

Therefore, the triangles that corresponded to the black areas in Figure 10 were extracted, and the angle of each triangle was calculated from equation (3). The pink triangles shown in Figure 12 are those corresponding to the black areas in Figure 10.

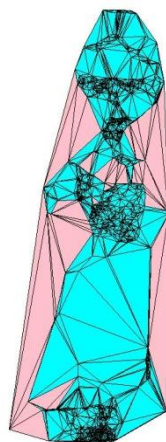


Figure 12. Corresponding triangles (pink).

Angle(°)	a	b	c	Angle(°)	a	b	c
2.5470	-0.0022	-0.0004	0.2254	69.7746	-0.0254	-0.0986	0.0409
6.4359	-0.0208	0.0205	0.3047	70.9886	-0.0435	0.0106	0.0162
7.4022	-0.0085	0.0412	0.3671	75.1923	-0.1220	-0.0970	0.0444
9.1170	0.0385	0.2886	2.0198	75.8923	-0.1195	-0.0215	0.0317
10.8133	-0.0086	0.0000	0.0472	76.2356	-0.1896	0.3109	0.1057
11.1733	-0.0084	-0.0001	0.0443	78.3137	4.8614	8.9446	2.3881
11.9364	0.0562	0.2478	1.3093	79.5486	-0.1652	3.7134	0.7826
12.2654	0.0003	0.0103	0.0523	80.7062	-0.1927	0.0213	0.0342
12.2978	0.0001	0.0103	0.0523	81.6252	-0.0751	-0.0543	0.0116
13.7689	0.0194	0.0206	0.1207	82.1881	-0.2219	-0.2557	0.0540
14.3143	0.0098	0.0103	0.0581	82.4537	-0.1527	-0.2456	0.0451
25.9379	-0.1245	-0.0221	0.2617	83.3925	0.0036	-0.1034	0.0096
28.7571	0.6171	-0.0559	1.0826	83.7692	1.9932	1.3772	0.3209
30.3892	0.0429	-0.0111	0.0723	84.3811	0.2757	-0.0023	0.0245
35.3569	0.1230	-0.0006	0.1727	84.5667	0.0112	-0.1300	0.0076
44.8325	-0.0557	0.0428	0.0751	86.8449	-0.1326	0.7182	0.0229
49.0949	-0.0371	0.0000	0.0323	87.3617	-0.0157	-1.6151	0.1094
49.3248	-0.0355	-0.0104	0.0320	87.3828	-0.0001	-0.1618	0.0013
51.3942	0.0676	-0.0983	0.0989	88.1997	0.0044	-0.3922	0.0206
55.7907	0.0189	-0.2171	0.1368	88.2233	-0.1326	2.6531	0.0132
58.4264	-0.3725	-0.7474	0.4957	88.3686	-0.7948	1.0965	0.0157
60.1019	0.0388	0.0000	0.0222	88.4519	0.2501	-0.2252	0.0127
63.1578	-0.4445	0.0002	0.2250	89.0306	-0.0578	0.2068	0.0087
65.6824	-0.0011	-1.5432	0.6307	89.0628	3.8930	-1.4334	0.0761
66.7888	-22.3536	27.9778	17.5167	89.9559	-2.0540	-0.2335	0.0206
67.4302	-0.6086	4.2634	2.0929	90.0225	-0.4047	-0.2962	0.0091
67.6171	-0.0886	0.0858	0.0561	90.1234	-2.6312	-0.8895	0.0415
67.9565	-0.1995	-0.5487	0.2129	90.5033	0.9592	-1.6206	0.0190

Table 2. Angle between eye vector and normal vector

Table 2 shows the angles and vector components that were calculated using the pink triangles that are shown in Figure 12. The vector components correspond to each axis of the coordinate system shown in Figure 11 (e.g., a corresponds to the X-axis). The vector components indicate the direction and magnitude corresponding to each axis. It is inferred that the angle represents the center image of the triplet.

In order to confirm that data can be supplemented, the angles and vector components were verified using the contents of Table 2 that are shaded in gray. These contents were selected because one angle (6.4359°) is near the top of the list. The other angle (9.1170°) was selected to supplement from the left and right side of the model shown in Figure 10. The sign of a for 9.1170° is opposite to that for 7.4022° , and the sign of a for 7.4022° is same as that for 6.4359° . The contents of the first row were excepted because the image that was searched by using the angle 2.5470° was same as that in Figure 3(c).

Figure 13 shows the images that were searched by the proposed imaging system. In this study, the turntable was rotated by increments of 5 degrees. Therefore, the search angle was selected close to the angles at which images were taken using this system.



(a) Image taken when turntable was rotated 5°



(b) Image taken when turntable was rotated 350°

Figure 13. Images searched.

As each image in Figure 13 is the center image of a triplet, the left and right images that corresponded to each center image were searched, and each difference image was constructed using the proposed method. Figure 14 shows the difference images that were constructed from these results.

7. CONCLUSION

In this study, the effectiveness and practicality of the proposed imaging system were investigated for the documentation of small objects. Consequently, results obtained in this study show that the proposed imaging system can be used effectively for the documentation of small objects, since the areas which are not obtained data are supplemented using the proposed method.

References

Martínez, J., Finat, J., Fuentes, L.M., Gonzalo, M., Vilorio, A., 2005. A COARSE-TO-FINE CURVED APPROACH TO 3D SURVEYING OF ORNAMENTAL ASPECTS AND SCULPTURES IN FACADES. Proc. XX CIPA Symposium, Turin, 26 Sep. – 1 Oct. <http://cipa.icomos.org/fileadmin/template/doc/TURIN/441.pdf> (14 Apr. 2011).

Chikatsu, H., Ohdake, T., 2006. UBIQUITOUS DIGITAL PHOTOGRAMMETRY BY CONSUMER GRADE DIGITAL CAMERA. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI, PART 5, ISSN 1682-1750 (on CD-ROM).

Jazayeri, I., Fraser, C.S., 2008. INTEREST OPERATORS IN CLOSE-RANGE OBJECT RECONSTRUCTION. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII, PART B5, pp.69-74.

Otsu, N., 1980. An Automatic Threshold Selection Method Based on Discriminant and Least Squares Criteria. *Trans. IECE of Japan*, Vol. 63-D, No.4, pp.349-356.

Sugiyama, F., Chikatsu, H., 2011. Improvement in the Accuracy of DEM for Orthophoto Generation by Reverse Projection Method. *Journal of ISPRS*, Vol. 50, No. 2, pp.80-89.



(a) Image taken when turntable was rotated 5° (b) Image taken when turntable was rotated 350°

Figure 14. Difference images that were constructed using search angle.

Figure 15 shows overlaid images of Figure 10 and Figure 14. In this figure, the black area in Figure 10 is supplemented using the image of Figure 14. Where the black areas are still not supplemented with data, it is possible to supplement using another angle in Table 2. Therefore, it is understood that requirement 7 of Section 2 is satisfied by this result.

From these results, all seven requirements that were described in Section 2 are satisfied.



(a) The image before supplementing data (b) The supplemented image

Figure 15. Supplemented difference image.