16th World Conference on Earthquake, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 679

Registration Code: S-Y1464668965

3D Model Reconstruction Using Aerial Photos and Application to Detection of Buildings Damaged due to Earthquakes

M. Matsuoka[(1)], H. Kawano[(2)], T. Satoh[(3)], H. Suzuki[(4)], I. Sato[(4)], and J. Miura[(4)]

[(1)] *Associate Professor, Tokyo Institute of Technology, Japan, matsuoka.m.ab@m.titech.ac.jp*
[(2)] *Formerly Graduate Student, Tokyo Institute of Technology, Japan, hy-kawano@ctie.co.jp*
[(3)] *Adjunct Associate Professor, Tokyo Institute of Technology, Japan, tuoost7017@pasco.co.jp*
[(4)] *Engineer, Aero Asahi, japan, {hideo-suzuki, isao-satou, jun-miura}@aeroasahi.co.jp*

Abstract

This paper present a disaster monitoring methodology based on the combination of Structure from Motion (SfM) and Multi View Stereo (MVS) techniques; the SfM-MVS technique is one of the three dimensional shape reconstruction technology based on the assembling of photos acquired from different sources. In order to study the potentiality of the technique for damage detection, first the SfM-MVS accuracy has been investigated. For such purpose, the dataset over the damaged areas in Hiroshima by a large-scale debris flow occurred on August 20th, 2014 was used. A number of photos were taken by an aerial vehicle just after the event, and LiDAR measurement was also carried out at the same time. The aerial photos were used for the creation of the surface model by SfM-MVS, and LiDAR data were used for the validation of the height values. For examining the height accuracy, we applied SfM to the dataset and observed the numbers of feature points extracted from input image. Then, we created point cloud by MVS and calculated the difference of vertical values and its error from the created model and actual height information from LiDAR data. Next, for the application to earthquake damage detection, we applied SfM-MVS to the affected areas by the 1995 Kobe earthquake and evaluated the extraction of severely damaged buildings.

Keywords: Three-Dimensional Model, SfM, MVS, the 2014 Hiroshima debris flow, the 1995 Kobe earthquake

1. Introduction

Large-scale natural disasters have caused human and property damages in the past. In the 1995 Kobe earthquake, the greatest cause of death was building collapses that occurred in a wide area. Remote sensing technologies are quick and effective solutions for gathering information on significant and widespread building damage. Most of the techniques use a vertical shot, taken from a satellite or an airplane, and the building damage is extracted using image-processing technologies [1]. However, in many cases, it is difficult to accurately analyze the damage to the middle and lower floors of the buildings with the existing structural analysis techniques that use only two-dimensional information such as photographs (taken vertically).

The SfM-MVS (Structure from Motion and Multi-view Stereo) technique can reproduce three-dimensional structures from the photographs taken at various angles from a variety of platforms. Using consumer-grade compact digital cameras, we can achieve the accuracy of the expensive three-dimensional measurement equipment even without the information obtained from the camera. This system excels for the low costs and for the simplicity of the post processing of data. The only issue is that we need high-resolution images to reproduce the information accurately.

Uchiyama et al. [2] and Obanawa et al. [3] made topography models, targeting the case study area where slope failure occurred, and the land-tied island where an on-site investigation was difficult, respectively, and illustrated the possibility of applying this technology to disaster investigation. However, the obtained aerial photographs of the disaster had various resolution qualities, and the altitude accuracy of the reconstructed three-dimensional model was not high. In addition, although this technology was applicable in damage investigation, it was not frequently used for collecting damage information after a disaster.

Therefore, in this study, we applied this technology to aerial photographs of disaster sites to validate the altitude accuracy of the three-dimensional models developed. In addition, we examined the relationship between the altitude accuracy and the spatial resolution of the images. Furthermore, we attempted to extract information about the building damage from the developed three-dimensional model of the disaster site.

2. Three-Dimensional Reconstructions and the Development of Models Using the SfM-MVS Technology

2.1 Three-dimensional structure reconstruction by SfM

The structure from motion (SfM) technology replicates the three-dimensional structure of an object and generates the detail information of the camera that was used for photographing, using the geometric relationship between the images. Pixels containing abundant shading information in the input image are detected as characteristic points, and the data association is performed between the images [4]. A constraint called epipolar geometry occurs at the points that correspond in the various images. Using this constraint, we can calculate the camera details, the photographing position and posture, and the three-dimensional positions of the corresponding points, which are all needed for the reconstruction of the structure. Generally, the image coordinates and the three-dimensional position coordinates are expressed by the following camera equations:

$$\lambda x = PX \tag{1}$$

$$P = A\begin{bmatrix} R & T \end{bmatrix} \tag{2}$$

Here, λ is a scalar value for normalization, x and X are a point on the image, and the three-dimensional point with respect to the corresponding point in the homogeneous coordinates, respectively, and P is the 3×4 camera matrix. A is a 3×3 calibration matrix for the details of the camera, R is a rotation matrix, and T is a translation matrix.

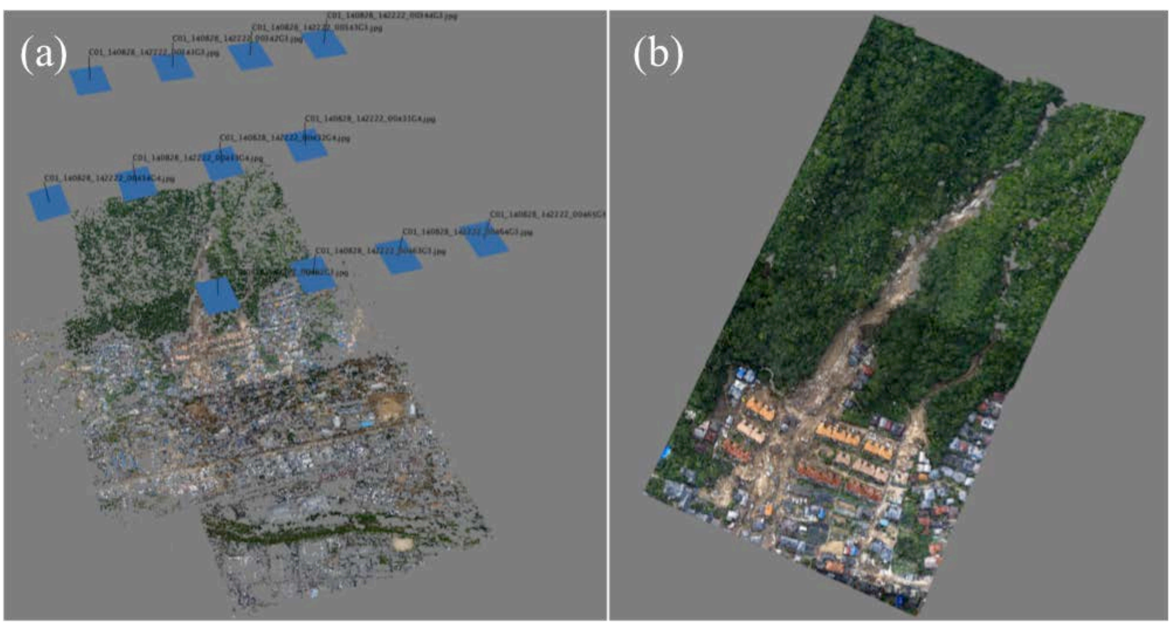

Fig. 1(a) – Estimated camera positions and postures, and the three-dimensional structure of the characteristic points; (b) – Dense three-dimensional model

Using A, R, and T, the camera matrix (P) is reconstructed by the SfM technology. As an example, Fig. 1(a) shows the three-dimensional replication of the estimated camera positions and postures and the common characteristic points, using the aerial photographs of the site of the Hiroshima debris flow disaster as the input images. The figure shows that the replicated three-dimensional points are not dense.

2.2 Three-dimensional model densely reconstructed using MVS

The technique for obtaining a three-dimensional, high-density, and high-accuracy reproduction from photographs, with known camera matrices, is known as multi-view stereo (MVS). First, a depth image is generated for the input image group. The depth is the depth of the pixel of the image measured from the camera. By searching for corresponding pixels among images, using normalized cross-correlation, and defining the correlation values around the pixel-of-focus as a function of depth, the depth of the pixel-of-focus is defined when the function has the maximum value. The accurate depth information of an object is calculated by integrating the depth images [5], and as a result, a dense three-dimensional model is reproduced. Finally, the model is converted to real coordinates using GCPs (Ground Control Points). As an example, Fig. 1(b) shows a dense three-dimensional model for the site of the Hiroshima debris flow disaster.

3. Validation of the Altitude Accuracy of the Three-Dimensional Models Constructed from Aerial Photographs

3.1 Area and data

We attempted to validate the altitude accuracy of the three-dimensional models for the Hiroshima debris flow disaster in 2014. Fig. 2 shows the whole observation area (white frame) and the mudslide area (red line) in Midorii, Asaminami-ku, Hiroshima City, where the damage was the heaviest. Twelve aerial photographs of the disaster area were taken on August 28, 2014 (after the disaster). Fig. 1(b) shows the constructed model (Model 1). The average density is 24 points/m^2. Table 1 shows the image resolution and the replicated point cloud. For accuracy validation, we used the DSM (Digital Surface Model) constructed using the light detection and ranging (LiDAR) observation data (the first pulse) of that day (to be called the LiDAR model in the following sections). The average density of the LiDAR model is 9 points/m^2, and the beam width is 0.19 m. The GCPs were chosen from the coordinates of the LiDAR data for the conversion to real coordinates.

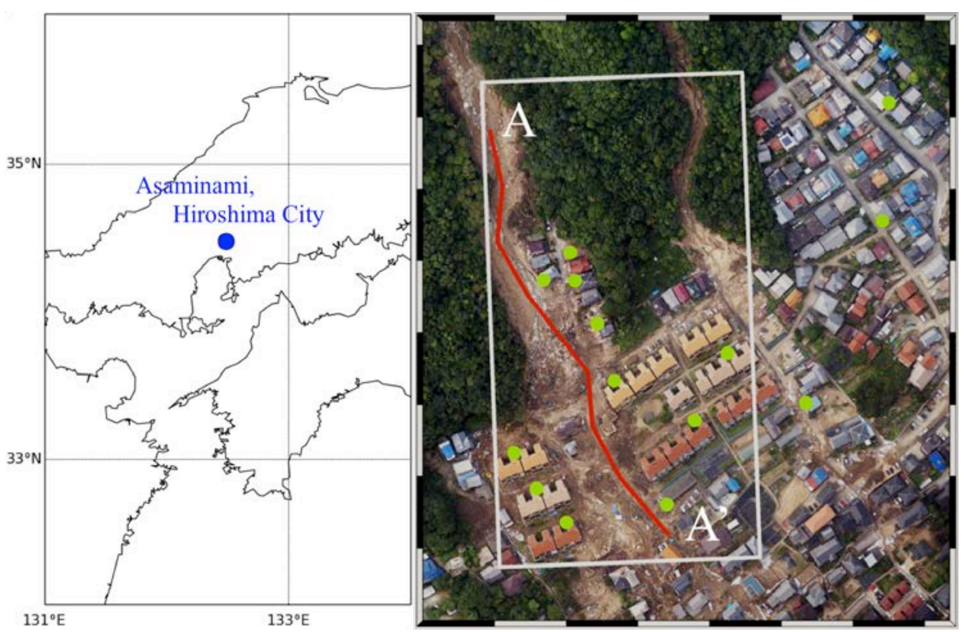

Fig. 2 – The whole observation area (white frame), the mudslide area (red line), and the GCP positions (green points); the background is an orthophotograph of the disaster area

Table 1 – Aerial photographs and the point cloud for the three-dimensional models

Model	Number of pixels	Spatial resolution (m)	Number of points
1	7,212×5,408	0.08	2,334,666
2	5,770×4,326	0.10	1,504,512
3	4,327×3,245	0.14	840,063
4	2,885×2,163	0.21	371,431
5	1,442×1,082	0.47	92,862

3.2 Validation of altitude accuracy

The difference between Model 1 and the LiDAR model is calculated for the validation of the altitude accuracy. Because the models have different spatial resolutions, we searched for the maximum point of Model 1 located within a range of less than 0.1 m (the radius of the beam width) around a point in the LiDAR model to get the altitude difference. Fig. 3 shows the results for the whole observation area. The mudslide area was validated using the point cloud located within a buffer range of 0.1 m from the red line. The RMSE (Root Mean Square Error) of the difference value was 1.75 m in the whole observation area, and 0.17 m in the mudslide area (Model 1 in Table 2). The difference is large in the area with vegetation and around the outer perimeters of the buildings because the altitude accuracy of the replicated three-dimensional model is degraded in areas with large altitude differences.

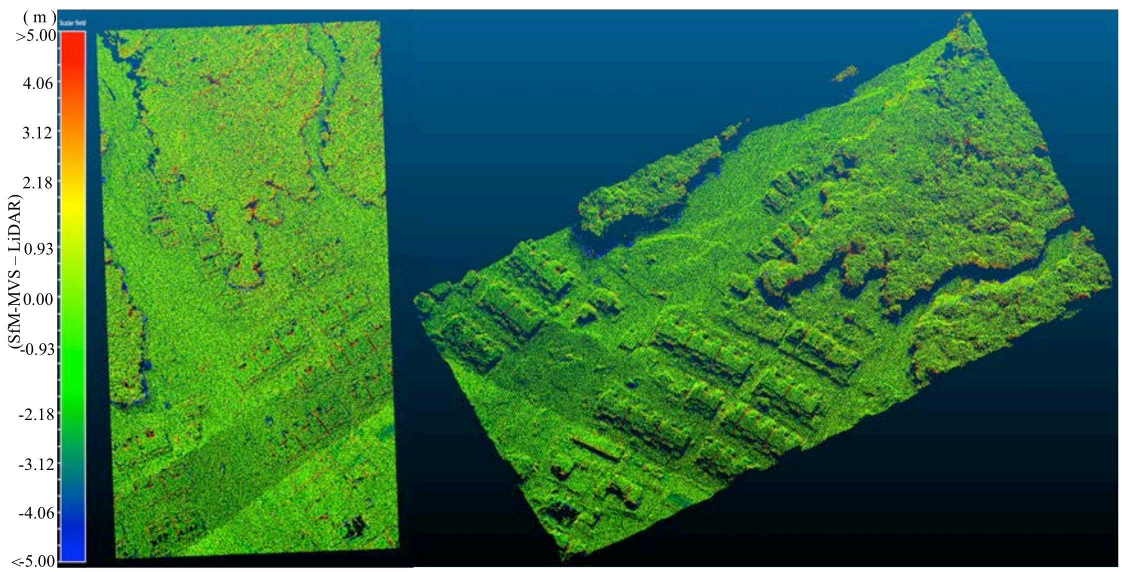

Fig. 3 – Difference in the whole observation area (Model 1 – LiDAR model)

It is known that the accuracy of a model created by the SfM-MVS technology depends, as discussed, on the resolution of the input image data set, and it is important to gain an understanding about this relationship. Therefore, by reducing the resolution of the aerial photographs, the altitude accuracy of the reconstructed three-dimensional model was evaluated to clarify the relationship between the spatial resolution of the image and the altitude accuracy. Four data, with reduced spatial resolutions, were created using the down-sampling process with bi-linear interpolations among the 12 aerial photographs used in Section 3.1. Table 1 shows the image resolution of Models 2 to 5, and the reconstructed point cloud.

The difference between the three-dimensional models created from the data set, and the LiDAR model was then calculated. Table 2 shows the RMSE of the differences in the whole observation area (upper table) and the mudslide area (lower table). It is found that the altitude accuracy of a model in the whole observation area decreases as the spatial resolution of the input photographs is degraded. In the mudslide area, considerable variations were found in the difference values because the abrupt altitude changes caused by the outflow rocks and the rubble could not be coped with effectively.

Table 2 – Altitude accuracy of the three-dimensional models with different levels of spatial resolution

Model		1	2	3	4	5
RMSE (m)	whole observation area	1.75	1.87	1.90	2.01	2.29
	mudslide area	0.17	0.24	0.28	0.43	0.34

4. Extraction of Building Damage Using the Three-Dimensional Models for the 1995 Kobe Earthquake

4.1 Area and data

The building damage caused by the 1995 Kobe earthquake is interpolated and extracted. The area consists of a group of buildings around the JR Uozaki Station in Kobe City. We used six aerial images taken at a vertical angle before the earthquake on May 8, 1994 and 30 photographs taken by a different institution after the earthquake on January 18 and 20, 1995. The two-day difference in the photographing dates after the earthquake is not significant in the disaster area. In addition, to help model the building exterior wall surface, still images were also clipped from a video of the affected area, taken at an obliquely downward angle, from a helicopter.

Table 3 – Sub-classification of building damage by visual interpretation (○: visual interpretation is possible, ×: visual interpretation is impossible)

Class	Classification by visual interpretation		Assessed damage level
	Texture change	Altitude change	
A	○	○	Complete destruction
B	○	×	Complete destruction
C	×	○	Complete destruction
D	×	×	Complete destruction
E	×	×	No damage

For the building damage data, in the verification process, we used the polygon data of the Building Research Institute [6], digitized by the Special Committee on Emergency Restoration of City from Earthquake using the results of assessing the degree of building damage. There are six degrees of building damage—complete destruction, major damage, minor damage, no damage, unexamined, and fire. Because there were clear differences in the damage degrees among the 'completely destructed' buildings in this study, the 'completely destructed' buildings were sub-classified by viewing the aerial photographs. Thereby, the combination of the two criteria—'texture change in the roof' and 'altitude change of the building'—were used for the sub-classification into four types (A to D). In addition, another sub-class—no damage (E) was added. Table 3 shows the results of the sub-classification and the distributions are mapped in Fig. 4. The altitude change of a building was obtained from several photographs of the concerned building.

Fig. 4 – Classification of damage by visual interpretation

Fig. 5(a) – Three-dimensional model with the images taken at a vertical angle before the disaster; (b) – Model with the images taken at a vertical angle after the disaster

4.2 Development of a model by the SfM-MVS technology

In the aerial image taken at a vertical angle, the unnecessary part associated with the scanning of a photograph was removed by mask processing. Since the photographs (with a two-day gap after the earthquake) were taken using different cameras, the three-dimensional model was developed using different focal lengths. The pre- and post-event three-dimensional models are shown in Figs. 5(a) and 5(b). Furthermore, the coordinates of the reference points established by Geospatial Information Authority of Japan (GSI) were used for the conversion to real coordinates. Although taken by different cameras, the two sets of photographs (taken after the earthquake) contain large numbers of images that are characterized by high spatial resolutions; hence, the buildings could be replicated quite well using a larger number of point cloud, when compared to the model before the earthquake.

4.3 Extraction of building damage

For extracting the building damage, for which visual interpretation is somewhat difficult due to altitude changes, the altitude difference of the building was calculated from the models constructed from the images taken with a vertical angle before and after the earthquake. The point cloud in the building polygon was identified as building footprint, and the comparisons between the points, in the images before and after the earthquake, were made. First, as a preprocessing step, using the center of gravity composed of the focus point and its neighboring points, smoothing was done to remove the outliers such as noise from the point cloud on the building roofs with inferior resolutions in the images taken before the earthquake. In addition, the point cloud was extracted in several flat areas such as playgrounds, and the offset was calculated using the most frequent values of altitude for kriging—a spatial interpolation method—to revise the model before the earthquake. The altitude difference between the point cloud on the roofs in the models before and after the earthquake was calculated by subtracting the maximum altitude after the earthquake (within a range of 0.5 m) from the altitude before the earthquake. Fig. 6 shows the difference distribution calculated from the three-dimensional models.

Fig. 6 – Altitude difference before and after the disaster (model before earthquake – model after earthquake)

The difference is found to be large in A and C with altitude changes. Large changes are not seen in B and D before and after the earthquake. Differences were found in the ‘no damage’ buildings E because the model after the earthquake replicated the building well, whereas the replication was insufficient in the model before the earthquake. An approximate regression equation was derived using the least-square method, from the RMSE values (in the lower part of Table 2) calculated for the accuracy verification in Section 3.2, to estimate the errors in the constructed model. The spatial resolution of the photographs before the earthquake was 0.517 m/pix, and the error estimated from the calculated regression equation was 2.361 m. Using this value as a threshold, the building was assessed as ‘damaged’ if at least 20% of the point cloud in the building footprint had differential values exceeding this threshold. The result is shown in Table 4.

Table 4 – Extraction results of building damage

Class	Damaged	No damage	Extraction accuracy (%)
A	95	23	80.5
B	22	35	38.6
C	23	10	69.7
D	23	73	24.0
E	28	123	81.5

The extraction rate was about 80% and a high-accuracy extraction of damage was possible in both ‘A’ (with severe damage) and ‘E’ (with no damage). In addition, the accuracy of extracting the damage in C was as high as 70%. It was shown that, even when the texture variation on the roof was small, the altitude change could be replicated well, as damage. Extraction errors were found to be more in the buildings with fewer reconstructed point groups and greater deflections, and the overall extraction accuracy was about 63%. Even though the model was three-dimensional, the accuracy of damage extraction was not high in ‘D’, where it was difficult to interpret the images taken at a vertical angle visually.

Fig. 7 – Three-dimensional model with the combined images taken at a vertical angle and at an obliquely downward angle

In order to verify the altitude changes and damages of building exterior wall, we re-constructed post-event three-dimensional model by adding oblique photographs taken from the helicopter. In the photographs, a streak-like noise was generated while creating still images. Therefore, to revise the images and construct a model, the noise was reduced by moving every pixel along the corresponding line by referring to the position of the building-wall surface in the image. In the model constructed from the images taken at an obliquely downward angle, there were no reference points within the image; therefore, the GCPs were chosen in flat areas such as playgrounds, and adjusted to the real coordinates using the laser measurement coordinates of the GSI. Then, using the iterative closest point algorithm (a position adjustment method for a point group model) [7], the model constructed from the aerial images taken at a vertical angle was integrated with the model constructed from the images taken at an obliquely downward angle. The results are shown in Fig. 7. It is shown that a three-dimensional model with information on building-wall surface can be constructed by incorporating the images taken at an obliquely downward angle. Using the three-dimensional model included the images taken at an obliquely downward angle (Fig. 7), we found that the visual interpretation of the building exterior wall damages was easy for even damage class ‘D’.

5. Conclusion

In this study, we examined the extraction performance of the three-dimensional building damage models, created by applying the SfM-MVS technology to images taken by a variety of platforms, for a detailed extracting of disaster damage in a wide area. First, by applying the SfM-MVS technology to the aerial photographs of the site of the Hiroshima debris flow disaster in 2014, a three-dimensional model was generated and the altitude accuracy of the model was examined by comparison with the LiDAR data. Then, we clarified the relationship between the spatial resolution and the accuracy of the aerial photographs. Furthermore, a three-dimensional model was created from the images taken at a vertical angle before and after the 1995 Kobe earthquake, and the building damage was extracted from the altitude differences in the buildings measured in these two different periods. As a result, it was found that we could extract the building damage from the altitude change that was otherwise difficult to interpret visually. It was also found that we could get information about the building exterior wall surface from the three-dimensional model constructed using the images taken from a helicopter at an obliquely downward angle. Hence, the three-dimensional information obtained by SfM-MVS technique is effective in assessing the damage.

6. Acknowledgements

We would like to express our gratitude to Dr. K. Horie of InterRisk Research Institute and Consulting, Prof. N. Maki of Kyoto University, and Prof. S. Tanaka of Tokoha University for providing the building damage data. A part of this study was supported in part by the grant-in aid for scientific research (KAKENHI research No. 25282117 and 15H05134). We would like to express our profound appreciation for this support.

7. References

[1] Minato T, Nariyuki Y, Abe M, Sato H. (2008): Fundamental study on automatic extraction method of areas with damaged houses considering direction characteristics of edge lines on aerial photo, *Journal of JSCE*, **17**, 233–242. (in Japanese)

[2] Uchiyama S, Inoue, H, Suzuki, H (2014): Approaches for reconstructing a three-dimensional model by SfM to utilize and apply this model for research on natural disasters, *National Research Institute for Science and Disaster Prevention report*, 81, 37–60. (in Japanese)

[3] Obanawa H, Hayakawa Y, Saito H, Gomez C. (2014): Comparison of DSMs derived from UAV-SfM method and terrestrial laser scanning, *Journal of JSPRS*, **53**, 67–74. (in Japanese)

[4] Lowe D (2004): Distinctive image feature from scale-invariant features, International Conference on Computer Vision, **60**(2), 91–110.

[5] Cambell NDF, Vogiatzis G, Hernndez C. Cipolla R. (2008): Using multiple hypotheses to improve depth-maps for Multi-View Stereo, *ECCV '08 Proceedings of the 10th European Conference on Computer Vision:* Part I, 766–779.

[6] Building Research Institute, Ministry of Construction (1996): *A final survey report for building damages due to the 1995 Hyogo-ken Nanbu Earthquake*, (in Japanese)

[7] Besl PJ, Jain RC (1986): Invariant surface characteristics for 3D object recognition in range images. Computer Vision, Graphics, and Image Processing, **33**(1), 33–80.