Estimation of Scale and Slope Information for Structure from Motion-based 3D Map

Yuyang Shao[1], Yonghoon Ji[1], Hiromitsu Fujii[1], Keiji Nagatani[1,2], Atsushi Yamashita[1], and Hajime Asama[1]

***Abstract*— Dealing with natural disasters, such as debris flow, typically involves teleoperation of an unmanned robot. Three-dimensional (3D) map information of an environment is especially useful for teleoperating the unmanned robot system to perform unmanned construction and auto-investigation. Recently, structure from motion (SfM) using an unmanned aerial vehicle (UAV) equipped with monocular cameras has been proposed as a general method to build 3D map information. However, absolute slope and scale information for the built 3D map cannot be obtained when applying general SfM technology without the deployment of ground control points (GCPs). Therefore, we propose a method that can recover absolute slope and scale information for a 3D map. To this end, we apply sensor fusion technology that combines inertial measurement unit (IMU) data from an unmanned ground vehicle (UGV) and image data from the camera mounted on a UAV. Experimental results show that our estimation scheme is able to successfully represent the slope and scale information along with the 3D map.**

I. Introduction

Recently, the collaboration of unmanned aerial vehicles (UAVs) and unmanned ground vehicles (UGVs) has attracted significant interest in unmanned construction [1]. Typical UGVs have some limitations in complex three-dimensional (3D) terrain investigations, because of ground barriers and obstacles such as cliffs and pits, which are difficult to detect. On the other hand, UAVs can effectively overcome these limitations owing to their small size and freedom of movement. Moreover, UAVs provide more multi-views over wider areas than UGVs; therefore, we can effectively investigate the surrounding environment of UGVs by exploiting UAVs.

Accurate 3D mapping can greatly help understand the environment for unmanned construction. The operator can make an appropriate operation plan according to the 3D map information. Kim *et al*. used a stereo camera system formed by a combination of two UAVs, each equipped with a monocular camera, to generate a 3D structure as shown in Fig. 1 [2]. The objective of this research is to generate a path for a UGV in an outdoor environment including obstacles. Since the obstacles are clearly represented in the generated 3D structure, the UGV can generate safe path. The relative position between the UAVs was calculated through recognition of markers attached to the UGV. UAVs provide environmental information around the UGV. However, this method has a limitation on the UAVs' activity area because the relative position between the UAVs and UGV cannot be calculated if the UGV goes beyond the visual range of the UAVs. To solve this limitation, several methods have been proposed to generate large-scale 3D map information based on simultaneous localization and mapping (SLAM) and structure from motion (SfM). Typical SLAM constructs a map of an unknown environment while simultaneously estimating pose using sensor information with odometry. Thus, the map built by SLAM contains absolute scale information. On the other hand, SfM constructs a 3D structure from 2D image sequences without odometry. Therefore, 3D map information built by SfM does not have an absolute scale, but rather has relative scale information. In other words, SLAM handles the real scale as well as inter-consistency [6, 7, 11],

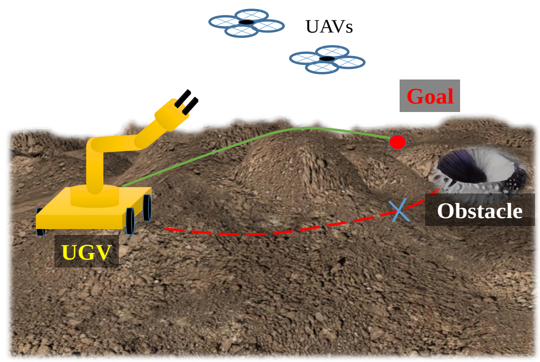


Fig. 1. Path planning of UGV by stereo camera system mounted on UAVs to generate 3D structure of surrounding environment. UGVs can avoid obstacles using information from 3D structure.

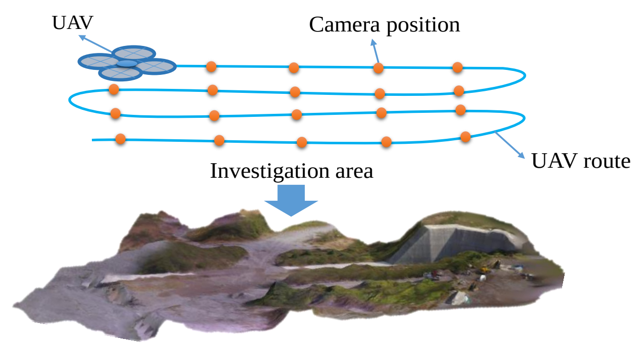


Fig. 2. 3D map generation by UAV. 3D environment can be reconstructed from 2D images taken by UAV. Blue line represents flight route and orange points represent positions of camera mounted on UAV.

*This work was in part supported by the 'Impulsing Paradigm Change through Disruptive Technologies Program (ImPACT)'.

[1]Department of Precision Engineering, The University of Tokyo, Tokyo, Japan. e-mail: {shao, ji, fujii, keiji, yamashita, asama} @robot.t.u-tokyo.ac.jp

[2]New Industry Creation Hatchery Center, Tohoku University, Sendai, Japan. e-mail: keiji@niche.tohoku.ac.jp

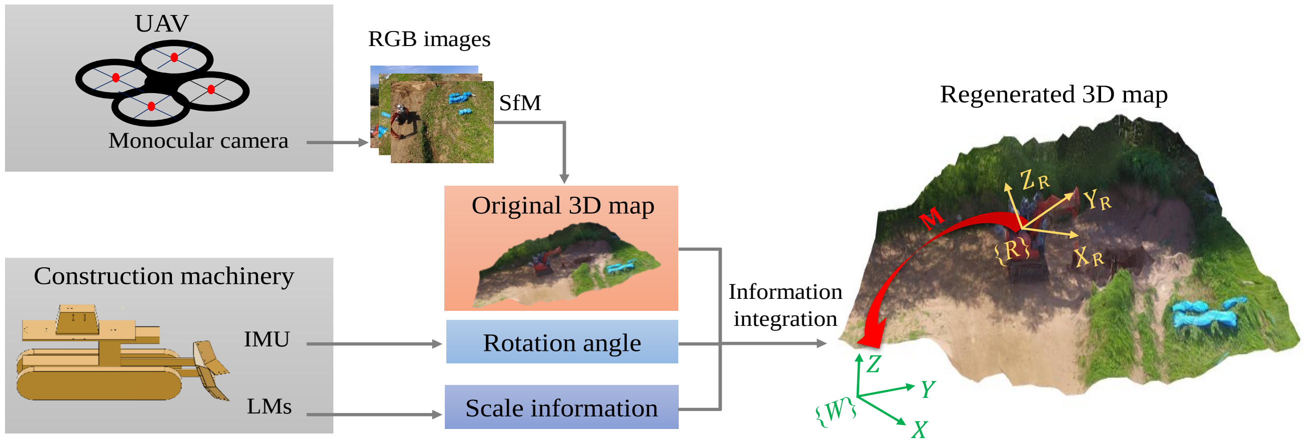


Fig. 3. Overview of proposed method. Firstly, original 3D map that has no absolute slope and scale information is generated from 2D images using SfM algorithm. Then, IMU data and marker information are used to obtain slope and scale information, respectively. Finally, 3D terrain map can be generated by integrating all of this information.

while SfM only considers the inter-consistency of the map [3]. Michael introduced a SLAM method to register the 3D structure computed by a UAV computed by a ground robot operating in a closer range. In this research, the UAV was equipped with a monocular camera while the ground robot relied on a range sensor. The 3D map is an integration of structures computed from the radically different viewpoints of the aerial and ground robots [1]. The 3D map can represent environmental structure on a broad scale, but is difficult to recognize what kind of obstacles are in the environment. SfM can effectively deal with image features in 3D space through low cost monocular camera [4], and provides a clearer 3D map than monocular visual SLAM in general; thus, SfM is a more suitable method for mapping a complex outdoor environment, even if we cannot calculate absolute scale information.

There has been some research based on SfM to generate clear 3D map information of an environment [5, 8, 9]. In Ref. [8], a series of images was acquired while the camera was moving in the scene, as shown in Fig. 2. A 3D environment can be reconstructed from 2D images, and a 3D map can represent an extensive area. However, the 3D map has no slope and scale information of the environment; thus, the 3D map is difficult to be used in operation of and investigation with the UGV [8]. The general method is using GCPs to get scale information, but it is difficult to configure GCPs in an emergency. And UAV-based SfM is not able to reconstruct absolute slope information of the environment, because orientation data cannot be measured by an inertia measurement unit (IMU) mounted on the UAV when it undergoes accelerating motion.

For the purpose of including slope information in a 3D map, Wendel *et al*. [10] proposed to align a 3D reconstruction built by a UAV-based SfM using a digital surface model (DSM). The alignment includes two steps: initial alignment provided by global positioning system (GPS) information and refined alignment computed by evaluating the correlation between a height map computed from the reconstruction and the DSM. This method can construct a 3D map that includes slope information; however, the DSM should be prepared in advance.

In order to overcome these disadvantages of the UAV-based SfM (i.e., no scale and slope information), we propose a novel UAV-based SfM method that combines IMU data from the UGV and image data from the camera mounted on the UAV to generate more useful 3D map information. Note that the 3D map built by SfM can provide color information meanwhile the map built by laser sensor does not include it. Furthermore, our method can also generate 3D map of the area where the UGV cannot reach as long as the UGV is within the UAV's field of view. Furthermore, our method can also generate 3D map of the area where the UGV could not reach as long as the UGV within the UAV field of view. The contribution of this research is as follows. Previous methods in which only relative scale information could be reconstructed have a significant limitation on unmanned construction. We can take using GPS sensor into consideration to reconstruct scale information. However, the GPS sensor has noises; therefore, the GPS-based method cannot estimate the scale information accurately. On the other hand, our UAV-based SfM that collaborates with UGV is able to construct not only absolute scale but also slope information. We combined IMU data from the ground plan and image data from the camera mounted on the UAV to calculate absolute slope and scale information.

The reminder of this paper is organized as follows. Section II provides an overview of our method and how to recover the absolute slope and scale information for a 3D map. In section III, we describe the experiment in a slope terrain. Finally, we present our conclusions in section IV.

II. Estimation of Scale and Slope Information

Figure 3 shows an overview of our method. The system consists of a UAV equipped with a monocular camera and construction machine equipped with an IMU. The UAV takes off from near of the construction machine and captures a set of images of the surrounding environment. The original 3D map of the environment, which has no scale and slope information, is generated by the SfM algorithm from 2D images. The processes required to build a modified 3D map including slope and scale information are as follows.

First, the absolute scale information is recovered by detecting marker information (i.e., location markers (LMs)) which is pasted on the construction machine. Next, slope information is calculated by coordinate transformation based on orientation data recorded by IMU. Finally, the 3D map information is regenerated by using the recovered scale and slope information as follows:

$$\mathbf{x} = \mathbf{M}\, s\, \mathbf{x}_O \tag{1}$$

Here, $\mathbf{x}_O$ and $\mathbf{x}$ denote 3D coordinates for the original 3D map and the regenerated 3D map including scale and slope information. $\mathbf{M}$ and s are transformation parameters which are described in the following subsections in detail.

A. Scale Information

It is difficult to recognize the size of obstacles using the original 3D map without scale information. Therefore, scale information is important for unmanned construction, with respect to safety. The 3D map built by SfM only has relative scale information. The absolute scale of the 3D map can be obtained from the distance between LMs that is measured in advance. Thus, the 3D map including scale information can be generated from the original 3D map information as follows:

$$\mathbf{x}_R = s\, \mathbf{x}_O \tag{2}$$

$$\mathbf{x}_O = \begin{bmatrix} x_O & y_O & z_O \end{bmatrix}^\top \tag{3}$$

$$\mathbf{x}_R = \begin{bmatrix} x_R & y_R & z_R \end{bmatrix}^\top \tag{4}$$

where s denotes a scale factor which can be calculated by comparing between measured distances in the real environment and the original 3D map. $\mathbf{x}_O$ and $\mathbf{x}_R$ respectively represent 3D coordinates for the original 3D map and the 3D map with scale information. Therefore, all points on the original 3D map are converted to the new points on the 3D map with scale information through Eqn. (2).

B. Slope Information

Aside from the scale information, it is also dangerous to perform unmanned construction based on a 3D map that has no slope information of the terrain. Therefore, it is necessary to include slope information in the 3D map. By coordinate transformation, the original 3D map can be modified. In our method, we set two coordination frames to generate slope information: the construction machine coordinate frame $\{R\}$ and the user-defined world coordinate frame $\{W\}$. The 3D map without slope information and the 3D map with slope information are defined based on each coordinate frames (i.e., $\{R\}$ and $\{W\}$). Therefore, the 3D map without slope information can be converted to the 3D map including slope information as follows:

$$\mathbf{x} = \mathbf{M}\, \mathbf{x}_R \tag{5}$$

$$\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix}^\top \tag{6}$$

where $\mathbf{x}_R$ and $\mathbf{x}$ respectively represent 3D coordinates for the 3D map without slope information and the 3D map with slope information. $\mathbf{M}$ is a 4×4 matrix used for coordinate transformation as follows:

$$\mathbf{M} = \begin{bmatrix} \cos\beta\cos\gamma & -\cos\alpha\sin\gamma+\sin\alpha\sin\beta\cos\gamma & \sin\alpha\sin\gamma+\cos\alpha\sin\beta\cos\gamma & t_1 \\ \cos\beta\sin\gamma & \cos\alpha\cos\gamma+\sin\alpha\sin\beta\sin\gamma & -\sin\alpha\cos\gamma+\cos\alpha\sin\beta\sin\gamma & t_2 \\ -\sin\beta & \sin\alpha\cos\beta & \cos\alpha\cos\beta & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

Here, α, β, and γ are rotation angles along the x-axis (roll), y-axis (pitch), and z-axis (yaw), respectively. (t_1, t_2, t_3) denotes a translation vector between the construction machine coordinate frame $\{R\}$ and the world coordinate frame $\{W\}$. The coordinate frames adopted in this study (i.e., the relationship between the construction machinery coordinate frame $\{R\}$ and the world coordinate frame $\{W\}$ in the 3D space) are shown in Fig. 3.

The rotation angles α, β, and γ are measured by the IMU attached to the construction machine. According to IMU data, the slope information of the terrain can be obtained. Through the coordinate transformation described above, all points on the 3D map without slope information can be transformed to the world coordinate frame $\{W\}$ and the transformed 3D map represents absolute slope information of the terrain.

III. Experiment

A. Experiment Design

In order to verify the feasibility and accuracy of our method, we conducted an experiment in a real environment that has sloped terrain and shallow pits, as shown in Fig. 4 (a). An IMU was installed on a construction machine and rotation angles of x-axis (roll), y-axis (pitch), and z-axis (yaw) were measured. The UAV took off from the side of the construction machine. A monocular camera was used to take aerial photos of the ground. An original 3D map was built based on these photos. In the experiment, the UAV took pictures from several points of view from two different heights (5 m and 20 m). Thus, the coverage area was approximately 70 m × 70 m when all the pictures had been processed. Eighty-one images were taken from the UAV-mounted monocular camera to generate a 3D map by the SfM algorithm. The experimental equipment is listed below.

- Construction machine (Fig. 4 (b)).
- UAV (DJI Phantom 4) (Fig. 4 (c)).
- IMU (RT-BT-9axisIMU).

(a)

(b) (c)

Fig. 4. Experiment environment and equipment: (a) sloped terrain environment (from UAV (DJI Phantom 4)) as red area, (b) construction machine, and (c) UAV.

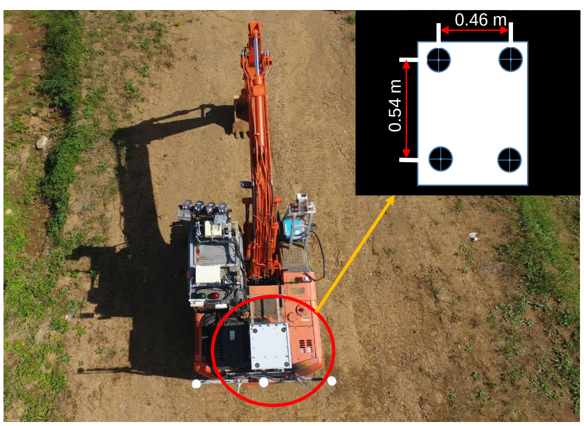


Fig. 5. Position and size of LMs attached to the construction machine. Fore LMs are attached in rectangular arrangement.

- Monocular camera (sensors: 1/2.3"(CMOS), effective pixels: 12.4 M, lens: FOV 94° 20 mm).
- Location markers (Fig. 5).

We set up several gauge points in the environment to evaluate the accuracy of this approach. The ground truths were measured using a laser range finder (BOSCH DLE40) in advance. Therefore, we could evaluate the validity of our proposed method by comparing the truth value with estimated value from the regenerated 3D map.

B. Experimental Results

As previously mentioned, our goal is to obtain an accurate 3D terrain map that includes scale and slope information. There were 81 photos, and memory size of each photo was approximately 5 MB. The original 3D map processing time was approximately 70 min.

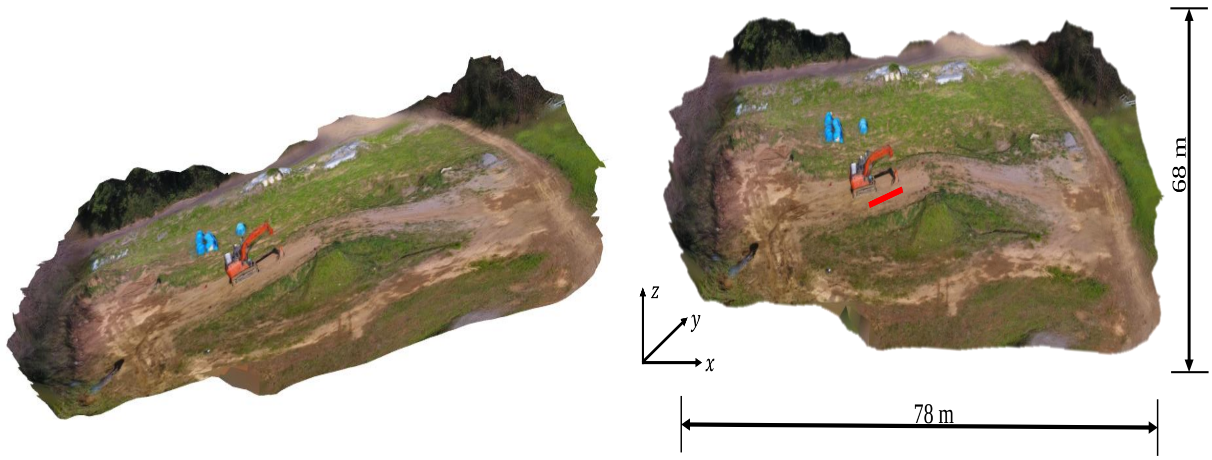
In order to acquire scale information, we used LMs on the construction machine, as shown in Fig. 5. Figure 6 (a) shows the original 3D map of the environment. The construction machine coordinate frame $\{R\}$ was set up using the LMs. Through the coordinate transformation described in Section 2, the 3D map was transformed to the world coordinate frame $\{W\}$. After applying our method based on LMs and IMU data to modify the 3D map information, the modified 3D map information is generated for the output, as shown in Fig. 6 (b). As a result, the regenerated 3D map can correctly represent the terrain.

TABLE I

TRUE DISTANCES AND ESTIMATED DISTANCES FROM GENERATED 3D MAP WITH SCALE INFORMATION

	Distance a	Distance b	Distance c
True value	6.78 m	5.90 m	7.60 m
Estimated value	6.97 m	5.94 m	7.56 m
Error	2.8 %	0.6 %	0.5 %

The distance between any two points in space can be measured by the regenerated 3D map. We measured a road width



(a) (b)

Fig. 6. Original 3D map and regenerated 3D map: (a) original 3D map built by SfM, which has no scale and slope information and (b) regenerated 3D map that has scale and slope information.

(a) (b) (c)

Fig. 7. Evaluated distances based on regenerated 3D map information: (a) distance a, (b) distance b, and (c) distance c.

and the size of a shallow pit in the environment according to the 3D map information, as shown in Fig. 7. Comparing the reconstructed distance from the regenerated 3D map with the ground truth, the results of reconstructed scale are presented in Table 1. The map was found to have an error in the range 0.04–0.19 m (i.e., average error is approximately 0.09 m) for a distance of approximately 7 m. According to the above results, if the construction machine goes forward 50 m, the error is approximately 0.65 m, which is close to the track width of the construction machine. Therefore, the accuracy of the scale is sufficient for a construction machine in outdoor path planning or navigation.

Next, we compared the real slope angle from the real environment with the estimated slope angle from the generated 3D map at the same area which has sloped terrain. The real average slope angle was 7.068 $^\circ$ and it was measured using inclinometer. On the other hand, estimated average slope angle was 7.202 $^\circ$ that was calculated from the 3D map information. The area where the slope evaluation was conducted is represented in dotted red line as shown in Fig. 6 (b). Note that very precise quantitative evaluation may be difficult in case of the evaluation for the slope angle since the slope is not a perfect plane in general; therefore, we need to calculate the average of the values measured at various locations. However, as the above result, we can conclude that the proposed SfM-based mapping method recovered accurate slope information in the 3D map with small error. We can determine whether other slopes in the regenerated 3D map are safe to climb by comparing them to the slope where the construction machine is located.

For these reasons, the regenerated 3D map is sufficient to ensure safety and reliability for unmanned operation of the construction machine.

IV. Conclusions

In this paper, we introduced a method to recover absolute slope and scale information for a 3D map by using a UAV and construction machine. In our system, the UAV was equipped with a monocular camera and the construction machine was equipped with an IMU sensor. We demonstrated that our method can solve the problems of the original 3D map without slope and scale information in different terrains.

The scale errors were sufficiently small as to be ignored when implementing unmanned operation of a construction machine in a dangerous area.

The experimental results showed that the regenerated 3D map built by our method is sufficiently accurate to provide terrain information when performing unmanned operation of a construction machine. However, generating a terrain map of a very large environment would require a long time. Therefore, the proposed method suffers from certain limitations in emergency circumstances. In the future, in order to deal with emergency operation in a dangerous area, the 3D terrain map should be generated in a short time. To this end, the next step for our research will be to develop a method to update 3D map information in a limited time based on pre-existing map information.

ACKNOWLEDGEMENTS

The authors would like to thank the members of Intelligent Construction Systems Laboratory, the University of Tokyo, for useful suggestions, especially Mr. Shingo Yamamoto and Mr. Takumi Chiba from Fujita Corporation and Dr. Kazuhiro Chayama from KOKANKYO Engineering Corporation.

REFERENCES

[1] Michael N, Shen S, Mohta K.: Collaborative mapping of an earthquake-damaged building via ground and aerial robots. Journal of Field Robotics 29(5), 832–841 (2012)

[2] Kim, J. H., Kwon, J. W., Seo, J.: Multi-UAV-based stereo vision system without GPS for ground obstacle mapping to assist path planning of UGV. Electronics Letters 50(20), 1431–1432 (2014)

[3] Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge university press. (2003)

[4] Furukawa, Y., Curless, B., Seitz, S. M., Szeliski, R.: Reconstructing building interiors from images. In: IEEE International Conference on Computer Vision, pp. 80-87 (2009)

[5] Mancini, F., Dubbini, M., Gattelli, M., Stecchi, F., Fabbri, S., Gabbianelli, G.:Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: the structure from motion approach on coastal environments. Remote Sensing 5(12), 6880-6898 (2013)

[6] Castellanos, J. A., Montiel, J. M. M., Neira, J., Tard ⊠ s, J. D.: The SPmap: A probabilistic framework for simultaneous localization and map building. IEEE Transactions on Robotics and Automation 15(5), 948-952 (1999)

[7] Montemerlo, M., Thrun, S., Koller, D., Wegbreit, B.: FastSLAM: A factored solution to the simultaneous localization and mapping problem. In: Aaai/iaai pp. 593-598 (2002)

[8] Sujit, P. B., Hudzietz, B. P., Saripalli, S.: Route planning for angle constrained terrain mapping using an unmanned aerial vehicle. Journal of Intelligent & Robotic Systems 1(11), 273–283 (2013)

[9] Turner, D., Lucieer, A., Watson, C.: An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds. Remote Sensing 4(5), 1392-1410 (2012)

[10] Wendel, A., Irschara, A., Bischof, H.: Automatic alignment of 3D reconstructions using a digital surface model. In: IEEE International Conference on Computer Society, pp. 29-36 (2011)

[11] Zhao, W., Nister, D., Hsu, S.: Alignment of continuous video onto 3D point clouds. IEEE transactions on pattern analysis and machine intelligence 27(8), 1305-1318 (2005)