IPSJ Transactions on Computer Vision and Applications Vol.7 89–93 (July 2015)

[DOI: 10.2197/ipsjtcva.7.89]

Express Paper

Rectification of Aerial 3D Laser Scans via Line-based Registration to Ground Model

Ryoichi Ishikawa[1,a)] Bo Zheng[1,b)] Takeshi Oishi[1,c)] Katsushi Ikeuchi[1,d)]

Received: March 13, 2015, Accepted: April 20, 2015, Released: July 27, 2015

Abstract: The aerial 3D laser scanner is needed for scanning the areas that cannot be observed from the ground. Since the laser scanning takes time, the obtained range data is distorted due to the sensor motion while scanning. This paper presents a rectification method for the distorted range data by aligning each scan line to the 3D data obtained from the ground. To avoid the instability and ambiguity of the line-based alignment, the parameters to be optimized are selected alternately, and the smoothness constraint is introduced by assuming that the sensor motion is smooth. The experimental results show that the proposed method has the good accuracy simulation and actual data.

Keywords: registration, smoothness constraint, 3D digital archive

1. Introduction

Recently, many large-scale cultural heritage assets have been urgently desired to be preserved because the deterioration states have become getting worse and worse by natural and man-made breaks. 3D laser scanning is able to play an important role in the preservation of the cultural heritage assets [1], the large structures over a hundred meters can be digitized in the accuracy of a few millimeters with the widely available laser range sensors. The digital 3D models obtained in this way can be used for the analysis and restoration of the structures.

The scanning only from the ground is not enough due to the large-scale structure of the target objects such as large buildings. It often happens that the surface on the top of the building cannot be observed from any location. **Figure 1** (a) shows an example of the range data taken from the ground, where the data missing appears on the top of the tower.

To obtain the complete 3D model of such a large structure, the aerial laser scanning system is required. The aerial laser scanning systems with inertial sensors or video cameras or both have been previously proposed. Generally, the aerial scans are distorted due to the sensor movement in the air (See Fig. 1 (b)). The sensors are used to rectify the distortion of the range data by estimating the motion of the sensor. However, there are limitations on the motion estimation accuracy by such sensors and cameras.

On the other hand, there must be the scans from the ground overlapped with the scans from the aerial scanning system. Figure 1 (b) shows an aerial scan that overlaps with the ground scan shown in Fig. 1 (a). We do not need to take inertial sensors into account to estimate the motions. Thus, we face the challenge of eliminating the distortion of the aerial scan by using the undistorted scans taken from the ground. Figure 1 (c) shows the rectified range data of Fig. 1 (b). Figure 1 (d) shows the merged result of (c) to (a), where the missing areas are filled with the aerial data.

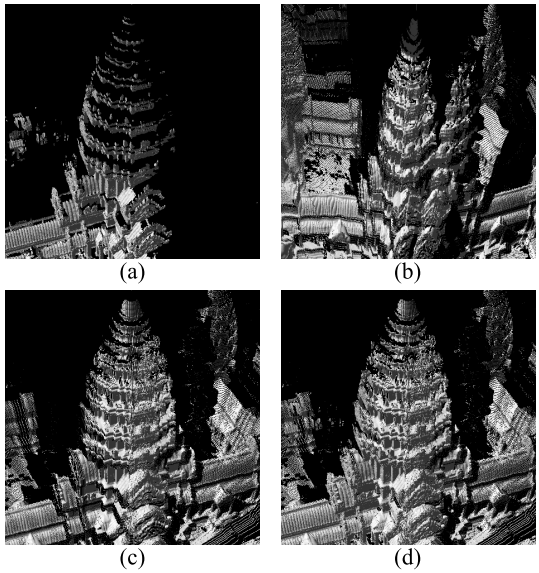

Fig. 1 Combination of the 3D data from the ground and aerial range data. (a) 3D data from the ground. Missing surfaces exist. (b) Distorted aerial range data. (c) rectified aerial range data. (d) 3D data from the ground and rectified aerial range data. Rectified aerial range data complement the missing surfaces in the data from the ground.

In this paper, we propose a method that rectifies the distorted range data based on a line-based alignment algorithm. We assume that range data consists of 3D scan lines obtained sequentially by the laser scanner. The proposed method aims to estimate the motion of each scan line by aligning it with 3D scans taken from the ground. To avoid the instability and ambiguity of the line-based alignment, the motion parameters are estimated based on a nonlinear alternative optimization. The smoothness constraint is also

[1] Institute of Industrial Science, The University of Tokyo, Meguro, Tokyo 153–8505, Japan
[a)] ishikawa@cvl.iis.u-tokyo.ac.jp
[b)] zheng@cvl.iis.u-tokyo.ac.jp
[c)] oishi@cvl.iis.u-tokyo.ac.jp
[d)] ki@cvl.iis.u-tokyo.ac.jp

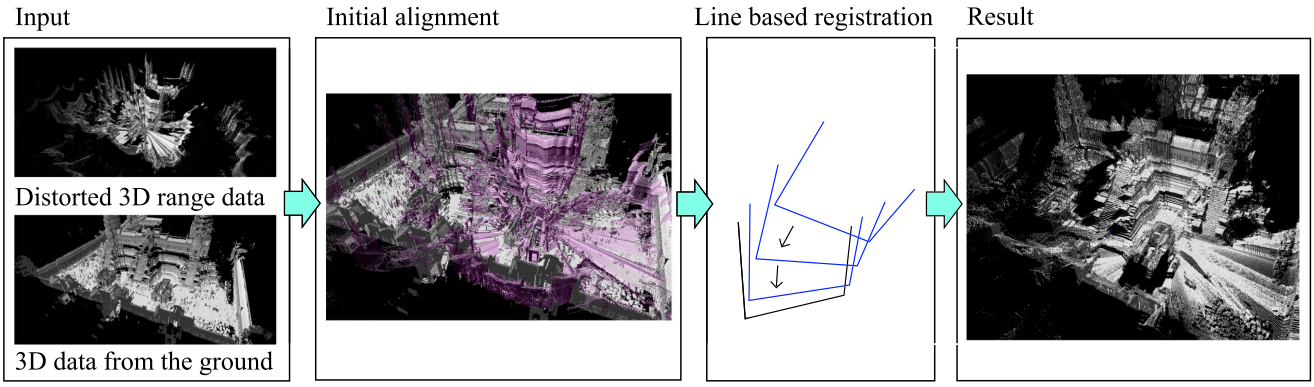

Fig. 2 Method overview. Inputs are aerial data and ground data. First, the initial position of aerial data is given by rigid alignment. Second, aerial data is processed to line based registration method and finally we obtain rectified result.

introduced by the assumption that the sensor motion is smooth because the sensor is mounted on the balloon. The experimental results show that the proposed method has good accuracy.

2. Related Work

Our work is related to four research streams in the literature.

Structure from Motion (SfM). To recover the sensor motion, SfM-based methods using 2D cameras are popular to use, since the relative pose and position of each frame can be calculated out by using 2D image features [2]. For example, a famous and robust implementation is proposed in Ref. [4]. However, in the case of digital archive, the absolute scale of the range data, which is not easy to get for SfM, is required.

SLAM++. SLAM++ [7] is proposed to robustly estimate the sensor motions using depth sensor, suppose 3D priors are observed online and used to improve the localization. This method is very useful for robotics or AR applications. However, our method focuses on the high-accuracy 3D line reconstruction. It faces more challenging problem as how to match a 3D line to a 3D prior.

Non-rigid registration. Many methods (e.g., Refs. [3], [6], [10], [11], [12]) have been proposed in the field of non-rigid registration, which geometrically solve the matching problem between the deformed models. However, in this paper, we take the temporal information into account of the deformation (rectified from sensor motion) which is more reasonable in laser scanner case. Another notable work is proposed in Ref. [9] which registers 2D curves to a 3D shape encoded by an implicit polynomial and shows efficiency for medical image processing.

Aerial Laser Scanning. The method most related to our work is the method proposed by Banno et al. [5], which combines video cameras and range sensor into the scanning system and SfM techniques are used in the motion estimation algorithm. In this paper, we only consider the information obtained from range sensor. Thus our method is independent to any other third-party sensors for saving hardware cost, and also it avoids the calibration problems which often happened in sensor fusion systems.

3. Method

3.1 Overview

The proposed method is based on three assumptions: 1) distortion of one scan line can be ignored because one line scanning time is short (approximate 0.05 sec); 2) 3D data measured from on the ground is given in advance. This assumption is reasonable because the purpose of this research is to complement the part of data that cannot be measured from the ground; 3) sensor motion is smooth. On these assumptions, position parameters are set to each scan lines and registration is processed. To avoid the instability and ambiguity of the line-based alignment, smoothness constraint is imposed on between scan lines.

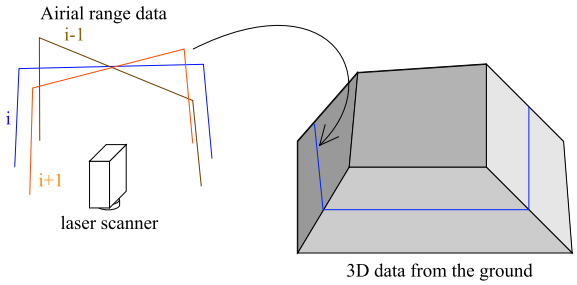

Fig. 3 Mosion model of scan line.

Figure 2 shows the overview of our method. Aerial scan range data and 3D data from the ground are inputs. First, the initial position of aerial data is given by rigid alignment [8]. Second, aerial data is processed to line based registration method and finally we obtain rectified result.

3.2 Line based registration

3.2.1 Initial alignment

Distorted range data is aligned to 3D data from the ground for obtaining initial position. This phase is rigid registration. We used fast alignment method presented in Ref. [8] in this phase. This method is fast because of searching closest point along to the view direction.

3.2.2 Motion Model of Scan Line

Figure 3 shows a motion model of scan line. Each line is located on the local coordinates of the range sensor at first. The sensor motion is not taken into consideration about raw data, therefore each line is deviating each other as moving amount of sensor. Each line is given 6 parameters indicating sensor position, translation parameters $\mathbf{T} = (x, y, z)$ and rotation parameters $\mathbf{r} = (\theta, \psi, \phi)$ (x, y ,z axis rotation). The i-th line is firstly trans-

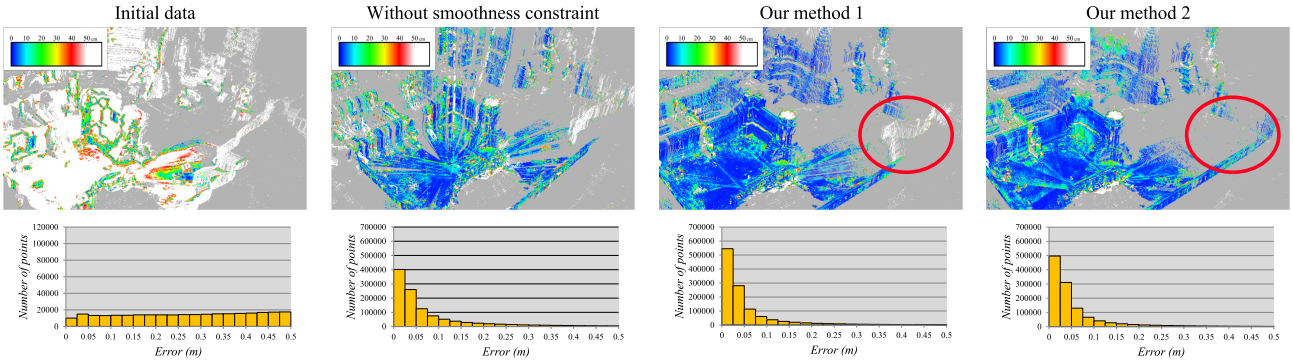

Fig. 4 Estimated motion of the sensor. First row: Error visualization images. Second row: Histogram of error. Our method 1: Optimizing 6 parameters. Our method 2: Optimizing selected 3 parameters alternately.

formed by rotation matrix $\mathbf{R_i}$ and translation vector $\mathbf{t_i}$ calculated from 6 parameters of i-th line, (x_i, y_i, z_i) and $(\phi_i, \theta_i, \psi_i)$, and second, is transformed by $\mathbf{R}$, $\mathbf{t}$ obtained by initial alignment.

3.3 Optimization

3.3.1 Cost Function

Our cost function consists of 2 terms: 1) the distance error term and 2) the smoothness constraint term. The distance error term is based on ICP algorithm. The distance error of i-th line E_{di} is as follow

$$E_{di} = \sum_j a^2 \rho\left(\frac{s_j}{a^2}\right), \tag{1}$$

where $\rho(x) = log(1 + x)$ is the loss function and a is a scale value,

$$s_j = |\mathbf{p'_{ij}} - (\mathbf{R}(\mathbf{R_i}\mathbf{p_{ij}} + \mathbf{t}_i) + \mathbf{t})|^2. \tag{2}$$

Here, $\mathbf{p_{ij}}$ is j-th point of i-th line, and $\mathbf{p'_{ij}}$ is closest point of transformed point of $\mathbf{p_{ij}}$ by $\mathbf{R}$, $\mathbf{t}$, $\mathbf{R_i}$, $\mathbf{t_i}$ in the data from the ground.

The smoothness constraint term E_{smooth} is defined as follow,

$$E_{smooth} = \int \left(\lambda_1 |\frac{d^2\mathbf{T}}{dt^2}|^2 + \lambda_2 |\frac{d^2\mathbf{r}}{dt^2}|^2\right) dt, \tag{3}$$

This equation is based on the assumption that sensor's acceleration fluctuates smoothly.

Consequently, Global cost function is

$$E = \frac{1}{N}\left(\sum_i E_{di}\right) + wE_{smooth}, \tag{4}$$

where N is the number of point, w is a coefficient determined experimentally.

3.3.2 Decremental Smoothness Constraint

Smoothness constraint is introduced for reducing instability and ambiguity. However, Eq. (3) indicates that the ideal is that variation of sensor's acceleration is *0*. Practical variation of sensor's acceleration is not *0* and a paradox exists between practical sensor movement and Eq. (3). Therefore coefficient w in Eq. (4) is gradually decreased.

3.3.3 Parameters Selection

Variables to be optimized are 6 parameters of each line position. However, in some phases, only 3 parameters are optimized simultaneously and parameters to be optimized are selected alternatively because the lower the stability and the ambiguity of the line-based alignment. (x, y, ϕ) are initially selected as parameters to be optimized because these parameters are tend to vary wider than other 3 parameters. This tendency proceeds from balloon movement.

Therefore it is reasonable that $3 \times$ *number of lines* parameters are optimized globally and simulteniously at once to minimize the cost function, and parameters to be optimized are selected alternately. This scheme is especially effective in case that ambiguity of scan line is large.

As a non-linear least square solution of this optimization problem, Levenberg-Marquardt Method is selected.

4. Experimental Results

In this experiment, we process 2 methods; 1) Optimizing 6 parameters simultaneously and 2) Optimizing selected 3 parameters alternately. In each method, the following steps are processed and each step finishes when cost function is convergence or iteration times reached the specified value. Optimizing 6 parameters of each lines simultaneously: 1. w is 3.64. 2. w is 0.0364. 3. w is 0.000364. 4. no smoothness constraint. Optimizing selected 3 parameters alternately: 1. (x, y, ϕ) are optimized, w is 3.64. 2. (z, θ, ψ) are optimized, w is 3.64. 3. (x, y, ϕ) are optimized, w is 0.0364. 4. (z, θ, ψ) are optimized, w is 0.0364. 5. (x, y, ϕ) are optimized, w is 0.000364. 6. (x, y, ϕ) are optimized, no smoothness constraint.

4.1 Simulation Data

In this experiment, we use two range data from the ground: one is used as a reference data and another is used as a distorted data. The distorted data is made by transforming each scan line with preliminary determined parameters. These parameters are manually set to imitate balloon motion. Therefore we can evaluate rectified data by comparing the estimated parameters and the preliminary determined parameters as true values.

Results are showed in **Fig. 7** and **Fig. 8**. Graphs in the first column show true values of line positions and graphs in the second and third columns show error of the estimated motion defined as (*True value*) − (*Estimated value*). Therefore these graph indicate that the closer to 0 error is, the more acculate estimated parameter is.

From Fig. 8, it is possible to say that our methods succeed in estimating true positions for the most part. Comparing our method

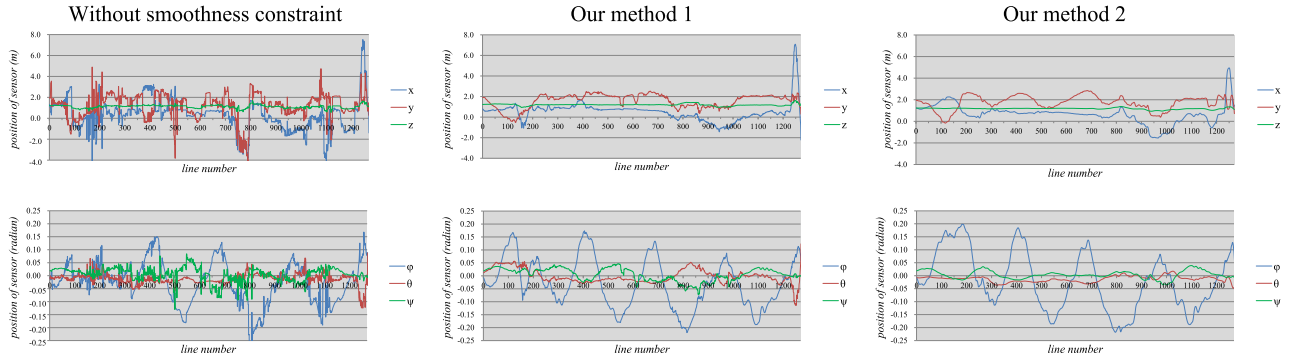

Fig. 5 Estimated motion of the sensor.

Fig. 6 Other results of rectification. From left to right, ground data, distorted data, rectified data, error visualization.

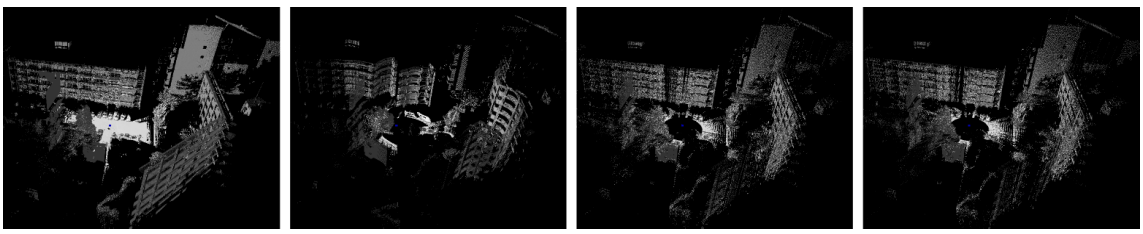

Fig. 7 3D image of simulation data. (a) reference data, (b) distorted data, (c) our method 1, (d) our method 2.

1 and our method 2, there is no large tendency that where lines which have large error is located. The results of method 1 are seemingly more accurate, however processing time of method 2 tend to be shorter because the number of parameter to be optimized simultaneously is half as large as the method 2, therefore optimization of selected 3 parameters alternately has the room for practical use.

4.2 Actual Aerial Data

As evaluation methods in this section, error visualized 3D images and histogram of errors are selected. The error is defined by Euclidean distance between a point in aerial data and its closest point in 3D data from the ground.

Figure 4 shows comparison results. In the error visualization image, each point is colored depending on its error (See a color bar at the top left of error visualization image). Blue point has small error and white point is large error point because of wrong estimation or missing ground data. Our method achieved the aerial data rectification with less error than without smoothness constraint. Bottom row of **Fig. 5** shows the position of line.

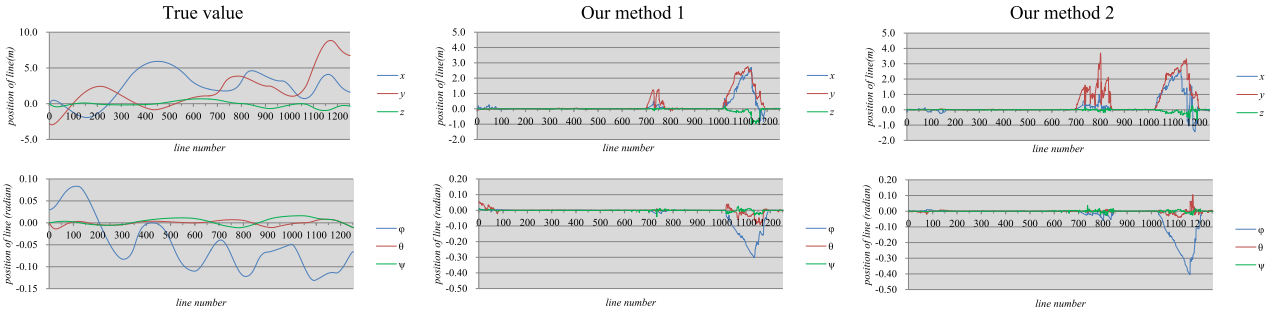

Fig. 8 Estimated position of the line.

Each line is scanned in time order, therefore this graphs indicate estimated crude sensor motions. From Fig. 5, in case of imposing smoothness constraint, estimated result is more smooth than without constraint.

Comparing to results by method 1 (optimizing 6 position parameters simultaneously) and method 2 (selected 3 parameters alternatively), we will see that the method 2 succeed to estimate position of lines which method 1 cannot rectified. (In red circle in Fig. 4). However, the result of method 1 has more points whose error is small than the result of method 2 (See the histograms in Fig. 4). Therefore, it may be possible to achieve more acculate result by finding appropriate combination of steps.

Figure 6 shows results of other data. The results in first 3 rows are successful cases. The result in the bottom row is comparable failure case. In this case, z axis rotation movement of sensor is comparably drastic and some points in aerial range data, in particular far points from sensor, are located on far coordinates from correct place initially.

5. Conclusion

With smoothness constraint considering time domain, our line-based alignment method was dramatically improved and achieved well-rectified results nevertheless only geometrical information is used.

Accuracy of rectified results is depends on 2 elements: one is shape characteristic and another is a magnitude of the distortion.

Shape characteristic is important information to reduce ambiguity and to determine position of line. Therefore uneven object, like Angkor-Wat, is compatible with our method compared to flat object.

A magnitude of the distortion is intensity of sensor motion during measurement and it depends on the intensity of the wind. Therefore it is of course ideal to do measurement under weak wind conditions. However, a lot of measurements are needed to archive a large cultural structure, hence it is difficult to always obtain data measured under the ideal conditions from a cost and schedule perspective. Thus it is needed to rectify even range data with large distortion measured under the strong wind conditions.

Our method shows good performance to an aerial data measured under the calm wind conditions. However when distortion of aerial data is drastic, our method cannot shows good performance.

For obtaining more accurate results,We are thinking of implementing another scheme involving sensor motion, or use other data, such as reflectance information.

Acknowledgments This work is supported in part by Next-generation Energies for Tohoku Recovery (NET), MEXT, Japan.

References

[1] Ikeuchi, K., Hasegawa, K., Nakazawa, A., Takamatsu, J., Oishi, T. and Masuda, T.: Bayon digital archival project, *Proc. 10th International Conference on Virtual Systems and Multimedia*, pp.334–343 (2004).

[2] Carlo, T. and Takeo, K.: Shape and motion from image streams under orthography: A factorization method, *International Journal of Computer Vision*, pp.137–154 (1992).

[3] Fujiwara, K., Nishino, K., Takamatsu, J., Zheng, B. and Ikeuchi, K.: Locally Rigid Globally Non-rigid Surface Registration, *Proc. IEEE International Conference on Computer Vision* (*ICCV*), pp.1527–1534 (2011).

[4] Furukawa, Y. and Ponce, J.: Accurate, Dense, and Robust Multiview Stereopsis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.32, No.8, pp.1362–1376 (2010).

[5] Banno, A., Masuda, T., Oishi, T. and Ikeuchi, K.: Flying laser range sensor for large-scale site-modeling and its applications in Bayon digital archival project, *International Journal of Computer Vision*, pp.207–222 (2008).

[6] Masuda, T., Hirota, Y., Nishino, K. and Ikeuchi, K.: Simultaneous determination of registration and deformation parameters among 3D range images, *Proc. 5th International Conference on 3-D Digital Imaging and Modeling* (*3DIM2005*), pp.369–376 (2005).

[7] Salas-Moreno, R.F., Newcombe, R.A., Strasdat, H., Kelly, P.H.J. and Davison, A.J.: SLAM++: Simultaneous Localisation and Mapping at the Level of Objects, *Proc. Computer Vision and Pattern Recognition* (*CVPR*), pp.1352–1359 (2013).

[8] Oishi, T., Nakazawa, A., Kurazume, R. and Ikeuchi, K.: Fast Simultaneous Alignment of Multiple Range Images using Index Images, *Proc. 5th International Conference on 3-D Digital Imaging and Modeling* (*3DIM2005*), pp.476–483 (2005).

[9] Zheng, B., Ishikawa, R., Oishi, T., Takamatsu, J. and Ikeuchi, K.: A Fast Registration Method Using IP and Its Application to Ultrasound Image Registration, *IPSJ Trans. Computer Vision and Applications*, Vol.1, pp.209–219 (2009).

[10] Brown, B. and Rusinkiewicz, S.: Gloabal Non-Rigid Alignment of 3-D Scans, *ACM Transaction on Graphics* (*Proc. SIGGRAPH*), Vol.26, No.3 (2007).

[11] Sagawa, R., Akasaka, K., Yagi, Y. and Hamer, H.: Elastic Convolved ICP for the registration of deformable objects, *IEEE International Conference on Computer Vision Workshops* (*ICCV Workshops*), pp.1558–1565 (2009).

[12] Zhou, Q., Miller, S. and Koltun, V.: Elastic Fragments for Dense Scene Reconstruction, *IEEE International Conference on Computer Vision* (*ICCV*), pp.473–480 (2013).

(Communicated by *Akihiko Torii*)