ROBUSTLY REGISTERING A NETWORK OF RANGE IMAGES OF URBAN OBJECTS

Huijing ZHAO , Ryosuke SHIBASAKI
Center for Spatial Information Science, Univ. of Tokyo, JAPAN
{chou,shiba}@skl.iis.u-tokyo.ac.jp

KEY WORDS: laser range image, registration, urban 3D object.

ABSTRACT

There is a potentially strong demand for detailed 3D spatial data of urban area. Ground-based laser range scanner is one of the promising devices to acquire range images of urban 3D objects. In this paper, the authors propose a method of automatically registering a network of range images into a well-balanced model for the reconstruction of 3D urban objects. Assuming relative transformation between neighboring range images have been estimated using a pair-wise registration method, multiple registration is achieved in two steps, sequentially aligning all range images into a global coordinate system and minimizing violation, which is caused by the accumulation of pair-wise registration error. Experimental results are presented, where 42 range images are registered to construct a 3D model of the buildings in the campus of University of Tokyo. Two sets of ground truth, a 1:500 scale digital map and the location of viewpoints measured by GPS, are used to examine the accuracy in multiple registration. Experiments are conducted in two levels. The first experiment examines registration accuracy when location and viewing direction of range images are unknown, while in the second experiment, assuming that several viewpoints has been measured using GPS, improvement to registration accuracy using GPS points is studied. Through the experiments, it is demonstrated that the method has accuracy in urban outdoor environment.

1 INTRODUCTION

In a variety of applications ranging from visualization of urban landscape to advanced automobile/pedestrian guidance systems for ITS (intelligent transportation system), accurate and detailed 3D urban spatial database is becoming strongly demanded. There are two approaches of 3D data acquisition; air-based and ground-based. With air-based data acquisition techniques, typically, aerial survey, can cover relatively wide area, but may fail to capture details of urban objects such as side walls (facade) of buildings. On the other hand, ground-based method such as vehicle-borne CCD camera can easily cover such details of urban objects, though the spatial coverage may be limited. Recently, reconstructing 3D urban objects using ground-based acquisition techniques are attracting more attentions because such details of urban objects which can be easily viewed from streets or on ground surface are found important for guidance system and radio disturbance analysis in telecommunication and so forth. In many applications of 3D GIS, viewpoints of users are on the street or ground surfaces, not in the air.

Several researches using ground-based CCD camera have demonstrated that 3D urban object can be extracted using motion and stereo vision technique (e.g. S.Ozawa et.al.1998). However, insufficient robustness in stereo matching, distortion from limited resolution and unstable geometry of CCD cameras are major obstacles to the operational uses of these method. On the other hand, range scanners are also used for 3D object modeling. Efficiency and accuracy of range scanner serving for spatial data acquisition has been demonstrated in Thorpe et.al.1988 in a mobile navigation system; Kamgar-Parsi et.al.1991 in obtaining a map of ocean floor; Chen and Medioni 1992, Champleboux et.al.1992, Shum et.al.1994 in modeling small objects such as teeth, status, mechanical parts, etc.; Ng et.al.1998 in constructing 3D models of indoor objects; Lemmens et al.1997, Haala et al.1998 in air-borne systems. In recent years, with the development of eye-safe laser range scanner, ground-based measurement using laser range scanner in urban environment becomes technically feasible. However, there is still no other research addressed on the 3D urban outdoor object modeling using ground-based laser range scanner. Since one snapshot can not cover the entire 3D object, one of the major challenges in applying ground-based laser range scanner to the reconstruction of 3D urban objects is the registration of multiple overlapping range images acquired at different locations (viewpoint) with different viewing angles. Registering multiple range images - i.e., correctly aligning range images by transforming them into a global (common) coordinate system - is typically solved as a two-step procedure, pair-wise registration and multiple registration. In pair-wise registration, neighboring range images, which has a degree of overlay, are registered to find the relative transformation between the coordinate system of two range images. In multiple registration, absolute

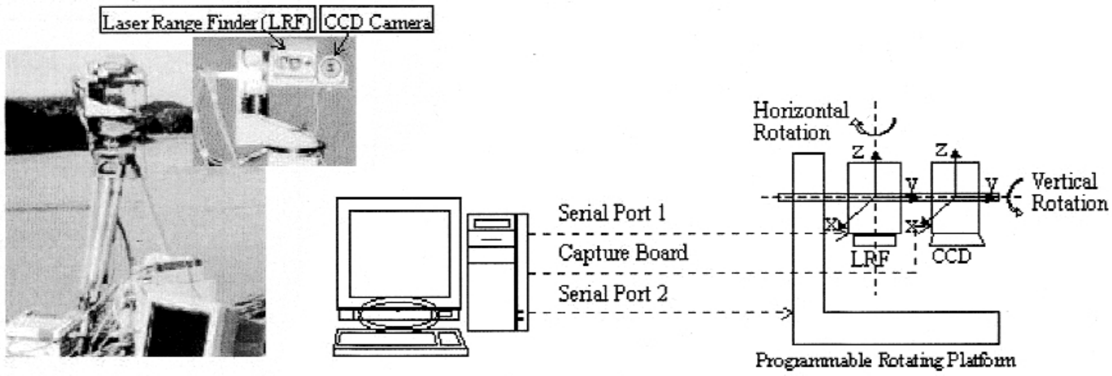


Figure 1: Architecture of sensor system.

transformations aligning all range images to a global coordinate system is obtained, where the accumulation of estimation error in pair-wise registration has to be solved.

A picture of the sensor system used in the research is shown in Figure 1. We assume the sensor's platform is set always horizontal to the ground surface, which is easy to be satisfied in urban survey. A registration with four degree of freedoms is studied in this research, where a horizontal rotation angle and three translation parameters from each range frame to a global coordinate system has to be estimated. A pair-wise registration method using "Z-image" has been developed in our previous study (Zhao and Shibasaki, 1999). In this research, we assume all the relative transformation parameters between neighboring range image pairs (two range images with a degree of overlay) have been obtained through the pair-wise registration. This paper contributes to a multiple registration method, where all range images are aligned to a global coordinate system, and simultaneously registered to achieve a well-balanced model. An experiment is conducted, 42 range images measured in the campus of the Univ. of Tokyo are registered. A sequence of examination is conducted to test the accuracy of the method. Two sets of ground truth are used to examine the registration accuracy. They are 1) a 1/500 scale digital map and, 2) location of viewpoints measured by GPS with an accuracy of ± 20cm.

2 REGISTRATION OF MULTIPLE OVERLAPPING RANGE IMAGES

2.1 Study review of multiple registration methods

A number of different methods have been developed to solve the error accumulation problem in multiple registration. Y.Chen and G.Medioni, 1992 partially solved the problem by registering the newly introduced range image with the integrated model consisting of all previously registered range images. Some of the researches simultaneously registered all range images using weight least square methods. R.Bergevin et al. 1996 equally distributed those accumulated registration errors by minimizing the squared sum of distance from control points to the corresponding tangent planes in other range frames. G.Blais and M.D.Levine 1995 evaluated the displacements between range images by the distance from a set of control points to their corresponding ones in other frames. However disadvantage of the methods is not only the accuracy of extracted control points might be significantly influenced by range error and range uncertainties, but also searching for global optimum can be very computational heavy. Shum et al. 1994 formulated the multiple registration as a problem of principal

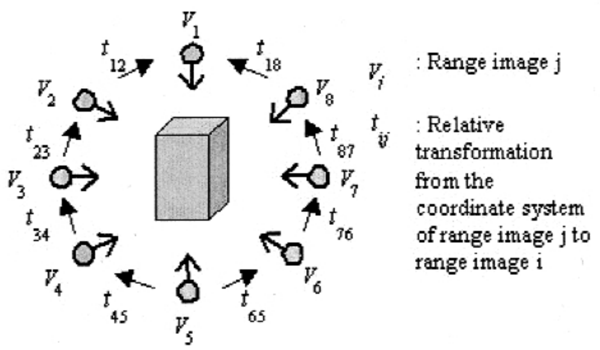


Figure 2: An example of range measurement from each facade of the building.

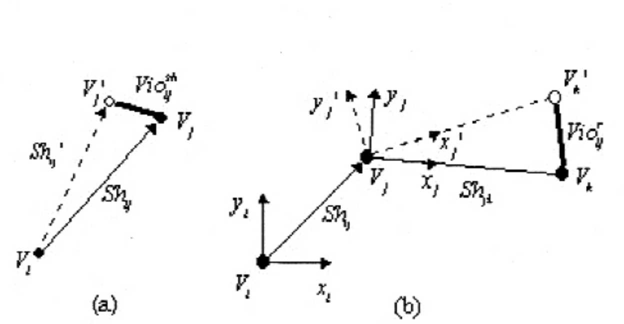


Figure 3: Evaluating the violation in shift vector and rotation matrix.

component analysis with missing data (PCAMD), where objective function is composed using the distance between corresponding planar faces in different range frames. However, enough number of planar faces is difficult to extract in urban area as has been addressed in Zhao and Shibasaki, 1999, and since contributions from all the planar faces are treated equivalently, multiple registration might be degraded by several unreliable planar faces. In this research, we follows the formulism in B.Kamgar-Parsi et al. 1991, where assuming that pair-wise registration yields locally optimized transformation between neighboring range image, multiple registration is formulated as a minimization of violations to local matching achieved in pair-wise registration. It is proved to be the most appropriate starting point in our effort to achieve a well-balanced network of range image.

2.2 Problem statement and outline of the method

To simplify the problem, we use homogeneous coordinates. Let $t_{ij} = \begin{pmatrix} R_{ij} & Sh_{ij} \\ 0 & 1 \end{pmatrix}$ be the relative transformation from the coordinate system of range image V_i to a neighboring range image V_j, where R_{ij} and Sh_{ij} are the rotation matrix and shift vector. Let $p_k = [x_{pk}, y_{pk}, z_{pk}, 1]^T$, $k = i, j$ be range points in V_i and V_j respectively. Then it has $p_i = t_{ij}p_j$. Suppose V_i has been registered to the global coordinate system with a transformation matrix T_i, the transformation matrix aligning V_j to the global coordinate system can be calculated by $T_i t_{ij}$. This kind of registration is called "sequential registration".

Figure 2 shows a motivational example of measuring a building using a network of range images $\{V_1, V_2, ..., V_8\}$. Suppose all range images are to be aligned to the coordinate system of V_1. V_6 can find its transformation matrix by

$$T_6 = t_{18}t_{87}t_{76} \tag{1}$$

We define "V_1-V_8-V_7" as the "registration path" of V_6. The number of range images in registration path is defined as the "length" of registration path. It is easy to show that pair-wise registration error is accumulated and propagated along the registration path in sequential registration. Although aligning all range images using the shorted registration path can somehow mitigate the error accumulation problem, inconsistencies happen if range images compose a looped network. In the above motivational example, V_5 has two optional registration paths. They are "V_1-V_2-V_3-V_4" and "V_1-V_8-V_7-V_6". Suppose the previous path is selected, then the transformation matrix aligning V_5 to the coordinate system of V_1 can be calculated as

$$T_5 = t_{12}t_{23}t_{34}t_{45} \tag{2}$$

V_5 and V_6 is said to be "independent in registration", since neither V_5 nor V_6 appears in the other's registration path. According to Eq.1 and Eq.2, it has $t'_{56} = T_5^{-1}T_6 = t_{54}t_{43}t_{32}t_{21}t_{18}t_{87}t_{76}$. In real situation, t_{56} is not equal to t'_{56} due to the accumulation of pair-wise registration errors. We define "violation" as the difference between t_{56} and t'_{56}. It can be easily shown that by sequential registration, violation exists only between the neighboring range images, which are independent in registration.

The multiple registration method developed in this research consists of two consecutive procedures. First, all range images are sequentially aligned to a global coordinate system by the shortest registration path. It is conducted in an iterative mode. In each iteration, a range image is registered to a global coordinate system, which has a shorter registration path than other unregistered range frames. Secondly, minimizing violations using a weight least square method. A quantitative definition to violation is given in next section. Arbitrarily specifying the global coordinate system as that of a reference range frame, by the above registration, a model of the spatial objects with respect to the reference frame is obtained. On the other hand, in urban outdoor environment, absolute locations of several range frames can be measured using GPS. When given the absolute locations as ground control points, a model of the spatial objects with respect to world coordinate system can be obtained. The additional work other than the above registration procedures is addressed in section 2.5.

2.3 Registration cost function

Let $t'_{ij} = T_i^{-1}T_j = \begin{pmatrix} R'_{ij} & Sh'_{ij} \\ 0 & 1 \end{pmatrix}$ be the relative transformation matrix obtained in multiple registration. Violation (Vio_ij) in range image pair V_i and V_j is evaluated by summing the violation in shift vector (Vio^{sh}_{ij}) and rotation matrix (Vio^{r}_{ij}) as follows.

$$Vio_{ij} = Vio^{sh}_{ij} + Vio^{r}_{ij} \tag{3}$$

In this research, we evaluate the violation by its effect on the location of its nearby range frames. Figure 3 shows the evaluation of violations in shift vector and rotation matrix. Due to the difference between and , the location of range frame is projected to (Figure 3(a)). Violation in shift vector is thus evaluated by the Euclidean distance between and , which is equal to . On the other hand, due to the difference between and , the location of range frame is projected to (Figure 3(b)). Violation in rotation matrix is evaluated by averaging its effect on all nearby range frames as follows.

$$Vio^r_{ij} = \frac{1}{n}(\sum_{(i,k)\in S} \omega_{jk} Vio^r_{ijk} + \sum_{(i,k)\in S} \omega_{ik} Vio^r_{jik}) \tag{4}$$

where, $Vio^r_{ijk} = \|R^{'}_{ij}Sh^{'}_{ij} - R_{ij}Sh_{ij}\|$; S is the set of all neighboring range images; n is the number of all range image pairs of V_i and V_j; $\omega_i j$ is a confidence value to the relative transformation obtained in pair-wise registration.

Multiple registration is to minimize the following cost function.

$$E_d = \sum_{(i,k)\in S} \omega_{ij} Vio^2_{ij} \tag{5}$$

Through the experiment in Zhao and Shibasaki 1999, we found that a failure or an unreliable pair-wise registration is always caused by extremely less overlay between range image pairs. An enough overlay between range image pairs is a premise for a reliable registration, whereas it might not be always satisfied in urban outdoor environment due to occlusions. To evaluate the reliability of pair-wise registration, in this research, the confidence value ω_{ij} is assigned as the number of the matched range points between V_i and V_j. After aligning V_i and V_j to a common coordinate system by the pair-wise registration result, matched range points are counted as the point pairs, which have distances smaller than a given threshold.

2.4 Multiple registration when location of several viewpoints in world coordinate system is known

When the world coordinates (e.g. GPS coordinate system) of several viewpoints are known, all range images can be directed aligned into the world coordinate system. Suppose $\{V_s|1 \quad s \quad n\}$ is the set of range images of which locations $\{p_s|1 \quad s \quad n\}$ in a world coordinate system is known. The addition work other than those addressed above is we have to find the transformation matrixes from $\{V_s|1 \quad s \quad n\}$ to the world coordinate system. It is conducted as follows. Arbitrarily select one range image V_k, $1 \quad k \quad n$, sequentially align all range images to the coordinate system of V_k. Let $\{p^{'}_s|1 \quad s \quad n\}$ be the viewpoints of $\{V_s|1 \quad s \quad n\}$ with respect to the coordinate system of V_k. A transformation matrix $\begin{pmatrix} R_k & Sh_k \\ 0 & 1 \end{pmatrix}$ transforming $\{p^{'}_s|1 \quad s \quad n\}$ to $\{p_s|1 \quad s \quad n\}$ can be obtained using least square method. Transformation matrix $\begin{pmatrix} R_s & Sh_s \\ 0 & 1 \end{pmatrix}$ from the coordinate system of V_s, $1 \quad s \quad n$ to the global coordinate system is obtained as follows.

$$\begin{array}{ll} \text{if } s = k & \text{then } R_s = R_k, Sh_s = p_s - p^{'}_s \\ \text{if } s \neq k & \text{then } R_s = R_k R_{ks}, Sh_s = p_s - p^{'}_s \end{array} \tag{6}$$

3 EXPERIMENT RESULTS AND DISCUSSIONS

42 overlapping range images were measured around a building in the campus of the Univ. of Tokyo. A map of the locations is given in Figure 4. Range images used in the experiment are measured by $[-180^o, +180^o]$ in horizontal rotation angle and $[-20^o, +40^o]$ in vertical rotation angle, with the resolution of 2 sample per horizontal degree and 1 sample per vertical degree. Laser range measurement has an accuracy of $\pm$ 5cm. In the followings, we present two sets of experimental results. The first experiment examines the accuracy of multiple registration when location and direction of range frame are totally unknown. We specify the global coordinate system as that of range image #1. All range images are registered, and a model of the buildings with respect to the coordinate system of range image #1 is obtained. The second experiment examines the registration accuracy when absolute location of several range frames are known. We compare the registration accuracy when given different number of absolute locations. Two sets of ground truth are used to examine the registration accuracy. They are 1) a 1:500 scale digital map and, 2) location of viewpoints measured by GPS with an accuracy of 20cm.

3.1 Multiple registration when location and direction of range images are unknown

Range images are first sequentially aligned to the coordinate system of range image #1. A conceptual structure of the alignment is shown in Figure 5. Since after sequential alignment, violation exists only in the neighboring

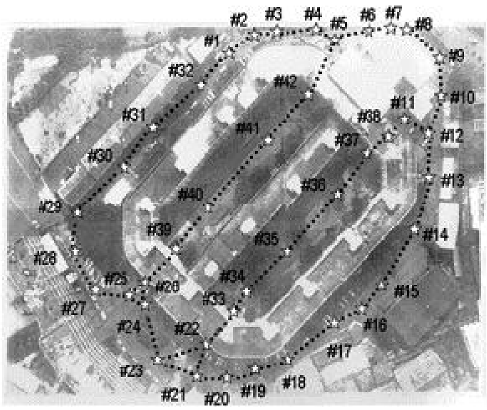


Figure 4: Location map of viewpoints of 42 range images in the campus of the Univ. of Tokyo.

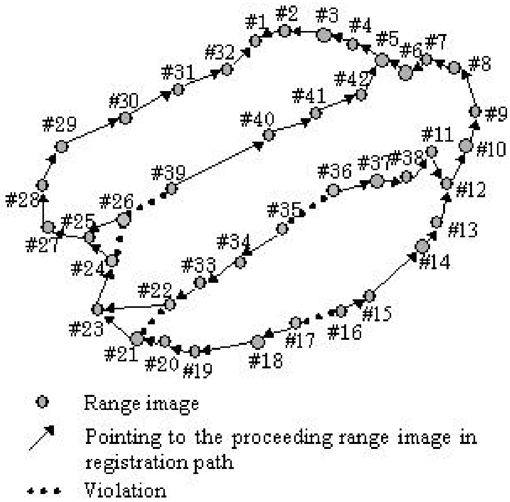


Figure 5: A conceptual structure of sequential alignment.

(a) Result of sequential alignment

(b) Result of minimizing violations

Figure 6: Result of multiple registration when location and direction of range images are unknown.

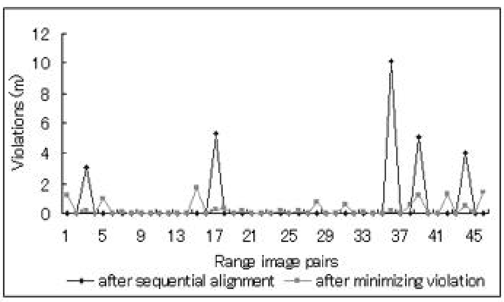


Figure 7: Location map of viewpoints of 42 range images in the campus of the Univ. of Tokyo.

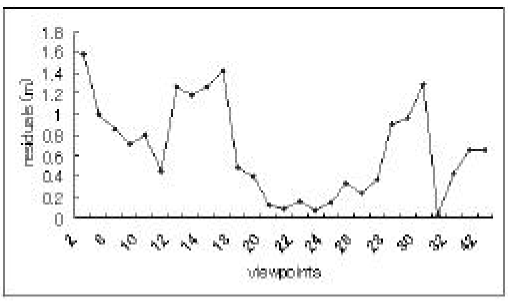
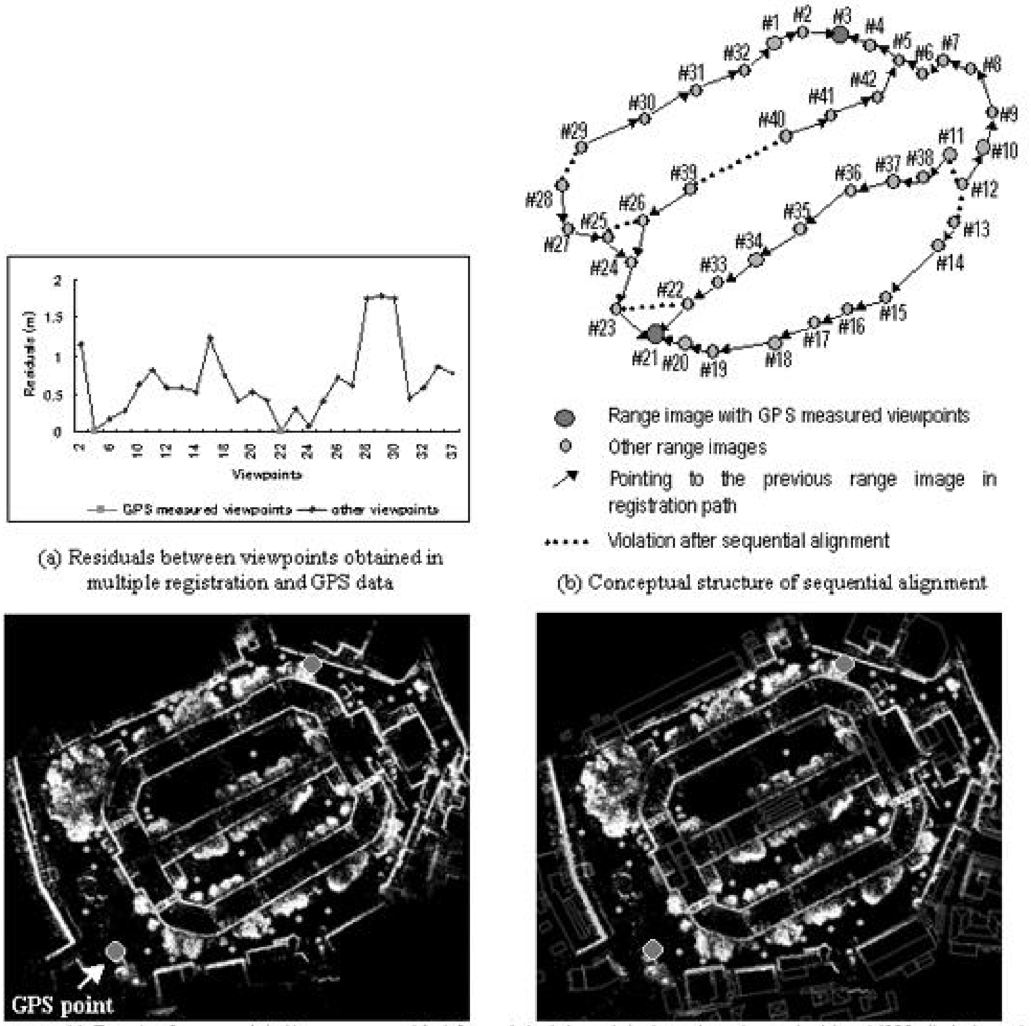


Figure 8: A conceptual structure of sequential alignment.



(a) Residuals between viewpoints obtained in multiple registration and GPS data

(b) Conceptual structure of sequential alignment

(c) Result of sequential alignment

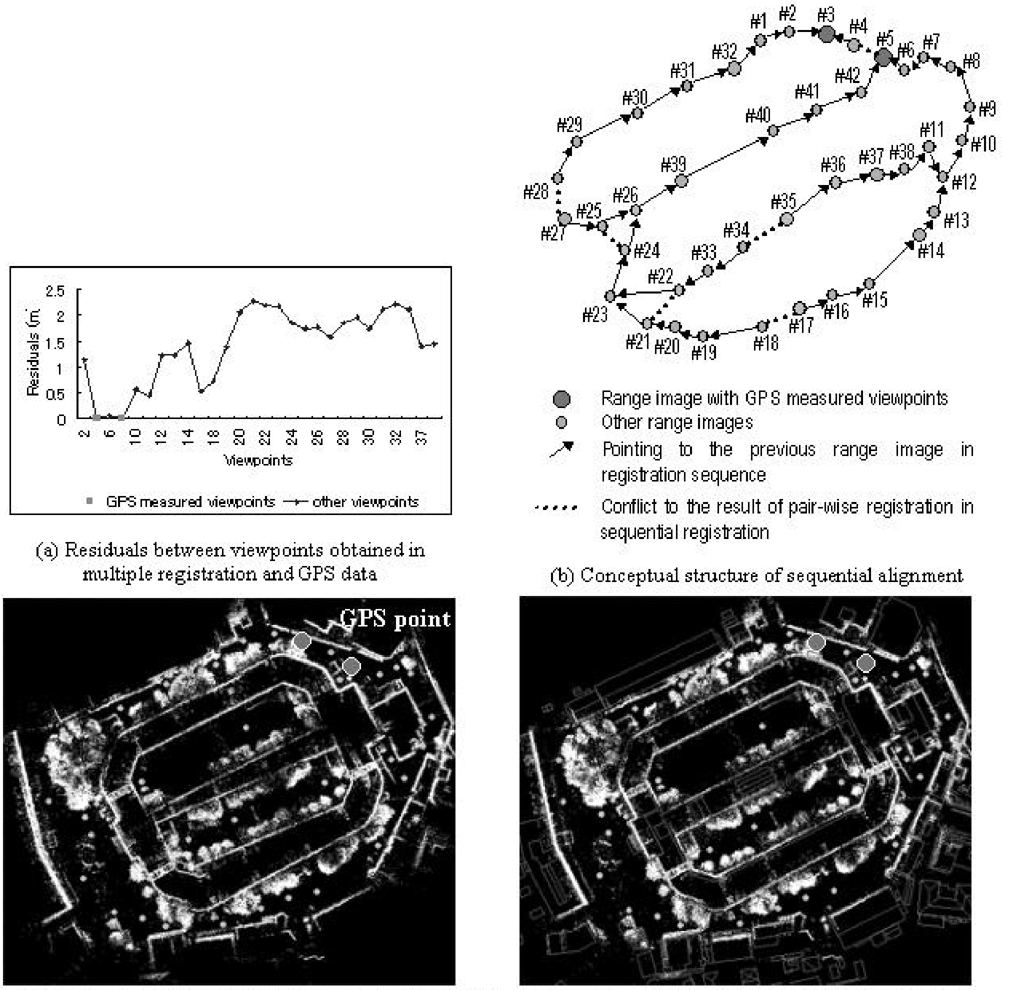
(d) After minimizing violations (overlapped with a 1/500 digital map)

Figure 9: Multiple registration using two GPS points which are located on the opposite side of the network.

range images, which are independent in registration. Violations are shown in Figure 5 as dotted lines. It can be found that violations exist in ve neighboring range image pairs, i.e. (#26, #39) (#24,#26) (#21,#22) (#16,#17), and (#35,#36), after sequential alignment. A result of sequential alignment is shown in Figure 6(a). The model is shown in Z-image as de ned in Zhao and Shibasaki, 1999. Pixel size of Z-image is 1m. Obvious inconsistency can be found in several places. Violations are minimized as addressed in section 2. A well-balanced model of range images is obtained as shown in Figure 6(b). Change of violations in each neighboring range image pairs is examined in Figure 7. The ve peaks after sequential alignment corresponds to the ve violations shown in Figure 5. It can be found that after minimization, violations are not disappeared but distributed into other range image pairs. Violations are not distributed equivalently, because a weight evaluating the reliability of pair-wise registration is assigned in minimization. To examine the relative accuracy of the model, a transformation matrix transforming the locations of range images in the model to the coordinate system of GPS is estimated, where the sum of squared residuals from registered location to GPS measured locations is minimized. Residuals between registration result and GPS data are graphed in Figure 8. Mean of residuals is 0.663m.

3.2 Multiple registration when absolute locations of several range images are known

We measure the absolute locations of several range images using GPS. With the absolute locations, all range images can be registered to the world coordinate system. Figure 11 and Figure 12 compares the results of registering 42 range images using two GPS points. Figure 11 shows the result where two GPS points are located on the opposite side of the network, while Figure 12 shows the result where two GPS points are located near to each other. In both of the cases, six violations exist in sequential alignment. They are shown as dotted lines in the conceptual structures (see Figure 11(b) and Figure 12(b). Results of sequential alignment are shown in



(a) Residuals between viewpoints obtained in multiple registration and GPS data

(b) Conceptual structure of sequential alignment

(c) Result of sequential alignment

(d) After minimizing violations (overlapped with a 1/500 digital map)

Figure 10: Multiple registration using two GPS points which are located near to each other.

Figure 11(c) and Figure 12(c). Results after minimizing violations are shown in Figure 11(d) and Figure 12(d). Obvious violations can still be found in Figure 12(d), while the model in Figure 11(d) has a better consistency. Both of the models are overlapped with a 1:500 digital map to show the matching between them. Registration accuracy is examined by evaluating the residuals from registered locations to the ground truth of GPS data as shown in Figure 11(a) and Figure 12(a). Mean of residuals in Figure 11(a) is 0.712m, and 1.525m in Figure 12(a). In both of the cases, mean of residuals is larger than the registration error without using GPS data, because a least square transformation is conducted in accuracy examination in later one. Conclusion drawn from this examination is balancing the distance between GPS points will help in reducing registration error.

Figure 13 and Figure 14 compares the results of registration accuracy when using two, four, six and nine GPS points. Locations of GPS points are shown in Figure 13, as well as the models of range images after minimizing violation. In the case of registration using two GPS measured viewpoints, the points are located the same with those shown in Figure 11, and so does the model generated. From Figure 13, we can not nd obvious di erences between the models generated, since the registration errors of most of the viewpoints are lower than 1m, which is the pixel size of the Z-images. Registration error of each viewpoint is shown in Figure 14. Mean of residuals is 0.712m when using two GPS points, 0.612m when using four GPS points, 0.436m when using six GPS points and 0.283m when using nine GPS points. Since the measurement of GPS points has an error of 20cm, registration using nine GPS points has almost the same order with GPS measurement. In most of the cases, as the number of GPS points used in multiple registration increased, higher accuracy in registration can be obtained. However, exceptions can be found when GPS data on the other hand upset the local consistency among the range images nearby, hence cause larger residuals. Reason for this might be inaccurate GPS data or extremely low overlapping between neighboring range image pairs.

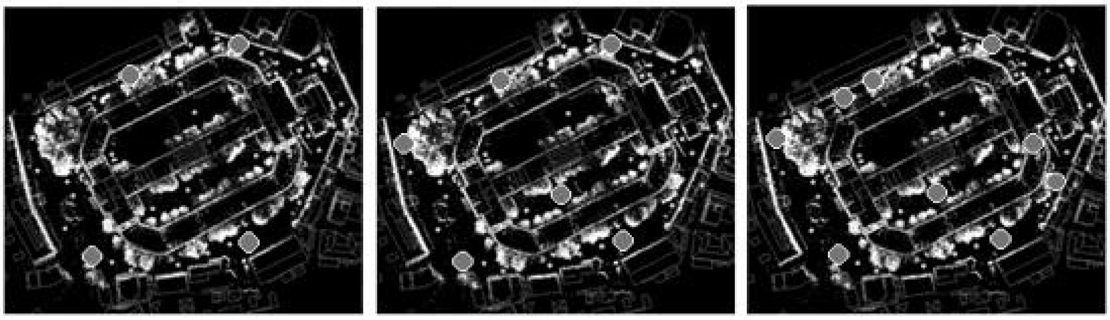


Figure 11: Multiple registration using four, six, and nine GPS points.

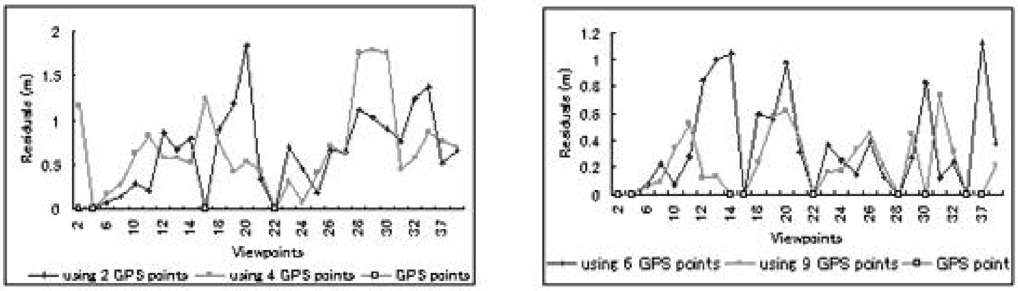


Figure 12: Multiple registration using four, six, and nine GPS points.

REFERENCES

Besl, P.J., N.D.McKay, A method for registration of 3-D shapes , IEEE Trans. PAMI 14(2), 1992.

Bergevin,R., M.Soucy, H.Gagnon, D.Laurendeau, Towards a general multi-view registration technique , IEEE Trans. PAMI 18(5), 1996.

Blais, G., M.D.Levine, Registering multiview range data to create 3D computer objects , IEEE Trans. PAMI 17(8), 1995.

Champleboux, G., S.Lavallee, R.Szeliski, L.Brunie, From Accurate Range Imaging Sensor Calibration to Accurate Model-Based 3-D Object Localization , Proc. CVPR, 1992.

Chen, Y., G.Medioni, Object Modelling by Registration of Multiple Range Images , Image and Vision Computing, vol.10, no.3, pp.145-155, 1992.

Haala, N., C.Brenner, C.Statter, An integrated system for urban model generation , Int. Archive of Photogrammetry and Remote Sensing, vol.XXXII, part 2, pp.96-103, 1998.

Kamgar-Parsi, B., J.L.Jones, A.Rosenfeld, Registration of multiple overlapping range image without distinctive features , IEEE Trans.PAMI, vol.13, no.9, 1991.

Lemmens, M.J.P.M., H.Deijkers, P.A.M.Looman, Building detection by fusion airborne laser-altimeter DEMs and 2D digital maps , Int. Archive of Photogrammetry and Remote Sensing, vol.32, Part3-4W2, pp.42-49, 1997.

Ng, K., V.Sequeira, S.Butter eld, D.Hogg, J.G.M.Goncalves, An Integrated Multi-sensory System for Photo-Realistic 3D Scene Reconstruction , Int. Archive of Photogrammetry and Remote Sensing, vol.XXXII, part 5, pp356-363, 1998.

Ozawa, S., M.Notomi, H.Zen, A Wide Scope Modeling to Reconstruct Urban Scene , Int. Archive of Photogrammetry and Remote Sensing, vol.XXXII, part 5, pp370-376, 1998.

Shum, H., K.Ikeuchi, and R.Reddy, Virtual reality modeling from a sequence of range images , Proc. IEEE/RSJ Intern. Conf on Intelligent Robots and Systems, pp.703-710, 1994.

Thorpe, M.H.Hebert, T.Kanade, S.A.Shafer, Vision and Navigation for the Carnegie-Mellon Navlab , IEEE Trans. PAMI, vol.10, no.3, pp.363-373, 1988.

Zhao, H., R.Shibasaki, A Robust Method for Registering 2.5D Laser Range Image of Urban Objects , IAPRS, Vol.32, Part2W1,1999.