MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

7—4 Reconstruction Textured Urban 3D Model by Fusing Ground-Based Laser Range Image and CCD Image

Huijing ZHAO Ryosuke SHIBASAKI
Institute of Industrial Science, University of Tokyo
Email: chou@skl.iis.u-tokyo.ac.jp

Abstract

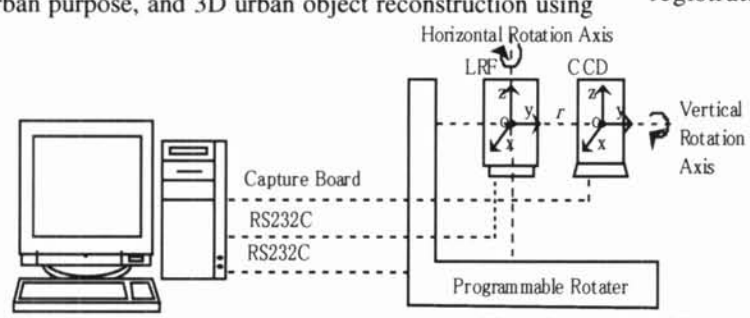
In this paper, a method of fusing ground-based laser range image and CCD image for the reconstruction of textured 3D urban object is proposed. An acquisition system was developed to capture laser range image and CCD image simultaneously from the same platform. The registration of laser range image is achieved by finding corresponding planar faces which are extracted from laser range images. Texture data are projected onto TIN-based object surfaces derived from laser range data. Through an outdoor experiment for reconstructing a building at IIS, University of Tokyo, it is demonstrated that textured 3D model of a building can be generated in an automated manner.

1 Introduction

By now, most of the researches on 3D urban object reconstruction have been devoted to the analysis of aerial based imageries. However, an often presence of occlusions, abrupt changes of depth and elevation in urban area, prevent full automation of this method. On the other hand, with the development of automobile navigation system and GIS applications, there is a growing demand for complete and accurate 3D urban database, which cannot be met by aerial based technique. Hence automated acquisition methods using ground-based sensors are attracting more and more attentions. Several researches using CCD camera have demonstrated that 3D information can be extracted using motion [10,12] and stereo [8] vision technique. Whereas, the difficulties in reliable stereo matching, distortion from limited resolution and unstable geometry of CCD cameras are the major obstacles to the accuracy of this technique. Range Sensors have also been used for 3D object acquisition, the efficiency and accuracy of the technique has been demonstrated using indoor objects [1-3,7,9,11,13]. In recent years, with the development of laser technique, range finder using eye safe laser, with long range measurement and high speed is being used for urban purpose, and 3D urban object reconstruction using ground-based laser range finder have got more concerned. In our previous research, a method of automated registration of two views of laser range image was proposed [4,5], and showed promising results with some preliminary experiment. The drawback of 3D spatial model generated only using laser range finder is it has poor understandability or interpretability, because it is rather difficult to find out correspondence between TIN-based 3D surfaces and the real world or what users can see.

In contrast to the previous approaches, this paper describes an attempt to reconstruct textured urban 3D model employing both ground-based laser range finder for 3D spatial data and CCD sensor for texture data of the surface. A sensor system is developed for this purpose jointly by Asia Aerial Survey Co., and the Univ. of Tokyo, where laser range finder (LRF) is setup together with a CCD camera on a programmable rotator (see fig.1). The platform of the sensor is fixed on the ground during data acquisition. There are two steps in the reconstruction of textured 3D model, 1) In each view, a spherical image is created by a sequence of CCD image patches, and a projection model is constructed to map the laser range points to the spherical image. Hence, after the acquisition of each view, laser range image and CCD image are fused, and a textured 3D model is obtained for each singular view; 2) Different views are registered using laser range image and an integrated textured 3D model is reconstructed. Multiple views' registration is based on our previous researches, where planar faces are extracted from laser range image in each view, transformation between sensor's local coordinate system is recovered by automatically identifying correspondences among planar faces (Zhao and Shibasaki, 1997).

An experiment is conducted using the buildings in IIS, Univ. of Tokyo as the target object. Several views of laser range image which have the resolution about 1 sample/degree (4 samples per square meter by average), as well as a sequence of CCD images are obtained. After registration of a sequential view pair, a textured 3D



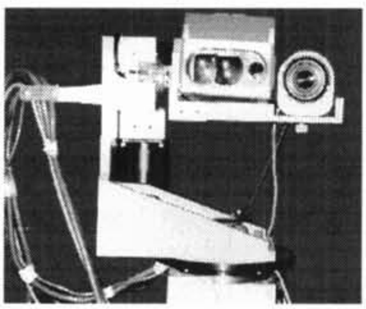


Fig.1 System Architecture

model on the integrated two views are constructed. In the followings, section 2 is devoted to the principle of sensor and acquisition system, experiment and result is addressed in section 3, finally conclusion is given in section 4.

2 Sensor and Acquisition System

2.1 Sensor alignment description

Laser range finder (LRF) is set together with a CCD camera on a programmable rotator. Both LRF and programmable rotator are controlled through a serial port by a DOS/V PC, while CCD camera is controlled using a capture board. A CCD image (patch image) is captured whenever the laser range finder measured a distance and angle of a target point. Alignment of the sensors and the platform is summarized as follows;

1) Sensor's coordinate system of LRF and CCD are parallel, where y-axes are coaxial, and a shift of r along y-axis exists.
2) Two rotation modes, vertical and horizontal are available. Vertical rotation axis is passing through the origin of both LRF and CCD 's coordinate system, while, horizontal rotation axis coincides to the z-axis of LRF. There is a displacement (r) from the horizontal rotation axis to the center of the CCD sensor. Thus the origin of the coordinate system of LRF is fixed, while the origin of CCD 's moves along a circle with the radius r when the sensor system rotates horizontally. Hereafter, a rotation is referred as (h,v), where h is for horizontal rotation angle, and v for vertical rotation angle.

2.2 Acquisition System

An acquisition system is designed to take a sequence of CCD image patches with an interval of (θ , ψ) (see fig.2.b,c). As a result, a spherical image is generated by sequentially assembling the image patches together. The horizontal displacement of the CCD camera from the center of rotation (i.e. the position of the laser range finder) yield a discontinuity between the CCD image patches. The discontinuity is caused by the difference of scale between neighboring images as shown in Fig.2 (a). To mitigate the discontinuity, viewing angle of each CCD scene have to be limited (Fig.2 (b)). Limiting the viewing angle of CCD scenes is also beneficial in reducing horizontal displacement of the texture information which are mapped onto the 3D surface, because the horizontal displacement of the texture information is proportional to the difference along depth direction between real 3D surface and reconstructed 3D surface as shown in Fig.2 (d).

On the other hand, vertical interval angle ψ can be any positive value not larger than camera's maximum visual angle because the center of both CCD camera and LRF are on the vertical rotation axis (Fig.2 (c)). In our outdoor experiment θ is set to be 3° , while ψ is set to be 30° (see fig.3).

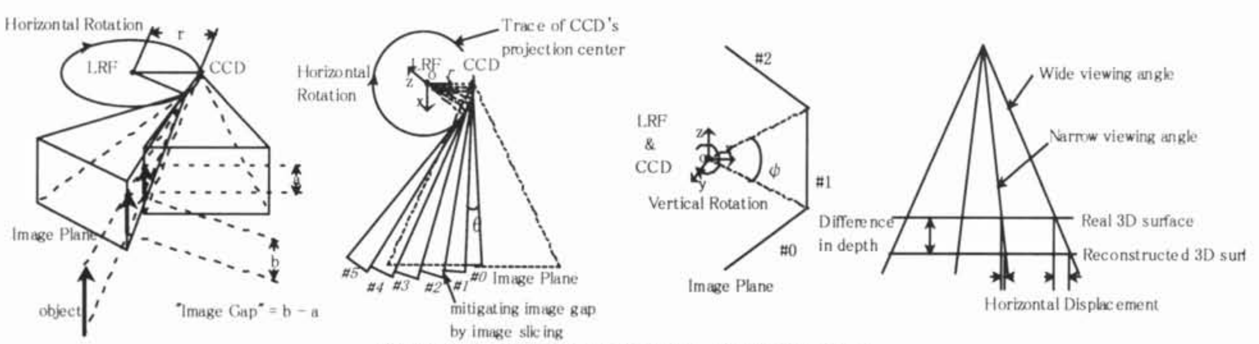


Fig.2 Geometric Model of Acquisition System

(a) Image Gap by the inconsistency of sensor's projection center (b) Image Gap is mitigated by image slicing in horizontal intersection plane (c) Geometric model in vertical intersection plane (d) Horizontal displacement can be reduced by limiting viewing angle.

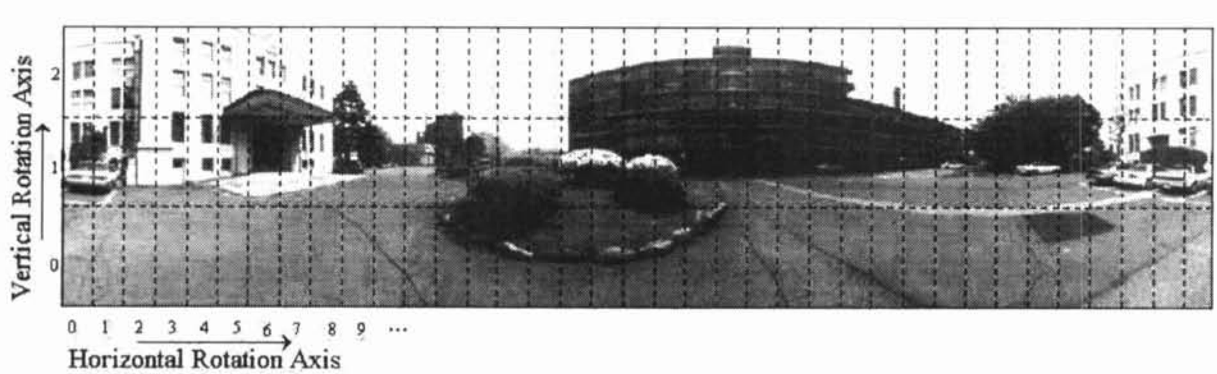
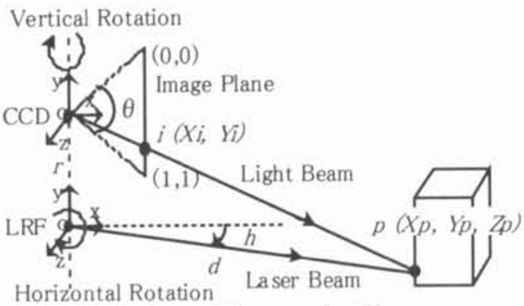
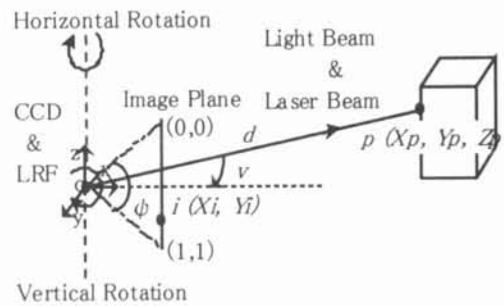


Fig.3 Spherical Image Generatedd from CCD Image Patches



(a) Horizontal Intersection Plane



(b) Vertical Intersection Plane

Fig.4 Projection Model from laser range point to CCD image

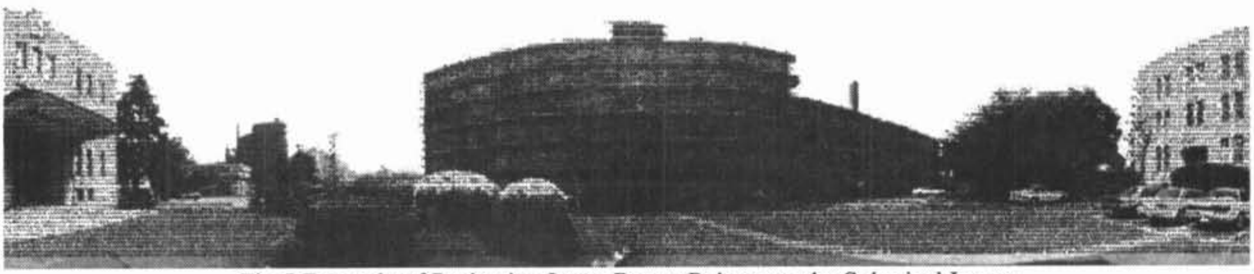
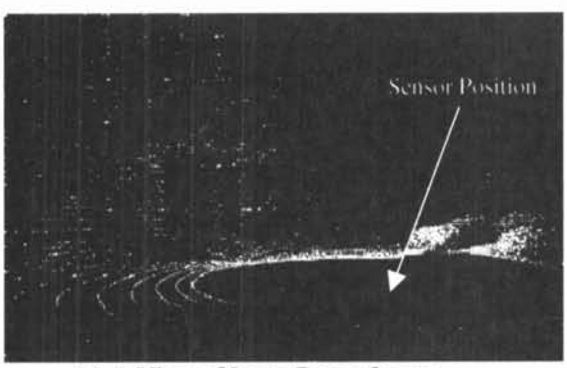
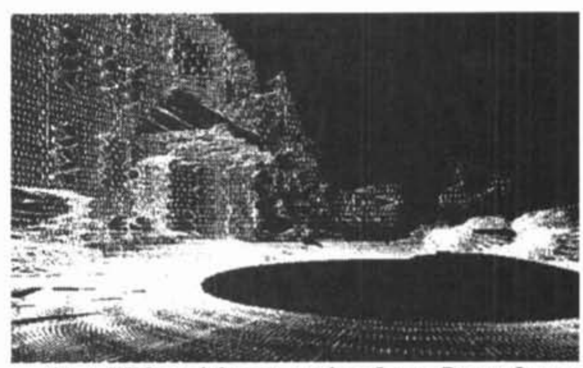


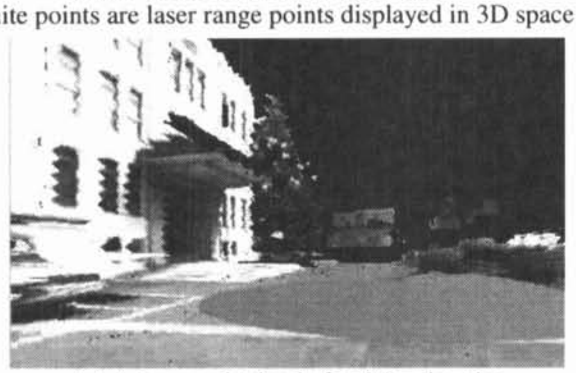
Fig.5 Example of Projection Laser Range Point onto the Spherical Image



(a) A View of Laser Range Image
white points are laser range points displayed in 3D space



(b) A TIN-model generated on Laser Range Image



(c) A Image obtained after Mapping the Spherical Image onto the TIN-model in (b)

(d) testing site
a buildings in IIS, Univ. of Tokyo

Fig.6 An example of reconstructing textured 3D MODEL on a singular view

2.3 Projection model from laser range image to CCD image (spherical image)

Fig.4 depicts the geometry of laser range point and CCD image data. Suppose laser range point $p(Xp,Yp,Zp)$ is obtained after a rotation (hp,vp) and that its image point $i(Xi,Yi)$ is found on an image patch *#s* taken at the rotation position (hs,vs). Suppose the target object of measurement are far enough that visual angles of the CCD camera can be approximated by its maximum visual angles, the corresponding position of $p(Xp,Yp,Zp)$ on the spherical image can be determined as follows (see Fig.4). Here let $h=hp-hs$, $v=vp-vs$. Fig.5 is an example of projecting laser range points onto the spherical image.

$$Xi = \frac{d * \sin h + r}{d * \cos h} + 0.5, \qquad 0 \le Xi < 1;$$

$$Yi = 0.5 - \frac{d * \sin v}{d * \cos v}, \qquad 0 \le Yi < 1. \quad \text{(formula 1)}$$

2.4 Textured 3D model on a single view

A textured 3D model is obtained by 1) constructing a TIN-based 3D surface model using laser range points, 2) mapping the spherical image data onto the triangular surfaces according to the inverse projecting model of laser range points (or triangles) to the spherical image as described in section 2.3. Fig.6 shows an example.

3 Reconstructing 3D Model From Multiple Views

Two sequential views of laser range images and CCD images were obtained in an outdoor experiment using the acquisition system described in the previous section. A textured 3D model from each singular view can be constructed directly after the acquisition process. In order to create an integrated 3D model from multiple views where the spatial relationship among sensor's local coordinate system are unknown, the registration algorithm proposed by the authors [5] are employed.

Laser range measurement in urban area has two characteristics, 1) long and wide measurement range yields a large amount of data in each view, 2) a lot of uncertainties exist which might be caused by window glass, electric cable, tree, passing car, or pedestrian. Considering the large amount of data, efficiency in registration is strongly required. Thus feature based registration method [7] is selected. Since laser range measurement of surface is comparably accurate than of an edge or a vertex, and many man-made urban object can be characterized by planar surface. In this research, we choose planar face as the primitive for views' registration. There are two steps in the registration process, 1) planar faces from each view are extracted, 2) corresponding planar faces from different views are identified, and transformation parameters between different sensor's local coordinate system are estimated.

3.1 Planar Face Extraction

Planar face extraction is accomplished in three steps. 1) Laser range images were first split into smaller range pieces in a similar way as Quad-Tree method. Each range piece represents a planar face with a regression variance lower than a given threshold. (see fig.7(a)) 2) Homogenous range pieces were merged together to minimize a MDL[6]-based cost function[4]. Parameter estimation of planar faces is based on a robust regression technique called *M-estimators*[14], where a *M-estimator* is defined to filter out outlying laser range points. In the formulation of MDL-based cost function, for each planar face, a histogram created from regression residuals is used to approximate the statistical distribution of laser range points. 3) Planar faces which have larger size than a given threshold are picked out for registration purpose. Fig.7(b) is an example of planar face extraction result. Planar faces are represented by their corresponding laser range points. Laser range points belonging to different planar faces are given a different intensity value.

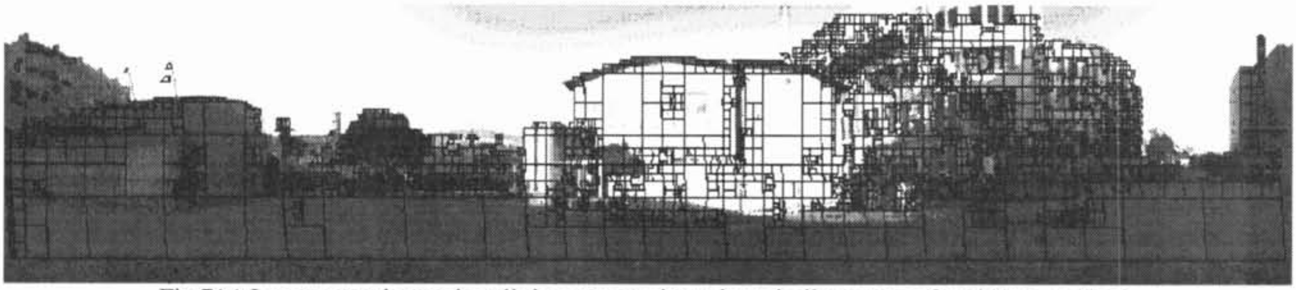


Fig.7(a) Laser range image is split into range pieces in a similar way as Quad-Tree method, until each range piece represents a planar face in a given threshold

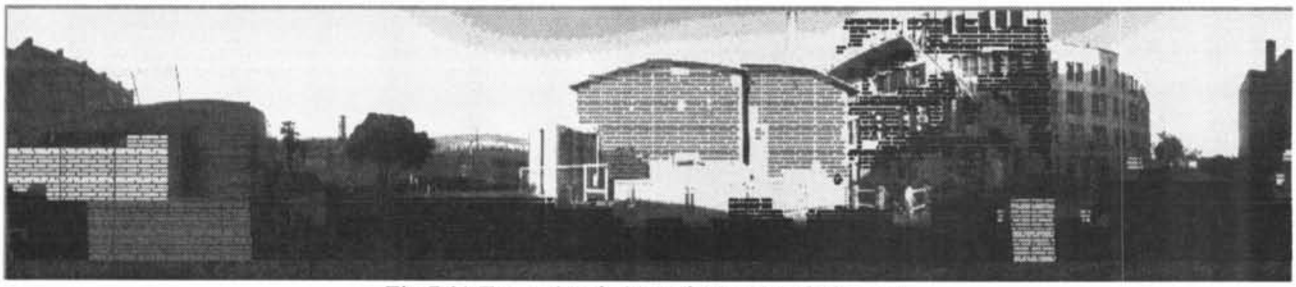


Fig.7(b) Example of planar face extraction result
Planar faces are represented by the laser range points belonging to them, different planar faces are given different intensity values.

3.2 Registration of Two Views

As results of the planar face extraction process, over 20 planar faces were extracted from each view. Since 3 nonparallel planar face pairs can determine transformation parameters between each local sensor coordinate systems, over 20 planar faces from each view means that there are more than ${}_{20}C_3 * {}_{20}P_3$ correspondence candidates to be checked. In order to find the most reliable one efficiently, the SRbPF (spatial relationship between planar faces) method developed by the authors[5] was employed. With the SRbPF method, two constraints, *Surface Normal constraint* and *Distance constraint*, are defined, corresponding candidates satisfying both two constraints are selected. The one yielding the largest overlay of the range points from different views is supposed to be the most reliable one. In this experiment, after the correspondence hunting on *Surface Normal Constraint*, there are about 500 candidates left; after the correspondence hunting on *Distance Constraint*, there are only 17 candidates left. To evaluate the reliability of each candidate, a "cloud" of range points is created first for view 1, where, the sequience of planar faces obtained in the process of range image splitting is used as the bone, a thickness is attached to create the "cloud". To count the overlapping range points of each correspondence candidate, after transforming view2 to view1 according to the candidate, a range point in view2 is said to be "matching to view 1", if and only if it is in the cloud of view1. The ratio of overlapping range points of two views is calculated by the percent of matched range points to total range points. In this experiment the largest overlay were count to 84.7% (see fig.8).

In fig.8, planar faces are represented by laser range points belonging to them. Same intensity(color) value is given to the corresponding planar face. In this experiment, several other candidates are also found as the correct correspondence relationship, and raising an overlay ratio more than 60%. On the other hand, overlapping ratio of wrong candidates is always lower than 50%.

3.3 A Textured 3D MODEL Generated from Two View

After registration, two views of laser range images are integrated, and by mapping the spherical images to the integrated laser range images, a textured 3D model is created (see fig.9). However, a method of removing the outlying range points to improve the accuracy of 3D model needs to be further studied.

4 Conclusion

In this paper, a method is presented of reconstructing textured urban 3D model by the fusion of ground-based laser range image and CCD image. An acquisition system was developed, and an outdoor experiment was conducted, where two view of laser range images and CCD images were obtained and registered automatically. After registration, a textured 3D model on integrated two views was reconstructed. The experiment proved a high performance of the system for real time 3D object reconstruction, and showed high accuracy in the planar face extraction and registration result.

Future research will have to be addressed on finer adjustment of registration result using the information from CCD image, optimizing texture data in overlaying area of different views, and mitigation of error accumulation in the integration of multiple views.

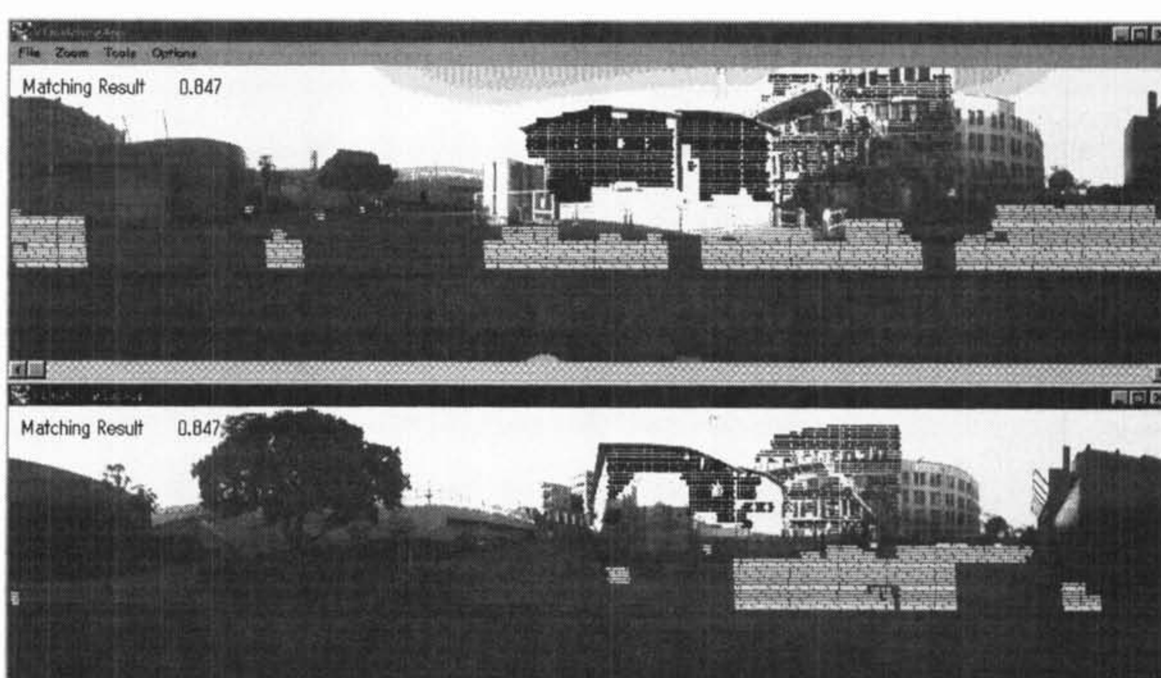
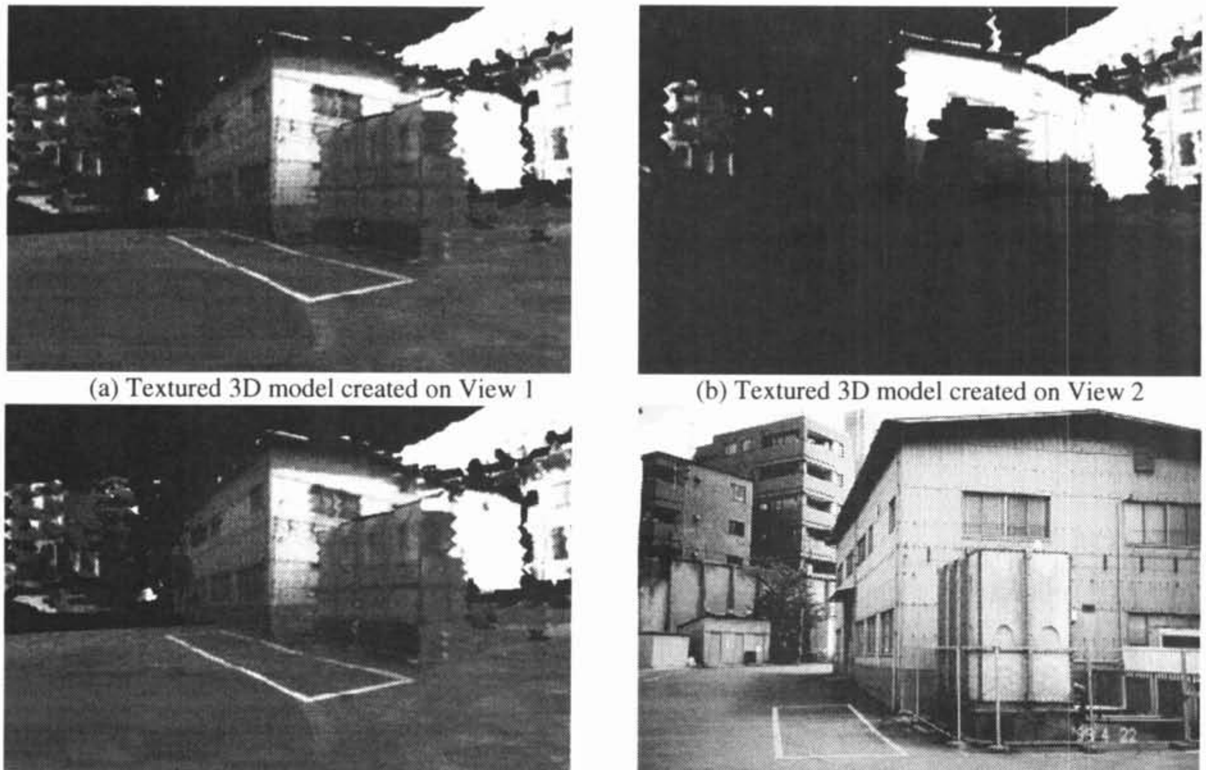


Fig.8 The most reliable correspondence candidate
Corresponding planar faces are given the same intensity value.



(a) Textured 3D model created on View 1

(b) Textured 3D model created on View 2

(c) Integrated 3D model on two views

(d) Testing site (a building of IIS, Univ. of Tokyo)

Fig.9 Example of textured 3D model constructed after registration of two views

Acknowledgement

The authors would like to express their gratitude to

College Women's Association of Japan, who have been supporting the work by a scholarship award to the first author.

Asia Aerial Survey Co., who have been providing the sensor system, and cooperation.

Reference

[1]D.S.Meek, D.J.Walton, Surface Alignment or Registration of Multiple Range Images, Congressus Numerantium 99(1994), pp.171-177.

[2]F. Stein, G.Medioni. TOSS-A System for Efficient Three Dimensional Object Recognition, Proc. DARPA Image Understanding Workshop
Pittsburgh, PN, Sept. 1990

[3]H.Shum, K.Ikeuchi, and R.Reddy. Principle Component Analysis with Missing Data and Its Application to Object Modeling. CMU-CS-93. 1993.

[4]H.Zhao, R.Shibasaki, Simulation on Automated Reconstruction of Urban 3D Object using Laser Range Finder (LRF) data, Journal of the Japan Society of Photogrammetry and Remote Sensing, vol.36, no.4, 1997

[5]H.Zhao, R.Shibasaki, Automated Registration of Ground-Based Laser Range Image for Reconstructing Urban 3D Object, IAPRS, Vol.32, Part3-4W2,1997,pp.27-34

[6]J.Rissanen, A universal prior for integers and estimation by minimum description length, Ann. Of Statist. 11, No.2, pp416-431, 1983.

[7]K.Higuchi,M.Hebert,and K.Ikeuchi, Building 3-D Models from Unregistered Range Images, Graphical Models and Image Processing, vol.57,No.4,July,pp.315-333,1995.

[8]M.Okutomi, T.Kanade, A Multiple-Baseline Stereo, IEEE Trans. PAMI, vol.15, no.4, 1993.

[9]Paul J. Besl, Neil D. McKay , A Method for Registration of 3-D Shape, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, Feb. 1992

[10]P.Wang, C.Li, H.Zen, M.Sakauchi, 3D Information Acquisition From Plane Spatio-Temporal Images with Complicated Camera Motion, Proc. of IAPR Workshop on Machine Vision Application, Tokyo.

[11]R.Krishnapuram, D.Casasent, Determination of Three-Dimensional Object Location and Orientation from Range Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, No.11, pp.1158-1167, Nov. 1989.

[12]S.Ozawa, M.Notomi, H.Zen, A Wide Scope Modeling to Reconstruct Urban Scene, IAPRS, Vol.XXXII, Part 5, pp370-376, 1998.

[13]Y.Chen, G.Medioni , Object modelling by registration of multiple range images, Image and Vision Computing,Vol. 10, No. 3, Apr. 1992.

[14]Z.Zhang, Parameter Estimation Techniques : A Tutorial with Application to Conic Fitting, Image and Vision Computing Journal, Vol.15, No.1, pp. 59-76, 1997.

IAPRS: International Archives of Photogrammetry and Remote Sensing