RECONSTRUCTING TEXTURED URBAN 3D MODEL BY FUSING GROUND-BASED LASER RANGE IMAGE AND VIDEO IMAGE

Huijing Zhao Ryosuke Shibaski
Institute of Industrial Science, University of Tokyo

KEYWORDS : Data Fusion, Range Image, Video Image, Laser Range Finder, 3D Object Reconstruction, Textured Urban 3D Model

ABSTRACT

In this paper, a method of combining ground-based laser range finder and CCD camera for the purpose of reconstructing textured 3D urban object is proposed, where the registration of laser range image is achieved by finding corresponding planar faces extracted from a pair of laser range images and texture data are projected onto TIN-based object surfaces derived from laser range data. An acquisition system was developed to capture laser range data together with CCD image simultaneously from the same platform. To reconstruct textured 3D objects, laser range images taken from different viewpoints are registered with each other by automatically identifying corresponding planar faces among those extracted from laser range images. Texture data from CCD images are mapped onto TIN-based object surfaces derived from laser range data, using sensor position and attitude estimated through the automated registration process. 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

As the market of automobile navigation and GIS applications evolves to a new stage such as ITS (Intelligent Transport System) and real-time mitigation of earthquake damages, detailed 3D urban spatial database are becoming demanded strongly. Since the acquisition of detailed 3D spatial data is still labor-intensive, automated acquisition methods using ground-based sensors are attracting more and more attentions. Ground-based laser range finder is one of the potential sensors for the automatic reconstruction of detailed 3D urban database. It is based on sequential range and angle measurements from an ground-based platform. To generate 3D urban space model from laser range images taken from different points, a method of automated registration of two views of laser range image is proposed by the authors [1,2], and showed promising results, but with a preliminary experiment. The drawback of 3D space model generated with the technique is the poor understandability or interpretability because it is rather difficult to find out corresponding between TIN-based 3D surfaces and the reality or what users can see. On the other hand, several researches also demonstrated that 3D information can be extracted from sequential video images. However the difficulty in reliable stereo-matching and distortion from limited resolution and unstable geometry of video cameras are the major obstacles to the implementation of this technique.

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 video 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 video image are fused, and a textured 3D model is obtained for each singular view; 2) Images from different views are registered and 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, the 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. Two views of laser range images which has the resolution about 1 sample/degree (4 samples per square meter by average), as well as a sequence of video images are obtained. A textured 3D model is reconstructed on the integrated two views. 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

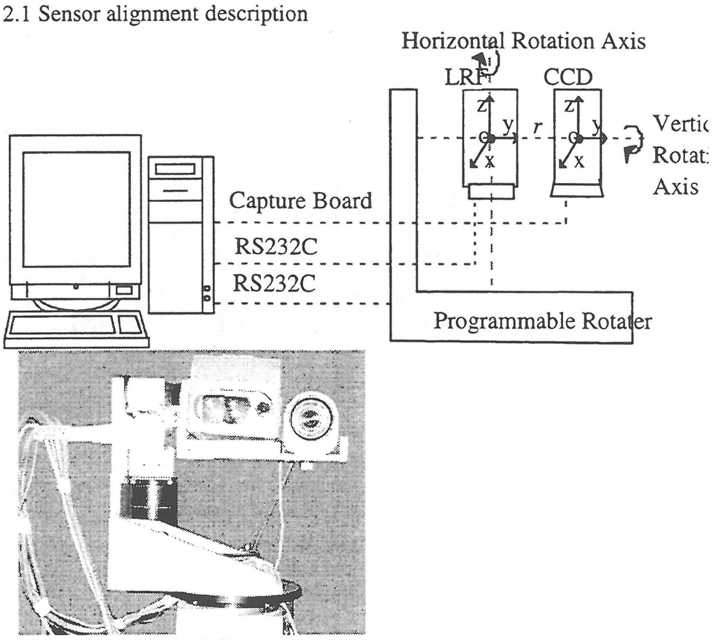


Fig.1 System Architecture

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 video image patches with an interval of (?,?) (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?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?is set to be 3°, while ?is set to be 30°(see fig.3).

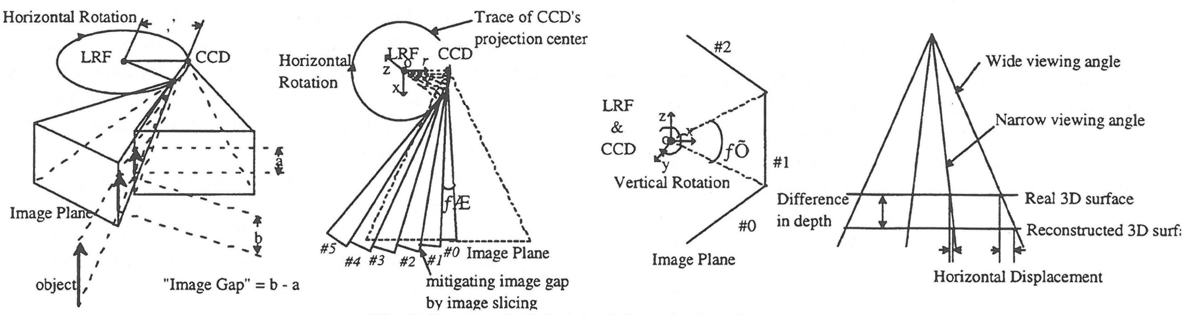


Fig.2 Geometric Model of Acquisition System

(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.

2.3 Projection model from laser range image to video 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

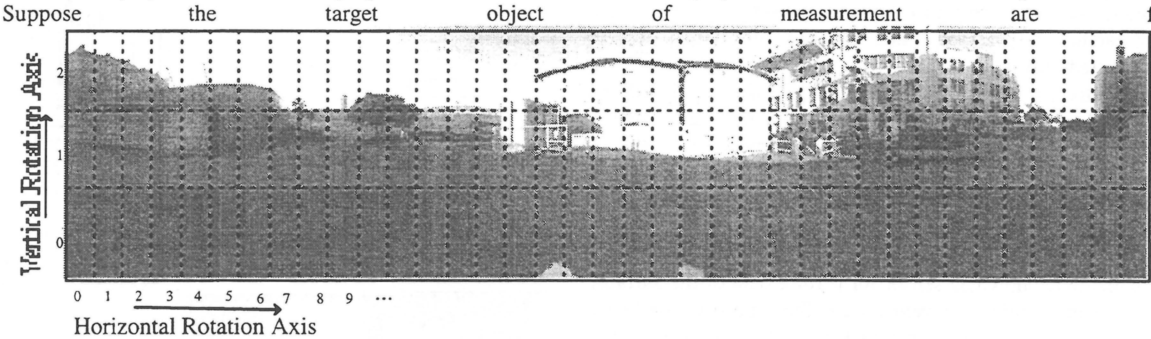
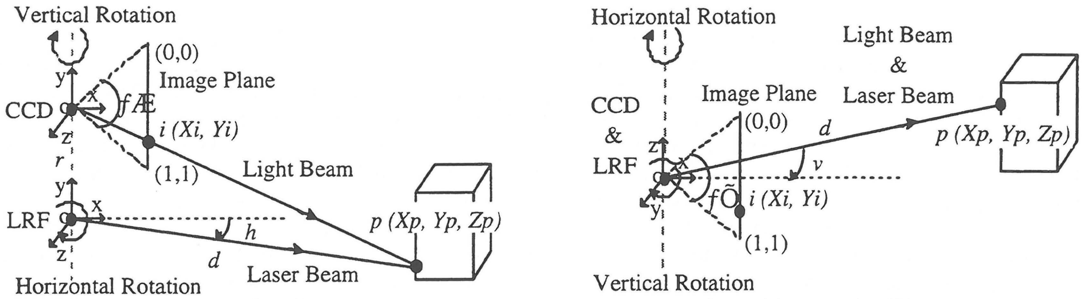


Fig.3 Spherical Image Generated from CCD Image Patches

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.

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

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



(a) Horizontal Intersection Plane (b) Vertical Intersection Plane

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

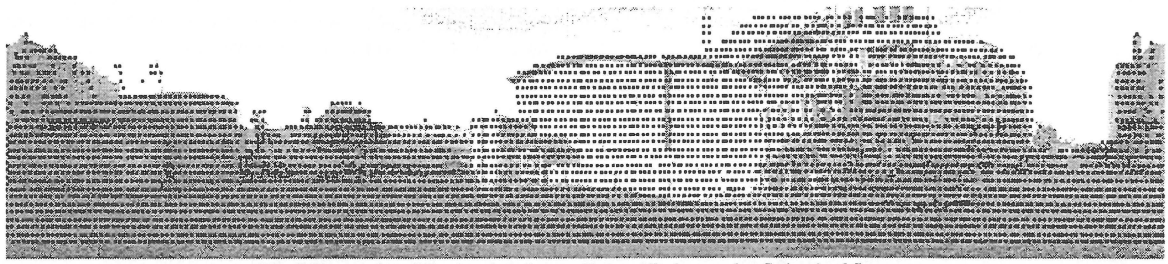
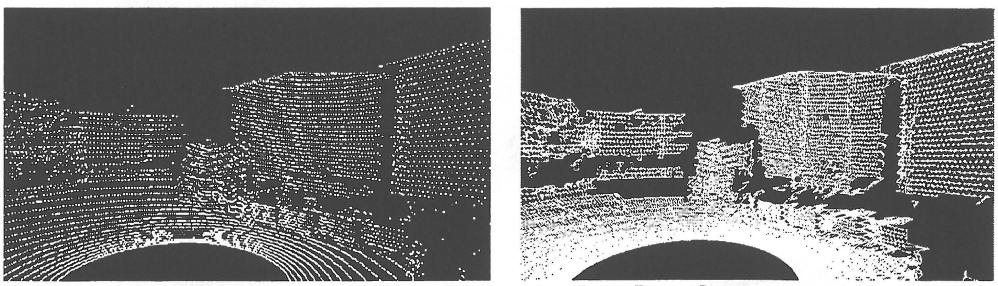


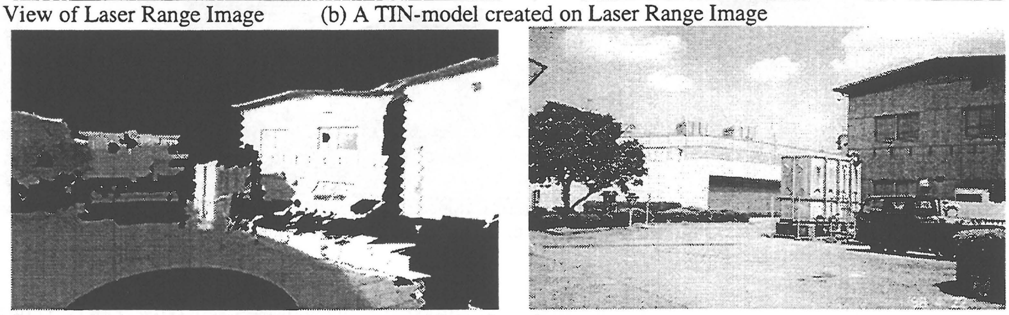
Fig.5 Example of Projection Laser Range Point to the Spherical Image

2.4 A Textured 3D model on a single view

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



(a) A View of Laser Range Image (b) A TIN-model created on Laser Range Image



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

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

3. RECONSTRUCTING TEXTURED 3D MODEL FROM MULTIPLE VIEWS

Two views of laser range images and video 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 registered 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 [2] are employed. There are two steps in the registration process, 1) planar faces from each view are extracted, 2) transformation parameters between different sensor's local coordinate system are estimated by automatically identifying the planar faces' correspondences.

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-based cost function. A planar face regression algorithm is based on *M-estimators* [3] which is rather robust to outlying range points. In the formulation of MDL-based cost function, for each planar face, a histogram created from the residuals of range points is used to approximate the statistical distribution of laser range points. 3) Planar faces which have larger size than a given threshold were picked out for the registration purpose. Fig.7(b),(c) are the planar face extraction result of view1 and view2 respectively. Planar faces are represented by their corresponding laser range points. Laser range points belonging to different planar faces are given a different intensity value.

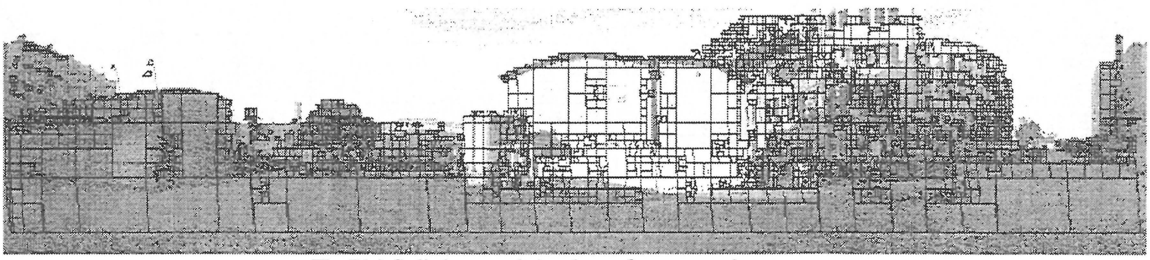


Fig.7(a) Splitting result in planar face extraction process

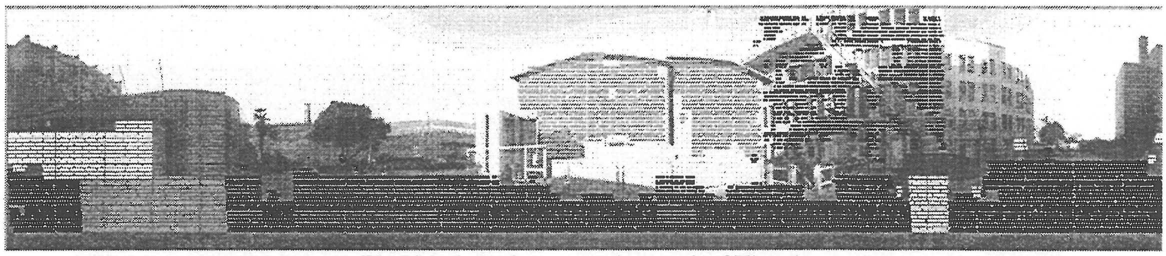


Fig.7(b) planar face extraction result of View 1

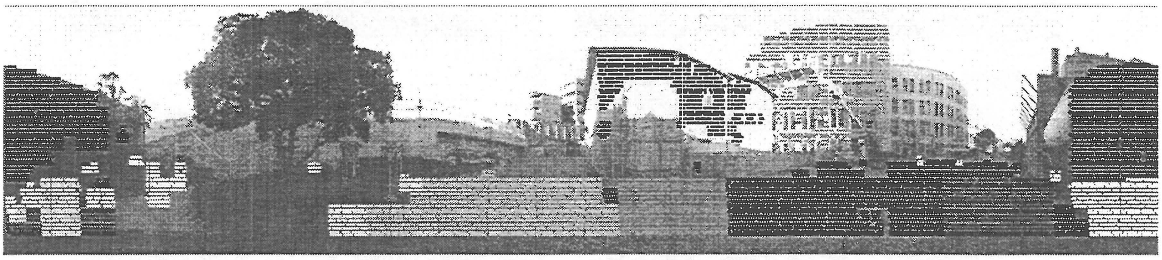


Fig.7(c) planar face extraction result of View 2

3.2 Registration of Two Views

As results of the planar face extraction process, over 20 planar faces were extracted from each views. 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[2] 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 a bone, a thickness is attached to create a "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).

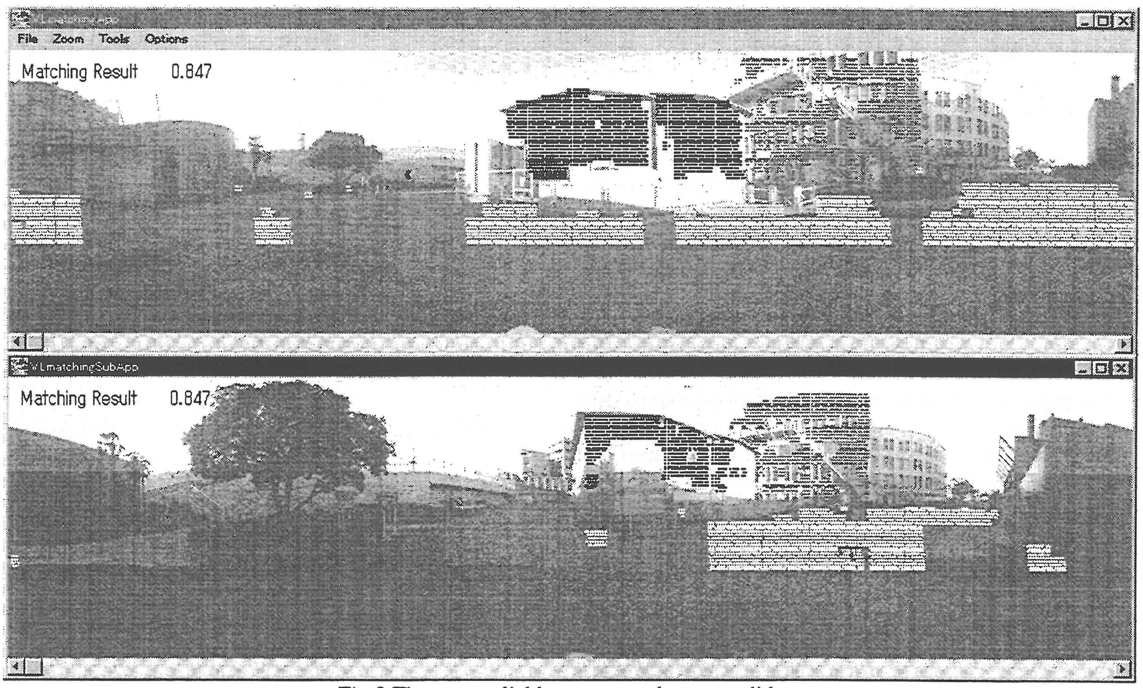
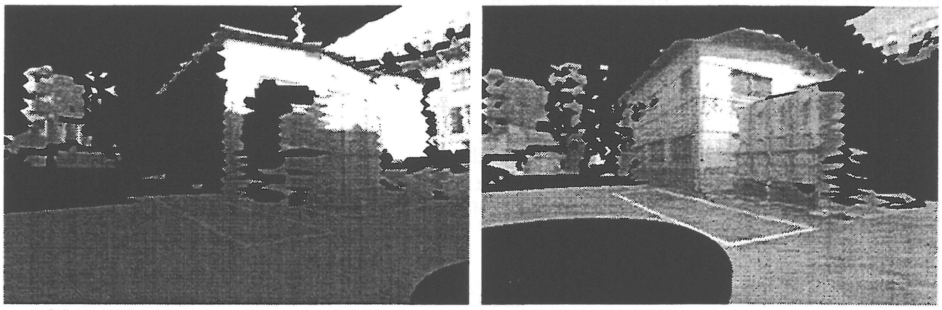
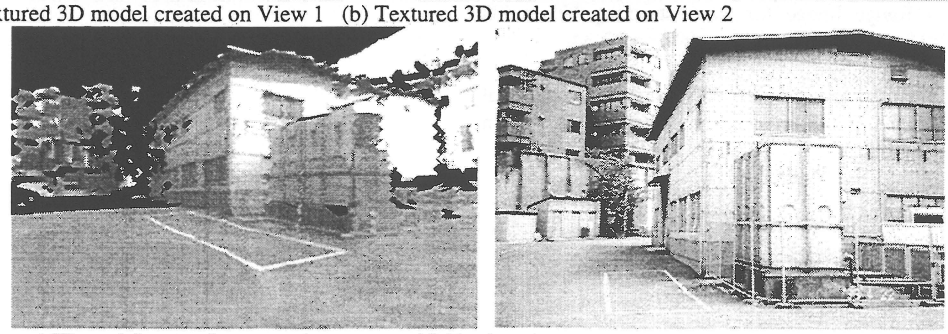


Fig.8 The most reliable correspondence candidate



(a) Textured 3D model created on View 1 (b) Textured 3D model created on View 2



(c) Textured 3D model created on two views (d) Testing site

Fig.9 An example of reconstructing textured 3D model on integrated two views

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 fusing ground-based laser range image and video image to reconstruct textured urban 3D model. An acquisition system was developed, and an outdoor experiment was conducted, where two view of laser range images and video 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 mapping, and showed high accuracy in the planar face extraction and registration result.

Future research will have to be made on finer adjustment using the information for video image, such as edges, feature points, etc. Besides, to improve the accuracy of 3D model by removing outlying points needs to be further studies.

5. REFERENCES

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

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

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.