PAPER *Special Issue on Machine Vision Applications*

# Reconstruction of Textured Urban 3D Model by Fusing Ground-Based Laser Range and CCD Images

**Huijing ZHAO**[†], ***Nonmember and*** **Ryosuke SHIBASAKI**[†], ***Regular Member***

**SUMMARY** In this paper, a method of fusing ground-based laser range image and CCD images for the reconstruction of textured 3D urban object is proposed. An acquisition system is developed to capture laser range image and CCD images simultaneously from the same platform. A registration method is developed using both laser range and CCD images in a coarse-to-fine process. Laser range images are registered with an assumption on sensor's setup, which aims at robustly detecting an initial configuration between the sensor's coordinate system of two views. CCD images are matched to refine the accuracy of the initial transformation, which might be degraded by improper sensor setup, unreliable feature extraction, or limited by low spatial resolution of laser range image. Textured 3D model is generated using planar faces for vertical walls and triangular cells for ground surface, trees and bushes. Through an outdoor experiment of reconstructing a building using six views of laser range and CCD images, it is demonstrated that textured 3D model of urban objects can be generated in an automated manner.

***key words:*** *laser range image, CCD image, registration, textured 3D model, urban outdoor object reconstruction*

## 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 vehicle-borne CCD camera have demonstrated that 3D information can be extracted using motion and stereo vision techniques, e.g. [15],[20]. 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, efficiency and accuracy of the technique has been demonstrated in [11] in obtaining a map of ocean floor; [4],[5],[18] in generating 3D model of small objects such as teeth, status, mechanical parts and so on; [7],[17] in creating textured model of indoor objects; [8] in extracting buildings using air-borne system. 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. The drawback of 3D model generated only using laser range finder is it has poor understandability or interpretability, because it captures only point samples of object surface.

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 data and CCD camera for texture. Since a panorama camera has always a low resolution while large distortion, it is not suitable for accurate 3D object modeling. In this research, a normal CCD camera is exploited. A sensor and acquisition system is developed, where a laser range and a panorama CCD image is measured simultaneously on the same platform, and fused according to sensor's geometry. The panorama CCD image is generated by sequentially assembling CCD image patches captured at different poses as the sensor system doing horizontal and vertical rotation. Since one snapshot can not cover the entire 3D urban object, one of the major challenges of the research is the registration of multiple overlapping laser range and CCD images obtained at different locations (views) into a common coordinate system. In order to find the relative configuration between two overlapping views of laser range and CCD images, a registration method exploiting both laser range and CCD images is developed. An initial transformation between the sensor's coordinate systems of two views is first detected by registering laser range images, then refined through matching CCD images. A textured 3D model is created on the integrated laser range and CCD images, where vertical walls are approximated by planar faces, ground surface and trees are represented using triangular cells.

An experiment is conducted to reconstruct a building in the campus of the Univ. of Tokyo. Six views of laser range and CCD images are measured, and sequentially aligned to a common coordinate system through pair-wise registering neighboring views. A textured 3D model on the integrated six views of laser range and CCD images are constructed. In the followings, sen-

Manuscript received October 4, 1999.
Manuscript revised March 13, 2000.

[†]The authors are with the Faculty of Center for Spatial Information Science, The University of Tokyo, Tokyo, 153–8508 Japan.

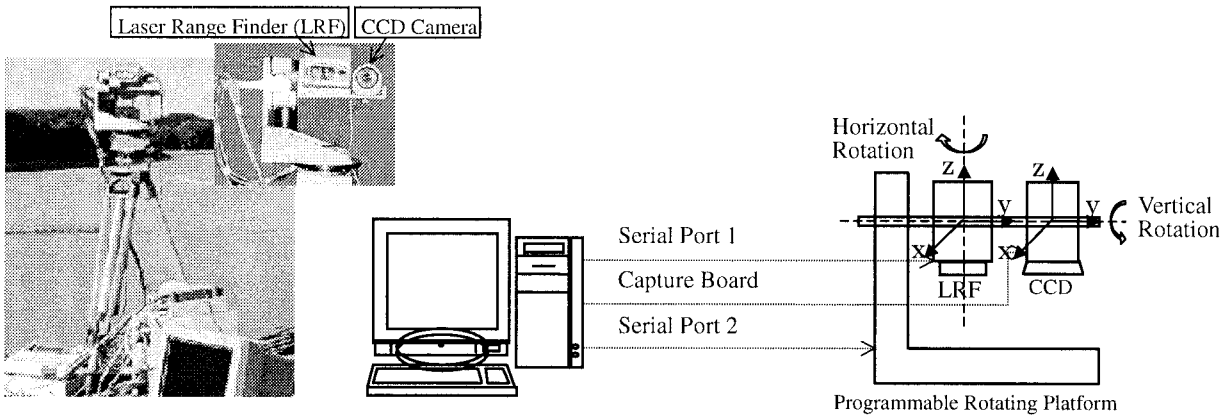



**Fig. 1** Architecture of sensor system.

sor and acquisition system is described in Sect. 2, registration of multiple overlapping laser range and CCD images is addressed in Sect. 3, textured model generation is described in Sect. 4, experimental results are presented in Sect. 4, followed by conclusion in Sect. 5.

## 2. Sensor and Acquisition System

### 2.1 Sensor Alignment

Laser range finder (LRF) is set together with a CCD camera (CCD) on a programmable rotator (Fig. 1). Both LRF and programmable rotator are controlled through serial ports on a PC, while CCD is controlled using a capture board. Alignment of sensors and platform is summarized as follows;

1. Origins in sensor's coordinate systems of LRF and CCD are coincident with their projection centers, where a shift of $r$ along Y-axis exists between them. X-axes and Z-axes of both coordinate systems are parallel, while Y-axes are coaxial.
2. Two rotation modes, vertical and horizontal are available. Vertical rotation axis is coaxial with the Y-axes in the sensor's coordinate systems of both LRF and CCD, while, horizontal rotation axis coincides to the Z-axis of LRF. Hereafter, a rotation is referred as $(h, v)$, where $h$ denotes horizontal rotation angle, and $v$ denotes vertical rotation angle.

### 2.2 Acquisition System

An acquisition system is designed as follows. Sensor system always does horizontal scanning. It changes vertical rotation angle only after the scanning of a horizontal line. Laser range finder keeps capturing distances and angles of 3D points on the surface of target objects, where in a specific rotation pose, only one 3D point is measured. CCD camera captures CCD images patches in an interval of $(\theta, \psi)$ as sensor system doing horizontal and vertical rotation, where $\theta$ denotes the interval in horizontal rotation, $\psi$ denotes the interval in vertical rotation. As the result, a laser range image consisting of 3D points is captured and a panorama CCD image is generated by sequentially assembling the image patches together on a panorama image plane (Fig. 3). Laser range and CCD images captured under the same sensor configuration is called a view. Since a displacement exists from CCD camera to the horizontal rotation axis of sensor system, the origin of CCD camera is shift on a circle around horizontal rotation axis when doing horizontal rotation. Hence CCD image patch taken in different horizontal rotation pose has a different scale (see Fig. 2). To mitigate the image gaps in panorama CCD image, horizontal viewing angles of each CCD image patches, as well as interval in horizontal rotation for CCD image capturing have to be limited. Panorama CCD images used in this research are taken by an interval of $\theta = 3^o$, and $\psi = 30^o$. In addition, to reduce the radical distortion of CCD image, only the central part of the CCD image captured at each pose is used to generate panorama CCD image.

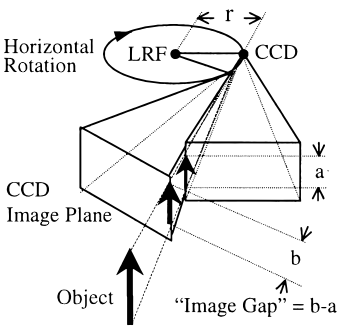



**Fig. 2** Image Gaps between the CCD image patches taken in different horizontal rotation poses.

### 2.3 Fusing Laser Range Image and Panorama CCD Image

In this research, we consider the simplest case in fus-

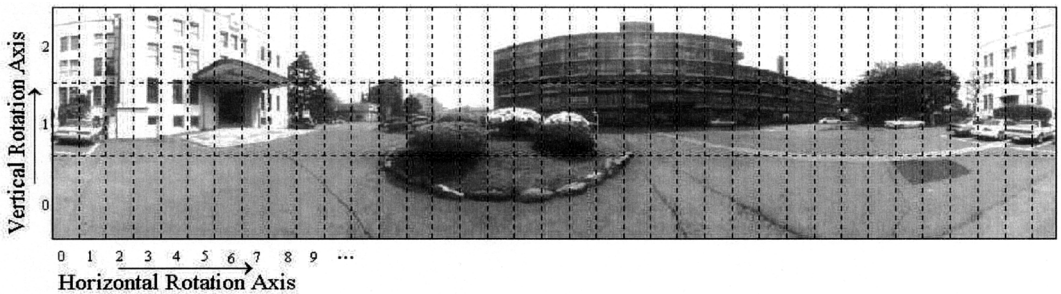
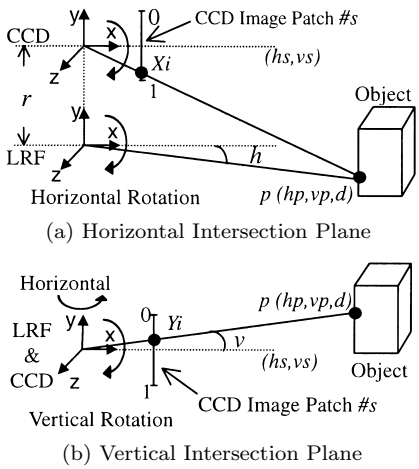



**Fig. 3** Panorama CCD image is generated by sequentially assembling CCD image patches.

**Fig. 4** A laser range image (projected onto panorama CCD image plane).




(a) Horizontal Intersection Plane

(b) Vertical Intersection Plane

**Fig. 5** Projection model from laser range image to CCD image.

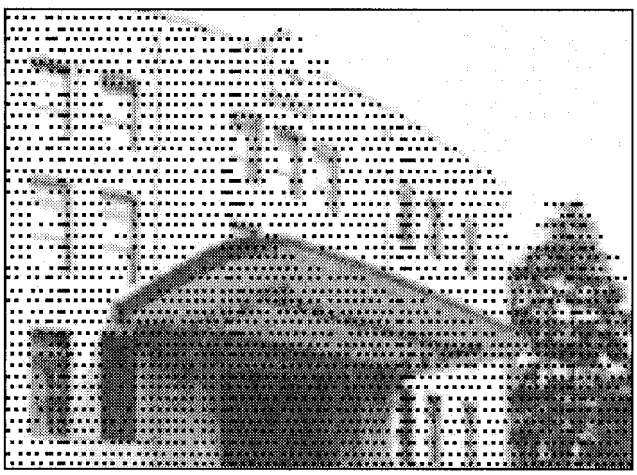
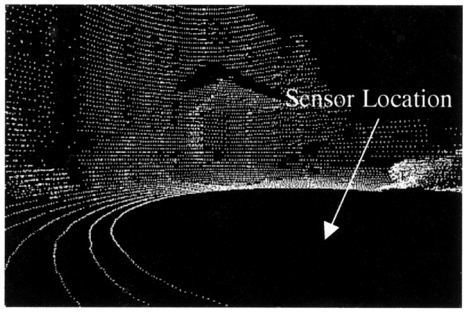


**Fig. 6** Example of projecting laser range points onto the panorama CCD image.

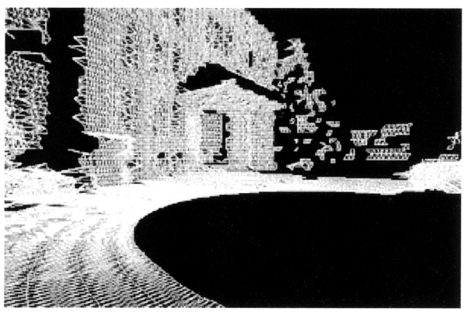
ing a laser range image with a panorama CCD image captured in the same platform, although the method proposed in [7] can be used for a more accurate calibration. It is supposed that sensor system has been aligned exactly the same with the address in Sect. 2.1, and the displacement $r$ between the origins of LRF and CCD is known. Figure 5 depicts the geometry between a laser range point $p$ and the corresponding CCD image patch $\#s$, where $p$ is measured in a rotation of $(h_p, v_p)$ with a distance of $d$, CCD image patch $\#s$ is taken in a rotation of $(h_s, v_s)$. Let $h = h_p - h_s$, $v = v_p - v_s$, and CCD image patches range from 0 to 1 in both width and length. The image point $i(X_i, Y_i)$ of $p$ on CCD image patch $\#s$ can be calculated by Eq. (1). The corresponding CCD image patch $\#s$ is found since only one image patch satisfies both $0 \le X_i < 1$ and $0 \le Y_i < 1$.

$$\begin{aligned} X_i &= \frac{d * \sin h + r}{d * \cos h} + 0.50 \le X_i < 1 \\ Y_i &= 0.5 - \frac{d * \sin v}{d * \cos v} \qquad 0 \le Y_i < 1 \end{aligned} \tag{1}$$

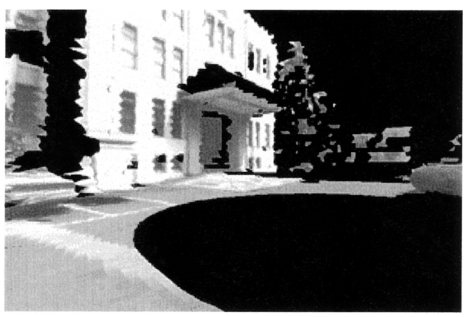
Figure 6 shows an example of projecting laser range points onto the background of panorama CCD image. It can be found that laser range image matches well enough with panorama CCD image. Range values of a laser range image are shown in Fig. 4 by projecting laser range points onto panorama CCD image plane and assigning pixel values to range distances. Linear interpolation is conducted in Fig. 4 to fill the




(a) Laser range image



(b) A TIN-model generated on laser range image



(c) Mapping Panorama CCD Image onto triangular cells

**Fig. 7** A textured 3D model generated on a view of laser range and CCD images.

holes among laser range points. On the other hand, a textured 3D model representation can be achieved by mapping panorama CCD image onto laser range image. Suppose a laser range point $p$ is measured in a rotation of $(h, v)$ with a distance of $d$. Coordinates of $p$ in the sensor's coordinate system of LRF is calculated as $X_p = d\cos v \sin h$, $Y_p = d\cos v \cos h$, $Z_p = d\sin v$. A laser range image can be shown in 3D mode as Fig. 7 (a). A TIN-based textured 3D model is obtained by first creating a TIN-based 3D model using laser range points (Fig. 7 (b)), then mapping panorama CCD image onto triangular cells according to the inverse projection model described in Sect. 2.3 (Fig. 7 (c)).

## 3. Registration of Multiple Overlapping Laser Range and CCD Images

Registration of multiple overlapping laser range and CCD images is typically solved as a two-step procedure, pair-wise registration and multiple registration. Pair-wise registration calculates relative configurations between sensor's coordinate systems of neighboring views, while multiple registration solves error accumulation problem in pair-wise registration to achieve a well-balanced network of views. In the followings, we will focus on pair-wise registration, while the novel multiple registration method developed in this research will be address in a different paper.

### 3.1 Problem Statement and Outline of the Method

Pair-wise registration of range images has been the core of many previous research efforts. In general, they can be divided into two complimentary categories. In the first category, researches invested in recovering a configuration when relative pose of sensor systems are totally unknown. Registration is solved as a correspondence problem, where geometric features such as boundary [16], contour [11], surface [18] etc., or virtual features, such as EGI (Extended Gaussian Image) [12], SAI (Spherical Attribute Image) [9] etc., are exploited. In the second category, researches aimed at refining an initial configuration, which is assumed according to sensor setup or using location modules [17]. Registration is solved as a least square problem, where distance between conjugated points [2],[3],[17], or distance from points to a model [4],[5] is minimized in an iterative mode.

Without assuming sensor's setup, a pair-wise registration of laser range images was investigated in our previous work [21]. An initial transformation between sensor's coordinate system of two views is first obtained using planar faces, then refined through a point-based registration, which minimize the distance from laser range points of one view to a model of another. However robustness of the method is still insufficient to achieve full automation in urban outdoor environment, because enough number of planar faces are difficult to extract, and accuracy of a point-based registration directly working on laser range points might be degraded by erroneous range measurement of urban outdoor objects. Registering laser range images of urban outdoor objects is required to be robust to highly noised laser range measurement, low spatial resolution of laser range points, and always presented occlusion. Noise in laser range measurement in urban outdoor environment comes from not only passing cars, pedestrians, but also window glasses, trees, and so on. Spatial resolution of laser range points reduces exponentially when the object surface becomes far from sensor's location. Occlusion in ground-based laser range measurement is heavy, because of limited vision field and concave object surface.

On the other hand, relative configuration between sensor's coordinate systems can be calculated using CCD images. A great deal of researches studied intensity images for registration purpose [14],[22]. When

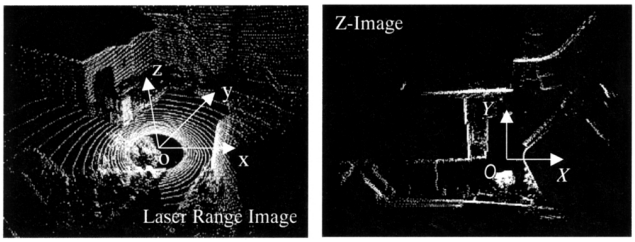



**Fig. 8** Generation of Z-Image from laser range image.

given a sequence of corresponding image points, coplanarity or epipolar condition is always exploited to calibrate the images [14],[22]. However when camera geometry are largely distorted, a good starting point — i.e. an initial transformation between camera coordinate system — is always required to solve the correspondence problem.

In this research, a pair-wise registration method is developed using both laser range and CCD images in a coarse-to-fine process. Laser range images are first registered with an assumption that sensor's platform is always set horizontal to the ground surface. Purpose of the step is to achieve robustness in calculating an initial transformation between two sensor's coordinate systems. Since laser range measurement is a sampling of discrete points on object surface, registration accuracy using laser range image is limited to the spatial resolution of laser range points. Whereas, CCD image captures surface information of spatial objects, and by fusing with laser range image, distortion in CCD image has almost the same order with those in laser range image, which is small enough and neglected in this research. A displacement between CCD images represents a mismatching of laser range images. A fine registration is then conducted using CCD images to refine the result in the registration of laser range image.

### 3.2 Coarse Registration Using Laser Range Image

To measure laser range and CCD images of urban objects, sensor system is located directly on the ground (Fig. 1). Sensor's platform is easy to be set horizontal to the ground surface. A coarse registration method is developed assuming that Z-axis of the sensor coordinate system is parallel to that of the world coordinate system. Unknown transformation parameters are reduced from six - i.e. relative positions $(\Delta x, \Delta y, \Delta z)$ and three rotating angles $(\omega, \psi, \kappa)$ - to four $(\Delta x, \Delta y, \Delta z$ and $\kappa)$. Registration can be achieved in two steps. In the first step, a "Z-image" is generated by projecting laser range points onto a horizontal plane in world coordinate system (Fig. 8). Value of each pixel in Z-image is the number of laser range points falling into the pixel. It is obvious that geometric features in Z-image are easier to detect than those in laser range images. By matching Z-images, a horizontal rotation angle and two translation parameters along X and Y-axes can be obtained. In the second step, elevation of the ground surface is tessellated using the laser range points on ground surface. Translation along Z-axis is obtained by matching the elevation maps of different views.

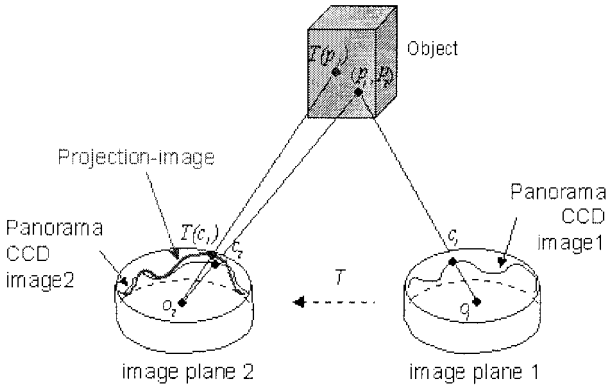



**Fig. 9** Generation of projection-image.

Matching Z-images is quite a two-dimensional problem, it is conducted by first extracting line segments using CFHT (Curve Fitting Hough Transform) [13], then examining the angles between intersecting line pairs and distance between parallel line pairs to find the best matching between Z-images. Two intersecting line pairs of different Z-images, which have the same intersection angles, or two parallel line pairs, which have the same distances between them, are used to find a transformation candidate. The candidate yielding the largest overlay of Z-images is selected as the optimum one. Through matching Z-images, a horizontal rotation angle $(\kappa)$ and two translation parameters in X-Y plane $(\Delta x, \Delta y)$ are obtained. On the other hand, an elevation map of the ground surface can be obtained by projecting laser range points onto a tessellated horizontal plane, where ground height in each pixel is assign to the minimal Z-coordinate of the laser range points falling into the grid cell. A linear interpolation is conducted to find the elevations of the grid cells near to sensor's location. After transforming the laser range images according to the parameters obtained in Z-image matching, there is only a displacement along Z-axis left. The translation parameter along Z-axis $(\Delta z)$ is determined by maximizing the number of the matched grid cells in elevation map.

### 3.3 Fine Registration Using Panorama CCD Image

Since corresponding image points are difficult to catch in the case camera geometry is largely distorted, a "Projection-image" is generated by transforming panorama CCD image1 to the image plane of panorama CCD image2 based on the initial transformation obtained in coarse registration (Fig. 9). A displacement between Projection-image and panorama CCD image2

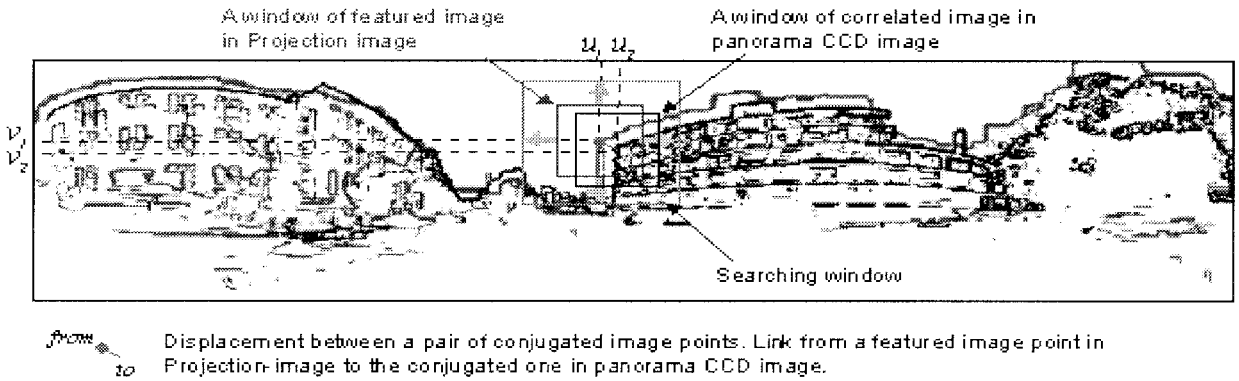



**Fig. 10** Matching image patches through correlation to find corresponding image points.

is caused either by erroneous transformation matrix or occlusion. Corresponding image point pairs between panorama CCD images can be found by correlation of Projection-image with panorama CCD image2, while the initial transformation can be refined by recovering the coplanarity condition between the corresponding image point pairs. Since not only different panorama CCD images, but also CCD image patches of the same panorama CCD image are captured under different illumination conditions, gradient images are matched instead of intensity values to find conjugated image points. It is believed that gradient may sharpen image features and reduce to some extent the affect from the dramatic change of illumination conditions. In this research, gradient is calculated by $g(u,v) = \sqrt{g_u^2 + g_v^2}$, where $g_u$ and $g_v$ are derivative operators. In addition, an adjustment with six degrees of freedom is considered for the sake when sensor's platform is not set exactly horizontal to the ground surface because of mistakes in manual operations, troubles in sensor system and so on.

To find the image points of interest, many researches suggested the use of high curvature points [22]. However in urban outdoor environment, a large mass of high curvature points might be caused by the data of trees, bush, telegram poles, etc., which are occluded and difficult to find reliable corresponding pairs. At the beginning, image points are simply picked at a regular grid. If the window of gradient image centered at the image point is not well featured, the image point is rejected. Given two values, $\lambda_p$ and $\lambda_w$, thresholding featured image points and featured image patches, a well-featured window of gradient image is defined as follows.

$$\frac{\sum_{v=1}^{R}\sum_{u=1}^{C} g(u,v)}{R*C} * N_f > \lambda \tag{2}$$

Where, $R$ and $C$ are the dimension of the window; $g(u,v)$ is the gradient value of at pixel $(u,v)$, $N_f$ is the number of featured image points. It defines a featured image patch to be the one, which has both large number of featured points and high average of gradient values. Featured image points are first selected on Projection-image, then the window of Projection-image centered at the image points are matched with panorama CCD image through correlation to find corresponding image points (Fig. 10). Let $I_1$ be the image patch centered at a featured image point $c_2'(u_1, v_1)$. Sliding $I_1$ on the panorama CCD image within a given searching range, the image patch $I_2$ in panorama CCD image centered at $c_2(u_2, v_2)$, which has the highest coefficient value with $I_1$ is selected as the best matching. Suppose $c_2'(u_1, v_1)$ has its original image point of $c_1(u_1, v_1)$ on panorama image1, $c_2(u_2, v_2)$ and $c_1(u_1, v_1)$ are treated as a pair of conjugated image points.

Suppose the range beam going through image point $c_1(u_1, v_1)$ on the panorama CCD image1 has a horizontal rotation angle $h$ and vertical rotation angle $v$ in the coordinate system of view1, $c_1(x_1, y_1, z_1)$ in sensor's coordinate system of view1 is calculated as $x_1 = \cos v \cos h$, $y_1 = \cos v \sin h$, $z_1 = \sin v$, since scaling factor of image points is free to coplanar condition.

Suppose $o_1$ and $o_2$ are the origins of view1 and view2, $c_1(x_1, y_1, z_1)$ of view1 and $c_2(x_2, y_2, z_2)$ of view2 are a pair of corresponding image points, $T$ is the transformation from the coordinate system of view1 to view2, where $(\Delta x, \Delta y, \Delta z)$ is the translation vector and $(\omega, \psi, \kappa)$ is the rotation angles. By integrating view1 to the coordinate system of view2, $o_2$ has a coordinate at $(\Delta x, \Delta y, \Delta z)$, while $c_1$ is transformed to $c_1'(x_1', y_1', z_1')$ by

$$\begin{pmatrix} x_1' \\ y_1' \\ z_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{pmatrix} \begin{pmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \tag{3}$$

According to the coplanar condition, we have

$$F = \begin{vmatrix} \Delta x & \Delta y & \Delta z \\ x_1' & y_1' & z_1' \\ x_2 & y_2 & z_2 \end{vmatrix} = 0 \quad (4)$$

When given an initial transformation $T_0$ and a set of conjugated image point pairs $\{(c_{1i}, c_{2i}) | 1 \leq i \leq n\}$, a sequence of $F_i(T_0, c_{1i}, c_{2i})$ is obtained, which might be values other than zero because of either erroneous $T_0$ or poor match of $c_{1i}$ and $c_{2i}$. Refining the initial transformation $T_0$ is conducted by recovering the coplanar geometry of $\sum_{i=1}^{n} F_i^2 \to 0$, where minimizing $\sum_{i=1}^{n} F_i^2$ can be achieved using a least square method.

## 4. Textured Model Generation

There are overwhelming redundancy among the laser range images and CCD images of different views. In order to create an integrated model representation, existing studies either re-triangulate range points to exploit all the details covered by range images [19] or merge several vertex into one to generate a simplified model [6],[10]. Since laser range measurement in urban area is highly noised, a model exploiting all laser range points might be distorted by erroneous range data. Still, laser range measurement on window glasses might fail because of mirror reflection of laser beam. A model generation filling the black holes on window area is required. In addition, model topology should be simplify to a desired balance of accuracy and speed. A surface model representation is developed in this research consisting of planar faces and triangular meshes. Planar faces are used to represent the surface of vertical walls, which can be extracted from the Z-image of integrated laser range images as line segments. Laser range points that are not belong to any vertical planar faces are triangulated to represent ground surface, trees, and other free-formed objects. Triangulation of laser range points is based on their neighboring relationships in each view measurement.

Since corner points in each triangular cell belong to the same laser range image, the texture projection model addressed in Sect. 2 can be directly exploited. Whereas, each planar face consist of a mix of laser range points belonging to different views. Re-triangulation of laser range points requires a *zippering* of different panorama CCD images, which is quite time consuming and noise sensitive. A simplified texture projection to planar face is generalized as follows.

1. Tessellating the planar face into square meshes;
2. In each square mesh, finding the laser range image which cast the largest number of laser range points onto the square mesh, the corresponding panorama CCD image is used in texture projection;
3. Interpolating the corners of square mesh to find corresponding position on panorama CCD image.

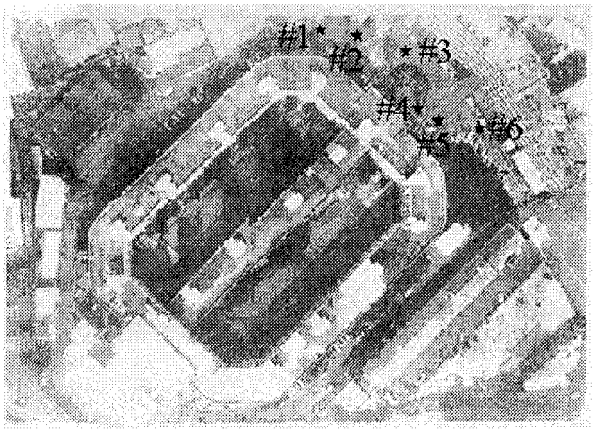
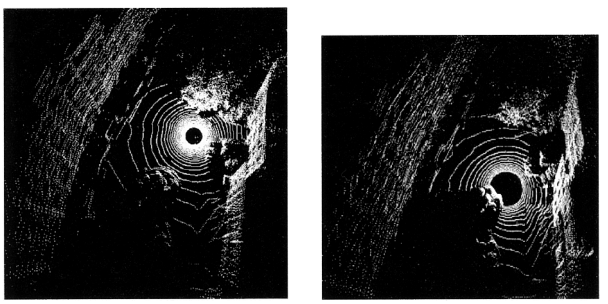



**Fig. 11** Sensor's locations.



(a) Laser range image #4 (b) Laser range image #5

**Fig. 12** Two laser range images.

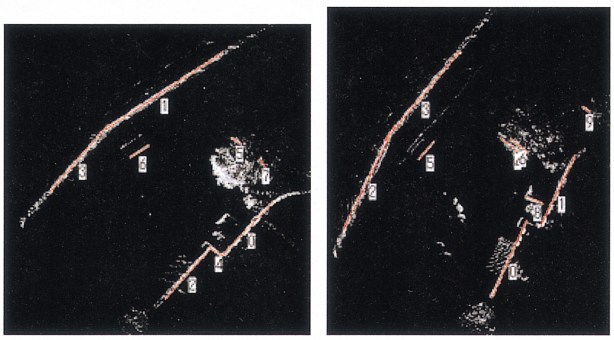
## 5. Experimental Results and Discussion

In this section, we present a set of experimental results in reconstructing a textured 3D model using six views of overlapping laser range and CCD images measured in the campus of University of Tokyo. Sensor's locations are shown in Fig. 11. Laser 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.

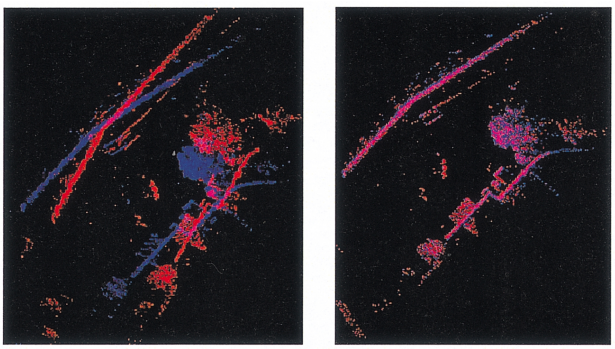
Since some major building surfaces are about 30m far from sensor's location, spatial resolution of laser range points on major building surfaces is about 4 points per square meter. Laser range measurement has an accuracy of 5 cm. Neighboring view pairs (e.g. view #1 and #2, view #2 and #3, etc.) are registered, and six views are integrated to the coordinate system of view1 in a sequential mode. As an example, we will first address the experimental results in registering view #4 and #5, then present the textured 3D model generated by integrating six views of laser range and CCD images.

### 5.1 Coarse Registration of Laser Range Image #4 and #5

Laser range image #4 and #5 are shown in Fig. 12.



(a) (b) Z-Image#4, Z-Image#5 and line extraction result



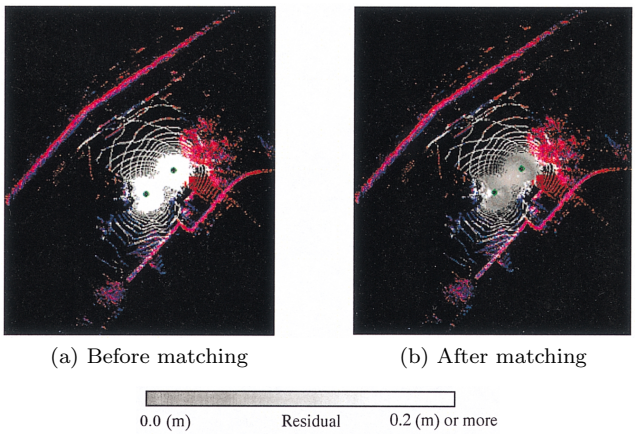
(c) Before Z-image matching (d) After Z-image matching

**Fig. 13** Matching Z-images.

Registration of laser range image #4 and #5 is conducted by first matching Z-images to find a horizontal rotation angle and two translation parameters in horizontal plane, then matching the laser range points on ground surface to find a translation along Z-axis. Figure 13 shows the results in Z-image matching. Line segments are first extracted from Z-images as shown in Figs. 13 (a) and (b). Z-images are matched using the line segment pairs of different Z-images which have the same intersection angles or orthogonal distances. The candidate which raises the largest overlay of Z-image is selected as the result of Z-image matching (Figs. 13 (c) and (d)). Figure 14 shows the result of interpolating and matching the elevation maps of ground surface using the laser range points on ground surface. Matching the elevation of ground surfaces is shown by intensity values, where good matching is shown in grey, poor matching which has a residual equal or larger than 0.2m is shown in white. Circular mark denotes the sensor's location in the acquisition of each view. Figure 15 show the final result in coarse registration of laser range image #4 and #5.

### 5.2 Fine Registration of Panorama CCD Image #4 and #5

Panorama CCD image #4 and #5 are shown in Fig. 16, as well as the Projection-image generated by transforming panorama CCD image #4 to the image plane of




(a) Before matching (b) After matching

**Fig. 14** Result of interpolating and matching the elevation of ground surface.

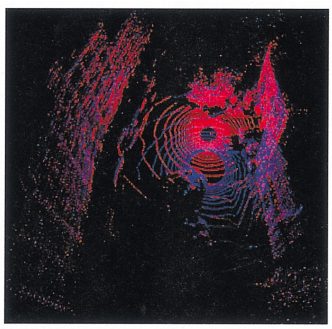
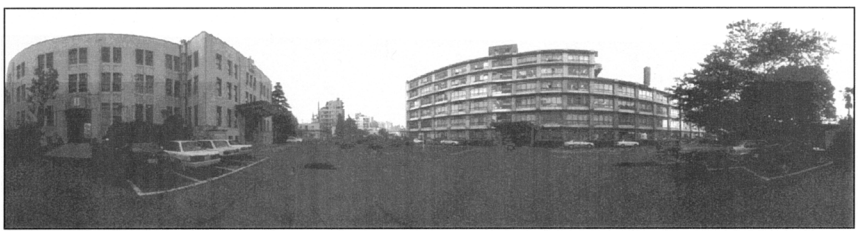


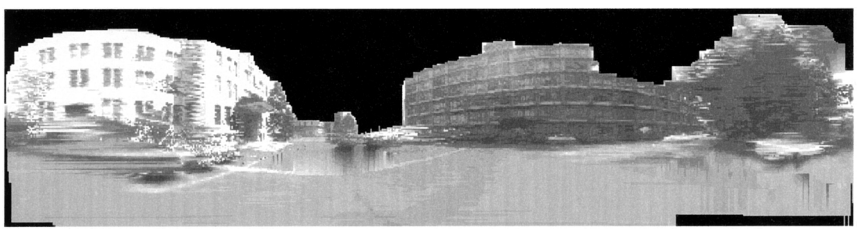
**Fig. 15** Result of registering laser range image #4 and #5.

panorama CCD image #5 based on the transformation parameters obtained in coarse registration. Displacement between the Projection-image and panorama image #5 can be found in Fig. 17 (a), which are demonstrated by overlapping the gradient images. Projection-image is shown in red, while panorama CCD image #5 is shown in blue. Image points of interest are first selected in Projection-image. Image patches in Projection-image centered at the image points of interest are matched on panorama CCD image #5 to find conjugated image points. Displacements between conjugated pairs are marked in the same way that defined in Fig. 10. Longer marks stand for larger displacement between the conjugated pairs. By recovering the coplanarity condition on conjugated image points pairs, the initial transformation obtained in coarse registration are adjusted, and a better matching between Projection-image and panorama CCD image #5 can be found in Fig. 17 (b). Figures 17 (c) and (d) give the enlarged results before and after fine registration. Comparing Figs. 17 (c) and (d), an obvious reduction in both the number and the size of green marks which stand for the displacements between two images can be found. Most of the green marks left in Fig. 17 (d) are either those of poor matchings or short ones.

(a) Panorama CCD image #4



(b) Panorama CCD image #5



(c) A Projection-image generated by transforming Panorama CCD image #4 to the image plane of Panorama CCD image #5 based on an initial transformation obtained in coarse registration

**Fig. 16** A pair of panorama CCD image and the Projection-image generated.

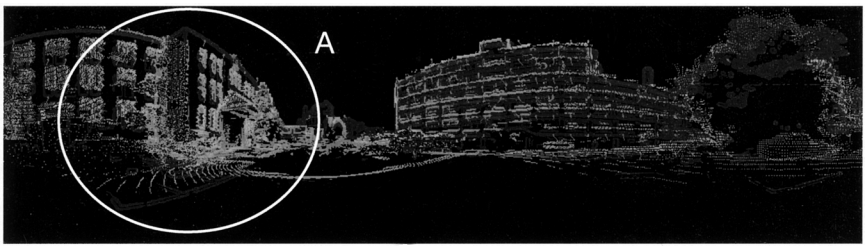
In addition, residuals between laser range images are examined in Fig. 18. A threshold value of 1.0 m is given as the maximum residual between matched laser range points. Before registration, matched laser range points account for 69.5%, while after registration, matched laser range points grow to 80.8%. On the other hand, the average residual of the matched laser range points is reduced from 0.636 m to 0.584 m by registration. Residual between matched laser range points is assigned a gray value, where good matching is shown in white, poor matching with a residual larger than 1.0 m is shown in black. It is obvious that the intensity values of ground surface points become whiter, and geometric feature of the object, such as the front gate of the building, central garden, trees become clearer after registration. This also tells that the residuals between the matched range points become lower after fine registration using panorama CCD image.

### 5.3 Textured 3D Model on Six Views of Laser Range Images and Panorama CCD Images

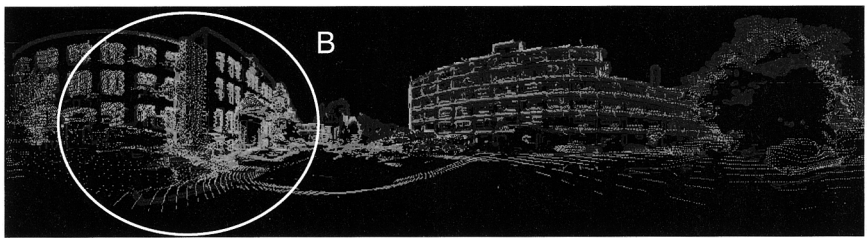
A Z-image is first generated on the integrated laser range images, and line segments representing the surface of vertical walls are extracted from Z-image (Fig. 19 (a)). Figure 19 (b) shows the integrated laser range images, where color points stand for the surface of vertical walls, white points represent ground surface, trees and bushes. Laser range points, which are not belonging to any planar faces, are triangulated. Panorama CCD images are projected onto both planar faces and triangular cells as shown in Fig. 19 (c).

## 6. Conclusion

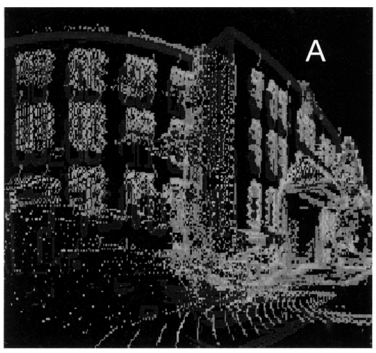
In this paper, a method of fusing ground-based laser range image and CCD image for the reconstruction of textured 3D urban object is proposed. An acquisition system is developed to capture laser range image and CCD images simultaneously from the same platform. A registration method is developed using both laser range and CCD images in a coarse-to-fine process. Laser range images are registered with an assumption on sensor's setup, which aims at robustly detecting an initial configuration between the sensor's coordinate system of two views. CCD images are matched to refine the accuracy of initial transformation, which might be degraded by improper sensor setup, unreliable feature extraction, or limited by low spatial resolution of laser range image. Textured 3D model is generated using planar faces for vertical walls and triangular cells for ground surface, trees and bushes. Through an outdoor experiment of reconstructing a building using six views of laser range and CCD images, it is demonstrated that textured 3D model of urban objects can be generated in an automated manner.

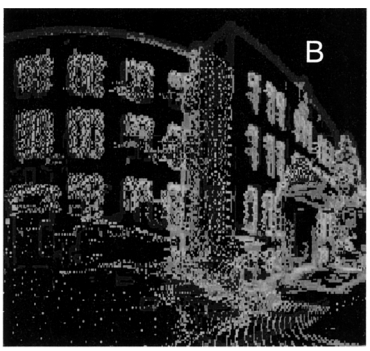



(a) Before fine registration using panorama CCD image



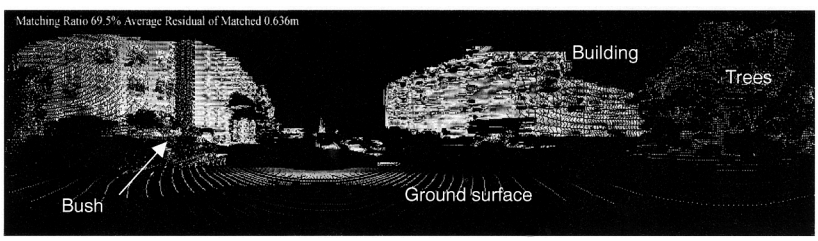

(b) After fine registration using panorama CCD image



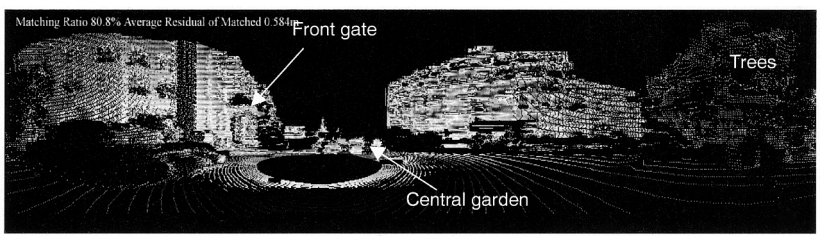




(c) An enlarged image before fine registration (d) An enlarged image after fine registration

**Fig. 17** Overlapping the gradient images of Projection-image with panorama CCD image.




(a) Before fine registration using panorama CCD image (matching ratio is 69.5%, mean of residual is 0.636 m)




(b) After fine registration using panorama CCD image (matching ratio is 80.8%, mean of residual is 0.584 m)

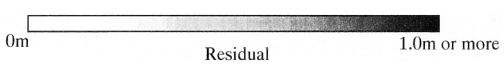
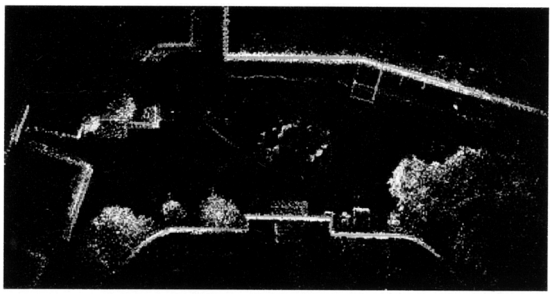



**Fig. 18** Residual between two laser range images.



(a) Z-image of the integrated laser range images and line extraction result

(b) Integrated laser range images

(c) Mapping CCD images onto planar faces and triangular cells

**Fig. 19** Textured 3D model generated on six views of laser range images and panorama CCD images.

Future research will have to be addressed on optimizing texture projection model according to different applications. Besides, developing a method of automatic route designing is required.

**References**

[1] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Towards a general multi-view registration technique," IEEE Trans. Pattern Anal. & Mach. Intell., vol.18, no.5, pp.540–547, 1996.

[2] P.J. Besl and N.D. McKay, "A method for registration of 3-D shape," IEEE Trans. Pattern Anal. & Mach. Intell., vol.14, no.2, pp.239–256, Feb. 1992.

[3] G. Blais and M.D. Levine, "Registering multiview range data to create 3D computer objects," IEEE Pattern Anal. & Mach. Intell., vol.17, no.8, pp.820–824, 1995.

[4] G. Champleboux, S. Lavallee, R. Szeliski, and L. Brunie, "From accurate range imaging sensor calibration to accurate model-based 3-D object localization," Proc. CVPR, pp.83–88, 1992.

[5] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," Image and Vision Computing, vol.10, no.3, pp.145–155, April 1992.

[6] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," SIGGRAPH 96, Computer Graphics Proceedings, 1996.

[7] P. Grandjean, "3-D modeling of indoor scenes by fusing of noisy range and stereo view," IEEE Int. Conf. on Robotics and Automation, pp.681–687, 1989.

[8] N. Haala, C. Brenner, and C. Statter, "An integrated system for urban model generation," IAPRS, vol.XXXII, part 2, pp.96–103, 1998.

[9] K. Higuchi, M. Hebert, and K. Ikeuchi, "Building 3-D models from unregistered range images," Graphical Models and Image Processing, vol.57, no.4, pp.315–333, 1995.

[10] H. Hoppe, T. DeRose, T. Duchamp, and J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," Computer Graphics, vol.26, no.2, pp.71–78, 1992.

[11] B. Kamgar-Parsi, J.L. Jones, and A. Rosenfeld, "Registration of multiple overlapping range image without distinctive features," IEEE Trans. Pattern Anal. & Mach. Intell., vol.13, no.9, pp.857–871, 1991.

[12] R. Krishnapuram and D. Casasent, "Determination of three-dimensional object location and orientation from range images," IEEE Trans. Pattern Anal. & Mach. Intell., vol.11, no.11, pp.1158–1167, 1989.

[13] P. Liang, "A new and efficient transform for curve detection," Journal of Robotic Systems, vol.8, no.6, pp.841–847, 1991.

[14] E. Nishimura, G. Xu, and S. Tsuji, "Motion segmentation and correspondence using epipolar constraint," Proc. 1st Asian Conf. Computer Vision, pp.199–204, 1993.

[15] S. Ozawa, M. Notomi, and H. Zen, "A wide scope modeling to reconstruct urban scene," IAPRS, vol.XXXII, Part 5, pp.370–376, 1998.

[16] B. Parvin and G. Medioni, "A constraint satisfaction network for matching 3D objects," Proc. Int. Conf. on Neural Networks, vol.2, pp.281–286, 1989

[17] V. Sequeira, K. Ng, S. Butterfield, J.G.M. Goncalves, and D. Hogg, "3D textured models of indoor scenes from composite range and video images," SPIE, vol.3313, pp.46–58, 1998.

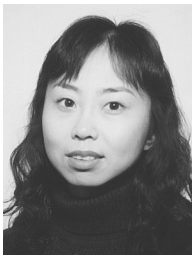
[18] H. Shum, K. Ikeuchi, and R. Reddy, "Virtual reality modeling from a sequence of range images," Proc. IEEE/RSJ Intern. Conf on Intelligent Robots and Systems, pp.703–710, 1994.

[19] M. Soucy and D. Laurendeau, "A general surface approach to the integration of a set of range views," IEEE Trans. Pattern Anal. & Mach. Intell., vol.17, no.4, pp.344–358, 1995.

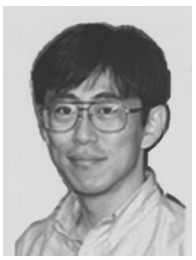
[20] C. Thorpe, M.H. Hebert, T. Kanade, and S.A. Shafer, "Vision and navigation for the Carnegie-Mellon navlab," IEEE Trans. Pattern Anal. & Mach. Intell., vol.10, no.3, pp.363–373, 1988.

[21] H. Zhao, and R. Shibasaki, "Automated registration of ground-based laser range image for reconstructing urban 3D object," IAPRS, vol.32, Part3-4W2, pp.27–34, 1997.

[22] 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.



**Huijing Zhao** obtained B.S in computer science in 1991 from Peking Univ. China. From 1991 to 1994. She was recruited by Peking Univ. in a project of developing a GIS platform. She obtained Ph.D in civil engineering in 1999 from the Univ. of Tokyo, Japan. She is now doing post-doctoral research in Center for Spatial Information Science, the Univ. of Tokyo, Japan.



**Ryosuke Shibasaki** obtained M.S in civil engineering in 1982 from the University of Tokyo. From 1982 to 1988, he worked in Public Works Research Institute, Ministry of Construction. From 1988 to 1991, he was an associate professor of civil engineering department, the Univ. of Tokyo. In 1991, he moved to Institute of Industrial Science, the Univ. of Tokyo. In 1998, he was promoted a professor, Center for Spatial Information Science, the University of Tokyo. His research interest covers 3D data aquisition for GIS, conceptual modeling for spatial objects and agent-based microsimulation in GIS environment.